



Entreculturas 13 (2023) pp. 143-168 — ISSN: 1989-5097

# Análisis de la calidad. El subtulado en vivo interlingüístico de YouTube en un programa de noticias estadounidense accesible a usuarios peruanos

*Quality analysis. YouTube interlingual real time captioning on an American news broadcast accessible for Peruvian users*

Laura Ríos Valero

Universidad de Valladolid

Recibido: 12 de septiembre de 2022

Aceptado: 25 de noviembre de 2022

Publicado: 27 de febrero de 2023

## ABSTRACT

This research aims to analyze the quality of interlingual live subtitling using automatic speech recognition and machine translation provided by YouTube. It is a qualitative descriptive study. In this case study, the instruments applied were an analysis sheet and a questionnaire. The corpus analyzed was an interview in English on mental health due to the COVID-19 pandemic and the population questioned was nine Peruvian users. It was possible to know the accuracy rate and the criteria to establish the quality of interlingual live subtitling by applying the NTR model. The study identified that machine translation generated significantly more errors than automatic speech recognition; these were mostly minor errors. Although the workflow used fell short of quality, participants understood more than half of the information presented.

**KEYWORDS:** quality, live subtitling, NTR model, automatic speech recognition, machine translation.

## RESUMEN

El objetivo de esta investigación fue analizar la calidad del subtulado en vivo interlingüístico, mediante el reconocimiento de voz automático y la traducción automática, ofrecido por YouTube. Tuvo un enfoque cualitativo y nivel descriptivo. El diseño fue el estudio de caso y los instrumentos aplicados fueron una ficha de análisis y un cuestionario. El corpus analizado fue una entrevista en inglés sobre la salud mental con relación a la pandemia de la COVID-19 y la población interrogada fue de nueve usuarios peruanos. Mediante la aplicación del modelo NTR, se logró conocer la tasa de exactitud y los criterios para establecer la calidad del subtulado en vivo interlingüístico. Se determinó que la traducción automática generó muchos más errores que el reconocimiento de voz automático; estos fueron, en su mayoría, errores menores. Aunque el flujo de trabajo no logró tener calidad, los participantes comprendieron más de la mitad de la información.

**PALABRAS CLAVE:** calidad, subtulado en vivo, modelo NTR, reconocimiento de voz automático, traducción automática.

## 1. Introducción

Tener acceso al Internet y a su contenido internacional debería ser un derecho para cualquier usuario en el mundo. El Internet ofrece la capacidad de lograr una comunicación desde cualquier parte del mundo, en cualquier momento, sin límites espaciales, de forma interactiva y en distintas modalidades (Castells, 2014). El formato preferido por la población para acceder a la información es el audiovisual, específicamente el vídeo, ya que se puede retener el 95 % de un mensaje, su acceso es más rápido y conlleva menos tiempo, como afirma Navimedia (s.f.). De acuerdo a Trajectory Partnership (2010), la accesibilidad a la información empodera a las personas y les da autonomía, sensación de tener libertad y control, lo que, a su vez, aumenta su bienestar personal. Esto se ha podido observar sobre todo en las poblaciones de países en vías de desarrollo, personas de bajos recursos o que cuentan con poca preparación académica.

Durante la pandemia de la COVID-19, encontrar información fiable sobre la salud en Internet ha sido de vital importancia no solo por ser una enfermedad de la que no se tenían conocimientos previos, sino también por las consecuencias psicológicas y por los tratamientos que se interrumpieron debido al confinamiento. Dicha información se puede obtener desde las páginas web gubernamentales, prestigiosas universidades de medicina, organizaciones médicas profesionales locales e internacionales, profesionales de la salud, noticias médicas actualizadas, entre otros (National Institute on Aging, 2018).

En el caso de las noticias, el periodismo ha contribuido a que la población adopte comportamientos positivos de higiene y cuidado personal (Puente, 2020). Asimismo, ha permitido que la población se mantenga al día con las últimas actualizaciones de la OMS y que estas lleguen rápida y masivamente a muchas personas. De acuerdo a Tarakini et al. (2021), los medios de comunicación han mejorado la comprensión de la crisis sanitaria por parte de la población, reduciendo la incertidumbre y dando tranquilidad. Por lo tanto, se puede decir que tener acceso a la información fue necesario para que la población mundial lograra estar más protegida durante la pandemia. En el caso de las noticias a través de medios audiovisuales, existe información relevan-

te de medios internacionales a través del Internet, como la que proviene de instituciones reconocidas como los CDC (Centros para el Control y Prevención de Enfermedades) de Estados Unidos o de profesionales de la salud que son referentes en el mundo. Sin embargo, como es natural, dicha información está en otro idioma.

Sería valioso poder acceder a la información provista por medios de comunicación internacionales para que la población esté informada no solo con aquello que los medios locales puedan limitadamente proveer. En el caso de Perú y de la televisión peruana, por ejemplo, Galarza Loayza (2018) indica que la televisión ha perdido su objetivo educativo y que actualmente solo existe la «televisión basura», que presenta morbo, sensacionalismo y escándalo para atraer a la audiencia. En consecuencia, la población peruana, que desee acceder a medios de comunicación internacionales en otro idioma en busca de un mejor contenido informativo, puede hacerlo a través del Internet en páginas como YouTube, en las que podría habilitar el subtítulo automático interlingüístico. Sin embargo, como es sabido, este subtítulo automático no incluye puntuación, lo que hace que el subtítulo se vea desordenado y, por lo tanto, poco comprensible, y a veces omite palabras.

La traducción audiovisual, a través del subtítulo en vivo interlingüístico, puede hacer que ese contenido sea accesible en tiempo real no solo para personas con discapacidad auditiva, sino también para las personas con discapacidad cognitiva o que no conocen el idioma original. Uno de los métodos para lograr un subtítulo en vivo es el rehablado, preferido frente a la estenotipia, ya que es menos costoso (Lambourne, 2007).

Romero-Fresco (2011) define el rehablado como la técnica por la que el rehablador escucha y rehbla el idioma original de un programa, incluyendo la puntuación, a un software de reconocimiento de voz que convierte ese rehablado en subtítulos que se proyectarán en pantalla con el menor retraso posible. Para evaluar la calidad del subtítulo en vivo interlingüístico, Romero-Fresco y Pöchlacker (2017) desarrollaron el modelo NTR que penaliza los errores de traducción y de reconocimiento para encontrar la tasa de exactitud. Además, recoge otros factores como las ediciones por parte del rehablador para lograr una comunicación

más efectiva, la velocidad, el retraso, la coherencia, entre otros. Entre los métodos para lograr el subtítulo interlingüístico en vivo están el rehablado interlingüístico (flujo de trabajo humano), la combinación de intérprete simultáneo y rehablador intralingüístico (flujo de trabajo humano - humano), la combinación de intérprete simultáneo y reconocimiento de voz (flujo de trabajo humano-máquina), la combinación de rehablador intralingüístico y traducción automática (flujo de trabajo humano-máquina) y la combinación de reconocimiento de voz y traducción automática (flujo de trabajo máquina-máquina) (Romero-Fresco y Alonso-Bacigalupe, 2022).

Por lo tanto, al aplicar el modelo NTR y analizar la comprensión y percepción por parte de los usuarios del subtítulo en vivo interlingüístico de YouTube, que usa un flujo de trabajo máquina-máquina, se podría conocer la calidad del mismo, evaluando la tasa de exactitud y otros factores, así como la comprensión y percepción por parte de dichos usuarios.

El presente trabajo de investigación tiene como objetivo analizar la calidad del subtítulo en vivo interlingüístico de YouTube utilizando el modelo NTR y todos los componentes que este incluye, así como analizar la comprensión y percepción por parte de los usuarios de dicho subtítulo. Este trabajo se justifica al estar basado en antecedentes teóricos como el modelo NTR que evalúa la calidad del subtítulo en vivo interlingüístico, al seguir ordenadamente las fases del método científico y, sobre todo, al buscar ser una referencia base para que las distintas plataformas en línea puedan contemplar soluciones accesibles a todos aquellos que no pueden informarse debido a barreras lingüísticas o discapacidades.

## 2. Consideraciones teórico-prácticas

### 2.1. Calidad del subtítulo en vivo y el modelo NTR

De acuerdo a Romero-Fresco (2011), el subtítulo en vivo es un tipo de subtítulo para sordos que usa en progra-

mas en vivo. A través de los años, este se ha podido realizar usando métodos como QWERTY, dual, velotipia, estenotipia y el rehablado (con reconocimiento de voz). Además de que el rehablado resulta ser menos costoso, también puede hacerse con una velocidad de habla de 160-190 palabras por minuto y con un retraso leve logrando una tasa de exactitud alta de 97-98 % (Lambourne, 2007). El subtítulo en vivo podría mostrar la información de sonidos como aplausos y ofrecer identificación del orador.

Entre los elementos que las diferentes guías reguladoras consideran para la calidad del subtítulo para sordos están la exactitud, la equivalencia, la claridad, la consistencia y la facilidad de lectura. Los conceptos de estos elementos son muchas veces similares y agregan alguna particularidad según cada guía.

Como indican Doherty y Kruger (2018), la calidad se relaciona con lograr la tasa de exactitud más alta posible. Esta se podría dar si no existen errores para lograr la equivalencia (Described and Captioned Media Program, s.f.). Es por eso que, en algunos casos, como la Canadian Association of Broadcasters (2012), se prefiere que los subtítulos sean *verbatim*, donde la edición será el último recurso por usar.

Estos subtítulos serán consistentes si son uniformes en identificación del orador, estilo, ubicación, formato, velocidad de visualización en pantalla, entre otros (Canadian Association of Broadcasters, 2012). Doherty y Kruger (2018) indican que se prefieren los subtítulos en bloques estáticos que los que aparecen palabra por palabra, mientras que la ubicación correcta, según la Canadian Radio-television and Telecommunications Commission (2016), sería donde no cubra ninguna información o elemento relevante para entender el mensaje. El Described and Captioned Media Program (s.f.) se refiere a la identificación del orador antes mencionada como la claridad que debe ofrecer el subtítulo.

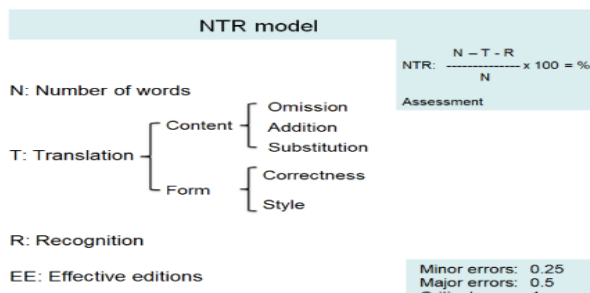
Eugeni (2020) afirma que el menor retraso posible es indicador de calidad. Por lo tanto, el subtítulo debería estar sincronizado con el audio para que su lectura sea fácil. Otros factores que hacen que el subtítulo sea fácil de leer son el tiempo apropiado que se queda en pantalla y cómo este no opaca ni se deja opacar por los elementos visuales (Described and Captioned Media Program, s.f.).

Romero-Fresco y Martínez (2015) incluyeron estos criterios para evaluar la calidad del subtítulado en vivo intralingüístico en el modelo NER, el cual es usado en la programación televisiva de distintos países europeos. Los componentes de este modelo NER son: el número de palabras y signos de puntuación obtenidos en el subtítulado (N), los errores de edición producto de las decisiones del rehablador o de mayúsculas, puntuación e identificación del orador si se trata de subtítulado automático (E), los errores de reconocimiento producto de la falta de reconocimiento por una mala pronunciación del rehablador o del software utilizado que produce los subtítulos (R), las ediciones correctas hechas para omitir información innecesaria que no conlleva a pérdida de información y por último, la evaluación de los resultados obtenidos junto a los factores mencionados en guías reguladoras como la velocidad, el retraso, la identificación del orador, la coherencia entre imagen y sonido, entre otros. Para hallar la tasa de exactitud, debe restarse N - E - R para luego multiplicar y dividir entre cien. Los errores de edición y reconocimiento están clasificados en errores leves, estándares y graves, que son penalizados con puntuación de 0,25, 0,50 y 1, respectivamente. Para poder establecer que existe calidad, la tasa de exactitud debería ser de un mínimo de 98 %.

## 2.2. Modelo NTR

El modelo NTR, realizado por Romero-Fresco y Pöchhacker (2017), se basa en el modelo NER, con la diferencia de que los errores de edición pasaron a ser errores de traducción.

Figura 1: Modelo NTR



Extraído de Romero-Fresco y Pöchhacker (2017)

Al igual que el modelo NER, la tasa de exactitud debe alcanzar un mínimo de 98 %, aunque en el modelo NTR, la calificación ya no es en porcentaje, sino en una escala del 1 al 10, donde el 98 % equivale a 5.

Entre los componentes del modelo NTR de la figura 1 están:

**N:** número de palabras.

**T:** errores de traducción. Estos errores se subdividen en errores de contenido (adiciones, omisiones y sustituciones) y de forma (gramática, terminología, pertinencia, naturalidad, registro).

**R:** errores de reconocimiento. Estos errores se basan en los malos reconocimientos por parte del software usado.

**EE:** ediciones efectivas. No son parte de la fórmula, pero sí de la evaluación integral, ya que se tratan de las ediciones por las que la comunicación se vuelve más efectiva y no conlleva a una pérdida de información.

**Evaluación:** comprende el análisis de la tasa de exactitud, la velocidad, el retraso, los comentarios de las ediciones efectivas, el flujo de subtítulos en pantalla, la identificación del orador, la coherencia entre sonido/imagen y el subtítulo, el tiempo perdido en correcciones y una conclusión final.

En el caso de la clasificación de los errores antes mencionados y su penalización, en el modelo NTR, los errores pasaron a llamarse errores menores, mayores y graves. Los errores menores tienen una puntuación de 0,25 y son los errores que los usuarios pueden distinguir y que no afectan el sentido. Los errores mayores tienen una puntuación de 0,50 y son los errores que los usuarios no pueden distinguir, causando pérdida de información y confusión, pero no crean un nuevo sentido. Los errores graves tienen una puntuación de 1 y son los errores que los usuarios no pueden distinguir porque se crea un nuevo sentido.

## 2.3. Reconocimiento de voz

De acuerdo a Romero-Fresco (2011), el reconocimiento de voz reconoce lo que se dice, no lo que se quiere decir; por lo tanto, no hay inteligencia artificial en los softwares.

Según Lu et al. (2020), el reconocimiento de voz automático convierte la voz en texto y es un área de investigación del procesamiento de lenguaje natural. Esta tecnología emplea tres modelos (Romero-Fresco, 2011). El primero es el modelo acústico, que es un conjunto de datos que tiene material auditivo, la transcripción corregida y la representación del material auditivo en ondas de sonido. El segundo es el modelo gramática/vocabulario/léxico, diccionario, que es la lista de palabras de las que el motor de reconocimiento se sirve para reconocer lo que dice el orador. El tercero es el modelo lingüístico, que genera el texto calculando la probabilidad de una cadena de palabras.

De acuerdo a Doherty y Kruger (2018), los subtítulos automáticos están creados con tecnología de reconocimiento al transcribir el diálogo a texto. Romero-Fresco y Martínez (2015) sostienen que estos sí tienen cierto nivel de calidad, mientras que Jurafsky y Martin (2020) afirman que el nivel de calidad sí puede ser alto si la tecnología de reconocimiento de voz tiene un vasto vocabulario. En el caso de YouTube, los subtítulos están sincronizados con el vídeo. Google provee la tecnología de reconocimiento de voz a YouTube (Harrenstien, 2009), aunque solo alcancen el 80 % de tasa de exactitud (Lockrey, 2015), lo que lo coloca en un nivel subestándar (Romero-Fresco, 2016). YouTube (s.f.), por su parte, justifica la naturaleza variable de su calidad debido a que los subtítulos son creados mediante algoritmos de aprendizaje automático. Además de lo antes mencionado, otros problemas son la segmentación del subtítulo en dos líneas y la velocidad de lectura, cuya solución mejoraría mucho la calidad (Chan et al., 2019).

Por su parte, Jurafsky y Martin (2020) indican que reconocer la voz en una conversación es muy difícil porque ambos interlocutores podrían hablar rápido. Asimismo, sostienen que el ruido, el canal, el acento o las características propias del orador (edad, dialecto no conocido) podrían presentar problemas para el correcto reconocimiento de voz.

## 2.4. Traducción automática

Rao (1998) afirma que la traducción automática es una subdisciplina de la inteligencia artificial que se encarga

de diseñar sistemas que traduzcan de un idioma a otro. Este sistema se compone de un análisis de la lengua origen, la transferencia de una representación interna de la lengua origen y el resultado de esta en la lengua meta. Estos componentes requieren a su vez de mucho léxico y corpus. Dentro de los problemas que tiene la traducción automática está la complejidad y ambigüedad que se genera de los niveles de lenguaje natural: semántico, léxico, sintáctico y pragmático.

Por otro lado, Do (2012) señala que la traducción automática se trata de una subdisciplina de la lingüística computacional que busca automatizar el proceso de traducción de un idioma natural origen a otro idioma meta a través del uso de la computadora. Asimismo, señala que el reto de la traducción automática yace en la forma de programar una computadora para comprender un texto en lengua origen y crear un nuevo texto en lengua meta, tal y como lo haría el hombre.

Abadou y Khadich (2019) indican que la traducción automática se divide en traducción automática fuera de línea (mediante un software instalado en la computadora) y la traducción automática en línea (que requiere conexión a Internet). El traductor automático más popular es Google Translate de Google. Este es un sistema de redes neuronales, por lo que utiliza traducciones que existen en la web para tomar mejores decisiones, también llamado traducción automática estadística en la investigación de Abadou y Khadich (2019). Esto ha hecho que este traductor automático sea uno de los de mejor calidad que existen (Rubio, 2020). Sin embargo, Abadou y Khadich (2019) advierten también que, dado que los sistemas de traducción automática aplican un análisis semántico-sintáctico, esto no es suficiente para asegurar la coherencia del texto y el contexto.

Dentro de las ventajas de la traducción automática que Alsan (2022) menciona están el uso fácil, los resultados rápidos, la confiabilidad (muchos softwares son evaluados por traductores expertos antes de ponerlos en uso), el bajo costo, las operaciones continuas.

De acuerdo con la página web Ulatius (2022), el proceso de aplicación de reconocimiento de voz y traducción automática se realiza primero con la transcripción generada



automáticamente y luego, esta transcripción pasa por la traducción automática, dando un resultado en el idioma meta. Dentro de los beneficios del uso de ambos en los contenidos audiovisuales de las empresas están la rápida accesibilidad para todos, la obtención fácil de contexto para los vídeos, el incremento de marketing y ventas que genera el vídeo, la reducción de gastos y el incremento de la eficiencia, así como el aseguramiento de una ventaja competitiva.

### 3. Metodología

El presente trabajo es un estudio de caso que estuvo basado en el análisis de la calidad mediante la aplicación del modelo NTR del subtítulo en vivo interlingüístico en un programa de noticias del canal en línea CBSN. Este subtítulo se obtuvo mediante el reconocimiento en voz de YouTube y traducción automática de Google.

El programa de noticias emitido el 9 de marzo de 2021 es una entrevista de la presentadora Elaine Quijano al Dr. Roger McIntyre, catedrático de Farmacología y Psiquiatría en la Universidad de Toronto. La entrevista, en el contexto del primer aniversario de la pandemia de la COVID-19, lleva de título «The mental health impacts of the pandemic one year later» (CBS News & Paramount+, 2021). En ella, el Dr. McIntyre conversa sobre los problemas de salud mental, como el estrés y el trauma, que han surgido durante el confinamiento por la pandemia. Por otro lado, si bien la entrevista fue hecha a un especialista, el lenguaje utilizado fue sencillo, pues no hubo términos o frases especializadas de farmacología o psiquiatría. La entrevistadora y el entrevistado eran nativos del inglés, con pronunciación clara y una velocidad de 171 palabras por minuto y nunca se interrumpieron. El vídeo elegido no muestra ruido de fondo y el audio no tiene fallas de sonido.

Dado que el subtítulo en vivo interlingüístico es accesible para diversos usuarios, se aplicó un cuestionario a una muestra intencional de nueve usuarios oyentes que no saben inglés. Estos usuarios fueron hombres y mujeres, de entre 17 y 66 años, que contaban con educación básica completa hasta educación superior universitaria completa (Cuadro 1).

Cuadro 1. Información de participantes

Participante	Sexo	Edad	Nivel educativo alcanzado
1	Femenino	60	Técnica completa
2	Femenino	34	Técnica completa
3	Masculino	26	Universitaria incompleta
4	Femenino	52	Técnica completa
5	Femenino	17	Secundaria completa
6	Masculino	22	Técnica incompleta
7	Femenino	31	Universitaria completa
8	Masculino	66	Universitaria completa
9	Femenino	50	Técnica completa

Con dichos nueve participantes voluntarios, se llevó a cabo una reunión mediante la plataforma Zoom. Se les dio a conocer el propósito de la investigación y se compartió el vídeo de YouTube subtítulo al español. Luego de ver el vídeo, se les dio el enlace de un cuestionario para que puedan llenarlo de manera anónima. Una vez obtenidas las respuestas de los participantes, se analizaron mediante la creación de una hoja de cálculo para localizar y observar las respuestas coincidentes.

El cuestionario se aplicó mediante la herramienta Formularios de Google, en el que se redactaron preguntas sencillas y claras, para que los participantes no tuviesen ninguna dificultad para responder. Dicho cuestionario se dividió en dos partes: la comprensión lectora del subtítulo y la percepción de la calidad de los subtítulos automáticos de YouTube.

En el caso de la comprensión lectora, se redactaron diez preguntas en total, entre abiertas y cerradas, que exploraron los niveles literal (información explícita y básica), inferencial (interpretación y conclusión) y crítico (juicio valorativo) de Cooper (1999). Estas fueron:

- ¿En qué contexto se da la entrevista? (Nivel literal, pregunta cerrada).
- ¿De acuerdo a la entrevista, está de acuerdo con la afirmación del doctor con respecto a que la pandemia fue

- un «experimento de trauma y estrés»? ¿Por qué? (Nivel crítico, pregunta abierta).
- ¿De acuerdo a las declaraciones del doctor, el científico Isaac Newton cometió un delito? (Nivel inferencial, pregunta cerrada).
- ¿Según la entrevista, para quiénes ha sido más difícil estar en casa y hacer todo durante la pandemia? (Nivel literal, pregunta cerrada).
- ¿Está de acuerdo con las recomendaciones del doctor, tales como controlar los tiempos para dormir, hacer ejercicios, así como moderar su consumo de alcohol? ¿Por qué? (Nivel crítico, pregunta abierta).
- ¿El doctor está a favor de las vacunas? (Nivel inferencial, pregunta cerrada).
- ¿Según la entrevista, cuál fue uno de los factores de estrés durante pandemia? (Nivel literal, pregunta cerrada).
- ¿El doctor está a favor de que los colegios hayan vuelto a la presencialidad? (Nivel inferencial, pregunta cerrada).
- ¿Según las declaraciones del doctor, está de acuerdo con sus afirmaciones con respecto a que a la pandemia de la COVID-19 le seguirá una ola de enfermedades mentales? (Nivel crítico, pregunta abierta).
- ¿De acuerdo a la entrevista, qué necesitan las personas con riesgo de tener enfermedades mentales? (Nivel literal, pregunta cerrada).

En el caso de las preguntas de la percepción de calidad de los subtítulos, se redactaron seis preguntas en total, entre abiertas y cerradas, siendo la última pregunta opcional, siguiendo los indicadores que corresponden a la evaluación del modelo NTR (Romero-Fresco y Pöchhacker, 2017), tales como la velocidad, el retraso, el flujo de subtítulos en pantalla, la identificación del orador, la coherencia entre sonido/imagen y el subtítulo. Estas preguntas fueron:

- ¿Es suficiente el tiempo para leer los subtítulos? (Pregunta cerrada).
- ¿Tardó mucho en aparecer el subtítulo en el video? (Pregunta cerrada).

- ¿Logró identificar a los hablantes en los subtítulos? (Pregunta cerrada).
- ¿Había coherencia entre la imagen/el sonido y el subtítulo? (Pregunta cerrada).
- ¿Qué calificación pondría a los subtítulos de YouTube del 0 al 10, donde 0 es muy malo y 10 es muy bueno? (Pregunta abierta).
- Si tiene algún comentario con respecto a la calidad de los subtítulos de YouTube, indíquelo. (Pregunta abierta y opcional).

Por otro lado, el procedimiento de aplicación del modelo NTR comenzó con la habilitación del subtulado en vivo de YouTube en inglés. Una vez que el vídeo mostró en pantalla «inglés (generado automáticamente)», se procedió a extraer el subtulado mediante la página web Downsub en formato SRT. Este archivo fue editado según el discurso del vídeo para obtener la versión original correctamente redactada. En segundo lugar, se configuró la traducción automática al español en el vídeo. Una vez que el vídeo mostró en pantalla «español (generado automáticamente)», se procedió a transcribir el subtulado en Microsoft Word, compuesto por un total de 368 subtítulos. Mediante la herramienta Control de cambios, se procedió a ubicar los errores de traducción y reconocimiento comparándolo con la versión original en inglés. Los errores fueron puestos en unas fichas para iniciar con el análisis. Una vez que se obtuvieron el número total de palabras del discurso original (N), el número de errores de traducción (T) y el número de errores de reconocimiento (R), así como si estos eran errores menores, mayores y graves, se aplicó la fórmula del modelo NTR y se obtuvo la tasa de exactitud.

Dichas fichas de análisis contaron con seis casilleros:

- número de ejemplo;
- tiempo de aparición del subtítulo;
- número de subtítulo;
- extracto del discurso original;
- subtítulo en vivo interlingüístico de YouTube;
- descripción del análisis realizado.

## 4. Resultados y discusión

### 4.1. Número de palabras

El número de palabras más los signos de puntuación de la traducción de los subtítulos en inglés al español fue de 2326.

### 4.2. Errores de traducción

Se encontraron 12 errores graves (12), 52 errores mayores (26,5) y 381 errores menores (95,25). La puntuación total de errores de traducción fue de 133,25. En los cuadros 2, 3 y 4, se presentan los subtítulos más representativos de cada tipo.

Cuadro 2. Errores graves de traducción

Ej.	Detalle
	Tiempo 00:00:45 – 00:00:47
	N.º de Subtítulo 18
1	Discurso Original So, you know, in some ways it is really
	Subtitulado interlingüístico de YouTube para que sepa que, de alguna manera, es realmente
	Descripción Falso sentido por no separar unidades de traducciones con coma «so, you know».
Ej.	Detalle
	Tiempo 00:00:51 – 00:00:53
	N.º de Subtítulo 21
2	Discurso Original the widespread state home orders put in
	Subtitulado interlingüístico de YouTube implementaron las órdenes de residencia estatales generalizadas
	Descripción Falso sentido por traducción literal de «home orders».
Ej.	Detalle
	Tiempo 00:01:12 – 00:01:16
	N.º de Subtítulo 31-32
3	Discurso Original a tremendous degree of stress, I mean we've all signed up
	Subtitulado interlingüístico de YouTube un grado tremendo de estrés. Quiero decir que todos nos hemos inscrito en
	Descripción Falso sentido por traducción literal de expresión «I mean».
Ej.	Detalle
	Tiempo 00:03:31 – 00:03:32
	N.º de Subtítulo 99
4	Discurso Original you know, the she-,
	Subtitulado interlingüístico de YouTube you know they sho
	Descripción Omisión total de traducción del subtítulo.



<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:03:47 – 00:03:52
	N.º de Subtítulo 108-109
<b>5</b>	Discurso Original So, I think, in fact, there needs to be a gift, a gift to ourselves
	Subtitulado interlingüístico de YouTube así que creo que, de hecho, es necesario ser un regalo un regalo para
	Descripción Falso sentido por traducción literal de «to be».
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:06:04 – 00:06:07
	N.º de Subtítulo 177
<b>6</b>	Discurso Original said so accurately, we have had, you know,
	Subtitulado interlingüístico de YouTube decir con tanta precisión, le hemos hecho saber
	Descripción Falso sentido por no separar la expresión «you know» del enunciado.
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:06:28 – 00:06:31
	N.º de Subtítulo 189
<b>7</b>	Discurso Original So, it's a stressor to take away
	Subtitulado interlingüístico de YouTube por lo que es un factor estresante quitarlo
	Descripción Falso sentido por pronombre «lo».
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:09:34 – 00:09:39
	N.º de Subtítulo 281-282
<b>8</b>	Discurso Original 19 come to an end, which we all want, only to be followed by spikes or
	Subtitulado interlingüístico de YouTube 19 encubierto llegue a su fin y todos queremos que solo sigan picos o
	Descripción Falso sentido por falta de separación de unidades de traducción. Además, presenta errores de puntuación (error menor).
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:11:49 – 00:11:52
	N.º de Subtítulo 344
<b>9</b>	Discurso Original etc. That's important, but I think, in fact,
	Subtitulado interlingüístico de YouTube etc.
	Descripción Omisión total de parte de una oración.

Ej.	Detalle
	Tiempo 00:12:33 – 00:12:44
	N.º de Subtítulo 365-366
<b>10</b>	Discurso Original number is 1-800-273-8255. Text “talk” to the number 741-741
	Subtitulado interlingüístico de YouTube número es 1-800-273-8255, envíe un mensaje de texto al número 741-741
	Descripción Falso sentido por omisión de «talk».

**Cuadro 3. Errores mayores de traducción**

Ej.	Detalle
	Tiempo 00:00:21 – 00:00:23
	N.º de Subtítulo 7
<b>1</b>	Discurso Original mental health experts say feeling lower
	Subtitulado interlingüístico de YouTube expertos en salud mental dicen que sentirse más bajo
	Descripción Sin sentido por traducción literal de «lower».

Ej.	Detalle
	Tiempo 00:02:04 – 00:02:07
	N.º de Subtítulo 56
<b>2</b>	Discurso Original thing is for sure going to end.
	Subtitulado interlingüístico de YouTube esto terminará con seguridad con
	Descripción Sin sentido por adición de «con» al final del segmento.

Ej.	Detalle
	Tiempo 00:02:09 – 00:02:11
	N.º de Subtítulo 58
<b>3</b>	Discurso Original negative effect on how we feel so it's a
	Subtitulado interlingüístico de YouTube Un efecto muy negativo en cómo nos sentimos, por
	Descripción Sin sentido por adición de «por» al final del segmento.

Ej.	Detalle
	Tiempo 00:02:12 – 00:02:15
	N.º de Subtítulo 60
<b>4</b>	Discurso Original understandable reaction
	Subtitulado interlingüístico de YouTube bajo. reacción tolerable
	Descripción Sin sentido por adición de «bajo». Además, presenta errores de puntuación (error menor).

Ej.	Detalle
	Tiempo 00:02:31 – 00:02:32
	N.º de Subtítulo 68
5	Discurso Original feeling that? What are some ways that
	Subtitulado interlingüístico de YouTube sentir que cuáles son algunas formas en las
	Descripción Sin sentido por sustitución de signo de interrogación por «que».
Ej.	Detalle
	Tiempo 00:03:15 – 00:03:17
	N.º de Subtítulo 90
6	Discurso Original and I would say that for some people in
	Subtitulado interlingüístico de YouTube un y diría que para algunas personas en
	Descripción Sin sentido por adición de «un» al inicio del segmento.
Ej.	Detalle
	Tiempo 00:03:32 – 00:03:34
	N.º de Subtítulo 100
7	Discurso Original the she-cession, the recession, that's
	Subtitulado interlingüístico de YouTube la sesión de ella la recesión que
	Descripción Sin sentido por falta de conocimiento de término «she-cession». Además, presenta error de puntuación (error menor).
Ej.	Detalle
	Tiempo 00:04:29 – 00:04:32
	N.º de Subtítulo 129
8	Discurso Original so drained, they're so stressed, in many
	Subtitulado interlingüístico de YouTube tan agotada que están tan estresados en muchos casos
	Descripción Sin sentido por sustitución de coma por «que».
Ej.	Detalle
	Tiempo 00:05:52 – 00:05:55
	N.º de Subtítulo 172
9	Discurso Original but as human beings, we are not designed
	Subtitulado interlingüístico de YouTube pero como seres humanos somos no está diseñado
	Descripción Sin sentido por adición de «somos» y mala conjugación del verbo «estar». Además, presenta error de puntuación (error menor).

Ej.	Detalle
	Tiempo 00:07:25 – 00:07:27
	N.º de Subtítulo 213
<b>10</b>	Discurso Original fatigue, they're Zoom-angry, they want to
	Subtitulado interlingüístico de YouTube fatiga del zoom están enojados quieren
	Descripción Sin sentido por omisión de «Zoom» en la segunda unidad de traducción «Zoom-angry». Además, presenta error de puntuación (error menor).

**Cuadro 4. Errores menores de traducción**

Ej.	Detalle
	Tiempo 00:00:08 – 00:00:11
	N.º de Subtítulo 1
<b>1</b>	Discurso Original In Healthwatch, this week marks one year
	Subtitulado interlingüístico de YouTube en healthwatch esta semana se cumple un año
	Descripción Error de forma por falta de mayúscula en «Healthwatch». Este segmento tiene otros errores de forma por falta de mayúscula al inicio de la oración y coma después de «Healthwatch».

Ej.	Detalle
	Tiempo 00:00:11 – 00:00:12
	N.º de Subtítulo 2
<b>2</b>	Discurso Original since the World Health Organization
	Subtitulado interlingüístico de YouTube desde que la organización mundial de la salud
	Descripción Error de forma por falta de mayúscula en «organización mundial de la salud».

Ej.	Detalle
	Tiempo 00:00:40 – 00:00:42
	N.º de Subtítulo 16
<b>3</b>	Discurso Original us. It's great to be with you. Thanks for
	Subtitulado interlingüístico de YouTube nosotros es genial estar con ustedes gracias por
	Descripción Error de forma por falta de punto después de «nosotros». Este segmento tiene otros errores de forma por falta de mayúscula en «es», por falta de punto después de «ustedes» y por falta de mayúscula en «gracias».

<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:00:53 – 00:00:56
	N.º de Subtítulo 22
4	Discurso Original place. There's been such tremendous loss
	Subtitulado interlingüístico de YouTube ,ha habido una pérdida tan tremenda
	Descripción Error de forma por inserción de coma al inicio del segmento y por falta de mayúscula en «ha».
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:01:50 – 00:01:53
	N.º de Subtítulo 50
5	Discurso Original I mean, how much stress do we need? All of
	Subtitulado interlingüístico de YouTube es decir, cuánto estrés necesitamos, todo
	Descripción Error de forma por falta de signo de interrogación antes de «cuánto». Este segmento tiene otros errores de forma por falta de signo de interrogación después de «necesitamos» y por falta de mayúscula en «todo».
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:02:18 – 00:02:20
	N.º de Subtítulo 62
6	Discurso Original concept of “pandemic guilt”,
	Subtitulado interlingüístico de YouTube concepto de culpa pandémica entre comillas
	Descripción Error de forma por falta de comillas en «culpa pandémica». Este segmento tiene otro error de forma por falta de coma al final del segmento.
<b>Ej.</b>	<b>Detalle</b>
	Tiempo 00:03:19 – 00:03:24
	N.º de Subtítulo 93-94
7	Discurso Original For example, people who are single parents. Our female population has
	Subtitulado interlingüístico de YouTube por ejemplo, las personas que son padres solteros, nuestra población femenina
	Descripción Error de forma por falta de punto al final de la primera línea. Este segmento tiene otro error de forma por falta de mayúscula al inicio de la segunda línea.



Ej.	Detalle
	Tiempo 00:04:20 – 00:04:22
	N.º de Subtítulo 124
8	Discurso Original need to respect public health recommendations
	Subtitulado interlingüístico de YouTube debemos respetar las recomendaciones de salud públicas,
	Descripción Error de forma por plural de adjetivo «pública».
Ej.	Detalle
	Tiempo 00:05:26 – 00:05:28
	N.º de Subtítulo 157
9	Discurso Original um, for instance, what are the... some of the
	Subtitulado interlingüístico de YouTube empleo, cuáles son algunos? de l
	Descripción Error de forma por falta de signo de interrogación antes de «cuáles». Este segmento tiene otros errores de forma por ubicación del signo de interrogación de cierre y palabras incompletas.
Ej.	Detalle
	Tiempo 00:06:26 – 00:06:28
	N.º de Subtítulo 188
10	Discurso Original is to be in the company of other people
	Subtitulado interlingüístico de YouTube es estar en el compañía de otras personas,
	Descripción Error de forma por uso erróneo de artículo masculino antes de sustantivo femenino.
Ej.	Detalle
	Tiempo 00:08:41 – 00:08:42
	N.º de Subtítulo 251
11	Discurso Original You mentioned schools.
	Subtitulado interlingüístico de YouTube usted mencionó las escuelas,
	Descripción Error de forma por falta de mayúscula al inicio del segmento. Este segmento presenta otro error de forma por falta de punto al final del segmento.
Ej.	Detalle
	Tiempo 00:09:30 – 00:09:32
	N.º de Subtítulo 279
12	Discurso Original to deal with these issues. The last thing
	Subtitulado interlingüístico de YouTube para tratar con estos problemas, lo último
	Descripción Error de forma por falta de punto después de «problemas» en lugar de coma. Este segmento tiene otro error de forma por falta de mayúscula en «lo».

Ej.	Detalle
	Tiempo 00:10:38 – 00:10:41
	N.º de Subtítulo 313
13	Discurso Original I think the younger population, especially,
	Subtitulado interlingüístico de YouTube y creo que la población más joven, especialmente,
	Descripción Error de estilo por ubicación de adverbio al final del segundo segmento.
Ej.	Detalle
	Tiempo 00:10:49 – 00:10:51
	N.º de Subtítulo 330
14	Discurso Original interview with meghan markle
	Subtitulado interlingüístico de YouTube de meghan markle durante
	Descripción Error de forma por falta de mayúscula en «Meghan». Este segmento tiene otro error de forma por falta de mayúscula en «Markle».
Ej.	Detalle
	Tiempo 00:11:42 – 00:11:44
	N.º de Subtítulo 341
15	Discurso Original have a public conversation about suicide.
	Subtitulado interlingüístico de YouTube tengamos una conversación pública sobre el suicidio
	Descripción Error de forma por falta de punto al final del segmento.

Así como en los estudios de subtitulado en vivo intralingüístico de Romero-Fresco (2016) y Moores (2020), donde se usó reablabado; de Fresno et al. (2020), donde se usó estenotipia; o de Ríos Valero (2022), donde se usó reconocimiento de voz automático, los errores más numerosos fueron los errores leves de edición encontrados a través del modelo NER, de la misma forma, en el presente estudio, los errores más numerosos encontrados a través del modelo NTR fueron los errores menores de traducción. Entre ellos, los errores de forma fueron los más numerosos, generalmente porque en el reconocimiento de voz automático, no se incluye puntuación apropiada, mayúsculas y existen errores morfológicos.

Estos resultados se parecen a los de Sandrelli (2021) y Korybski et al. (2022). En el primero, se utilizó un flujo de trabajo de reablabado intralingüístico con traducción auto-

mática. Esta última causó errores de forma. En el segundo estudio, donde también se usó el mismo flujo de trabajo, se encontraron errores de puntuación.

Los errores mayores de traducción ocupan el segundo lugar, especialmente los errores de contenido por sustitución y, en menor cantidad, por omisión y adición. Las sustituciones, reflejadas en el desorden de palabras o en las palabras cortadas, causaron sin sentidos. Por el contrario, en el estudio de Sandrelli (2020), se encontraron sobre todo omisiones en el flujo de trabajo de reablabado interlingüístico. Por el contrario, en el estudio de Dawson y Romero-Fresco (2021), donde en los flujos de trabajo participaron intérpretes con y sin conocimiento de subtitulado, así como subtituladores con y sin conocimiento de interpretación, hubo más errores de contenido y menos errores de forma.

Al igual que el estudio de Sandrelli (2021), los errores graves de traducción ocuparon el tercer puesto en cantidad, casi exclusivamente por errores de contenido, como omisiones y sustituciones. Estos últimos causaron falsos sentidos.

### 4.3. Errores de reconocimiento

Se encontraron 1 error grave (1), 2 errores mayores (1) y 3 errores menores (0,75). La puntuación total de errores de traducción fue de 2,75. En los cuadros 5, 6 y 7, se presentan los subtítulos encontrados.

Cuadro 5. Errores graves de reconocimiento

Ej.	Detalle
	Tiempo 00:09:13 – 00:09:15
	N.º de Subtítulo 269
	Discurso Original and everyone's gotten a real heavy dose
1	Subtitulado interlingüístico de YouTube y todos han tenido un trabajo realmente pesado se
	Descripción Error de reconocimiento de «dose». El reconocimiento de voz automático reconoció «work» en el inglés, lo que provocó que en la traducción automática al español se coloque «trabajo». Este segmento además tiene un error menor de traducción de forma por selección lexical de «pesado» y un error mayor de traducción de contenido por adición de «se».

Cuadro 6. Errores mayores de reconocimiento

Ej.	Detalle
	Tiempo 00:09:01 – 00:09:04
	N.º de Subtítulo 263
	Discurso Original I think on coming out of this COVID-19, this
1	Subtitulado interlingüístico de YouTube hacer creo que salir de esta codicia 19 esto
	Descripción Error de reconocimiento de «COVID». El reconocimiento de voz automático reconoció «covet» en el inglés, lo que provocó que en la traducción automática al español se coloque «codicia». Este segmento además tiene errores de traducción de forma y contenido, menores y mayores por corrección reflejada en la falta de puntuación y gramática adecuadas, así como por sustitución por sinonimia parcial entre «creer» y «pensar», respectivamente.
Ej.	Detalle
	Tiempo 00:10:18 – 00:10:20
	N.º de Subtítulo 301
	Discurso Original socially close. And thirdly,
2	Subtitulado interlingüístico de YouTube socialmente cerrado y, en tercer lugar,
	Descripción Error de reconocimiento de «close». El reconocimiento de voz automático reconoció «closed» en el inglés, lo que provocó que en la traducción automática al español se coloque «cerrado». Este segmento además tiene errores menores de traducción de forma por corrección reflejada en la falta de puntuación y mayúsculas adecuadas.

Cuadro 7. Errores menores de reconocimiento

Ej.	Detalle
	Tiempo 00:08:18 – 00:08:19
	N.º de Subtítulo 239
	Discurso Original it's actually the younger population
1	Subtitulado interlingüístico de YouTube en realidad es una población más joven a la
	Descripción Error de reconocimiento del artículo definido «the». El reconocimiento de voz automático reconoció el artículo indefinido «a» en el inglés, lo que provocó que en la traducción automática al español se coloque «una». Este segmento además tiene un error de forma por falta de coma después de «realidad».
Ej.	Detalle
	Tiempo 00:08:46 – 00:08:50
	N.º de Subtítulo 255 - 256
	Discurso Original children's social development, obviously their... their educational development
2	Subtitulado interlingüístico de YouTube desarrollo social de los niños, obviamente son su desarrollo educativo
	Descripción Error de reconocimiento de «their». El reconocimiento de voz automático reconoció «they're» en el inglés, lo que provocó que en la traducción automática al español se coloque «(ellos) son».
Ej.	Detalle
	Tiempo 00:12:33 – 00:12:44
	N.º de Subtítulo 365 - 366
	Discurso Original number is 1-800-273-8255 or text <i>talk</i> to the number 741-741
3	Subtitulado interlingüístico de YouTube número es 1-800-273-8255, envíe un mensaje de texto al número 741-741
	Descripción Error de reconocimiento de «or». El reconocimiento de voz automático no reconoció la conjunción en inglés, lo que provocó que en la traducción automática esta sea sustituida por una coma.

Con respecto a la cantidad de errores de reconocimiento, en primer lugar, se encontraron errores menores; en segundo lugar, errores mayores; en tercer lugar, errores graves. Este orden se evidencia, de la misma forma, en el estudio de Dawson y Romero-Fresco (2021). A comparación de los errores de traducción, los errores de reconoci-

miento son muy pocos, como cuando se tiene una pronunciación poco clara como en el estudio de Sandrelli (2021).

Como señalan Levis y Suvorov (2012) y Jurafsky y Martin (2020), dado que a YouTube le falta incorporar más vocabulario, esto hizo que no reconociera correctamente neologismos como «she-cession».

#### 4.4. Resumen de errores de traducción y reconocimiento

En el cuadro 8 se reúnen todos los errores y su puntuación según el modelo NTR.

Se observa que, en gran mayoría, hay muchos más errores menores, mayores y graves de traducción que de reconocimiento. Los errores más numerosos en este caso han sido los errores menores de traducción. Tras aplicar la fórmula del modelo NTR, se obtuvo una tasa de exactitud de 94,13 %.

$$\text{Exactitud} = \frac{2\ 326 - 133,75 - 2,75}{2\ 326} \times 100 = 94,13\ \% = 0/10$$

De acuerdo a Romero-Fresco (2016), el nivel de exactitud es subestándar. Si los errores menores de traducción no hubiesen ocurrido en el subtítulo en vivo, la tasa de exactitud hubiese alcanzado el nivel mínimo de calidad. Se pueden mencionar otros estudios como el de Sandrelli (2021), en el que se logró una tasa de 94 % de exactitud mediante rehabladores interlingüísticos, o el de Robert y Re-

mael (2017), en el que solo uno de los cuatro episodios del corpus analizado alcanzó el 98 %. En este último estudio, se utilizó una adaptación del modelo NER para analizar el subtítulo en vivo interlingüístico mediante un intérprete rehablador y un corrector. Romero-Fresco y Alonso-Bacigalupe (2022) obtuvieron una tasa de exactitud de 97,2 % (3/10) cuando analizaron el flujo de trabajo de reconocimiento de voz automático con traducción automática. Si bien dicho flujo logró la peor tasa de exactitud, fue el menos costoso y el que ofrecía menor retraso que todos los demás flujos. Por último, solo el estudio de Korybski et al. (2022) logró más del 98 % de tasa de exactitud al usar rehablado intralingüístico y traducción automática, con lo que se observa que sí es posible llegar a la tasa mínima en subtítulo en vivo interlingüístico.

#### 4.5. Ediciones efectivas

Se encontraron solo 2 ediciones efectivas, las cuales se evidencian en el cuadro 9.

Cuadro 8. Resumen del total de errores

Tipo de error		Errores de traducción	Errores de reconocimiento	Total de errores y puntuación
<b>Graves</b>	Nº	12	1	13
	Puntuación	12	1	13
<b>Mayores</b>	Nº	52	2	54
	Puntuación	26	1	27
<b>Menores</b>	Nº	381	3	384
	Puntuación	95,25	0,75	96
<b>Puntuación Total</b>		133,25	2,75	136

Cuadro 9. Ediciones efectivas

Ej.	Detalle
	Tiempo 00:01:09 - 00:01:11
	N.º de Subtítulo 29
<b>1</b>	Discurso Original Well, I think people should be aware of
	Subtitulado interlingüístico de YouTube ? Creo que la gente debería ser consciente
	Descripción Edición efectiva por omisión de interjección innecesaria «well» en la traducción. Este segmento también tiene un error de forma menor por inserción de signo de interrogación al inicio del segmento.



Ej.	Detalle
Tiempo	00:11:04 – 00:11:07
N.º de Subtítulo	324
Discurso Original	Well, it reminds us that suicidal
<b>2</b> Subtitulado interlingüístico de YouTube	esto nos recuerda que el
Descripción	Edición efectiva por omisión de interjección innecesaria «well» en la traducción. Este segmento también tiene errores de forma menor por falta de punto después de «esto» y falta de mayúscula en «nos».

El subtitulado en vivo de YouTube es *verbatim* y dado que se hace de manera automática, no se autoedita. Para considerar las ediciones, el subtitulado en vivo no debería ser *verbatim* (Fresno et al., 2020). Por eso, en este caso, se han considerado como ediciones efectivas a aquellas omisiones que no generaron pérdida del mensaje.

## 4.6. Evaluación

De acuerdo a Romero-Fresco y Pöchhacker (2017), la evaluación se basa en una conclusión final, la cual no se trata solo de la tasa de exactitud analizada, sino también de los otros aspectos de calidad propios del subtitulado.

### 4.6.1. Velocidad

La velocidad de los oradores en la entrevista fue de 171 palabras por minuto. A diferencia de los estudios de Sandrelli (2020, 2021) y Romero-Fresco y Alonso-Bacigalupe (2022), la velocidad de los oradores de la entrevista analizada fue mayor. A pesar de ser una entrevista y de que existía la posibilidad de que las voces de los oradores se sobrepongan una con la otra, esto no ocurrió, ya que dicha entrevista se realizó por una plataforma de reuniones, lo que hizo que ambas partes respetaran su turno. Por lo mismo, el reconocimiento de voz automático logró reconocer prácticamente toda la entrevista, aunque sí presentó algunos errores.

### 4.6.2. Retraso y flujo general de subtítulos

El reconocimiento de voz automático y la traducción automática de YouTube no generaron retraso dada su naturaleza automática, la falta de edición y/o corrección de los subtítulos y, además, porque el subtitulado apareció palabra por palabra en la pantalla. Solo en minuto 06:56, hubo un retraso de 4 segundos, pero, a pesar de eso, el subtitulado continuó siendo *verbatim*. Al no tener retraso, podría tener mayor aceptación por la audiencia, ya que, según Mikul (2014), los sordos, por ejemplo, se quejan mucho del retraso. Por lo mismo, siempre hubo coherencia entre la imagen y el subtitulado en el presente estudio.

Sin embargo, esta misma falta de retraso hizo que el subtitulado desapareciera rápido, causando así problemas para la lectura. La Asociación Española de Normalización (2012) y Ofcom (2005) indican que, en el caso de subtitulado para sordos, el máximo de caracteres por segundo debería ser de 15 cps (caracteres por segundo), mientras que Karakanta et al. (2021) mencionan que el límite máximo de un bloque de subtítulos de palabra por palabra debería tener un límite máximo de 84 caracteres. En este estudio, apenas se cubre, y solo en pocos casos, la condición del número máximo de caracteres por bloque de dos líneas. El bloque más corto tiene 51 caracteres con 17 cps y el bloque más largo tiene 115 caracteres con 16 cps.

De acuerdo a Romero-Fresco (2010), en el subtitulado que se visualiza como palabra por palabra, las personas

leen el 90 % del tiempo y solo ven el 10 % de la imagen, lo que podría afectar su comprensión, aunque esta forma tenga menor retraso que un bloque estático (Karakanta et al., 2021). De la misma forma, en los estudios de Romero-Fresco (2020) y Karakanta et al. (2021), se observa también que, en los flujos de trabajo donde hay intervención de la máquina, el retraso no supera los 4 segundos, a pesar de que la calidad de la traducción sea pobre.

#### 4.6.3. Identificación del orador

En el caso del presente estudio, no se identificó a los oradores. La Asociación Española de Normalización (2012) indica que se debe identificar a los oradores mediante colores, etiquetas o guiones, pero es importante que se haga para facilitar la trama del vídeo.

#### 4.6.4. Comprensión y percepción de los participantes

Con respecto a la comprensión, los participantes respondieron correctamente seis de diez preguntas (60 %). A nivel literal, los participantes respondieron correctamente dos de cuatro preguntas, las cuales trataban del contexto del primer año de la pandemia COVID-19 y de la pérdida del contacto social como uno de los factores de estrés durante la pandemia. A nivel inferencial, los participantes respondieron correctamente dos de tres preguntas, las cuales trataban de la posición del doctor sobre las vacunas y sobre el regreso a la presencialidad de las escuelas. A nivel crítico, los participantes respondieron, coherentemente y con base en la información de la entrevista, dos de tres preguntas, las cuales trataban sobre el experimento de trauma y estrés que fue la pandemia COVID-19 y la posibilidad de una ola de enfermedades mentales pospandemia.

Con respecto a la percepción, los participantes aprobaron la calidad del subtítulo en vivo interlingüístico de YouTube al 50 %. Los participantes calificaron positivamente dos criterios. El primero es que los subtítulos no se retrasaron o se retrasaron un poco en aparecer en pantalla. El segundo es que sí hubo o a veces hubo coherencia entre la imagen y los subtítulos. Los participantes calificaron

negativamente tres criterios: la falta de tiempo para leer los subtítulos, la falta de identificación de los oradores y la ubicación de los subtítulos. Algunos de los comentarios que dejaron los participantes fueron:

«Se debe mejorar la posición, el tamaño del texto, para que se pueda leer completo». (Participante 2)

«La forma de YouTube se presta a la confusión, también hay partes sin sentido. [...] Además, hay palabras que están incompletas». (Participante 9)

Sin embargo, en la última pregunta en la que los participantes tenían que colocar una nota del 0 al 10 al subtítulo en vivo interlingüístico de YouTube, la nota promedio fue de 7 puntos.

## 5. Conclusiones

Dado el escaso conocimiento con respecto al estudio del subtítulo en vivo, esta investigación contribuye con el aporte de resultados de un caso, cuyo flujo de trabajo fue enteramente automático para el subtítulo en vivo interlingüístico, donde la calidad es más difícil de lograr que en el intralingüístico, ya que existe un cambio de idioma. Asimismo, este estudio incluyó la recepción de participantes para evaluar la calidad del subtítulo desde su punto de vista. Asimismo, cabe mencionar que, en el proyecto inicial del presente estudio, se contemplaba trabajar también con un flujo de trabajo con intervención humana. Sin embargo, dicha idea no prosperó porque no se encontraron rehabilitadores interlingüísticos en la región latinoamericana, ni se tuvo acceso a rehabilitadores interlingüísticos de otras partes del mundo.

En el caso de la tasa de exactitud del subtítulo en vivo interlingüístico mediante el reconocimiento de voz automático y la traducción automática de YouTube, este estudio alcanzó el 94,13 %. De acuerdo a Romero-Fresco y Pöschhacker (2017), todas las tasas de exactitud menores a 96 % equivalen a 0/10 y, por lo tanto, calificarían como subestándar, como en el modelo NER de Romero-Fresco y Martínez (2015). Sin embargo, dado que dicha tasa sobrepasa

el 90 %, se podría afirmar que, aunque no se cumplan los criterios mínimos de calidad, el reconocimiento de voz automático y la traducción automática sí servirían para que el usuario tenga una comprensión general del video.

Los errores de traducción fueron mucho más numerosos que los errores de reconocimiento. Entre los errores de traducción, los más numerosos son los errores menores y son en su mayoría, por errores de forma por corrección y, en segundo lugar, están los errores de contenido por sustitución. La traducción automática al español sí incluía signos de puntuación, aunque no siempre utilizados de forma correcta. Este uso inapropiado de signos de puntuación estaría asociado a las pausas realizadas por los oradores en el habla espontánea. El presente estudio de caso no intenta generalizar, pero se evidencia que la traducción automática, a pesar de sus muestras de haber alcanzado una mejor calidad (Rubio, 2020), no tendría la suficiente calidad para traducir contenidos orales, pues esta ha sido el mayor problema para lograr la tasa de calidad esperada.

Los errores de reconocimiento fueron muy escasos. De igual forma, los más numerosos fueron los errores menores. Se observó que, básicamente, la baja tasa de exactitud alcanzada en este flujo de trabajo se debió más a la traducción automática y no tanto al reconocimiento de voz automático. Esto supondría que, a diferencia de las bajas tasas de exactitud que señala la página web de University of Minnesota Duluth (s.f.), el reconocimiento de voz automático de YouTube sí sería de buena calidad, como el caso de la tasa de 95 % de exactitud de Lockrey (2015).

Dada la naturaleza automática del reconocimiento de voz y la traducción automática, no hubo ediciones efectivas, por lo que, en este flujo de trabajo, se podría considerar de esa forma a aquellas omisiones que hicieron que la comunicación sea más asertiva. En este estudio, se omitió dos veces una interjección y, por lo tanto, no se generaría pérdida de información relevante.

Con respecto a la evaluación de la calidad, a pesar de la velocidad de la entrevista, el reconocimiento de voz automático fue prácticamente completo porque los oradores no se interrumpieron. Por esa misma razón, el subtítulo fue *verbatim* y por ello, a los usuarios les resultó difícil

leer todos los subtítulos porque pasaban muy rápido. En el estudio de Romero-Fresco y Alonso-Bacigalupe (2022), el flujo de reconocimiento de voz automático y traducción automática alcanzó una tasa de precisión de 97,2 % (3/10), el cual fue el peor de los flujos, pero también fue el que presentaba el menor retraso y el que era menos costoso. En el presente estudio, solo hubo un retraso de 4 segundos en un segmento, mientras que el resto de la entrevista no lo tuvo. Por lo tanto, el retraso sería un factor importante para tomar en cuenta en la evaluación de calidad, tanto que fue calificado como positivo por el grupo de usuarios, ya que eso también les permitió encontrar coherencia entre el subtítulo y la imagen.

Con respecto a la identificación de los oradores, YouTube no la tiene ni en el reconocimiento de voz automático, ni en la traducción automática. Asimismo, YouTube coloca el subtítulo en la parte inferior central del vídeo, el cual no se mueve a no ser que el usuario mueva el cursor sobre el vídeo. Dicha identificación de los oradores, así como la ubicación de subtítulos, son prueba de mala calidad del subtítulo en vivo interlingüístico de YouTube, puesto que causaron confusión a los participantes. Estos no sabían a quién correspondían las declaraciones y porque hubo momentos en los que los subtítulos cubrían las cintas informativas de la entrevista que se colocan en la parte inferior.

A pesar de que la tasa de exactitud fue subestándar, los participantes aprobaron la calidad del subtítulo con una nota de 7 (10 era la nota máxima) porque comprendieron más del 50 % de la información presentada. Se observa entonces que los estudios con mejores resultados son aquellos donde el flujo de trabajo incluye al hombre y a la máquina (Eugeni, 2020), precisamente porque cuando solo hay actividad de la máquina, surgen más errores de forma y, en el caso de los errores de contenido, estos se deben más a sustituciones. Cuando solo hay actividad del hombre, surgen más errores de contenido, sobre todo por omisiones. La combinación de hombre y máquina podría ser la solución para lidiar con el retraso y los errores de sentido, ya que, según Romero-Fresco (2020), es difícil lograr ambos. Quizás los estudios que incluyan comparaciones de

flujos de trabajo hombre-máquina y máquina-máquina podrían ofrecer datos para encontrar cuál sería la forma ideal para realizar el subtítulo en vivo interlingüístico.

## Bibliografía

- Abadou, Fadila y Khadich, Saleh (2019). Coherence in machine translation output. *Traduction et Langues*, 18(2), 138-153.
- Alsan, Merve (12 de junio de 2022). *The best machine translation software you can try in 2022*. Weglot. <https://weglot.com/blog/machine-translation-software/>
- Asociación Española de Normalización. (2012). *Subtitulado para personas sordas y personas con discapacidad auditiva (UNE 153010:2012)*. [www.une.org](http://www.une.org)
- Canadian Association of Broadcasters. (2012). *Closed Captioning Standards and Protocol for Canadian English Language Television Programming Services*. [www.cab-acr.ca](http://www.cab-acr.ca)
- Canadian Radio-television and Telecommunications Commission. (2016). *English-language closed captioning quality standard related to the accuracy rate for live programming*.
- Castells, Manuel (2014). El impacto de internet en la sociedad: una perspectiva global. In BBVA (Ed.), *C@mbio 19 ensayos fundamentales sobre cómo internet está cambiando nuestras vidas*. OpenMind/BBVA.
- CBS News, & Paramount+. (9 de marzo 2021). *The mental health impacts of the pandemic one year later*. YouTube. <https://www.youtube.com/watch?v=p2gzCJhcuT8>
- Chan, Wing Shan, Kruger, Jan-Louis y Doherty, Stephen (2019). Comparing the impact of automatically generated and corrected subtitles on cognitive load and learning in a first and second-language educational context. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 18, 237-272.
- Cooper, David (1999). *Cómo mejorar la comprensión lectora*. Editorial Visor.
- Dawson, Hayley y Romero-Fresco, Pablo (2021). Towards research-informed training in interlingual respeaking: an empirical approach. *The Interpreter and Translator Trainer*, 15(1), 66-84.
- Described and Captioned Media Program. (s.f.). *Captioning Key*. [https://www.captioningkey.org/about\\_c.html#1](https://www.captioningkey.org/about_c.html#1)
- Do, Thi Ngoc Diep (2012). *Extraction de corpus parallèle pour la traduction automatique* [Tesis doctoral no publicada]. Université de Grenoble
- Doherty, Stephen y Kruger, Jan-Louis (2018). Assessing Quality in Human- and Machine-Generated Subtitles and Captions. En J. Moorkens, S. Castilho, F. Gaspari, S. Doherty (eds), *Translation Quality Assessment. Machine Translation: Technologies and Applications* (pp. 179-197). Springer, Cham.
- Eugeni, Carlo (2020). Respeaking: aspects techniques, professionnels et linguistiques du sous-titrage en direct. *ESSACHESS-Journal for Communication Studies*, 13(25), 21-35.
- Fresno, Nazaret, Sepielak, Katarzyna y Krawczyk, Maciej (2020). Football for all: the quality of the live closed captioning in the Super Bowl LII. *Universal Access in the Information Society*, 1-12.
- Galarza Loayza, Katteryn (12 de diciembre de 2018). *La televisión peruana ya no es útil para la educación*. Chiqaq News. <https://medialab.unmsm.edu.pe/chiqaqnews/la-televison-peruana-ya-no-es-util-para-la-educacion/>
- Harrenstien, Ken (2009). *Automatic captions in YouTube*. Google blog. <https://googleblog.blogspot.com/2009/11/automatic-captions-in-youtube.html>
- Jurafsky, Daniel y Martin, James H. (2020). *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*. Stanford University. <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>
- Karakanta, A., Papi, S., Negri, M., y Turchi, M. (2021). Simultaneous speech translation for live subtitling: from delay to display. *MTSUMMIT*. <https://arxiv.org/pdf/2107.08807.pdf>
- Korybski, T., Davitti, E., Orăsan, C. y 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). European Language Resources Association. <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.468.pdf>



- Lambourne, A. (2007). Real-time Subtitling: Extreme Audio-visual Translation. Presentado en conferencia *Multidimensional Translation: LSP Translation Scenarios*. Vienna, mayo 2007.
- Levis, Jhon M. y Suvorov, Ruslan (2012). Automatic Speech Recognition. En Carol Chapelle (ed.), *Encyclopedia of Applied Linguistics* (pp. 1-8). Blackwell.
- Lockrey, Michael (2015). *YouTube automatic captions score an incredible 95% accuracy rate!*. Medium. <https://medium.com/@mlockrey/youtube-s-incredible-95-accuracy-rate-on-auto-generated-captions-b059924765d5>
- Lu, Xugang, Li, Sheng y Fujimoto, Masakiyo (2020). Automatic Speech Recognition. En Yutaka Kidiwara, Eiichiro Sumita y Hisashi Kawai (eds.), *Speech-to-speech translation* (pp. 21-38). SpringerBriefs in Computer Science.
- Mikul, Chris (2014). *Caption Quality: Approaches to standards and measurement*. Media Access Australia. <https://mediaaccess.org.au/research-policy/white-papers/caption-quality-international-approaches-to-standards-and-measurement>
- Moore, Zoe (2020). Fostering access for all through respeaking at live events. *The Journal of Specialised Translation*, 33, 207-226.
- National Institute on Aging. (2018). *¿Es confiable la información sobre la salud que se encuentra en el Internet?*. National Institute on Aging. <https://www.nia.nih.gov/espanol/confiable-informacion-sobre-salud-se-encuentra-internet>
- Navimedia. (s.f.). *¿Por qué el cerebro prefiere vídeo vs. texto?*. Navimedia <https://navimedia.es/por-que-el-cerebro-prefiere-el-video-vs-texto/>
- Ofcom. (2005). *Subtitling - An Issue of Speed?*
- Puente, Alejandra (2020). *El rol de los medios de comunicación en el Perú durante la pandemia*. Konrad Adenauer Stiftung. <https://www.kas.de/documents/269552/0/El+rol+de+los+medios+de+comunicaci%C3%B3n+en+el+Per%C3%BA+durante+la+Pandemia.pdf/d457a599-7f4a-99ca-3230-b8f4ec620bdd?version=1.0&t=1595436508362>
- Rao, Durgesh D. (1998). Machine translation. A gentle introduction. *Resonance*, 3(7), 61-70.
- Ríos Valero, Laura (2022). Análisis de la calidad. El subtitulado para sordos de YouTube en un programa de televisión peruano. En Miguel Ibáñez Rodríguez, Carmen Cuéllar Lázaro y Paola Masseur (eds.), *De la hipótesis a la tesis: Traductología y lingüística aplicada*. Comares.
- Robert, Isabelle R. y Remael, Aline (2017). Assessing quality in live interlingual subtitling: A new challenge. *Linguistica Antverpiensia, New Series-Themes in Translation Studies*, 16, 168-195.
- Romero-Fresco, Pablo (2010). Standing on quicksand: hearing viewers' comprehension and reading patterns of respoken subtitles for the news. En Jorge Díaz Cintas, Anna Matala y Josélia Neves (Eds.), *New insights into audiovisual translation and media accessibility* (pp. 175-199). Brill.
- Romero-Fresco, Pablo (2011). *Subtitling through speech recognition: Respeaking*. St Jerome.
- Romero-Fresco, Pablo (2016). Accessing communication: The quality of live subtitles in the UK. *Language and Communication*, 49, 56-69.
- Romero-Fresco, Pablo (2020, February 13). *The State of Live Captioning Today - An Expert's Perspective*. AppTek. <https://www.apptek.com/post/the-state-of-live-captioning-today-an-experts-perspective-accessibility-series-part-4>
- Romero-Fresco, Pablo y Alonso-Bacigalupe, Luis (2022). An empirical analysis on the efficiency of five interlingual live subtitling workflows. *XLinguae*, 15(2), 3-16.
- Romero-Fresco, Pablo y Martínez, Juan (2015). Accuracy Rate in Live Subtitling-the NER Model. *Audiovisual Translation in a Global Context: Mapping an Ever-Changing Landscape*, 28-50.
- Romero-Fresco, Pablo y Pöschhacker, Franz (2017). Quality assessment in interlingual live subtitling: The NTR model. *Linguistica Antverpiensia*, 16, 149-167.
- Rubio, Isabel (29 de mayo de 2020). ¿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>
- Sandrelli, Annalisa (2020). Interlingual respeaking and simultaneous interpreting in a conference setting: a comparison. En Nicoletta Spinolo y Amalia Amato (eds.), *inTRAlinea Special Issue: Technology in Interpreter Education and Practice*. Università degli Studi Internazionali di Roma



- Sandrelli, Annalisa (2021). Eventi dal vivo e accessibilità: Uno studio di caso sul respeaking interlinguistico. *Lingue e Linguaggi*, 43, 145-168.
- Tarakini, G., Mwedzi, T., Manyuchi, T. y Tarakini, T. (2021). The Role of Media During COVID-19 Global Outbreak: A Conservation Perspective. *Tropical Conservation Science*, 14, 1-13.
- Trajectory Partnership. (2010). *The Information Dividend: Why IT makes you "happier."* Trajectory. <https://trajectorypartnership.com/wp-content/uploads/2013/09/BCS-InformationDividend-UK.pdf>
- Ulatius. (20 de junio de 2022). *How Automatic Speech Recognition & Machine Translation are Revolutionizing Subtitling*. Ulatius. <https://www.ulation.com/translation-blog/how-automatic-speech-recognition-machine-translation-are-revolutionizing-subtitling/>
- University of Minnesota Duluth (Media Hub) (s.f.). *Correcting YouTube Auto-Captions*. <https://itss.d.umn.edu/centers-locations/media-hub/media-accessibility-services/captioning-and-captioning-services/correct>
- Youtube. (s.f.). *Ayuda de YouTube*. <https://support.google.com/youtube/?hl=es-419#topic=9257498>

## ANEXO 1. MATRIZ DE INSTRUMENTO: CUESTIONARIO

Antes de responder las preguntas, se explicó a los participantes que el objetivo del estudio era evaluar la calidad del subtítulo de YouTube.

### I. DATOS GENERALES

Sexo

- Femenino
- Masculino

¿Qué edad tiene?

¿Qué grado educativo has alcanzado?

- Sin estudios
- Primaria incompleta
- Primera completa
- Secundaria incompleta
- Secundaria completa
- Superior técnico incompleto
- Superior técnico completo
- Superior universitario incompleto
- Superior universitario completo

### II. NIVEL DE COMPRENSIÓN

Este nivel solo intenta medir su comprensión solo para saber si los subtítulos de YouTube son de calidad.

1) ¿En qué contexto se da la entrevista?

- a) Al inicio de la pandemia.
- b) Un año después del inicio de la pandemia.
- c) Cuando no había pandemia.

2) ¿De acuerdo a la entrevista, está de acuerdo con la afirmación del doctor con respecto a que la pandemia fue un «experimento de trauma y estrés»? ¿Por qué?

3) ¿De acuerdo a las declaraciones del doctor, el científico Isaac Newton cometió un delito?

- a) Sí
- b) No

4) ¿Según la entrevista, para quiénes ha sido más difícil estar en casa y hacer todo durante la pandemia?

- a) Padres y niños
- b) Padres solteros y mujeres
- c) Mujeres y hombres casados

5) ¿Está de acuerdo con las recomendaciones del doctor, tales como controlar los tiempos para dormir, hacer ejercicios, así como moderar su consumo de alcohol? ¿Por qué?

- 6) ¿El doctor está a favor de las vacunas?
- a) Sí
  - b) No
- 7) ¿Según la entrevista, cuál fue uno de los factores de estrés durante pandemia?
- a) La falta de vacunas
  - b) Las medidas del gobierno
  - c) La pérdida de contacto social
- 8) ¿El doctor está a favor de que los colegios hayan vuelto a la presencialidad?
- a) Sí
  - b) No
- 9) ¿Según las declaraciones del doctor, está de acuerdo con sus afirmaciones con respecto a que a la pandemia de la COVID-19 le seguirá una ola de enfermedades mentales?
- 10) ¿De acuerdo a la entrevista, qué necesitan las personas con riesgo de tener enfermedades mentales?
- a) un sistema de salud de mejor calidad
  - b) salir más con amigos
  - c) hablar más del tema.

## NIVEL DE PERCEPCIÓN DE LA CALIDAD DE LOS SUBTÍTULOS AUTOMÁTICOS DE YOUTUBE

Este nivel intenta saber su opinión con respecto a la calidad de los subtítulos automáticos de YouTube.

- 11) ¿Es suficiente el tiempo para leer los subtítulos?
- a) Sí, es suficiente.
  - b) Puede mejorar
  - c) No, no es suficiente.
- 12) ¿Tardó mucho en aparecer el subtítulo en el video?
- a) No, no tarda nada.
  - b) Tarda un poco.
  - c) Sí, tarda mucho.
- 13) ¿Logró identificar a los hablantes en los subtítulos?
- a) Sí, sí pude identificar a cada hablante.
  - b) A veces pude identificarlos.
  - c) No, no pude identificar quién estaba hablando.
- 14) ¿Había coherencia entre la imagen/el sonido y el subtítulo?
- a) Sí, sí había coherencia entre la imagen/sonido y el subtítulo.
  - b) A veces tenían coherencia.
  - c) No, la imagen/el sonido no tenía nada que ver con el subtítulo.
- 15) ¿Qué calificación pondría a los subtítulos de YouTube del 0 al 10, donde 0 es muy malo y 10 es muy bueno?
- 16) Si tiene algún comentario con respecto a la calidad de los subtítulos de YouTube, indícalo. Si no, puedes colocar terminar.