

**TRABAJO DE FINAL DE GRADO EN TRADUCCIÓN E
INTERPRETACIÓN**

TREBALL DE FI DE GRAU EN TRADUCCIÓ I INTERPRETACIÓ

Departament de Traducció i Comunicació

TÍTULO / TÍTOL

**Estudio de caso de la subtitulación automática en
YouTube**

Autora: Laura Corrochano Muñoz

Tutora: Laura Mejías Climent

Fecha de lectura/Data de lectura: junio de 2023



Resumen/ Resum:

En este trabajo se estudia la subtítulos automática que presenta un vídeo de YouTube. Se analizan la totalidad de los subtítulos que conforma el vídeo, de 51 minutos de duración.

En primer lugar se explican las convenciones de subtítulos propuestas por Díaz Cintas, que se utilizarán posteriormente en el análisis.

En segundo lugar, se buscarán las limitaciones de la subtítulos automática y si presenta algún aspecto positivo. Para ello, se recurrirá a dos baremos de catalogación de errores, uno para subtítulos y otro para traducciones.

Por último, expondremos las conclusiones del análisis con el objetivo de discutir hasta qué punto la subtítulos automática de YouTube es aceptable.

Palabras clave/ Paraules clau: (5)

Subtítulos, subtítulos automática, traducción audiovisual, YouTube, limitaciones

Contenido

1. Introducción.....	5
1.1. Justificación y objetivos.....	5
1.2. Metodología.....	5
1.3. Estructura.....	6
2. Marco teórico.....	6
2.1. Traducción audiovisual.....	6
2.1.1. La subtitulación.....	7
2.1.1.1. Pautado y velocidad de lectura.....	9
2.1.1.2. Segmentación.....	10
2.1.1.3. Formato del subtítulo.....	10
2.1.1.4. Convenciones ortotipográficas.....	11
2.1.1.5. Síntesis de la información.....	12
2.1.2. Subtitulación automática.....	12
3. Análisis del corpus.....	14
3.1. Pautado y velocidad de lectura.....	15
3.2. Segmentación.....	17
3.3. Formato del subtítulo.....	17
3.4. Convenciones ortotipográficas.....	18
3.5. Síntesis de la información.....	20

4.	Conclusiones.....	20
5.	Bibliografía.....	23
6.	Anexos.....	24
6.1.	Ficha de análisis.....	24

1. Introducción

1.1. Justificación y objetivos

A lo largo de la carrera hemos tratado las diferentes modalidades de traducción y, a pesar de haber escogido especializarme en la traducción jurídica, la traducción audiovisual siempre ha movido algo en mí, porque el mundo audiovisual forma una parte muy importante de mi vida.

Con ello en mente, en un primer lugar pensé en analizar la subtitulación automática de la plataforma Twitch, porque consideré que los subtítulos podrían adaptarse para ser accesibles para las personas sordas. Sin embargo, tuve que descartar la idea, ya que no encontré ningún *streamer* de habla hispana que tuviera activados los subtítulos, ya que no es un elemento que la plataforma Twitch ofrezca, sino que el creador de contenido tiene que instalarse una extensión para ofrecer este servicio.

Después de plantearme diferentes alternativas, decidí centrar mi Trabajo de Fin de Grado en la subtitulación automática de YouTube porque, en ocasiones, consumo estos subtítulos y ya me había percatado de ciertos errores que presentan. Por ello, pensé en analizarlos para comprobar cuántas limitaciones y aspectos positivos pueden identificarse de forma sistemática.

Una vez que tenía decidido el material con el que trabajar, quise cambiarlo, por ciertas informaciones que se publicaron sobre el *youtuber* en quien había pensado inicialmente, y me acabé decantando por el episodio 33 de la temporada 4¹ del podcast [Buenismo Bien](#) (M. Burque, Q. Peinado y H. Álvarez, 2018-), porque lo había visto recientemente, se trata de un contenido que consumo habitualmente y presenta una duración que consideré adecuada (51:03 minutos).

1.2. Metodología

Para caracterizar los subtítulos que acompañan a este podcast, se llevará a cabo un análisis de carácter descriptivo de dichos subtítulos automáticos generados por YouTube para el

¹ *Buenismo Bien*, 4x33, «Andrea Compton, la mejor», disponible en <https://www.youtube.com/watch?v=rWqhEjb3zbc> (acceso en mayo de 2023).

vídeo en cuestión. En dicho análisis se prestará especial atención a los errores que se presenten. Para determinar los diferentes tipos de errores utilizaremos el baremo de corrección de subtítulos propuesto por el grupo TRAMA en el año 2020. Además, complementaremos esta taxonomía utilizando el baremo de traducción que emplea el Departamento de Traducción y Comunicación de la UJI para la revisión de traducciones del grado en Traducción e Interpretación. Pese a no estar analizando subtítulos interlingüísticos, en una valoración inicial de estos subtítulos se detectaron ya errores de sentido que era necesario catalogar y el baremo de corrección de subtítulos resultaba insuficiente para errores de contenido; de ahí la combinación de ambos baremos como herramientas de análisis.

1.3. Estructura

La estructura de nuestro Trabajo Fin de Grado cuenta con seis capítulos.

En el primer capítulo, se presenta el tema y la justificación de su elección.

El segundo capítulo se centra en el marco teórico y se divide en dos apartados: la subtitulación y la subtitulación automática. El primer apartado explica qué es la subtitulación y las convenciones de esta. A continuación, el segundo apartado presenta un resumen breve de diversos proyectos de subtitulación automática.

El siguiente capítulo se corresponde con el análisis del vídeo que nos concierne y los resultados obtenidos basándonos en las convenciones presentadas en el segundo capítulo.

Los tres últimos capítulos presentan las conclusiones del análisis, la bibliografía y los anexos con un enlace a la ficha de análisis.

2. Marco teórico

2.1. Traducción audiovisual

Actualmente, tenemos contacto con productos audiovisuales a diario. Podemos disponer de entretenimiento mediante cualquier medio, ya sea las plataformas de *streaming*, la televisión, los videojuegos, las redes sociales, etc. Gracias a la globalización, además, tenemos la posibilidad de consultar cualquier contenido de cualquier país en cualquier

idioma y la traducción audiovisual facilita ese consumo globalizado (Bywood, Panayota, Etchegoyhen, 2017). Pero, primero hay que entender qué es.

La traducción audiovisual es una variedad de traducción que, de manera intra o interlingüística, se ocupa de los textos audiovisuales definidos como aquellos que se transmiten mediante, al menos, dos canales de comunicación: el canal acústico, con códigos de significación tales como la voz, el contenido paralingüísticos o los efectos especiales, y el canal visual, con códigos de significación tales como las imágenes, el movimiento o los carteles, entre otros. (Chaume, 2004). Desde el inicio del cine mudo, considerado como el origen de la traducción audiovisual tal y como la conocemos ahora, este campo de especialización de la traducción se ha denominado de diversas formas, como, por ejemplo, *Screen translation* (Mason, 1989) o traducción cinematográfica (Hurtado, 1994).

Dentro de la traducción audiovisual, el doblaje y la subtitulación son las modalidades más conocidas. No obstante, no son las únicas. Podemos dividir las en categorías según en su modo de ejecución: *revoicing*, *captioning* y otras. Dentro del *revoicing*, las modalidades que podemos observar son el doblaje, el *voice-over*, el doblaje parcial, la narración, la interpretación simultánea, la traducción a la vista, el comentario libre, la interpretación consecutiva, los *fundubs* y la audiodescripción. Por otro lado, el *captioning* cuenta con la subtitulación, la sobretitulación, el rehablado, la subtitulación para sordos y el *fansubbing*. Por último, más recientemente, ha habido autores que han señalado la existencia de un tercer grupo donde clasificar modalidades con otras particularidades como Hernández Bartolomé y Mendiluce Cabrera (2005). Este último grupo estaría compuesto por la interpretación de lengua de signos, la traducción a la vista de guiones, la escritura de diálogos para textos de animación y la traducción multimedia (Chaume, 2012).

No obstante, este trabajo se centrará únicamente en la subtitulación y, concretamente, en la subtitulación automática. Por tanto, nos ocuparemos de la modalidad más prominente del grupo conocido como *captioning*.

2.1.1. La subtitulación

Entre otros autores, Chaume (2012) define la subtitulación de la siguiente manera:

La subtitulación, como su nombre indica, consiste en incorporar un texto escrito (subtítulos) en la lengua meta a la pantalla en donde se exhibe una película en versión original, de modo que estos subtítulos coincidan aproximadamente con las intervenciones de los actores de la pantalla (p. 33).

Otros autores como Díaz Cintas y Remael (2021) definen la subtitulación de la siguiente forma:

A translation practice that consists in presenting a written text, generally on the lower part of the Screen, that aims to recount the original dialogue exchanged among the various speakers, as well as all the other verbal information that is transmitted visually (letters, inserts, graffiti, text messages, inscriptions, placards, and the like) and aurally (songs, voices off, voiceover narration). [...] The interaction of these three components [the spoken word, the image and the subtitles], along with the viewer's ability to read both the images and the written text at a particular speed, and the actual size of the screen, determine the basic characteristics of the audiovisual medium. Subtitles must appear in synchrony with the images and dialogue, provide a semantically adequate account of the SL dialogues, and remain displayed on Screen long enough for the viewers to be able to read them (p. 9).

Por otro lado, los subtítulos pueden presentar diversas realizaciones y es por ello por lo que Díaz Cintas y Remael (2021) propone una clasificación de subtítulos de acuerdo con cinco aspectos: según la lengua, según el tiempo de preparación, según parámetros técnicos, según la forma de aparición y según los métodos de proyección.

A su vez, estos parámetros se pueden subdividir según varios aspectos. En los parámetros lingüísticos, se pueden distinguir tres tipos. Los subtítulos intralingüísticos están redactados en la misma lengua que se escucha en el producto audiovisual. Este tipo se utiliza para el aprendizaje de lenguas, para sordos, para entender dialectos de una misma lengua o como un efecto karaoke. Por otro lado, los subtítulos interlingüísticos se pueden emplear para el aprendizaje de otras lenguas y para sordos. Estos consisten en la traducción de una lengua a otra. Por último, existe una modalidad bilingüe, que es una variante del tipo interlingüístico y que se producen zonas donde se hablan varias lenguas,

por lo que se traduce a dos lenguas meta simultáneamente, como se muestra en la figura 1.

Según el tiempo de preparación, se diferencian dos tipos de subtítulos: preparados y en directo. La única diferencia es que los preparados se confeccionan antes de que el



Imagen 1. Ejemplo de subtítulos bilingües (Nuria Cabezas, 2018)

producto se emita y los subtítulos en

directo se hacen en el mismo momento en el que producto se emite, mediante el rehablado o la traducción automática.

Respecto a los aspectos técnicos, existen los subtítulos abiertos y los cerrados. Los abiertos son los que están siempre visibles porque están sobreimpresos en la imagen, mientras que los cerrados se encuentran visibles si el espectador así lo quiere.

Los subtítulos pueden aparecer de dos formas: en bloque o en desplazamiento. Los subtítulos en bloque consisten en una línea o dos que van apareciendo y desapareciendo. Por otro lado, el texto de los subtítulos en desplazamiento va apareciendo de izquierda a derecha o de abajo hacia arriba.

Finalmente, entre los métodos de proyección se pueden distinguir seis tipos de subtítulos: manuales, químicos, ópticos, electrónicos, por láser y sobreproyección.

En relación al trabajo que nos concierne, analizaremos subtítulos intralingüísticos; abiertos, puesto que YouTube da la opción al usuario de elegir si quiere verlos o no, en directo y en desplazamiento.

2.1.1.1. Pautado y velocidad de lectura

Uno de los aspectos más importante de los subtítulos es el tiempo en que aparecen y desaparecen para acoplarse sincrónicamente al personaje que habla. Así, es muy importante que el subtítulo aparezca en el *frame* exacto y dure lo necesario para que el espectador tenga el tiempo suficiente para leerlo (Díaz Cintas y Remael, 2021). La duración del subtítulo vendrá determinada por el espacio del que se disponga, la longitud

de las líneas y el tiempo de la intervención del personaje. Se ha venido recomendando que como máximo dure un total de seis segundos, pero últimamente se recomienda que el subtítulo aparezca cinco segundos como máximo. Por otro lado, para que esté el suficiente tiempo en pantalla como para que se pueda leer completo una sola vez, se recomienda que esté al menos un segundo (Díaz Cintas y Remael, 2021). Una de las implicaciones más importantes de la subtitulación es que captamos la información de manera más rápida mediante el oído que con los ojos. Por ello, se tendrá que limitar la cantidad de información que los subtítulos ofrecen y, por tanto, la longitud de estos (Cerezo Merchán, 2019).

Además de estos parámetros hay que tener en cuenta la velocidad de lectura, que es la relación que existe entre la cantidad de texto del subtítulo y el tiempo que aparece en pantalla. Este aspecto se mide en caracteres por segundo y palabras por minuto (Díaz Cintas y Remael, 2021).

Asimismo, hay más elementos audiovisuales que marcan en qué momento se debe cambiar de subtítulo. Estos otros elementos son las pausas, los cambios de escena, el texto en pantalla y el cambio de plano (Díaz Cintas y Remael, 2021). Los resultados de la encuesta llevada a cabo por el grupo TRAMA muestran que las empresas determinan el cambio de línea del subtítulo en base al ritmo, las pausas y la cantidad de diálogo (Tamayo Masero, 2019)

2.1.1.2. Segmentación

Resulta crucial para el espectador que el subtítulo tenga sentido completo y para ello, es imprescindible que no se dividan palabras ni se corte la unidad de sentido a final de línea o subtítulo (Díaz Cintas y Remael, 2021). Algunas de las categorías oracionales que los traductores audiovisuales evitan que se queden al final de una línea son preposiciones, artículos, posesivos (Tamayo Masero, 2019), sintagmas nominales o verbos auxiliares (Cerezo Merchán, 2019).

2.1.1.3. Formato del subtítulo

Los subtítulos, como norma general, deben aparecer en la parte baja de la pantalla, en el centro, y pueden ocupar hasta un máximo dos líneas. En el caso de que el texto solo

ocupara una línea, tiende a colocarse en el espacio de la segunda línea según los últimos estudios (Díaz Cintas y Remael, 2021).

En razón a la fuente del subtítulo, puede variar según el cliente o la plataforma en la que se muestre. Por ello se puede decir que no hay establecida una fuente normalizada que sigan todos los subtítulos, sino que se tiende a utilizar fuentes neutrales (Díaz Cintas y Remael, 2021). No obstante, sí que se está de acuerdo en que se coloquen sobre un fondo negro y que las letras sean de color blanco (Díaz Cintas y Remael, 2021).

Cada línea del subtítulo tiene que tener un espacio determinado para no ocupar toda la pantalla. No hay una regla clara sobre cuantos caracteres por línea puede haber, aunque tradicionalmente se ha optado por 37 caracteres contando espacios en blanco, hasta un máximo de 40, aproximadamente, según el cliente. No obstante, dependerá de la fuente que se utilice, puesto que el espacio entre caracteres varía (Díaz Cintas y Remael, 2021).

2.1.1.4. Convenciones ortotipográficas

Es imprescindible comentar que se emplean las reglas de ortotipografía general, con algunas excepciones. Es el caso de versalitas y el subrayado, ya que no se utilizan nunca en subtitulación. En razón a las abreviaturas y siglas, estas deberían evitarse y solo escribir las que sean muy conocidas, como el ejemplo de las unidades de medida (Díaz Cintas y Remael, 2021).

Asimismo, hay casos especiales, como son:

- La cursiva: se emplea para personas y voces en *off*, pero no para las personas fuera de plano. En el caso de los narradores de documentales tampoco se emplea. Los otros casos en los que se utilizan son pensamientos y sueños, canciones traducidas y las voces que provienen de aparatos electrónicos (Díaz Cintas y Remael, 2021).
- Los guiones: se utilizan sin espacio detrás, para no perder un carácter, con el objetivo de marcar diálogos y existe la opción de colocarlo tanto en las dos líneas como solo en la segunda (Díaz Cintas y Remael, 2021).

- Los puntos suspensivos: se emplean con el fin de marcar duda, pausas o frases inacabadas y conectar subtítulos, aunque estos últimos son opcionales (Díaz Cintas y Remael, 2021).
- Las mayúsculas: pueden usarse en títulos, rótulos y el texto que aparezca en pantalla (Mejías-Climent, 2019)
- Las comillas: es muy importante que aparezcan comillas de inicio en el primer subtítulo y de cierre en el último. Hay más casos en los que se emplean como son las citas, el intertexto, las palabras extranjeras, para marcar una palabra y algo que debería aparecer en cursiva, pero que está en un subtítulo en cursiva (Díaz Cintas y Remael, 2021).

2.1.1.5. *Síntesis de la información*

Aunque se busca subtitular todo el contenido expresado por los personajes en sus diálogos, en ocasiones esto no puede ser así, dadas las restricciones espacio-temporales del texto escrito, y es por ello por lo que prima el principio de la relevancia. Es decir, habrá que subtitular los aspectos más importantes de cada frase si no se puede subtitular todo (Díaz Cintas y Remael, 2021). En todo caso, cabe mencionar que, en ocasiones, la información omitida en los subtítulos se puede recuperar mediante el canal visual (Martínez Sierra, 2019) Se trata de aplicar estrategias de condensación, omisión y reformulación para reducir el texto. Algunas recursos de condensación de la información son el uso de sinónimos más cortos, la sustitución de sintagmas nominales por pronombre o la escritura de números en guarismos, entre otros (Cerezo Merchán, 2019). Por otro lado, hay elementos como los nombres, las repeticiones, los vocativos o las interjecciones, entre otras, que se pueden evitar (Díaz Cintas y Remael, 2021).

2.1.2. Subtitulación automática

En las últimas décadas ha habido un aumento de la demanda de traducción audiovisual y para responder a ello se ha confiado en la tecnología con el objetivo de optimizar la labor del traductor. Sin embargo, la subtitulación no ha visto tantos resultados positivos como el resto de la industria (Bywood, Panayota, Etchegoyhen, 2017).

La industria se ha resistido más a incorporar tecnologías de traducción automática en el ámbito de la traducción audiovisual, principalmente, por cuestiones como el amplio rango

de vocabulario que se puede utilizar en una producción audiovisual, o la falta de lenguajes controlados; la creatividad necesaria para trasladar convenientemente los diálogos de una lengua a otra. Por lo tanto, una de las estrategias esenciales para incorporar tecnologías de traducción automática y hacer su resultado algo más adecuado es poseer el texto que producen estos sistemas. El estudio del proyecto SUMAT (Bywood, Georgakopoulou, Volk, & Fishel, 2012) ha mostrado que para este tipo de posesión los profesionales deben prepararse más exhaustivamente. No obstante, cabe mencionar que la traducción automática con posesión crea mejores traducciones cuanto más cercanas son las lenguas (Bywood, Panayota, Etchegoyhen, 2017).

Las primeras herramientas desarrolladas para los traductores fueron correctores gramaticales y ortográficos. La industria de la subtitulación ha evolucionado desarrollando *softwares* de subtitulación que han facilitado la tarea de la subtitulación que han convertido a las empresas desarrolladoras en empresas líderes hasta hoy en día (Georgakopoulou, 2019). Asimismo, se han creado varios proyectos para intentar crear un sistema de subtitulación automática que pueda realizar el mismo trabajo que un traductor. Uno de los primeros fue *SUMAT*, financiado por la Unión Europea (Bywood, Panayota, Etchegoyhen, 2017). Otro sistema de subtitulación automática fue *ALTO* creado para subtitular en español retransmisiones estadounidenses, cuyos resultados los hispanoparlantes encontraron favorables reconociendo el 70 % de las traducciones como aceptables o correctas. Posteriormente, se patentó *TranslateTV*, que lleva en uso desde 2003 para subtítulos en español de programas en directo norteamericanos (Bywood, Panayota, Etchegoyhen, 2017).

Los proyectos de subtitulación automática se interesaron en introducir el reconocimiento de voz para subtitulación en directo. Los avances en la calidad de los sistemas de reconocimiento de voz automáticos hicieron que fuera más barato y más rápido de entrenar que el regrabado. No obstante, sus niveles de precisión distan mucho de los niveles de los subtítulos producidos por humanos (Georgakopoulou, 2019). Un ejemplo práctico de este sistema es el *software* desarrollado por un grupo de investigadores de la universidad Carlos III. Este *software* permite subtitular automáticamente los informativos territoriales de RTVE mediante una inteligencia artificial y técnicas de aprendizaje automático y, además, el envío de los subtítulos se produce de manera más rápida (La Vanguardia, 2020).

En el caso que nos concierne, YouTube aplica un sistema desarrollado por Google de reconocimiento de voz para la lengua inglesa (Chaume, 2019). Sin embargo, ha sido imposible localizar información que afirme que programa utiliza YouTube para el idioma español. La única información localizada ha sido gracias a una pregunta en el foro de la comunidad de YouTube donde se afirma que YouTube utiliza sistemas propios de Google como los que se pueden consultar en Google Cloud. No obstante, en Google Cloud no se ofrece ninguna información al respecto.

3. Análisis del corpus

A continuación, analizaremos los subtítulos automáticos que se pueden ver en un vídeo del sitio web YouTube, basándonos en el baremo de corrección de subtítulos del grupo TRAMA (2020) para catalogar errores y valorar el producto final considerando las características de la subtitulación explicadas en el apartado anterior. Inicialmente, se pretendía trabajar solo con este baremo, matizando que los errores de ortografía provienen, en realidad, de una mala transcripción. Sin embargo, tras la detección de repetidos errores de esta naturaleza y tratándose de un baremo sencillo, también recurrimos al baremo de traducción, pese a no ser una traducción interlingüística, sino unos subtítulos intralingüísticos en español peninsular.

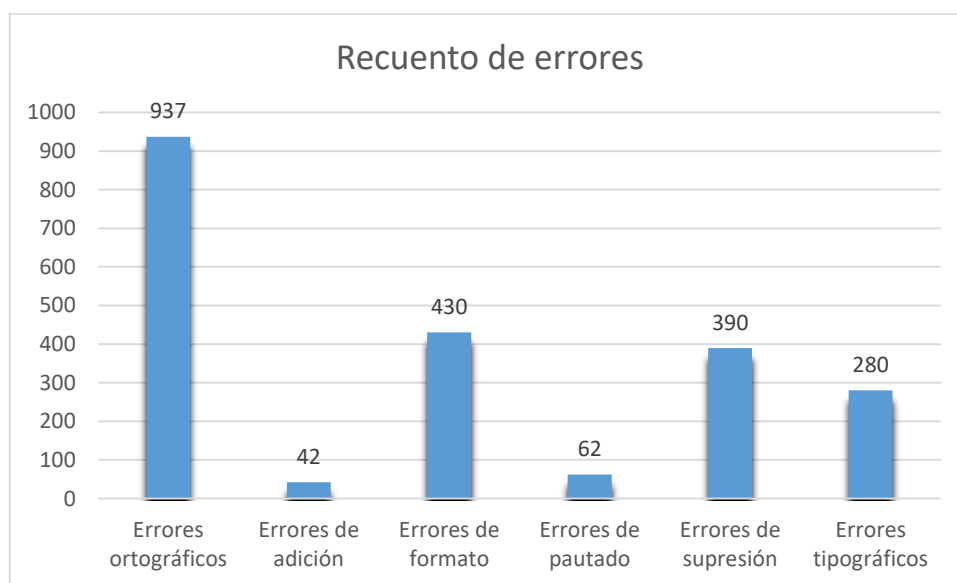
Para la realización de dicho análisis, hemos elaborado una tabla (Tabla 1, a continuación) en la que incluimos el número de subtítulo, el TCR, el subtítulo, el tipo de error y los comentarios que se estimen necesarios.

Número de subtítulo	TCR	Subtítulo	Tipo de error	Comentario
Tabla 1. Ficha de análisis				

Además, hemos confeccionado una segunda tabla (Tabla 2, a continuación) con los mismos apartados para los aciertos en cuestión de síntesis.

Número de subtítulo	TCR	Subtítulo	Comentario
Tabla 2. Ficha de análisis: aciertos de síntesis			

Para analizar los subtítulos hubo que descargarse los subtítulos mediante la web [downsub.com]. Al introducirlos en Subtitle Workshop surgió un problema, puesto que el programa interpretaba cada línea de los subtítulos que vemos en pantalla como un único subtítulo. Así pues, hemos utilizado esta segmentación para analizar cada línea independientemente e identificarla con el número que proporciona el programa. Por ello, hablaremos de los errores de las distintas líneas en vez de errores en subtítulos completos. Al tratarse de subtítulos en desplazamiento y no en bloque, consideramos que la catalogación de errores por línea es más ajustada que por subtítulo. En total, tras analizar los subtítulos se ha obtenido un recuento de 1112 líneas, cuyas características se explicarán a continuación.



Gráfica 1. Recuento de errores

3.1. Pautado y velocidad de lectura

En nuestro caso, los subtítulos automáticos producidos para YouTube son subtítulos en desplazamiento, lo que complica el cálculo del tiempo que estos permanecen en pantalla. Siguiendo los parámetros mencionados en el marco teórico respecto a la duración de los subtítulos en pantalla, cabe mencionar que no hay ninguna ocasión en la que el subtítulo permanezca menos de un segundo y que son bastantes las ocasiones en las que permanece más de 5 segundos.

Además, hay que mencionar que en dos ocasiones los subtítulos desaparecen como si fueran en bloque. Es el caso del subtítulo conformado por las líneas 339 y 340 y el subtítulo formado por las líneas 865 y 866.

El cálculo de la velocidad de lectura es más impreciso, ya que el texto aparece de forma secuencial según se va escuchando la intervención. En los casos en los que el emisor habla más rápido, el texto aparecerá más rápido, lo que implicará mayor velocidad y que, por tanto, la línea sea más difícil de leer. Por otro lado, los casos en los que el emisor habla más despacio, el texto aparece de manera más lenta, lo que implicará que sea más fácil de seguir por presentar una velocidad de lectura inferior. No obstante, cabe destacar que YouTube ofrece la opción de seleccionar la velocidad de reproducción en el menú de configuración, que también afecta a la velocidad en las que el texto va apareciendo. Por ello, no parece necesario dar cuenta de los caracteres por segundo a los que se muestran los subtítulos, ya que este valor es variable según las preferencias de configuración de los usuarios.

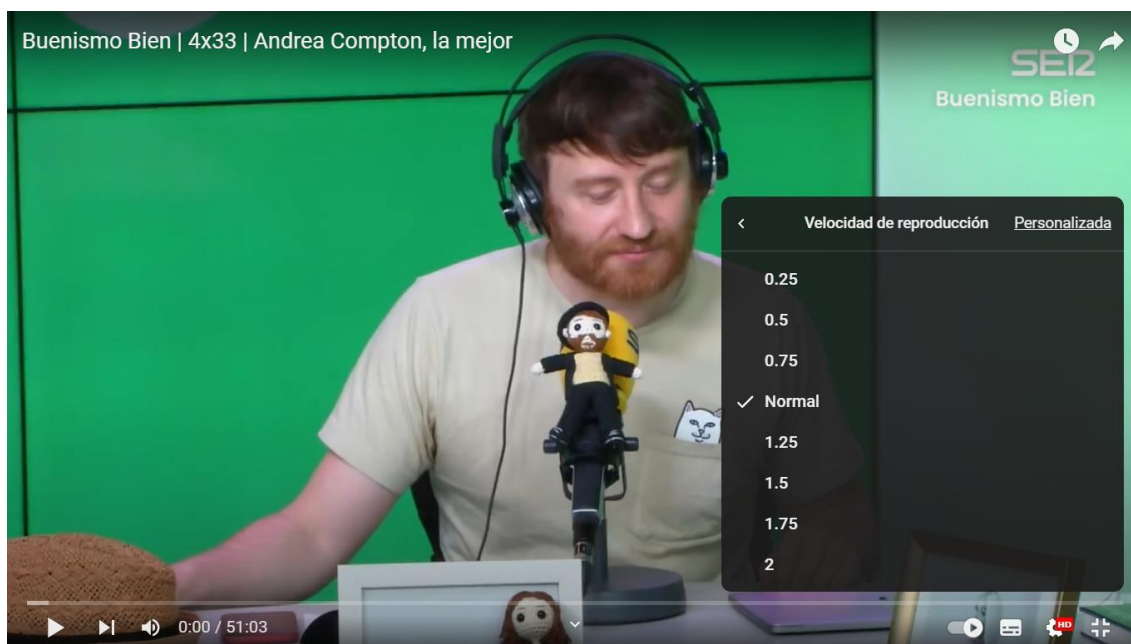


Imagen 2. Configuración de la velocidad de reproducción de un vídeo de YouTube

La aparición del texto de forma secuencial lleva a pensar que los subtítulos entraran en sincronía. Sin embargo, de las 1112 líneas de subtítulos que conforman el vídeo, son 60 el número de subtítulos que no inician en el *frame* exacto.

Por último, respecto a los elementos audiovisuales que determinan el cambio de subtítulo, no se respetan pausas, cambios de escena, texto en pantalla ni cambios de plano, excepto que estos ocurran de manera fortuita.

3.2. Segmentación

Tomando como referencia la afirmación de Díaz Cintas sobre el respeto de la unidad de sentido, hemos podido contar 393 errores de segmentación de subtítulos en un total de 1112 líneas de subtítulos. Para determinar estos errores, hemos tomado como referencia que el subtítulo no pueda acabar en artículo o determinante, preposición, conjunción o marcador del discurso, verbos auxiliares o adverbios.

Por otro lado, no se respeta la segmentación de subtítulo cuando se produce el cambio de un hablante a otro nada más que de manera fortuita en alguna ocasión. Además, en estos casos, lo que suele ocurrir es que la transcripción de la primera línea del subtítulo se corta antes de que la persona termine su intervención y la siguiente intervención aparece en la siguiente línea.

3.3. Formato del subtítulo

La ubicación del subtítulo de manera predeterminada es la parte inferior central del vídeo en cuestión. Sin embargo, YouTube ofrece la posibilidad de poder mover su posición libremente por el espacio que ocupa el vídeo en pantalla y, de esta manera, colocarlo donde cada espectador lo considere necesario.

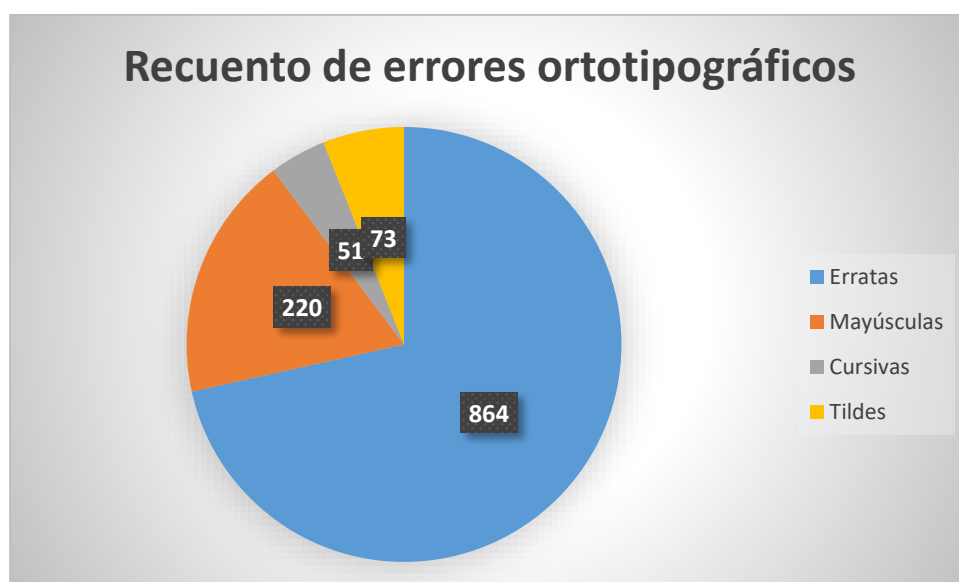
En cuanto a la fuente, la que se puede ver es la fuente predeterminada de YouTube. No obstante, también se da la libertad de elegir cambiarla pulsando sobre la rueda de configuración que se ve en la barra inferior del vídeo. Las fuentes que se pueden seleccionar son Serif monoespaciada, Serif proporcional, Sans-serif monoespaciada, Sans-serif proporcional, Casual, Cursiva, Versalitas. El hecho de que una fuente sea cursiva nos supondrá un problema que analizaremos más adelante.

Sin embargo, esta no es la única elección posible. Se puede personalizar el color tanto de la fuente como del fondo del subtítulo y de la ventana que enmarca el subtítulo, que de manera predeterminada no se ve. La opacidad de la fuente, del fondo y de la ventana también son una alternativa que se puede modificar con el fin de hacer partícipe al

espectador. Por último, se puede cambiar el estilo del borde de los caracteres y el tamaño de la fuente, pudiendo ampliarlo hasta un 400 %. Además, el tamaño de la fuente solo afecta al espacio que ocupa en pantalla y no al número de caracteres que aparecen por línea.

Por último, tomando como referencia el máximo de 40 caracteres, podemos decir que solo hay siete líneas de un total de 1112 que superen los 40 caracteres. Además, seis de estas siete líneas se pasan del número máximo de caracteres porque se transcriben las palabras malsonantes con el formato [___], lo que supone un añadido de espacios de más y una transcripción que no suele verse con frecuencia en los medios.

3.4. Convenciones ortotipográficas



Gráfica 2. Recuento de errores ortotipográficos

Contrariamente a las indicaciones proporcionadas por Díaz Cintas, los subtítulos automáticos de YouTube que nos conciernen no parece seguir ninguna norma ortotipográfica clara. Los elementos de tipografía tampoco se respetan, pues no se pueden apreciar:

- Cursiva: para los nombres de obras artísticas como lo son los nombres de canciones, libros y series, ni para señalar palabras extranjeras.
- Puntos suspensivos: para marcar que un subtítulo está inacabado.
- Mayúsculas: para subtítular los rótulos de las secciones ni las mayúsculas iniciales para nombres propios. Además, cabe mencionar que, de manera fortuita, ya que

tenía los subtítulos activados mientras consultaba otros vídeos de YouTube, como el denominado «YA VI GUARDIANES DE LA GALAXIA 3! | Reacción SIN SPOILERS» del canal Cámara en mano, sí que hacen uso de las mayúsculas.

- Comillas: para referencias intertextuales.
- Guiones: para marcar el inicio de las diferentes intervenciones de los hablantes dentro de un mismo subtítulo.
- Puntos: para determinar el final de una oración.

Cabe señalar que sí se puede observar el uso de guiones, aunque su uso es incoherente, puesto que en dos de las cuatro ocasiones que se observan, en vez de transcribir lo que se oye, aparecen como si el guion fuera un símbolo de la supresión, lo cual podría considerarse un anglicismo ortotipográfico. En una ocasión se transcribe como – porque lo que se escucha es la palabra *menos* y en la última ocasión, se trata de un error ortográfico por ser una mala transcripción (véase Anexo 1).

Retomando los aspectos ortográficos, hemos podido observar un total de 73 errores debidos a la falta de tildes. Dichos fallos, en su mayoría, son palabras que utilizan la tilde diacrítica para distinguir los significados de palabras que se escriben igual.

Como mencionábamos anteriormente, se han considerado los errores de transcripción como errores ortográficos. Este tipo de fallos asciende a 864 repartidos entre 480 líneas. En varias ocasiones, son errores que pueden ser comprensibles por la falta de vocalización. Un ejemplo de este tipo es la transcripción de *nos* en lugar de *no os*, es decir, la transcripción del singular en vez de las palabras originales, o bien, la transcripción de una palabra en masculino mientras que lo que se escucha es el femenino. Los nombres propios son otro caso de malas transcripciones. Son varias veces las que podemos ver que el apellido *Burque* transcribirse como *porque*.

Aunque, en un inicio, las adiciones se habían considerado como errores de ortografía, esta catalogación se matizó al recurrir al baremo de traducción complementario. Con este, pudimos añadir como error la categoría de *adiciones*. Hemos detectado 41 adiciones, de las cuales varias se deben a la subtitulación de música cuando no se escucha música o la transcripción de la letra de la canción que está sonando cuando no es necesario.

3.5. Síntesis de la información

Aunque inicialmente no considerábamos que se hubieran aplicado técnicas de síntesis, al recurrir a los baremos y al análisis detallado, se detectan 146 omisiones y reformulaciones que no cambian el sentido del texto original y que, por tanto, consideraremos formas de síntesis. Sin embargo, también se detectan muchos casos de supresión, concretamente, 377 errores de supresión en un total de 1112 líneas por la falta de, al menos, una palabra que afecte al sentido del subtítulo. Además, estas supresiones, en determinadas ocasiones, que se pueden consultar en la ficha de análisis en el anexo 1, provocan errores de sentido, errores gramaticales, errores semióticos, errores textuales o errores pragmáticos. De estas 377 supresiones, 51 se deben a la falta de subtitulación durante unos segundos.

4. Conclusiones

Como hemos mencionado en la introducción, el objetivo principal de este trabajo era analizar los subtítulos intralingüísticos generados automáticamente con el fin de reflexionar sobre las limitaciones y los posibles aspectos positivos de esta tecnología.

Hemos podido observar que hay alrededor de 860 erratas debidas a una mala transcripción repartidas entre 480 líneas, casi la mitad de las líneas que conforman el vídeo. Esto implica que la interpretación de la mitad del vídeo podría ser errónea porque en ocasiones, incluso, la errata cambia el sentido a la línea. Además, hay otras líneas que carecen de sentido y son incomprensibles, como se muestra en la imagen 3.



Imagen 3. Error de SS



Imagen 4. Diferencia de error de transcripción

ambas veces se pronuncia la palabra *Camela*, pero en una ocasión no se transcribe correctamente. No obstante, no hemos sido capaces de identificar el motivo de esta errata. Además, en ninguna de las dos ocasiones se utiliza la mayúscula inicial.

En relación con el uso de mayúsculas y nombres propios, podemos pensar en los referentes culturales, ya que el programa de subtitulación automática que emplea YouTube no los reconoce. Esto supone un problema, porque son bastantes las ocasiones en las que aparecen referentes de este tipo y algunos están bien escritos, aunque no emplean la mayúscula cuando deberían, mientras que otros están mal transcritos.

Es bastante problemático que no se pueda observar ningún signo de puntuación, porque dificulta la lectura y, en algunas ocasiones, cambia el sentido de los subtítulos, tal y como se puede comprobar en la segunda línea del subtítulo de la figura 5.



Imagen 5. NMS por error tipográfico

Como comentábamos anteriormente, el cálculo de la velocidad de lectura es impreciso porque el texto aparece de manera secuencial. Esto implica que, dependiendo de la velocidad a la que se pronuncie la intervención, el subtítulo aparecerá de manera más

Otro ejemplo de errata muy repetido es la transcripción de los nombres propios, porque resulta curioso que algunos estén mal transcritos en varias ocasiones y, en otras, se transcriba bien. Sin embargo, en ninguna de las ocasiones aparecen con mayúsculas inicial y, además, se podrían omitir como técnica de síntesis para reducir la extensión de la línea. Esto se puede comprobar en la imagen 4. Se transcribe mal el nombre del grupo Camela con una diferencia de solo unos 20 segundos. En este caso, concretamente, se puede escuchar como

rápida o más lenta. Esto puede resultar cómodo, pero dependerá de la persona que lo lea y la capacidad de lectura de esta. No obstante, la velocidad se podría modificar si el espectador decidiera cambiar la velocidad de reproducción del vídeo, lo que provoca que los subtítulos se adapten a una velocidad diferente.

La segmentación es un aspecto que, realmente, es molesto a la hora de leer los subtítulos porque corta el sentido de las oraciones e impide seguir la lectura. Además, el no marcar de ningún modo las intervenciones de personas distintas en un mismo subtítulo dificulta la lectura, porque, en ocasiones, otra persona empieza a hablar y ni siquiera está en pantalla.

Es interesante el hecho de que se pueda modificar la ubicación de los subtítulos, porque permite trasladarlos a otra parte de la pantalla si algo sucediera en la parte inferior central. Al igual que es buena idea que el tamaño del texto pueda modificarse, aspecto que facilita la lectura a personas con problemas de visión, pero esto tiene una desventaja y es que ocupa un espacio mayor en la pantalla. Sin embargo, el hecho de que la fuente se pueda modificar a una fuente en cursiva o versalitas provoca errores de tipografía, ya que en la fuente predeterminada no aparecen ni la cursiva ni la versalita y, en el vídeo analizado, la cursiva debería aparecer para señalar los nombres de obras artísticas y los extranjerismos.

Otro aspecto que dificulta la lectura es la falta y la adición de tildes diacríticas. Carece de sentido como el programa acierta en ocasiones a la hora de colocar la tilde, pero a la vez se equivoca en múltiples ocasiones.

Las supresiones impiden seguir el vídeo tanto por la falta de información en medio de un subtítulo como los segundos en los ni siquiera se subtitula nada y se deja al espectador sin ningún subtítulo.

Respecto a las canciones, el programa no distingue cuándo se trata de una persona hablando o de una canción y, por ello, subtitula la letra de la canción cuando no es necesario. Esto implica la adición de subtítulos innecesarios que impiden la transcripción de la conversación que se está teniendo.

Basándonos en todo lo anterior, podemos decir que, en el caso del vídeo analizado, la subtitulación automática presenta demasiados errores que dificultan enormemente seguir

el vídeo mediante los subtítulos de forma cómoda. Este trabajo podría servir como punto de partida para futuros estudios de recepción.

5. Bibliografía

Burque M., Peinado Q. y Álvarez H. (2021). Buenismo Bien | 4x33 | Andrea Compton, la mejor[YouTube]. Recuperado de <https://www.youtube.com/watch?v=rWqhEjb3zbc>

Bywood, L., Georgakopoulou, P., y Etchegoyhen, T. (2017). Embracing the threat: machine translation as a solution for subtitling. *Perspectives: Studies in Translatology*, 25 (3), 492-508. doi: <https://doi.org/10.1080/0907676X.2017.1291695>.

Cabezas, N. (2018) *La traducción audiovisual. Breve contextualización de la especialidad*. Recuperado de <https://www.universidadviu.com/es/actualidad/nuestros-expertos/la-traduccion-audiovisual-breve-contextualizacion-de-la-especialidad>

Cabrera, G. M., y Bartolomé, A. I. H. (2005). New trends in audiovisual translation: the latest challenging modes. *Miscelánea: A journal of english and american studies* (31), 89-104. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=2010052>.

Cerezo Merchán, B., Chaume Varela, F., De Higes Andino, I., De los Reyes Lozano, J., Granell Zafra, J., y Martí Ferriol, J. L. (2019). *La traducción para la subtitulación en España. Mapa de convenciones*. Castellón:Universitat Jaume I.

Chaume, F. (2004). *Cine y traducción*. Madrid: Cátedra.

Chaume, Frederic (2012) «La traducción audiovisual: Nuevas tecnologías, nuevas audiencias», en 12º *Congresso dell'Associazione Italiana di Linguistica Applicata*, Perugia, Guerra Edizioni: 143-159.

Díaz Cintas, J., y Remael, A. (2021). *Subtitling: Concepts and Practices*. Nueva York: Routledge.

Georgakopoulou, P. (2019). Technologization of Audiovisual Translation. En L. Pérez González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 516-539). Londres: Routledge Taylor and Francis Group.

Hagström, H., y Pedersen, J. (2022). Subtitles in the 2020s: The Influence of Machine Translation. *Journal of Audiovisual Translation*, 5(1), 207-225. doi: <https://doi.org/https://doi.org/10.47476/jat.v5i1.2022.195>

Hurtado, A. (1994). Modalidades y tipos de traducción. *Vasos Comunicantes* (4), 19-27

La Vanguardia, (15 de julio de 2020), RTVE implanta un software de la Carlos III para el subtulado en directo de informativos, *La Vanguardia*. Recuperado de: <https://www.lavanguardia.com/television/20200715/482325404406/rtve-software-subtitulado-informativos-en-directo-carlos-iii-inteligencia-artificial-rehablado.html>

Mason, I. (1989). Speaker meaning and reader meaning: preserving coherence in screen translating. En R. Kölmel y J. Payne (Eds.), *Babel. The Cultural and Linguistic Barriers between Nations* (p. 13-24), Aberdeen: Aberdeen University Press.

Subtitulación. Atrae.org. Recuperado de <https://atrae.org/subtitulacion/>

6. Anexos

6.1. Ficha de análisis

https://docs.google.com/spreadsheets/d/1nxLjpTVAXBr_mvOr1OuYQE-KZi4vVTBH/edit?usp=sharing&ouid=102367388144774885031&rtpof=true&sd=true