



Universidad de Valladolid

FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN
GRADO EN TRADUCCIÓN E INTERPRETACIÓN

TRABAJO DE FIN DE GRADO

La calidad en interpretación automática
inglés-español en el ámbito agroalimentario:
un estudio de caso

Presentado por D.^a Sofía Calvo del Barrio

Tutelado por la Dra. María Teresa Ortego Antón

Soria, junio de 2024

ÍNDICE

RESUMEN	1
ABSTRACT	1
1. Introducción	2
2. Competencias	3
3. Objetivos	4
3.1. Objetivo principal	4
3.2. Objetivos secundarios	4
4. Marco teórico	5
4.1. El concepto de interpretación	5
4.2. Las tecnologías de la interpretación.....	7
4.2.1. Definición y clasificación.....	7
4.2.2. La interpretación remota.....	8
4.2.3. La interpretación automática	9
5. Metodología de análisis	13
5.1. Sistema de interpretación automática	13
5.2. Muestras de análisis	14
5.3. Modelo de evaluación de interpretación simultánea.....	14
6. Análisis y resultados	19
6.1. Muestra de análisis 1.....	19
6.2. Muestra de análisis 2.....	21
6.3. Contraste de resultados	25
7. Conclusiones	27
8. Bibliografía	29

Índice de ilustraciones

Ilustración 1. Interfaz de Interpretfy.....	13
Ilustración 4. Gráfico de resultados de la muestra 1	20
Ilustración 5. Gráfico de los resultados de la muestra 2.....	24
Ilustración 6. Gráfico para el contraste de resultados de M1 y M2.....	25

Índice de tablas

Tabla 1. Modelo de análisis con código de colores.....	15
Tabla 2. Errores detectados en la muestra de análisis 1	19
Tabla 3. Errores detectados en la muestra de análisis 2	21

RESUMEN

Dados los últimos avances en materia de inteligencia artificial y redes neuronales, la interpretación automática ha experimentado una rápida transformación que ha permitido que se pueda aplicar en la práctica. Estos sistemas, que se enfrentan a los grandes retos del lenguaje humano, generan un interesante debate sobre la amenaza o la utilidad de estas herramientas para la profesión. En este contexto, el presente Trabajo de Fin de Grado pretende evaluar la calidad de uno de estos sistemas de interpretación automática mediante el análisis de la interpretación al español de dos fragmentos de audio originales en inglés de distinto grado de especialización. Para dicho análisis, elaboraremos un modelo de evaluación para el que combinaremos y adaptaremos los modelos de Collado Aís (2007) y Ortiz Boix (2016). La identificación de errores según dicho modelo nos permitirá detectar la frecuencia de aparición de dichos errores y concluir cuáles son los aspectos en los que se esperan mejoras en estos sistemas.

Palabras clave: interpretación automática, traducción automática, reconocimiento de voz automático, error, parámetro.

ABSTRACT

Along with the rapid developments in terms of artificial intelligence and neural networks, machine interpreting is advancing so rapidly that it can be employed in real scenarios. These systems facing the big challenges of human communication unleash so interesting a debate about whether they are a threat or useful tool for the professionals. Taking all this into consideration, this research project aims to assess the performance of a machine interpreting system by analysing two pieces of audio interpreted into Spanish. Both are original in English and they have a different level of specialization. Two models proposed by Collado Aís (2007) and Ortiz Boix (2016) will be combined and adapted for this project to create a new model to carry out the analysis. The results will provide us with information about the frequency of mistakes, as well as about the improvements that are required.

Keywords: machine interpreting, machine translation, automatic speech recognition, mistake, parameter.

1. Introducción

En la actualidad la tecnología está presente en todos los ámbitos, disciplinas y sectores, especialmente con el reciente auge de la inteligencia artificial. Los Estudios de Traducción e Interpretación no han sido ajenos a los cambios que ha provocado la irrupción de la tecnología, en labores como la redacción, traducción y la corrección, entre otras; y hemos podido observar que recursos tecnológicos, tales como diccionarios digitales, sistemas de traducción asistida por ordenador o sistemas para elaborar bases de datos terminológicas, han llegado a ser parte fundamental del proceso de traducción. Como consecuencia de la importancia que han cobrado estos recursos y herramientas en la profesión, la competencia tecnológica se ha añadido a otras que se esperan de traductores e intérpretes, como el dominio de las lenguas de trabajo. De hecho, se está otorgando un lugar muy relevante a la competencia tecnológica en la formación de los profesionales (Moreno Villanueva, 2023, p. 115-121).

Dado el rápido avance de la tecnología en la disciplina de interpretación, la importancia que esta competencia tiene en la formación y el rasgo novedoso de esta área de conocimiento, además del interés que despierta en mí la disciplina en sí, será la interpretación automática el tema que abordaremos en este Trabajo de Fin de Grado, con el objetivo de obtener una aproximación del producto resultante en dos muestras de diferente grado de especialización en un sistema de interpretación automática. Para ello, en primer lugar, en el Capítulo 3 dedicado al marco teórico, definiremos los conceptos que vertebran este trabajo. En el siguiente capítulo (Capítulo 4), se expondrá la metodología empleada para analizar los productos finales de un sistema de interpretación automática, para continuar con el análisis y resultados (Capítulo 5), que irá seguido de las conclusiones (Capítulo 6) y, por último, las referencias (Capítulo 7).

Por otro lado, nos gustaría dejar constancia de que este trabajo se realiza en el marco de una Beca de Colaboración en Departamentos Universitarios concedida por concurrencia competitiva por parte del Ministerio de Educación y Formación Profesional y dirigida a estudiantes de último curso de grado y máster. La temática del proyecto de esta beca es precisamente la evaluación de sistemas de interpretación automática, por lo que ambos trabajos, Trabajo de Fin de Grado y proyecto de Beca de Colaboración, se enriquecen mutuamente y me permiten alcanzar una mayor y más detallada profundización en el tema.

2. Competencias

En el presente Trabajo de Fin de Grado se han aplicado diferentes conocimientos teóricos y prácticos adquiridos a lo largo del Grado en Traducción e Interpretación, así como una serie de competencias específicas y competencias transversales. Todas ellas figuran en la guía docente para el Trabajo de Fin de Grado del curso 2023-24, conforme a la descripción recogida en la Memoria Verifica del Grado en Traducción e Interpretación de la Universidad de Valladolid¹.

En primer lugar, las competencias de carácter general que se han puesto en práctica en la realización de este trabajo son CB2, CB3, CB4 y CB5.

En lo que se refiere a competencias específicas, estas se han ido adquiriendo en las diferentes asignaturas del grado y las que se han aplicado a este trabajo son: E1, E2, E5, E6, E10, E13 y E16.

Por último, en este Trabajo de Fin de Grado se han trabajado T1, T3 Y T5 como competencias transversales.

¹ https://www.facultadtraduccionaria.es/wpcontent/uploads/2022/02/01_Memoria_Verifica_GradoTraduccion_Ministerio_.pdf
(Fecha de consulta: 15/05/2024)

3. Objetivos

En este capítulo, procedemos a describir el objetivo general de este Trabajo de Fin de Grado en el primero de los apartados, además de dos objetivos secundarios, los cuales quedan explicados en el segundo apartado de este capítulo.

3.1. Objetivo principal

- Verificar la calidad del producto resultante del sistema de interpretación automática.

3.2. Objetivos secundarios

- Detectar errores en los dos productos resultantes del sistema de interpretación automática.
- Obtener una aproximación de la evolución del empleo de la tecnología en la interpretación.

4. Marco teórico

4.1. El concepto de interpretación

Son muy diversas las definiciones que podemos hallar para el concepto de interpretación. En este caso, haremos alusión a dos definiciones diferentes con el objetivo de obtener una idea completa y precisa de esta disciplina.

La definición de Herbert (1952, p. 3) pone el foco en la figura del intérprete como comunicador y aleja la idea de una simple y rápida búsqueda de equivalentes. El intérprete hace posible que el conocimiento se pueda transmitir de un individuo o grupo a otro e incluso colabora en que estos lleguen a un acuerdo.

The mission of the interpreter is to help individuals and communities to acquire a fuller knowledge and a deeper understanding of one another, and, what is still more important, a greater respect for one another. Also to come to an agreement if they should want to do so. [...] The conference interpreter is an assistant whose intelligent contribution is an indispensable factor in any international gathering (Herbert, 1952, p. 3).

En el caso de la definición que ofrece Nord (1991, p. 6), la idea principal versa sobre el hecho de que en la situación comunicativa de un servicio de interpretación participan, normalmente, dos culturas diferentes, lo que hace que el intérprete deba tener un conocimiento pleno de las lenguas de trabajo y de las culturas implicadas para poder tomar las decisiones adecuadas en base a aspectos, como los conocimientos previos que se asumen en cada cultura. Por otra parte, apunta a una clara diferencia con respecto a la traducción (además de que en el caso de la interpretación se trata de lenguaje oral y no escrito, como en la traducción): en la interpretación, todas las partes implicadas en la situación comunicativa (emisor, intérprete y receptor) están presentes al mismo tiempo, en el mismo lugar y empleando el mismo medio. Sin embargo, cabe mencionar que, dados los últimos avances tecnológicos, tales como la interpretación remota, veremos que esta realidad no siempre se cumple en su totalidad.

Interpreting is a special form of translation, because the situation requires the presence of the ST recipient (ST-R) as well as the translator and the TT recipient (TT-R). Since interpreting is a form of face-to-face communication, both the sender and the recipient are present -as is shown in the usual models of communication- together with the translator in the role of TT producer. All the participants communicate in the same place at the same time using the same medium, and the function of the text is the same for all of them except the translator. However, what is different is the cultural background of the ST-S and the ST-R on the one hand, and the TT-R, on the other (Nord, 1991, p. 6).

En la anterior cita se pone de manifiesto que la figura del intérprete nació con la necesidad de comunicación y es tan antigua como el ser humano. En el antiguo Egipto y en la antigua Grecia y Roma, la labor del intérprete era requerida en el ámbito administrativo, comercial, religioso y militar. En los siglos posteriores, la profesión va ganando reconocimiento y prestigio hasta llegar a profesionalizarse a principios del siglo XX (Universidad de Vigo, s.f.). De hecho, es en este siglo cuando se producen dos hitos muy importantes para la disciplina en los que la tecnología juega un papel fundamental.

Primeramente, los juicios de Núremberg (1945) tras la Segunda Guerra Mundial (1939-1945) suponen el inicio de la interpretación simultánea con el empleo de sistemas inalámbricos de transmisión de voz. Con ello, se propició que la demanda de los servicios de interpretación incrementara con la situación geopolítica tan convulsa y cambiante del siglo XX y que la labor del intérprete experimentara mejoras y se profesionalizara. En concreto, la interpretación simultánea es la modalidad en la que el intérprete reformula en lengua meta el discurso del orador principal en tiempo real, esto es, el discurso original y la interpretación se superponen en el tiempo (Universidad de Vigo, s.f.). Se trata de una actividad muy compleja en la que entran en juego una serie de procesos al mismo tiempo (lingüístico, pragmático, cognitivo, cultural...) y, de hecho, es la actividad cognitiva humana más compleja (Ahmed Ahmed, 2022, p. 336). Tras un estudio neurológico que forma parte de la tesis de Kalderon (2023), *Neurofisiología de la interpretación simultánea*, se ha observado que esta modalidad implica una mayor activación de ciertas zonas del cerebro, aunque participan menos áreas cerebrales en comparación con otras prácticas, como el *shadowing*.

En segundo lugar, el acceso a internet a partir de la década de 1990, el mayor repositorio textual disponible en diferentes idiomas y sobre diferentes temas, significó otro hito en la profesión, dado que internet ha permitido a los intérpretes llevar a cabo una exhaustiva labor de documentación previa a la interpretación, desde creación de corpus, investigación del tema en cuestión, búsqueda de equivalentes, etc. (Fantinuoli, 2018, p. 2-3).

Una vez que el concepto de interpretación ha quedado delimitado, procedemos a abordar el que es el segundo pilar de este trabajo, las tecnologías de la interpretación.

4.2. Las tecnologías de la interpretación

4.2.1. Definición y clasificación

En cuanto al empleo de la tecnología en interpretación, es necesario puntualizar que no existe consenso entre los investigadores en lo que se refiere a la clasificación de los recursos tecnológicos en interpretación y a la propia denominación de estas tipologías. En función de la clasificación por la que se opte, la definición de cada término puede verse modificada. En consecuencia, ante las múltiples clasificaciones existentes, en este trabajo nos hemos decantado por la clasificación que propone Prandi (2023, p. 25-38) y que procedemos a describir a continuación.

En primer lugar, diferenciamos las herramientas tecnológicas que median en la interpretación (*technology mediating interpreting*), que ofrecen grandes posibilidades para el intérprete, el orador y los oyentes. Se trata de herramientas que han cambiado los escenarios tradicionales de interpretación o que ha aportado alguna solución dentro de la disciplina. La siguiente categoría corresponde a las tecnologías que generan el resultado de la interpretación (*technologies generating interpreting*), marcadas por los avances en materia de inteligencia artificial (IA) y redes neuronales, lo que las ha convertido en más que una amenaza teórica para los profesionales. El ejemplo más claro es la interpretación automática (MI, por sus siglas en inglés), en la que nos centraremos más adelante (Prandi, 2023, p. 27-31). Por último, distinguimos las tecnologías que ofrecen apoyo durante el proceso (*technologies supporting interpreting*). En este caso, haremos referencia a tres ejemplos de herramientas, que pueden ser *software* o *hardware*. El primer ejemplo son las tabletas para intérpretes, que se han empleado en mayor medida dentro de la modalidad de interpretación consecutiva para la toma de notas, frente al empleo de papel y bolígrafo. Prandi (2023, p. 31-34) hace referencia a las ventajas de esta herramienta frente a los materiales tradicionales, dada la posibilidad de emplear estas tabletas cómodamente en cabina o la batería de larga duración. El segundo de los ejemplos es el bolígrafo digital para interpretación consecutiva y simultánea, que da lugar, precisamente, a una modalidad híbrida de interpretación, puesto que el intérprete toma notas y estas quedan almacenadas de forma sincronizada con el audio, de forma que al realizar el ejercicio de interpretación el profesional puede seguir las al mismo tiempo que reproduce el audio de nuevo. El tercer ejemplo son las herramientas de interpretación asistida por ordenador (CAI, por sus siglas en inglés), las cuales se encuadran dentro de *process-orientated technologies*, ya que el objetivo es reducir la carga cognitiva tanto como sea posible (Fantinuoli, 2023c, p. 50-52). Prandi (2023, p. 34-38) las define como *software* diseñados especialmente para brindar soluciones y satisfacer las necesidades del intérprete a nivel de terminología y de gestión del conocimiento. Se trata de herramientas que pueden servir de ayuda al intérprete no solo en el ejercicio de interpretación en sí mismo, sino también en la fase previa y posterior al propio ejercicio. Existen herramientas

tecnológicas dentro de esta tipología a disposición del intérprete en la fase de adquisición de conocimiento y de preparación, es decir, familiarización con el tema, glosarios multilingües y documentación, todo ello desde un punto de vista lingüístico y del tema. En el caso del propio ejercicio de interpretación, se puede recurrir a una herramienta CAI para recuperar información útil que se ha extraído en la fase anterior, por ejemplo, para resolver una duda terminológica. Por último, en la fase posterior al ejercicio de interpretación, son útiles las herramientas de gestión del conocimiento para reorganizar y sistematizar la información. El intérprete puede, entre otras cosas, ampliar sus bases terminológicas y sus apuntes de documentación de información, añadir notas, etc., lo cual puede ayudar a mejorar la calidad de futuras interpretaciones (Fantinuoli, 2023c, p. 50-52).

Una vez delimitado el campo de las tecnologías de la interpretación, nos gustaría prestar atención a la interpretación remota, dado que desencadenó el desarrollo de la interpretación automática.

4.2.2. La interpretación remota

El concepto de interpretación remota (RI, en sus siglas en inglés), como se ha mencionado anteriormente, queda clasificado dentro las herramientas tecnológicas que intervienen en el proceso de interpretación (*technology mediating interpreting*). Esta tecnología cambia por completo el escenario tradicional de un servicio de interpretación, puesto que elimina la necesidad de que las partes implicadas (intérprete, orador y público) deban estar presentes en el mismo lugar, lo que lo hace más flexible y permite el ahorro de costes y de equipos. Un orador emite su discurso y su imagen y su voz llegan al intérprete a través de una plataforma en la nube diseñada para ello. El intérprete necesitará simplemente un ordenador con cámara, un acceso fiable a internet y unos cascos con micrófono para poder emitir la interpretación a los oyentes, que generalmente estarán también conectados a la plataforma (Interprefy, 2022). Este tipo de interpretación se puede emplear en conferencias, seminarios, pequeñas reuniones e incluso en servicios públicos, en los que se necesita interpretación bilateral y donde el método preferido es el teléfono, especialmente en el ámbito médico y social, seguido del sector de los negocios. Resulta interesante mencionar que hasta ahora el teléfono es el método de interpretación remota más empleado y el más conocido en Europa (Cabrera Méndez *et al.*, 2018, p. 54). Estos autores hacen alusión a una encuesta en la que se recoge que el 44 % de las personas o instituciones que recurren a los servicios de interpretación han hecho uso de interpretación remota y el 79 % lo valora como positivo. Gracias al avance en tecnologías más fácilmente accesibles y fiables, numerosas herramientas hacen posible esta modalidad de interpretación (Cabrera Méndez *et al.*, 2018, p. 55), que se inició en la década de los 70 del siglo pasado con el objetivo de minimizar los costes a nivel económico, acabar con la dificultad que puede surgir a la hora de encontrar intérpretes disponibles en el lugar

en el que se necesita el servicio en un momento determinado y reducir la carga de trabajo a nivel de organización (Ziegler y Gigliobianco, 2018, p. 119-123). En concreto, existen empresas que ponen a disposición de intérpretes *software* de pago que les permiten dar este servicio con un equipo más sencillo, desde casa y conectados con personas a grandes distancias sin necesidad de invertir en desplazamiento ni estancias. Por otra parte, estas empresas cuentan con una serie de intérpretes entre los que el cliente puede elegir para el servicio que necesita con cita previa o, en caso de urgencias, la propia empresa proporciona al cliente el intérprete más adecuado que esté disponible en ese momento (Interprefy, 2022). Por último, es preciso indicar que, desde sus inicios, se han llevado a cabo numerosos experimentos para evaluar esta modalidad y, en muchas ocasiones, los profesionales han coincidido en que el hecho de no estar presentes en la situación comunicativa como tal les provocaba una sensación de no tener bajo control la situación y experimentar más estrés, dolor de cabeza, dificultad de concentración y fatiga visual (Ziegler y Gigliobianco, 2018, p. 123-139).

4.2.3. La interpretación automática

Una vez revisado el concepto de interpretación remota, nos ocuparemos ahora de la interpretación automática, que, como se ha explicado anteriormente, es el claro ejemplo de herramientas tecnológicas que generan interpretación (*technologies generating interpreting*), es decir, sistemas que ofrecen el producto final de la interpretación (Prandi, 2023, p. 28-31). Antes de atender a la definición que nos ofrece Fantinuoli para «interpretación automática», cabe destacar que el término «interpretación automática» (*machine interpreting*) es el término empleado en el ámbito de los Estudios de Traducción, mientras que en ciencias de la computación el término empleado es «traducción de voz a voz» (*speech-to-speech translation*) (Fantinuoli, 2024). De ahora en adelante, fieles al campo que nos ocupa, emplearemos el término «interpretación automática», definida como: «...an automated language translation process that converts spoken content from one language to another in the form of speech» (Fantinuoli, 2023b, p. 10).

Este tipo de interpretación puede realizarse en directo o no. En caso de no ser en directo, se lleva a cabo la interpretación de un discurso que ha sido grabado con anterioridad, pero la interpretación automática se produce propiamente cuando el discurso tiene lugar en directo (Fantinuoli, 2024).

La interpretación automática puede ser consecutiva o simultánea. En la modalidad de consecutiva, existe la ventaja de que se cuenta con el texto por adelantado y que no se transmite con inmediatez absoluta, por lo que es posible revisar y editar el resultado final que nos ofrece el sistema. En cambio, en el caso de la interpretación automática simultánea el reto es mayor, puesto que el sistema no conoce con antelación el texto y la interpretación se produce al mismo tiempo

que el discurso original. Esto significa que se debe buscar constantemente el equilibrio entre latencia y adecuación (Fantinuoli, 2024). Para lograrlo y sufragar esa complejidad, el discurso se divide en segmentos con significado en función, por ejemplo, de pausas del orador o de análisis sintácticos en tiempo real (Fantinuoli, 2023b, p. 10). En el presente trabajo, nos centraremos en la modalidad de simultánea.

Por otro lado, es posible distinguir al menos dos tendencias dentro de la interpretación automática (Fantinuoli, 2023a):

1. Sistemas en cascada: reciben este nombre porque en esta tendencia tienen lugar tres procesos que suceden uno después del otro:
 - Reconocimiento automático de voz (ASR, por sus siglas en inglés): transcribe el mensaje oral a escrito.
 - Traducción automática (MT, por sus siglas en inglés): traduce el mensaje escrito de una lengua a otra.
 - Conversión de texto a voz (TTS, por sus siglas en inglés): da una versión oral de la traducción escrita.

Se trata del procedimiento que encontramos en la aplicación de Google Translator. Sin embargo, estos sistemas implican algunas deficiencias, ya que emplean grandes modelos de traducción entrenados con recursos escritos, lo que hace que puedan generarse ciertos errores como consecuencia de no resolver correctamente fenómenos del lenguaje hablado en el primer paso, errores que se arrastran en el siguiente proceso de traducción automática y sucesivamente (Fantinuoli, 2023b, p. 10).

2. *End-to-end approach*: a diferencia del modelo anterior, este traslada el mensaje oral de una lengua a un mensaje oral en otra lengua sin necesidad de transcripción, algo que hace que el proceso sea más rápido y se disminuya el riesgo de errores. Por el momento, se trata de un método experimental (Fantinuoli, 2023a).

Anteriormente se ha mencionado que la interpretación simultánea es una actividad de gran complejidad para el cerebro humano, de modo que resulta interesante atender a los procesos fisiológicos que tienen lugar en la figura del intérprete y al funcionamiento de un sistema de interpretación automática en cascada, que será la tendencia elegida para este trabajo. La primera fase es la de escucha y comprensión, en la que el mensaje codificado llega en forma de ondas al oído del intérprete y el sistema nervioso lo conduce al cerebro, donde se selecciona la información más importante al circular por las neuronas, cuando el mensaje todavía está codificado. Los sistemas de traducción e interpretación automática tratan de imitar estos mecanismos y, a pesar de la complejidad del proceso, se han sucedido importantes avances en este sentido. En particular, la memoria es un elemento fundamental para la interpretación y es precisamente un punto

importante que los sistemas de traducción e interpretación automática pretenden imitar con las memorias de traducción. Mediante un proceso de reconocimiento automático de voz a texto, la máquina convierte el mensaje oral a escrito y lo descodifica para prepararlo para la fase de traducción automática. Por el contrario, en el caso del intérprete humano, el mensaje permanece codificado en esta fase. En el siguiente proceso, el de reformulación, el mensaje se descodifica y se transporta en forma de señales desde el área Broca y el cortex prefrontal del hemisferio izquierdo a través de neuronas y nervios hasta llegar al aparato fonador para dar lugar al producto final. En el caso de la interpretación automática, es una máquina la que se encarga de traducir el mensaje para estar listo para convertirse en un mensaje oral (Ahmed Ahmed, 2022, p. 336 - 340).

Se han producido rápidos avances en este ámbito, lo que ha permitido que la interpretación automática se aplique en la práctica, no solo en ocasiones puntuales del plano social, sino también en escenarios más profesionales, como una conferencia. Precisamente se están desarrollando avances que pretenden abordar los retos que supone la complejidad del lenguaje humano para la interpretación automática. Sin embargo, es preciso mencionar que hay elementos del lenguaje, como el lenguaje no verbal, la intención del orador, el contexto o segundos sentidos, además de situaciones que requieren empatía, que hacen que la presencia de un intérprete humano sea fundamental. (Fantinuoli, 2023c, p. 58)

A lo largo de la historia, los profesionales de los diferentes sectores han sido en menor o mayor medida reacios al empleo de la tecnología en sus labores, al tomarla como una amenaza para sus puestos de trabajo. Se trata de una reacción que también se ha podido observar en el área de Traducción e Interpretación. De hecho, el surgimiento de la interpretación simultánea y el empleo de herramientas tecnológicas que ello generó cierta reticencia entre los intérpretes que venían practicando la interpretación consecutiva, ya que consideraron que ese cambio podría suponer una pérdida de calidad en la interpretación y en el proceso de comunicación. Sin embargo, la interpretación simultánea ha llegado a ser la predominante en la interpretación de conferencias (Prandi, 2023, p. 26). En la actualidad, conviven diferentes posturas sobre el empleo de los diferentes recursos tecnológicos mencionados en interpretación simultánea, en especial con respecto a la interpretación automática. Hay profesionales que asumen que la tecnología supone una amenaza que empieza por desprofesionalizar la disciplina, afectar a la calidad del servicio y convertirla en una profesión más descualificada y que podría terminar por sustituir a los intérpretes. La tecnología y el uso de herramientas CAT y CAI podría afectar negativamente al *status* y al valor que se otorga a los profesionales. Hay posturas que indican que el uso excesivo de la tecnología en la disciplina puede acabar por minar la capacidad de pensamiento crítico, analítico y creativo de traductores e intérpretes. Por otro lado, otros profesionales hacen alusión a las ventajas que brinda la tecnología a la interpretación simultánea, dado que evita tener que desplazarse, libera a los intérpretes de mayores cargas de trabajo y reduce los costes en general.

Sin embargo, existen opiniones más neutrales que aseguran que, al menos en un futuro cercano, los sistemas tecnológicos no sustituirán a los traductores ni intérpretes humanos, sino que servirán como herramienta útil para la profesión. Los profesionales seguirán ocupando un papel fundamental y dando lugar a traducciones e interpretaciones de gran calidad, aunque se redefinan sus roles, por ejemplo, ayudando a mejorar las herramientas CAT (Ahmed Ahmed, 2022, p. 327-329). A raíz de esta polémica, Fantinuoli (2019, p.13-14) aboga por la investigación en esta área con el objetivo de formar a futuras generaciones de intérpretes y que estos estén preparados para posibles cambios en la profesión desencadenados por la tecnología.

Una vez definidos los conceptos que resultan fundamentales para la realización de este trabajo, procedemos a explicar en el capítulo número cinco, *Metodología de análisis*, el procedimiento que se va a emplear para analizar los productos resultantes.

5. Metodología de análisis

En este apartado se recoge la selección del sistema de interpretación automática, las muestras que serán objeto de análisis, los criterios de evaluación de la interpretación de las muestras escogidas realizada por el sistema de interpretación de automática y el procedimiento de recogida y análisis de datos.

5.1. Sistema de interpretación automática

En primer lugar, resulta interesante realizar una descripción del sistema de interpretación automática que se va a emplear para analizar las dos muestras de análisis. En este caso, el *software* elegido es Interprefy, una plataforma que pone a disposición del usuario un servicio de interpretación remota e interpretación automática. Se trata del único sistema de interpretación remota del que la Facultad de Traducción e Interpretación tiene licencias y este ha sido el principal criterio de selección. A través de su plataforma en la nube, los intérpretes pueden trabajar de forma remota y simultánea para clientes a grandes distancias y en contextos multilingües, como seminarios, pequeñas reuniones o conferencias. Además de proporcionar al cliente el intérprete más adecuado para cada situación según la combinación de idiomas y el grado de especialización del profesional en el tema del encargo, el sistema también ofrece formación a los profesionales. Por último, cabe mencionar que, desde su creación en 2015, Interprefy ha prestado sus servicios a empresas como la gigante alemana Siemens o la biofarmacéutica británica GSK Glaxosmithkline, a instituciones como la Casa Blanca y se ha empleado en eventos como la Uefa Euro 2020 (Interprefy, 2024).

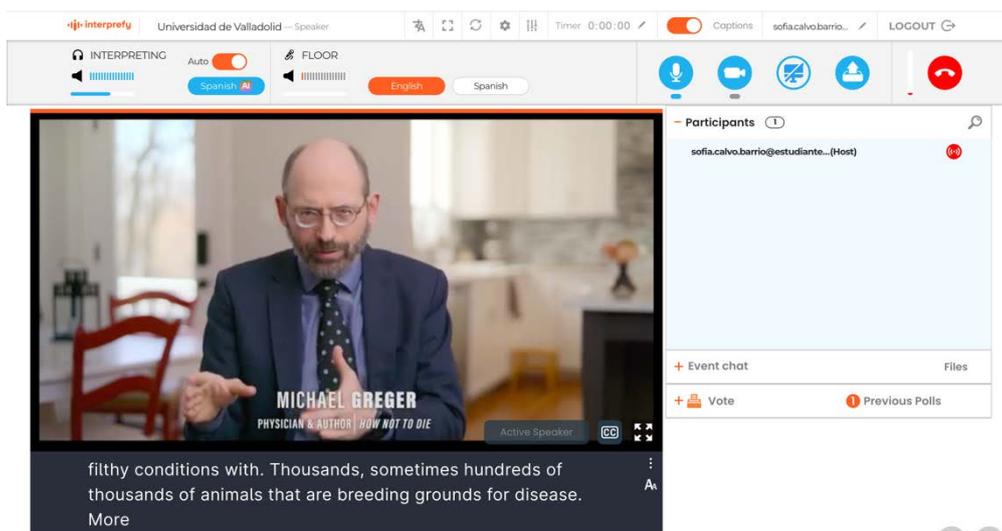


Ilustración 1. Interfaz de Interprefy

5.2. Muestras de análisis

Las muestras de análisis están constituidas por dos fragmentos de audio en lengua inglesa (versión original) extraídas del episodio número tres de la serie documental de cuatro episodios *Somos lo que comemos*, emitida desde 2023 en la conocida plataforma de *streaming* Netflix. Esta serie documental da a conocer varias parejas de gemelos idénticos que cambian su dieta y su estilo de vida durante ocho semanas como parte de un estudio científico que pretende examinar el impacto de la alimentación en la salud (Netflix, 2023). A la hora de decantarnos por esta miniserie para obtener las muestras, ha sido fundamental el hecho de que la alimentación relacionada con temas medioambientales y de salud es un asunto muy actual que podemos encontrar de forma frecuente en los medios de comunicación.

Hemos escogido dos muestras de audio de diferente grado de especialización de aproximadamente 30 segundos de duración cada una. Como muestra de audio con mayor grado de especialización (Muestra 1), se ha seleccionado el fragmento desde 7'31'' hasta 8'02''. En este fragmento, una abogada y consejera por la defensa de los recursos naturales y un físico y escritor realizan cada uno una intervención individual sobre el uso de antibióticos en ganadería. Puesto que se trata de expertos y dada la presencia de términos especializados, hemos elegido finalmente esta parte del episodio. La segunda de las muestras (Muestra 2) se corresponde con el fragmento desde 19'06'' hasta 19'39'' como muestra con menor grado de especialización. En este caso, uno de los gemelos de las 22 parejas que protagonizan esta miniserie mantiene una conversación con una amiga que ha seguido una dieta vegana desde niña mientras cocinan y le pide consejos para iniciarse en la cocina vegana. En este contexto, una nutricionista experta en materia de salud pública aporta algunas explicaciones a modo de recomendaciones y parte de su intervención es la que ocupa este fragmento de audio que será objeto de análisis. Se trata de una intervención con vocabulario cotidiano, en su mayoría relacionado con el mundo de los alimentos y la cocina, como parte de una explicación sencilla.

5.3. Modelo de evaluación de interpretación simultánea

A la hora de determinar los parámetros que se tendrán en cuenta en la evaluación de la interpretación de las dos muestras de análisis, se ha tomado como referencia la propuesta de Ortiz Boix (2016) y la de Collado Aís (2007).

En primer lugar, los parámetros que sugiere Ortiz Boix (2016) en su tesis doctoral *Implementing machine translation and post-editing to the translation of wildlife documentaries through voice-over and off-screen dubbing* están orientados a la evaluación de traducciones realizadas con traducción automática, de modo que hemos procedido a adaptar la propuesta a la modalidad de interpretación automática. A pesar de que se trata de un baremo para traducción

automática, es preciso tener en cuenta que este es uno de los tres escenarios presentes en la interpretación automática en cascada, de modo que tiene cabida aplicar algunos de estos aspectos a este análisis, aunque con algunas modificaciones.

En segundo lugar, tomamos como referencia los parámetros que sugiere Collado Aís (2007) en su obra *La evaluación de la calidad en interpretación simultánea. La importancia de la comunicación no verbal*. Dado que este modelo está diseñado para evaluar el trabajo de un intérprete profesional y con el objetivo de adaptarlo a las características y a la situación de un sistema de interpretación automática, tomaremos como referencia una serie de aspectos, añadiremos algunos y modificaremos o eliminaremos otros para ajustar los parámetros a la interpretación automática.

A continuación, se recogen todos los parámetros que formarán parte del modelo de evaluación que seguiremos en el presente trabajo. Por un lado, figuran en la columna de la izquierda los parámetros lingüísticos y, por otro lado, los parámetros que atañan al plano de la paralingüística se encuentran en la columna de la derecha. Además, mostramos el código de colores empleado para señalar los errores y obtener una visión global de los mismos. Por último, comentaremos brevemente cada uno de ellos y justificaremos su presencia, así como las razones para modificar algunos aspectos y prescindir de otros.

Aspectos lingüísticos		Aspectos paralingüísticos
Terminología		Entonación
Traducción incorrecta	Demasiado literal	Fluidez
	Calco	Pronunciación
	Falso sentido	Latencia de habla
	Cifra	
	Nombre propio	
	Traducción innecesaria	
Omisión		
Adición innecesaria		
Elemento no traducido		
Uso incorrecto de la gramática		
Registro o estilo inadecuado a la situación comunicativa		

Tabla 1. Modelo de análisis con código de colores

En lo que se refiere a aspectos lingüísticos, el primero que distinguimos es la terminología. Catalogaremos como error de terminología la elección de un término que no es el

adecuado por no ajustarse al campo, a la situación comunicativa o por no ser el que el cliente había especificado. Posteriormente, encontramos la categoría de traducción incorrecta en la que se incluyen seis parámetros diferentes, todos ellos relacionados con un desajuste de contenido entre el texto original y el texto meta. Una traducción demasiado literal que recuerde a una traducción palabra por palabra sin prestar atención al sentido se considerará un error. En caso de que en la interpretación se haga uso de una palabra o una estructura que se basa en el texto original o es similar al original, se catalogará como error de calco. En cuanto a las cifras, se espera que coincidan las cifras que se ofrecen en el original con las de la interpretación, ya sean fechas, porcentajes, etc. Algo parecido se espera de los nombres propios, por ejemplo, de instituciones o lugares, pues deben coincidir con los que figuran en el original o, en caso de que sea necesario traducirlos, la traducción debe ser la adecuada o la oficial. En ambos casos, es imprescindible que no exista desajuste entre el original y la interpretación para garantizar el cumplimiento del principio de fidelidad. Por último, dentro de esta categoría, la traducción de elementos que no se deberían haber traducido y deberían permanecer en lengua origen en el discurso meta se tomará como error de traducción innecesaria.

Antes de continuar con la descripción del resto de parámetros, nos detendremos a comentar brevemente qué cambios se han realizado con respecto al modelo propuesto por Ortiz Boix (2016). Para ello, retomamos el parámetro de calco, ya que en el modelo original este parámetro recibe del nombre de *false friend*, un anglicismo que sirve para referirse a aquellas palabras a las que, por su parecido ortográfico a otra palabra de una lengua diferente, se les atribuye un significado erróneo. Se ha decidido ampliar este parámetro a calco con la finalidad de tener la posibilidad de abarcar un mayor número de elementos más allá de los heterosemánticos, como por ejemplo los calcos de estructuras sintácticas. Por otra parte, otro parámetro presente en el modelo de referencia es el de conversión de unidades de medida. Dado el rasgo de inmediatez que caracteriza a la interpretación, no se espera ni de un intérprete humano ni de un sistema de interpretación automática que sea capaz de adaptar las unidades de medida a la situación comunicativa, pues esto añade una carga cognitiva excesiva. Por lo tanto, se ha decidido prescindir de este aspecto en nuestro modelo, como parte de la necesidad de adaptarlo a la modalidad de interpretación. El siguiente parámetro que se ha tomado del modelo de Ortiz Boix (2016) es el de omisión, en el que se incluyen todas aquellas informaciones presentes en el discurso origen que no se trasladan al discurso meta, lo que hace que no se transmita el mensaje de forma completa. Por otra parte, clasificaremos como error por adición innecesaria aquellas informaciones que se incluyan en el discurso meta, pero que no se mencionen en el texto origen y no sea necesario que se añadan. El siguiente parámetro lingüístico que forma parte de nuestro modelo hace referencia a los elementos no traducidos, es decir, aquellos que permanecen en lengua origen en el discurso meta, a pesar de que se deberían haber traducido. Asimismo, se

tendrán en cuenta aquellos errores que tengan que ver un uso incorrecto de la gramática de la lengua meta, ya que se trata de una expectativa que, con seguridad, comparten la mayoría de los receptores y que en gran medida contribuye a la transmisión correcta del mensaje, puesto que en ocasiones los errores gramaticales pueden llevar a incurrir en errores a nivel de contenido. Por último, se prestará atención a los errores por falta de adecuación de estilo y registro a la situación comunicativa, un elemento de adición propia cuyo objetivo es subrayar la importancia de identificar y analizar la situación comunicativa y, más en concreto, los receptores del mensaje para seleccionar el registro y estilo adecuados.

En base al modelo propuesto por Collado Aís (2007), se han tomado una serie de aspectos paralingüísticos para adaptar el nuestro al rasgo oral de la interpretación. En primer lugar, se ha tomado del modelo de referencia el parámetro de entonación, así como el de fluidez. Por su parte, la entonación alberga casi tanta carga de significado como las palabras, de modo que resulta fundamental reproducir la entonación del discurso original para favorecer a dejar clara la intención del orador. La fluidez tiene que ver con la doble faceta que tiene un intérprete, en este caso el sistema de interpretación automática: por un lado, la función fundamental de comunicador y, por otro lado, la labor de generar un discurso agradable y entretener al público, para lo que es imprescindible que haya fluidez en la emisión del discurso. Uno de los elementos que puede contribuir a restar fluidez al discurso son las pausas innecesarias, sobre todo si su duración es considerable. Hemos optado por prescindir del aspecto del acento nativo, presente en el modelo de la autora, y sustituirlo por pronunciación. El acento resulta ser algo muy abstracto y más bien propio de un intérprete humano que de un sistema de interpretación automática, ya que cabe esperar un acento neutro de un sistema automático. Además, una correcta pronunciación y vocalización son de mayor relevancia a la hora de contribuir a la comprensión del mensaje y a un aspecto que, precisamente, procedemos a introducir a continuación. Otros de los parámetros que se ha decidido no emplear en este modelo son la impresión de fiabilidad y seguridad, así como la agradabilidad de la voz, dado que se trata de aspectos que se relacionan más fácilmente con las actitudes humanas y, puesto que este trabajo se centra en la labor de interpretación que realiza un sistema automático, resulta innecesario atender a estos parámetros. Por último, relacionado con la velocidad de habla, incorporamos un nuevo aspecto al análisis: latencia de habla. A propósito de los aspectos paralingüísticos que se han señalado anteriormente, la velocidad a la que se transmite la interpretación y cómo se adecúa esta a la velocidad de transmisión del original son aspectos dignos de analizar, ya que nos ofrecerán mucha información sobre la gestión que realiza el sistema del tiempo y las pausas.

Con todos los parámetros que constituyen este modelo de evaluación de un sistema de interpretación automática ya explicados, es necesario indicar cuál va a ser el método de análisis que se va a seguir para dicha evaluación. Una vez se han obtenido las muestras de análisis de la

interpretación de los dos fragmentos de audio, dividiremos la transcripción del discurso origen y del discurso meta en segmentos alineados en un documento Excel de Microsoft para poder señalar sobre la propia transcripción aquellos errores de carácter lingüístico y paralingüístico que se detecten, para lo que será fundamental tener muy presentes las muestras de análisis en audio en lengua origen y meta. Para señalar los errores, se empleará un código de colores, de forma que cada color corresponderá a un parámetro diferente y será más fácil diferenciar los distintos errores y la categoría a la que pertenecen.

6. Análisis y resultados

En este apartado, ofreceremos un análisis de los errores identificados en la interpretación automática al español de las dos muestras de audio originales en inglés. Para ello, como se ha especificado en el capítulo anterior, emplearemos un código de colores que permita señalar cada error según el parámetro del modelo con el que se corresponda.

6.1. Muestra de análisis 1

En la Tabla 2 se muestran los errores detectados en la primera de las muestras, que se explicarán a continuación.

EN	ES
<i>More antibiotics are fed to farm animals than are used in all of human medicine</i>	Se administran más antibióticos para formar (1) animales que los que se utilizan en medicina humana (2)
<i>just to promote growth and prevent disease in such a stressful, unhygienic environment.</i>	solo para promover (3) el crecimiento y prevenir enfermedades en un entorno tan estresante y antihigiénico.
<i>One of the side effects of giving animals low levels of antibiotics in their feed or injected</i>	Uno de los efectos secundarios de dar a los animales niveles bajos de antibióticos en su alimentación o (4) inyectado (5)
<i>is that they become a breeding ground for antibiotic resistant bacteria,</i>	es que se convierten en un caldo de cultivo (6) para bacterias resistentes a los antibióticos (7),
<i>things that can cause infections that are hard or impossible to treat.</i>	cosas que (8) pueden causar infecciones que son difíciles o imposibles de tratar.

Tabla 2. Errores detectados en la muestra de análisis 1

(1) Se trata de un error de falso sentido, que seguramente se ha producido como consecuencia de un error en la fase de reconocimiento de voz a texto (*form*, «formar», en lugar de *farm*). Cabe puntualizar que los antibióticos no se administran para «formar animales», sino para fomentar su crecimiento. En cualquier caso, se emplea *farm* para hacer referencia a los «animales de granja» o «ganado». Por último, resulta interesante mencionar que *farm animals* es un *cluster*, una de las mayores dificultades que supone la interpretación inglés-español, ya que es necesario esperar a que se mencione el último elemento para poder empezar a traducir el conjunto, lo cual puede ocasionar un problema a nivel de latencia de habla.

(2) Estamos ante un error de entonación, puesto que el tono refleja que estos dos segmentos son unidades independientes, cuando ambas conforman una oración y una unidad de significado completa.

(3) Como se ha podido comprobar en textos y recursos sobre este tema, el equivalente más adecuado para *promote* es «estimular». Se trata, por tanto, de un error de terminología.

(4) Se realiza una pausa innecesaria que aísla los elementos unidos por la conjunción disyuntiva «o». En consecuencia, se puede catalogar como un error de fluidez.

(5) Nos encontramos nuevamente ante una pausa innecesaria demasiado larga, que, además, separa el sujeto del resto del predicado, lo cual afecta a la fluidez del discurso.

(6) Dado que nos encontramos ante la intervención de un experto y en la parte del episodio en la que se transmite información de corte más especializado, quizá sería interesante adaptar el registro de esta expresión idiomática («caldo de cultivo») a un tono más formal y sin connotaciones metafóricas. En este caso, sería posible hablar del concepto de propensión.

(7) En este caso, se trata de un error de entonación, pues el tono no refleja la intención con la que se explica la siguiente información, que es una aclaración a modo de información adicional de lo que se ha dicho anteriormente. Sin embargo, la entonación en la interpretación realizada por el sistema podría dar a entender que no son informaciones relacionadas.

(8) Diferenciamos un error por traducción demasiado literal, ya que se ha reproducido la estructura del original, en una traducción palabra por palabra, y, aunque esto no supone ningún error en la parte intermedia y final de la oración, en la parte del principio hace que no se adecúe al registro y carezca de idiomática. Para subsanar este error, es posible recurrir a partículas de relativo como «lo que» o «lo cual».

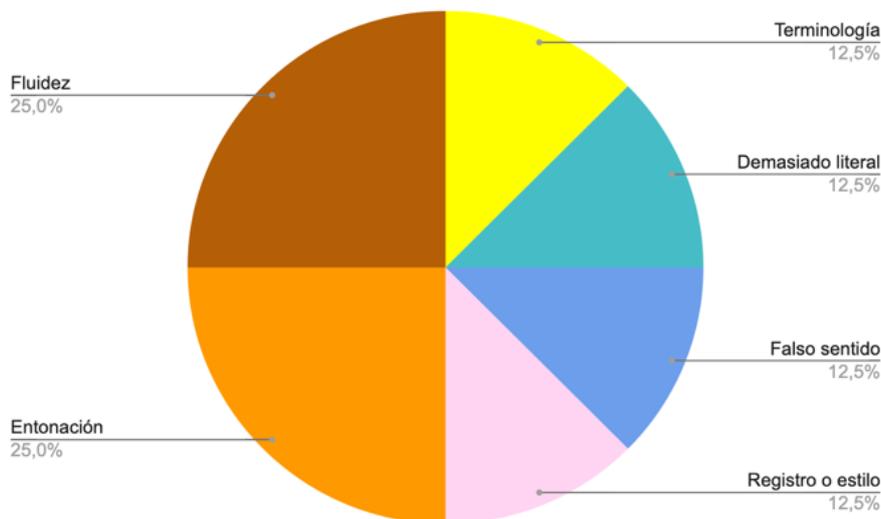


Ilustración 2. Gráfico de resultados de la muestra 1

La siguiente gráfica ofrece un análisis cuantitativo de los errores que hemos detectado en esta primera muestra de análisis.

A partir de los datos que nos ofrece este gráfico, se desprende que los principales errores se corresponden con los parámetros de entonación (25,0 %) y fluidez (25,0 %) de nuestro modelo de análisis y el resto tienen la misma representación con respecto al total de los errores.

6.2. Muestra de análisis 2

A continuación, procedemos a replicar la misma metodología de análisis en la segunda de las muestras.

EN	ES
<i>What you're doing is using the same ingredients, but changing the foundation of the meal.</i>	Hacer es usar (1) los mismos ingredientes, pero cambiando la base de la comida (2).
<i>For example, if you have a nine inch plate, you might have bean Stew, right? Or a chili.</i>	Por ejemplo, si tiene una placa (3) de nueve pulgadas (4), es posible que la tenga (5). (6) Guiso de frijoles, ¿verdad? O un chile (7).
<i>And that might be about a third of your plate.</i>	Y eso podría ser aproximadamente (8) un tercio de su (9) plato.
<i>And then half of your plate would be some form of vegetables.</i>	Y luego la mitad de su plato sería un tipo de verduras, (10)
<i>Sauteed in a salad. Steamed.</i>	Salteado (11) en una ensalada, (12) al vapor.
<i>And the other part you want to be a whole grain brown rice, black rice, wild rice, quinoa, Millet, bulgur.</i>	Y la otra parte quieres que sea (13) un (integral integral (14), (15) arroz negro, arroz salvaje, quinoa, mijo, bulgar (16).

Tabla 3. Errores detectados en la muestra de análisis 2

(1) Esta traducción carece de sentido completamente y cabe afirmar que al receptor le resultaría imposible comprender el mensaje, aún contando con el contexto de los minutos de audio anteriores. Una alternativa sería la siguiente: «Se trata de usar los mismos ingredientes, pero cambiar el planteamiento del plato». En este caso, resulta complicado encuadrar este error en alguno de los parámetros del modelo, ya que va mucho más allá de un falso sentido y el error no tiene su origen en ajustarse o traducir el original palabra por palabra. Por ello, podría clasificarse como traducción incorrecta, como categoría general.

(2) Con el objetivo de favorecer la comprensión de este concepto, realizamos la siguiente propuesta para *the foundation of the meal*: «el planteamiento del plato». Por un lado, el contexto nos permite saber que *meal* hace referencia al modelo de plato que explica, entendido como

«comida preparada para ser consumida»². Por otra parte, *foundation* se refiere a la idea y los objetivos con los que se seleccionan unos determinados ingredientes. Teniendo en cuenta todo esto, «el planteamiento del plato» sería una opción más precisa, clara y coherente con el original.

(3) En este caso, diferenciamos un error de terminología, ya que el equivalente de *plate* es «plato» y no «placa», por tratarse de un contexto de alimentación y cocina.

(4) Diferenciamos un error de latencia de habla, pues el tiempo de desfase es demasiado largo. La oradora principal ya ha pronunciado la primera oración y parte de la segunda cuando el sistema comienza a interpretar la primera, de modo que llega un punto que aumenta en exceso la latencia de habla para poder adecuarse al original, pero sin que en el discurso original se aumente la velocidad de habla.

(5) Se trata de un error considerablemente grave de falso sentido, dado que, además de no reproducir el sentido del original, no se ha enlazado con los elementos siguientes y ambas partes se han interpretado como unidades de significado diferentes. Para ayudar a eliminar el falso sentido, es necesario llevar a cabo una reformulación que aleje la estructura oracional de una oración condicional. Una propuesta de traducción correcta es la siguiente: «Por ejemplo, en un plato de nueve pulgadas, puede incluir / puede comenzar con un guiso de frijoles o chili».

(6) Este error, que se corresponde con el parámetro de fluidez, está estrechamente relacionado con el siguiente error de entonación. En este caso, se ha realizado una pausa innecesaria entre el verbo principal de la oración y su complemento directo, lo cual afecta no solo a la fluidez del discurso, sino que también puede suponer una dificultad de comprensión para los receptores.

(7) Como se ha adelantado en el comentario anterior, este séptimo error tiene que ver con la entonación. El tono con el que se pronuncian las sugerencias de legumbre para este plato no es el adecuado, pues se espera que la entonación en esta parte de la oración de a entender que se trata de una pequeña enumeración a modo de sugerencia, con el rasgo de oralidad de *right*, como coletilla interrogativa.

(8) Distinguimos un error de pronunciación: «*aproximadamente».

(9) Primeramente, es preciso tener en cuenta que el empleo de artículos posesivos es mucho más común en inglés que en español. Por lo tanto, estamos ante un error de calco, pues al reproducir la estructura de inglés con el uso del artículo posesivo, se ha incurrido en un uso innecesario del posesivo equivalente en español. Este error se repite justo en la siguiente oración.

² <https://dle.rae.es/plato> (Fecha de consulta: 15/05/2024)

(10) En este caso, identificamos un error de fluidez, puesto que la información que figura después de la coma es una aclaración sobre el sintagma «tipo de verduras», por lo que no es necesario realizar una pausa tan larga y la entonación debe dejar clara esa idea de aclaración.

(11) Se trata de un error de uso incorrecto de la gramática, ya que el adjetivo «salteado» no concuerda en género ni número con el sustantivo al que modifica, «verduras».

(12) Se realiza una pausa innecesaria demasiado larga entre estos dos elementos de la enumeración, lo que resta fluidez al discurso y no ayuda a dejar claro que se está realizando una enumeración de opciones.

(13) Diferenciamos un nuevo error de falso sentido. A pesar de que se emplea el verbo *want*, lo que se quiere expresar es una propuesta de ingredientes para esta última parte del plato. Por ello, proponemos la siguiente alternativa: «La otra parte del plato podría constar de...».

(14) Se trata de un error de terminología, pues el equivalente de *whole grain* es «grano integral». De hecho, para facilitar la comprensión y dado que es el hiperónimo de los distintos tipos que se señalan a continuación, podría añadirse «algún tipo de grano integral»

(15) Como ocurría en un caso anterior, la entonación que se da en este segmento no es la adecuada, pues no ayuda a entender la idea de que se van a introducir varios ejemplos de grano integral a modo de sugerencia.

(16) Estamos nuevamente ante un error de terminología. El equivalente de *bulgur* no es «bulgar», sino «bulgur».

Una vez comentados los errores detectados en esta segunda muestra de análisis, ofrecemos el siguiente gráfico con el objetivo de tener una visión general y cuantitativa de los errores identificados.

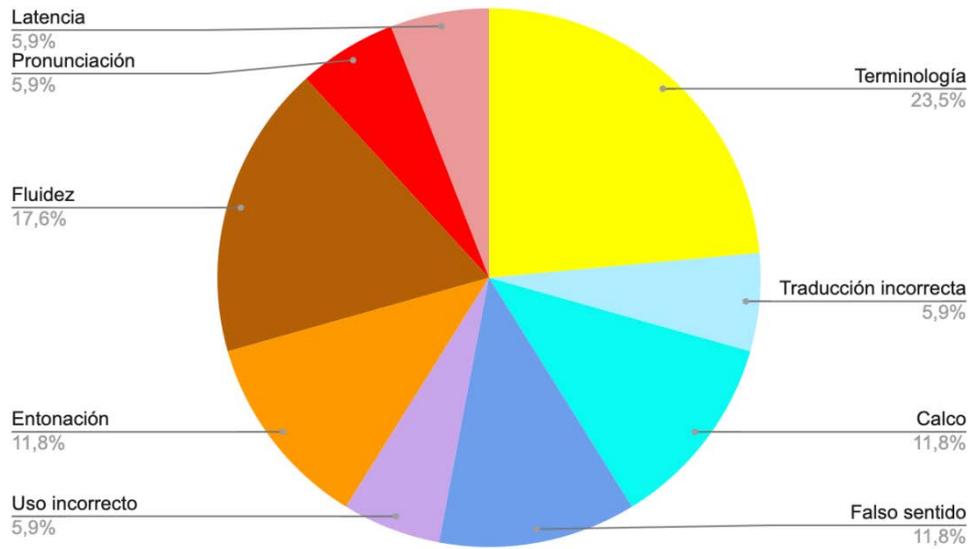


Ilustración 3. Gráfico de los resultados de la muestra 2

Los datos recogidos en este gráfico nos permiten concluir que la categoría de error predominante es la terminología (23,5 %), seguida de la fluidez (17,6 %). Por otra parte, los errores de calco, falso sentido y entonación representan cada uno un 11,8 %, mientras que los errores relacionados con la latencia de habla, la pronunciación, la traducción incorrecta y el uso incorrecto de la gramática están representados en un 5,9 % cada uno.

6.3. Contraste de resultados

Para poder realizar una comparación entre ambas muestras de análisis y obtener una visión en conjunto de los resultados, observamos y comentamos el siguiente gráfico.

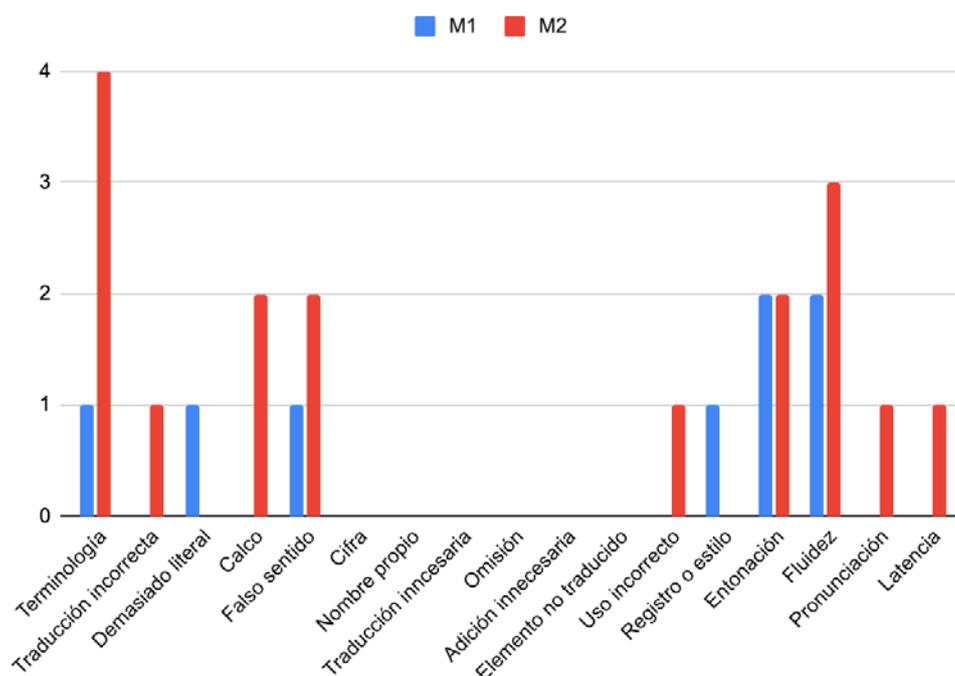


Ilustración 4. Gráfico para el contraste de resultados de M1 y M2

Tras haber realizado el análisis de la interpretación al español de ambas muestras de análisis de acuerdo con el modelo propuesto en este trabajo, resulta interesante, en primer lugar, reconocer la importancia de cada proceso del modelo en cascada de interpretación automática para el resultado final.

En concreto, a partir del error de falso sentido de la primera muestra de análisis, hemos podido comprobar que un rendimiento adecuado en la fase de reconocimiento automático de voz es de vital importancia para que la información que llega a la fase de traducción automática sea rigurosamente fidedigna con respecto al discurso original para la optimización del resultado final. De lo contrario, un error en este primer proceso repercutiría en los siguientes y, por consiguiente, en el producto final.

Por otra parte, hemos observado que la mayoría de errores en ambas muestras están relacionados con los parámetros de entonación, fluidez y latencia de habla. Cabe destacar que, en la mayor parte de los ejemplos, hemos podido identificar que una entonación errónea o la realización de pausas innecesarias puede dificultar el proceso de comprensión, puesto que no se arroja claridad sobre cuál es la relación entre los elementos de la oración o incluso se da a entender

que la oración ha finalizado, cuando todavía tiene continuación. Todo esto resta fluidez al discurso y hace que pueda resultar menos agradable para el receptor. Además, es imprescindible señalar que el tiempo de desfase, esto es, los segundos que pasan desde que el orador principal pronuncia una unidad de significado hasta que el intérprete comienza a interpretarla en lengua meta, es en la mayoría de los casos excesivo. Como resultado, no se aprecia regularidad en la latencia de habla, ya que, como consecuencia del tiempo de desfase excesivo, se debe aumentar posteriormente la latencia de habla para poder seguir el ritmo del discurso, aunque la latencia de habla del discurso original se mantenga homogénea sin sufrir aceleraciones, en cuyo caso sí sería necesario que el intérprete aumentara también la velocidad de habla. En el caso concreto de estas dos muestras de análisis, esta irregularidad en la latencia de habla se hace más evidente en la muestra 2, aunque los errores de entonación y fluidez son comunes en ambos casos.

Otros de los errores más frecuentes en ambas muestras tienen que ver con desajustes en la traducción. Estos son más numerosos en la muestra de análisis con menor grado de especialización. En la primera muestra de análisis, al tratarse de una intervención de mayor especialización y no encontrar muletillas ni demasiados rasgos de oralidad, como sí ocurre en la segunda, la dificultad a la hora de estructurar las oraciones, así como el riesgo de poder ajustarse a la estructura del original sin caer en errores es menor. Por otra parte, como consecuencia de una falta de conocimiento temático y sobre la situación comunicativa, que permitieran alejarse del texto origen para sumar idiomática y naturalidad al discurso meta, se producen distintos casos de traducción literal, calco, falso sentido, así como traducción incorrecta. Asimismo, como resultado nuevamente de la falta de conocimiento del contexto, seleccionar el registro y el estilo adecuados también supone una dificultad para el motor de traducción automática, lo cual se hace evidente únicamente en la muestra 1.

Con respecto a la terminología, hay algunos casos en los que directamente no se ha optado por el término correcto, mientras que en otras se podría haber recurrido a opciones más adecuadas o más frecuentes en uso. Sin embargo, se detecta un mayor número de desajustes terminológicos en la segunda muestra en comparación con la primera, de mayor especialización.

7. Conclusiones

Una vez finalizado el proceso de análisis y reflexión, es posible afirmar que el resultado obtenido de la interpretación automática de estas dos muestras de análisis es aceptable, aunque sujeto a mejoras, sobre todo, en materia de entonación, latencia de habla, fluidez y de aspectos relacionados con el proceso de traducción propiamente. En consecuencia, podemos confirmar que hemos alcanzado tanto el objetivo principal propuesto, como los objetivos secundarios.

Primeramente, centraremos nuestra atención en los aspectos paralingüísticos señalados en el párrafo anterior y para ello haremos referencia al proceso de reconocimiento automático de voz. Como señala Downie (2020, p. 38-39), los problemas que se produzcan en este primer proceso son una fuente de errores principal en el proceso de traducción propiamente. Cabe tener en cuenta que las palabras por sí solas pueden producir ambigüedad, de modo que otros aspectos paralingüísticos, como la entonación, el tono o incluso los gestos o las expresiones faciales, pueden jugar un papel muy relevante a la hora de comprender correctamente un determinado mensaje. Esta ambigüedad es uno de los aspectos que puede arrastrarse desde el primero de los procesos hasta el de conversión de texto a voz, de modo que reconocer y analizar en el primer paso del modelo en cascada aquellos aspectos que ayuden a evitar ambigüedades sería de vital importancia. Para ello, la solución podría consistir en alimentar al sistema con gran variedad de acentos y entrenarlo para ser capaz de reconocer distintas entonaciones y tonos (Downie, 2020, p. 38-49).

La importancia de un buen rendimiento en el proceso de reconocimiento automático de voz no acaba en el plano paralingüístico, sino que atañe de igual forma al plano traductológico, dado que, si lo que se proporciona al sistema de traducción automática desde el sistema de reconocimiento automático de voz no tiene una calidad óptima, no se puede esperar que la calidad del producto final de la traducción sea aceptable. No obstante, es posible solventar estas limitaciones si se proporciona al sistema de reconocimiento automático de voz una mayor cantidad de datos, por ejemplo, en materia de nuevas palabras y expresiones idiomáticas (Downie, 2020, p. 38-49). En la línea del proceso de traducción propiamente, resulta lógico pensar que el conocimiento sobre el contexto del discurso, el conocimiento cultural y la profundización en la intención y el tono del orador principal serían de gran importancia para optimizar el producto final y aumentar las posibilidades de emplear estos sistemas de interpretación automática en escenarios reales (Fantinuoli, 2023c, p. 56-58). Además, cabe añadir que todo este conocimiento mencionado anteriormente constituye un pilar fundamental no solo para solventar los desajustes en la traducción o los errores en los aspectos paralingüísticos señalados, sino también a la hora de adoptar el registro y el estilo más adecuados. Según se ha podido comprobar en varios experimentos, así como los sistemas de traducción automática destacan en adecuación, el estilo

es precisamente el aspecto en el que los sistemas automáticos deben mejorar con respecto al rendimiento de profesionales humanos (Downie, 2020, p. 57-59). Por otra parte, en lo que concierne a aumentar la calidad del producto que obtenemos de la traducción automática a nivel terminológico, cabría la posibilidad de alimentar a la máquina con glosarios en los que se incluyan términos ya previamente contrastados o aquellos términos que el cliente requiera para un determinado encargo.

Tras haber comprobado el potencial de este sistema de interpretación automática y dado el rápido avance que la interpretación automática está experimentando, consideraríamos muy interesante volver a abordar y analizar el rendimiento de estos sistemas desde diferentes perspectivas, metodologías y con nuevos objetivos en futuros trabajos de investigación.

8. Bibliografía

- Ahmed Ahmed, Safa'a. (2022). Technology and artificial intelligence in simultaneous interpreting: a multidisciplinary approach. *Occasional Papers*, 78. https://opde.journals.ekb.eg/article_249945_b9781cf0b9712533721caeeb6fe0ff2c.pdf (Fecha de consulta: 15/05/2024).
- Cabrera Méndez, Gabriel, Capiozzo, Enrico y Danese, Lisa (2018). Parties, factors and instrumentalities involved. En A. Amato, N. Spinolo y M. J. González Rodríguez (Eds.), *Handbook of Remote Interpreting*. Bologna: SHIFT in Orality Erasmus+ Project, pp. 53-63.
- Collado Aís, Ángela. (2007). *La evaluación de la calidad en interpretación simultánea. La importancia de la comunicación no verbal*. Granada: Editorial Comares.
- Fantinuoli, Claudio. (2018). *Interpreting and technology (Translation and Multilingual Natural Language Processing 11)*. Berlin: Language Science Press.
- Fantinuoli, Claudio. (2023a). Machine Interpreting. *EasyAI*. <https://easyai.uni-mainz.de/html/machine-interpreting.html> (Fecha de consulta: 21/01/2024)
- Fantinuoli Claudio. (2023b). The Emergence of Machine Interpreting. *European Society for Translation Studies*, 62. https://www.claudofantinuoli.org/docs/ESTNL_May_2023.pdf (Fecha de consulta: 15/05/2024).
- Fantinuoli, Claudio. (2023c). Towards AI enhanced computer-assisted interpreting. En Gloria Corpas Pastor y Bart Defrancq (Eds.), *Interpreting Technologies – Current and Future Trends*. Ámsterdam / Filadelfia: John Benjamins, pp. 47-71.
- Fantinuoli, Claudio. (24 de enero de 2024). Defining Machine Interpreting. *claudofantinuoli.org blog* <https://claudofantinuoli.org/blog/2024/01/24/defining-machine-interpreting/> (Fecha de consulta: 15/05/2024).
- Herbert, Jean. (1952). (Ed. 2). *The Interpreter's Handbook: How to Become a Conference Interpreter*. Gêneve: Librairie de l'Université.
- Interprefy. (29 de agosto de 2022). *How does remote simultaneous interpreting work?* <https://www.interprefy.com/resources/blog/how-does-remote-simultaneous-interpretation-work> (Fecha de consulta: 15/05/2024).
- Interprefy (2024). <https://www.interprefy.com/> (Fecha de consulta: 06/02/2024)

- Kalderon, Eliza. (6 de mayo de 2023). *Neurofisiología de la interpretación simultánea*. (Webinar). Colegio Mexicano de Intérpretes de Conferencias. <https://www.youtube.com/watch?v=G0wAgUK8YHw> (Fecha de consulta: 15/05/2024).
- Ziegler, Klaus y Gigliobianco, Sebastiano. (2018). Present? Remote? Remotely present! New technological approaches to remote simultaneous conference interpreting. En Claudio Fantinuoli (Ed.), *Interpreting and technology*. Berlín: Language Science Press, pp. 119–139. DOI:10.5281/zenodo.1493299
- Moreno Villanueva, José Antonio. (2023). Recursos lexicográficos para la resolución de dudas: hábitos de consulta de los profesionales de la traducción. En Sara María Torres Outón y Ramón Méndez González (Eds.), *Realidades y desafíos de la tecnología aplicada a la traducción e interpretación*. Berlín: Peter Lang, pp. 115-134.
- Netflix (2024). *Somos lo que comemos*. <https://www.netflix.com/es/title/81133260> (Fecha de consulta: 02/05/2024)
- Nord, Christiane. (1991). *Text Analysis in Translation. Theory, Method, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Ámsterdam, New York: Rodopi.
- Ortiz Boix, Carla. (2016). *Implementing Machine Translation and Post-Editing to the Translation of Wildlife Documentaries through Voice-over and Off-screen Dubbing. A Research on Effort and Quality*. Tesis doctoral. Barcelona: Universidad Autónoma de Barcelona. <https://tesisenred.net/bitstream/handle/10803/400020/cob1de1.pdf?sequence=1&isAllowed=y> (Fecha de consulta: 15/05/2024).
- Prandi, Bianca. (2023) *Computer-assisted simultaneous interpreting: A cognitive-experimental study on terminology (Translation and Multilingual Natural Language Processing 22)*. Berlín: Language Science Press.
- Universidad de Vigo. (s.f.). *Historia*. <https://linkinterpreting.uvigo.es/historia/> (Fecha de consulta: 19/01/2024)