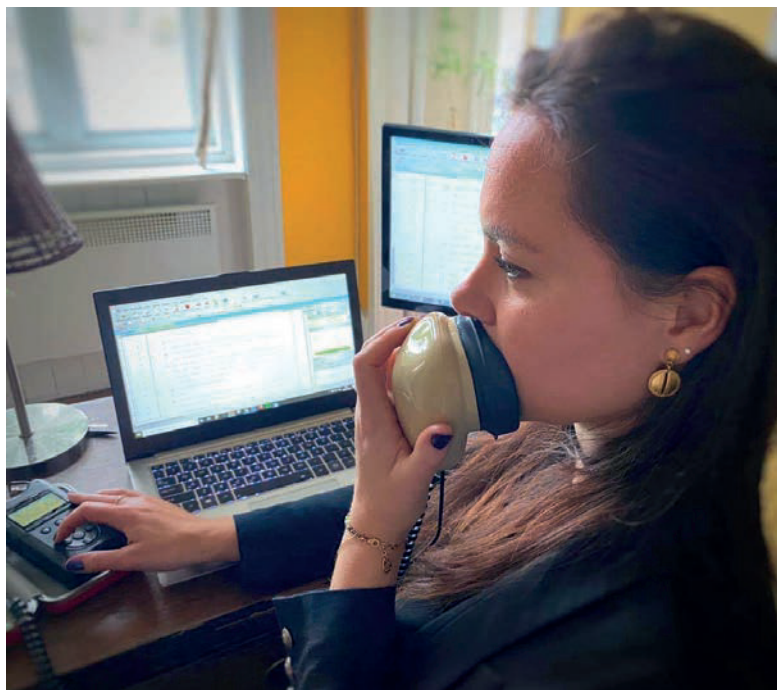


Subtitulación a través de la técnica del rehablado

Marimar Rufino Morales está cursando su doctorado en estudios hispánicos en la Université de Montréal e imparte clases de traducción español-francés. Se graduó en la EUTI de Granada y se especializó en traducción audiovisual en Canadá, donde lleva más de 25 años subtitulando para la televisión en diferido, y 15 en directo. Dentro de su línea de investigación -la variación del español-le interesa optimizar la transcripción de la lengua hablada con ayuda de la tecnología inteligente.



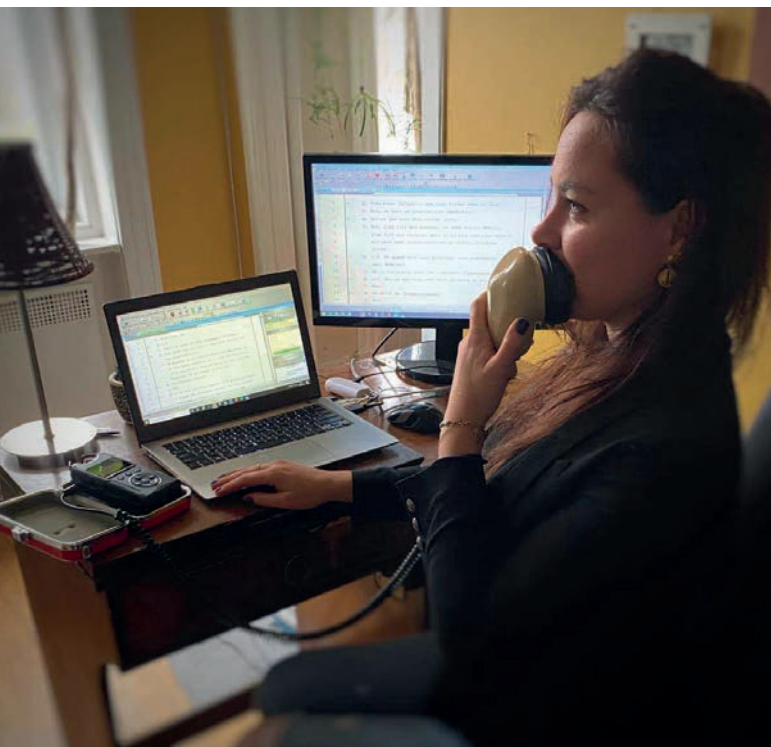
¿Qué es la técnica del rehablado? ¿En qué se diferencia de la estenotipia?

En 2012, la AENOR definía el rehablado como una técnica para producir subtítulos mediante el reconocimiento automático del habla. En una definición actualizada podemos establecer que con esta técnica el rehablador optimiza la conversión al escrito de cualquier discurso oral, repitiéndolo o parafraseándolo al tiempo que lo escucha, con ayuda de un programa (software) de dictado automático. El rehablado y la estenotipia son dos procedimientos para transcribir de forma simultánea cualquier comunicación oral. Ambos persiguen el mismo objetivo, aunque con distintos profesionales y herramientas propias de cada gremio.

Aunque el rehablado suele vincularse a la accesibilidad a través de subtítulos en televisión, lo cierto es que el

concepto nació en el cine. En sus orígenes, el cine silente era accesible para la comunidad no oyente, pero con el estreno de *The Jazz Singer* (1927) surgieron los primeros problemas: también para el público que no entendía el idioma en el que se pronunciaban los diálogos, pues las películas sonoras se proyectaban en versión original (de ahí la necesidad del doblaje, del que no nos ocupamos aquí). El actor cubano sordo Emerson Irving Romero, cuyo nombre artístico fue Tommy Albert, consiguió editar varias películas insertándoles textos que explicaban los diálogos. Por sus grandes esfuerzos por acercar el cine sonoro a su comunidad se le considera el pionero de la subtitulación.

Hoy, rodeados de tecnología digital, es probable que nos resulte difícil imaginar los obstáculos franqueados, desde aquellos experimentos iniciales de Emerson Romero, hasta que en nuestras pantallas aparecieron los subtítulos de for-



ma simultánea a su producción. Cuando se transmitieron los primeros programas de televisión, entre 1936 y 1939, la imagen y el sonido debían tratarse por separado; dos retos de peso a los que, para subtítular, había que añadir tanto la producción como el envío del texto escrito. Las películas con subtítulos configurados para el cine no podían usarse para televisión, por tener esta menos contrastes. Con la llegada del teletexto, países como Reino Unido o Estados Unidos lograron emitir subtítulos desde 1979. En España, los subtítulos a través del teletexto llegaron por la línea 888 en 1990, primero en la CCRTV y, dos meses más tarde, en TVE. Se transmitían en circuito cerrado (las iniciales CC vienen de la denominación en inglés: Closed Caption), de modo que, para leerlos o descifrarlos, siempre y cuando estuvieran disponibles, hacía falta un decodificador. En 1993 se obligó a los fabricantes a incorporarlos en los aparatos de televisión, pero hasta entonces había que adquirirlos por separado.

Para entender el rehablado, hemos de interesarnos en cómo se hacen los subtítulos. En general, el proceso de creación de subtítulos, no solo en cine y en televisión, implica tres etapas. Primero hay que transcribir el contenido sonoro en forma de texto escrito; después hay que sincronizarlo con la imagen y el sonido; finalmente, para que el texto pueda ser leído, debemos insertarlo en el producto original o enviarlo al tiempo que las imágenes y el sonido.

Cuando disponemos del material para subtítular con antelación (programas grabados o emitidos en diferido, ya

que se emiten posteriormente a su grabación, como una película, un documental, un reportaje), los subtítulos suelen generarse realizando una transcripción manual, es decir, mecanografiando lo que se quiere plasmar por escrito a medida que se va visionando el material. Pero, ¿qué pasa cuando lo que debemos subtítular es un programa que transcurre y se emite en vivo? ¿Cómo se subtítulan informativos, eventos deportivos, ruedas de prensa, debates, entrevistas y demás programas en directo?

En los años 80, las cadenas de televisión que empezaron a subtítular de forma regular probaron distintas soluciones para los programas en directo: ninguna de ellas llegaba a ser totalmente satisfactoria. En varios países, como los Estados Unidos, se recurrió a taquígrafos (*court reporters*). Podían subtítular entre 180 y 200 palabras por minuto, pero necesitaban ser relevados de la toma de notas cada cuarto de hora. En Reino Unido, el canal ITV ideó un sistema en el que las teclas de un teclado QWERTY correspondían a abreviaturas, y aunque este método se usó para subtítular el Mundial de fútbol o la visita del papa en 1982, no era lo bastante rápido como para subtítular las noticias diariamente, por lo que se optó por el teclado silábico holandés (Velotype) que permitía a los subtítuladores, después de doce meses de entrenamiento, producir entre 90 y 120 palabras por minuto. Varios países lo adoptaron en tándem de dos a cinco personas, entre ellos España.

Sin embargo, hasta finales de los 90, para subtítular en tiempo real lo más común era la estenotipia, llamémosla "manual" para diferenciar esta forma de la informatizada que surgiría más adelante. La estenotipia es un sistema que permite a una persona, entrenada en el método, escribir en un papel testigo mediante un teclado reducido (normalmente la máquina francesa Grandjean) de veintiuna teclas distribuidas en dos filas que corresponden a pulsaciones de sílabas o palabras completas. La velocidad del discurso puede llegar a 300 palabras por minuto en noticias, programas de entretenimiento o eventos deportivos. Los mejores estenotipistas alcanzan a transcribir más de 200 palabras por minuto, pero hay que tener en cuenta que queda por hacer la corrección o transcripción final del papel testigo que sale de la máquina de estenotipia. Además, la formación es larga –alrededor de dos años– y costosa, de manera que la estenotipia manual desde el primer momento en que empezó a usarse en televisión apareció como un método vulnerable.

Retomando el hilo conductor de la cronología, llegamos a los años 2000, momento en el que las cadenas de televisión comenzaron a subtítular con la tecnología del reconocimiento automático del habla. Desde que se em-

pezó a subtítular en tiempo real se sopesó la posibilidad de usar la interacción persona-máquina a través del habla para desempeñar esta tarea. Recurrir a estenotipistas y taquígrafos (o estenógrafos) conllevaba, como ya hemos mencionado, numerosos problemas para las empresas entre los que destacan la formación larga y costosa y la cantidad de personal que implica.

Los primeros intentos realizados unos años antes, cuando la tecnología tuvo un coste asequible como para que se pudiera experimentar con ella en televisión, no funcionaron. En uno realizado en la Universidad de Southampton, en 1985, se subtítuló usando un teclado QWERTY mientras se efectuaban varias tareas con ayuda de comandos vocales. Sirvió para demostrar que para el subtítulador suponía el 9 % más de manipulaciones con este nuevo método; además el ruido del teclado mermaba el rendimiento del reconocedor automático. Y cuando se intentó subtítular automatizando todo el proceso de transcripción con un programa de reconocimiento del habla tampoco se pudo, ya que el resultado final contenía demasiados errores, muchos provocados por interferencias de ruidos, o músicas de fondo, o por la simultaneidad de varios discursos. Tampoco se podían posicionar los subtítulos ni cambiar su apariencia para adecuarlos a las convenciones establecidas.

Las conclusiones de estos trabajos iniciales apuntaron que una posible solución para paliar las imperfecciones del reconocimiento del habla por ordenador podría venir de una mayor interacción humano-computadora: intervendría una persona que, entrenada al uso del programa, a su vez entrenaría a este y mejoraría su vocabulario en un lugar

exento de ruido. Primero se puso a prueba un sistema que recurría a dos operadores. Mientras uno escuchaba el programa de televisión en vivo con un auricular, repetía lo que se estaba diciendo e iba añadiendo los signos de puntuación, el otro debía editar el texto del discurso rehablado que le aparecía en una ventana en un mínimo de tiempo y después enviarlo en forma de subtítulos. Sobre esta base, a medida que los niveles de exactitud alcanzados con un "hablador" experimentado fueron lo bastante aceptables, pudo prescindirse del corrector.

Fue este el método que se expandió en las televisiones y productoras de subtítulos de varios países. La razón: frente a los demás que se probaron, representaba un ahorro de hasta el 40 % de los costes globales de subtitulación en un momento en el que los difusores debían respetar las legislaciones en materia de accesibilidad. Estamos ante la génesis del rehablado.

¿Cómo se rehabla y qué tipo de programas de software se utilizan?

En el ámbito de la televisión, el rehablador, instalado en un entorno exento de ruido, a la vez que escucha la emisión a subtítular, la repite o parafrasea verbalmente a través de un micro conectado a un programa (*software*) de dictado automático de manera que genere una transcripción legible. El texto resultante es enviado casi simultáneamente, en forma de subtítulos, y la emisión se hace accesible para aquellas personas que no pueden oírlo.

Para rehalar hay que escuchar y hablar al mismo tiempo, no distraerse con el sonido de su propia voz, repetir sin

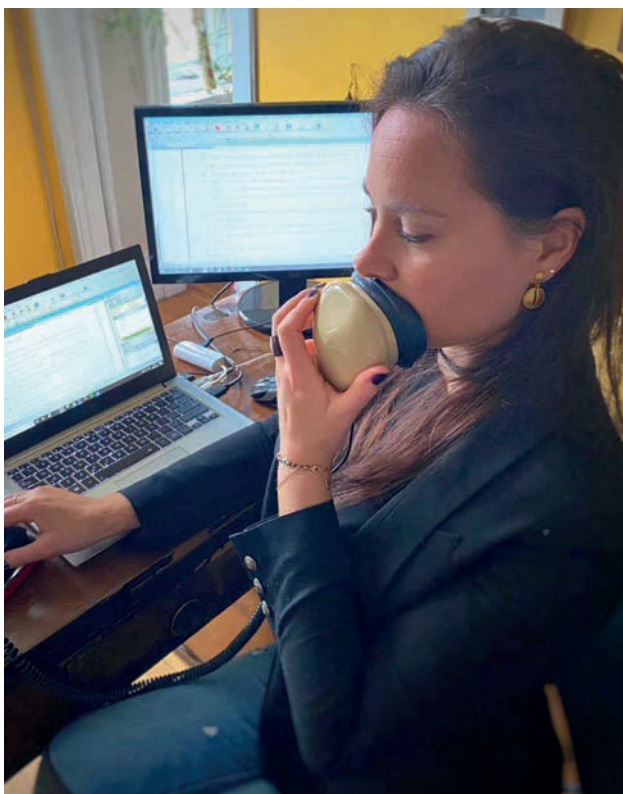


OTOLOGÍA, AUDIOLOGÍA Y AUDIOCIRUGÍA (Tratamiento de problemas auditivos)
OTONEUROLOGÍA Y OTONEUROCIURUGÍA (Tratamiento del vértigo y la parálisis facial)
IMPLANTES DE OÍDO MEDIO Y OSTEOINTEGRADOS (Tratamiento de pérdida de audición)
IMPLANTES COCLEARES (Prótesis auditiva para la sordera profunda)
FONIATRÍA Y LOGOPEDIA (Reeducación de la voz, Labiolectura,...)
UNIDAD DE ACÚFENOS (Ruido de Oído)
PSICOLOGÍA Y AUDIOLOGÍA INFANTIL (Niños sordos)
PATOLOGÍA DE LA NARIZ, FARINGE Y LARINGE

C/ Arturo Soria, 119 – 28043 Madrid
Telf. 91 4130514 – Fax. 91 5195085

info@institutoorl-iom.com
www.institutoorl-iom.com

demora y prestar atención a lo que se rehabla (forma y contenido). La complejidad cognitiva de la tarea requiere numerosas estrategias. No basta con repetir o parafrasear lo que se oye, también se debe añadir –o dictar– los signos de puntuación, la información paralingüística (ruidos, música, aplausos) y todas las palabras ausentes del *software* (expresiones, nombres propios). Además, se necesita anticipar y paliar los posibles errores de reconocimiento, como los homófonos (*haber/a ver*), al tiempo que se vigila el texto resultante para detectar y corregir incoherencias o errores, con el fin de mantenerse fiel al mensaje original y, por encima de todo, proporcionar un discurso escrito inteligible.



Los programas de conversión voz-texto a los que nos referimos se basan en la tecnología de reconocimiento automático del habla (*Automatic Speech Recognition*), la cual ha logrado aumentar de forma significativa la precisión gracias a los avances de la Inteligencia Artificial (IA) y el aprendizaje profundo. La aplicación de redes neuronales artificiales en el tratamiento automático del lenguaje fue disminuyendo progresivamente las tasas de error hasta llegar, a finales de 2016, a un punto de ruptura: el ordenador superó al humano. Rápidamente, pudimos comprobar cómo mejoraban las interacciones con asistentes personales inteligentes y asistentes virtuales como Siri o Cortana. A pesar de ello, hasta el día de hoy, cuando se trata de transcribir con un programa de dictado automático la lengua hablada, la conversación espontánea o la variación, los resultados fluctúan. Tanto los experimentos como la

propia práctica muestran que el reconocimiento aún no es estable y siguen reclamando, al menos de momento, la presencia del rehablador para mediar ante la computadora.

Hoy día, la mayoría de los *softwares* de dictado automático de uso profesional y personal están diseñados sobre la base de modelos genéricos independientes de locutor; esto es, no necesitan entrenamiento. En televisión, las empresas subtituladoras suelen tener reconocedores propios, con modelos acústicos y de lenguaje personalizados. Si dejamos de lado los hechos a medida, actualmente, Dragon NaturallySpeaking, comercializado por Nuance parece ser uno de los *softwares* de dictado automático voz-texto más populares.

¿El rehalado también se utiliza para traducciones simultáneas y para programas grabados?

El rehalado se emplea no solo en televisión, no solo en tiempo real y no solo en un idioma.

El subtitulado y la audiodescripción son dos instrumentos para eliminar barreras de acceso a la comunicación, pero existen otras. Independientemente de la forma en que leamos las transcripciones, con el rehalado en tiempo real se hace más accesible cualquier creación escénica y espectacular (teatro, ópera, conciertos), pero también actos públicos, reuniones, conferencias y seminarios presenciales o virtuales, actos en iglesias, museos y demás manifestaciones culturales.

Por otro lado, el rehalado hace la educación más inclusiva. Un ejemplo de ello es Sudáfrica, donde hay once idiomas oficiales reconocidos (y otras nueve lenguas indígenas más no oficiales), pero en las universidades, la mayoría de las clases se imparten solo en inglés. El departamento de idiomas de la North West University brinda la transcripción en tiempo real en inglés del discurso del profesor, gracias a un micro conectado al auricular del rehablador. Por otra parte, la compañía australiana Ai-Live (presente también en Estados Unidos y Canadá) colabora con universidades para hacer accesibles las clases. Los alumnos que lo requieran solo necesitan una conexión a Internet para ir leyendo en su portátil, tableta o teléfono inteligente la transcripción rehalada de lo que dice el profesor y de toda la interacción entre los asistentes.

Por analogía, el rehalado permite producir una transcripción en un idioma distinto al del contenido sonoro. Rehablar en otra lengua no es más ni menos que ofrecer una nueva forma de accesibilidad a la comunicación eliminando barreras lingüísticas. Por ejemplo, en programas de televisión

en directo en Taiwán surgen diálogos en taiwanés, pero el subtítulo en vivo solo suele hacerse hacia el chino mandarín, que es el único idioma oficial.

En todos estos casos se trata del **rehablado on-line**, ya que la transcripción obtenida en tiempo real se envía simultáneamente para que un tercero tenga acceso a lo que se está diciendo en el mismo momento en que se produce el discurso. Pues bien, también es posible utilizar el sistema de dictado para hablar con el fin de crear una transcripción al tiempo que se produce el discurso, pero no para que sea leída de forma simultánea. En este caso, se emplea la técnica del hablado para optimizar el proceso de transcripción en el ámbito académico, médico, judicial, en los distintos sectores de la administración pública, ya sea para transcribir actas, testimonios, entrevistas, investigaciones o informes. Lo llamamos **rehablado off-line** para incluir a los escritores de voz (*voice writers*), estenotipistas de formación que transcriben con ayuda de un sistema automático de dictado, generalmente a través de un silenciador de voz.

Además, muchas de las empresas que subtitulan en directo para la televisión mediante hablado y que cuentan con la infraestructura necesaria, rentabilizan equipos y personal subtitulando programas 'enlatados' o pregrabados.

Por último, el hablado se emplea para acelerar la transcripción de cualquier contenido audiovisual, destinado al cine, a la televisión o a la radio, como al resto de las plataformas multimedia, distribuidas a través de la web y por otros soportes: CD, CD-ROM, consolas de videojuegos, etc.

¿Se dictan los signos de puntuación en el hablado? ¿Qué pasa con las palabras que no están en el diccionario? ¿El software de reconocimiento de voz las reconoce igual?

La particularidad distintiva de nuestra profesión –y su razón de ser– radica en esta relación de interdependencia entre el hablador y el programa de dictado. Los habladores necesitamos familiarizarnos con el *software* que vayamos a usar. Los resultados dependen directamente del conocimiento que tengamos de todas sus posibilidades de rendimiento y, proporcionalmente, de nuestra capacidad para anticipar su reconocimiento paliando sus carencias con las estrategias más adecuadas.

Estos *softwares* para el dictado se basan en la tecnología de reconocimiento automático del habla, que es la que logra hacer una representación escrita de la señal de la voz continua. El término "reconocimiento de voz" se refiere a la tecnología que identifica a la persona que habla (también

llamada de reconocimiento del locutor), mientras que el "reconocimiento del habla" se ocupa de lo que dice.

Veamos qué ocurre entre el momento en que el programa de dictado recibe el *input* suministrado por el hablador y su transmisión en forma de texto escrito; es decir, cómo funciona la tecnología de reconocimiento automático del habla. El contenido del mensaje debe procesarse para obtener una versión simbólica del mismo que facilite su transmisión física (análisis acústico y extracción de características). A diferencia de la comunicación entre seres humanos, el programa informático debe producir la misma representación cada vez que trata la misma señal comparándola a los modelos que posee. Una de las mayores dificultades del análisis del lenguaje natural reside en cómo crear una gramática capaz de generar todas las frases posibles y a la vez únicamente las frases posibles: una suerte de falacia, pues la característica principal del lenguaje natural es su constante evolución. Si se crea una gramática de cobertura completa, con muchas reglas, resultará limitante en numerosos aspectos. A su vez, una gramática de baja cobertura puede producir múltiples combinaciones inexistentes. Los programas modernos se basan en el análisis probabilístico de las hipótesis generadas, las enfrentan al *input* lingüístico y seleccionan la mejor probabilidad que poseen (las posibles combinaciones de fonemas y de sílabas que pueden darse al producirse el habla). Por su parte, los modelos de lenguaje se encargan de evaluar las probabilidades de una secuencia de palabras enfrentándola al conjunto de textos que poseen. La combinación de información de los modelos acústicos y lingüísticos se hace gracias al diccionario de pronunciación creado a partir de la selección del vocabulario y de la descripción de cada una de estas entradas léxicas mediante unidades acústicas elementales.

Así, volviendo a la subtitulación en televisión, solo podrán aparecer en la pantalla aquellas palabras, o combinaciones de palabras, que se le hayan mostrado previamente al reconocedor, bien a través de las bases de datos utilizadas para su entrenamiento, bien porque hayan sido añadidas de forma manual. Es lo que suele hacer el hablador en la fase preparatoria previa al hablado. Dependiendo del programa con el que se trabaje, cabe la posibilidad de alimentarlo no solo con palabras, listas enteras de vocabulario o textos, sino incluso con pronunciaciones y atajos.

Y, sin embargo, pese a todas las precauciones, siempre puede aparecer en el discurso original algún vocablo de forma inesperada ausente en el reconocedor (un neologismo como Covid19, una palabra en otra lengua, el nombre de un deportista o de un actor extranjero). También puede que el

programa cometa un error de reconocimiento (que escriba "ir a les cuela" en lugar de "irá a la escuela"; "oler" en lugar de "o leer") o que el rehablador lo induzca en error (porque reable "Juan" en lugar de "Julián" o porque se quede en blanco). Incluso es posible que varios interlocutores hablen a la vez y haya un discurso simultáneo. Ante cualquiera de estos ejemplos, el rehablador suele tener delante el texto que va a salir en pantalla y desde el punto de vista técnico, es posible corregirlo. Sin embargo, hay que entender que la edición del error (borrado, sustitución, adicción) para enmendar un texto no se hace con el programa de dictado; se lleva a cabo a través de la infraestructura encargada de convertir el texto en subtítulos y enviarlos. La intervención del rehablador, o de un corrector, en este instante, provocará inevitablemente cierto retardo en la aparición de los subtítulos.

El tratamiento de los signos de puntuación es otra de las manifestaciones que usaremos para ejemplificar la interacción entre el rehablador y el software y que permiten obtener una transcripción de calidad. Con algunos programas, el rehablador tiene las manos libres para otras tareas, como con Dragon NaturallySpeaking o Web Captioner. Pero en ese caso, hay que dictar todos los signos de puntuación: "punto", "coma", "dos puntos", "puntos suspensivos", "abre interrogación", "cierra interrogación". Se encuentran en la misma categoría los programas que son (más o menos) capaces de interpretar ciertas pausas en el dictado como un punto. Por otro lado están los programas en donde la puntuación se añade accionando ciertas teclas del ordenador o, como en el caso de STDirect, con un mando de consola de videojuego, ofreciéndole al rehablador una micropausa en su discurso.

¿Cuál ha sido el primer rehablado de la historia?

La idea de transcribir un discurso oído usando la propia voz se le atribuye al inventor del silenciador de voz. Según relata la *National Verbatim Reporters Association*, en los años cuarenta, Horace Webb llevaba apenas dos años transcribiendo en el tribunal de Chicago cuando revolucionó el oficio. Hasta entonces, para obtener las actas, se procedía a la toma de notas en el transcurso de la sesión y, una vez concluida, había que dictarlas y mecanografiarlas antes de obtener el texto final correspondiente a la transcripción del discurso original. A Webb se le ocurrió que podía dictar directamente en el tribunal, pero para ello necesitaba eliminar el ruido ambiente y, sobre todo, pasar desapercibido en el entorno. Insertó un micrófono en una caja de cigarrillos, y después en una lata de café, que amortiguaba el sonido mientras repetía y grababa lo que oía en el tribunal. Puesto que consiguió hablar más rápido

de lo que se puede escribir y eliminar el ruido ambiental durante la grabación, transcribía palabra por palabra el discurso original.

Quizá el auténtico precursor del rehablado pudiera ser este otro procedimiento que también se empleó en el mismo contexto en Estados Unidos allá por 1901. Se colocaba un tubo en la sala principal que conducía a cuatro tubos más, cada uno en una habitación con una persona que iba transcribiendo según un plan determinado. Se necesitaban cuatro personas para lograr la transcripción de una sesión.

A mediados de los noventa, los escritores de voz que trabajaban con silenciador y los taquígrafos-estenógrafos que se grababan adoptaron rápidamente los programas de dictado automático. En Norteamérica, la escritura de voz con reconocimiento automático del habla se extendió rápidamente en los años 2000. Christine Ales y Betty Keyes son dos pioneras de la enseñanza de este método.

En el mismo momento, la tecnología del reconocimiento del habla irrumpía en las televisiones de los países que ya contaban con un servicio de subtitulación en directo como estrategia de inclusión, conforme al principio de igualdad que recogen todas las constituciones a partir de la Declaración de los derechos humanos (1994). Sin embargo, como hemos apuntado, los resultados obtenidos cuando se intentaba subtítular de forma automática no eran halagüeños. Yo trabajaba entonces como subtuladora en TVA, una cadena de televisión francófona canadiense. Estuvimos casi un año practicando con un programa creado por el *Centre de Recherche Informatique de Montréal* (CRIM), STDirect, antes de darles salida para permitir que los televidentes nos leyeran en directo. Al principio, solo durante el pronóstico del tiempo, que era la parte de los informativos donde obteníamos mejores resultados. En 2007, apenas seis años después, todas las cadenas de televisión francófonas y anglófonas tenían la obligación legal de subtítular el 100 % de la programación en Canadá.

Con las altas velocidades de dicción en los medios, para no demorar (a mayor literalidad, mayores retardos): ¿un rehablador interpreta y/o condensa las unidades de significado?

La velocidad óptima de producción oral se sitúa entre 170 y 190 palabras por minuto. Según un estudio realizado en 2005, donde se midió el manejo de la voz de los presentadores del telediario en tres cadenas de televisión, la locución de Ana Blanco es de 234 palabras por minuto. Las altas velocidades en las voces de televisión es una realidad y cabría preguntarse si la comprensión del mensaje vehicu-

lado con la voz a esas velocidades está garantizada incluso para quienes lo escuchan. Suponiendo que así fuera, nosotros queremos saber si podemos leer tantas palabras por minuto y si es posible rehablar y subtítular a tal velocidad.

Desde que hicieron su aparición en televisión, la velocidad de los subtítulos no ha dejado de aumentar a medida que se han ido eliminando barreras tecnológicas, que también han reducido los costes de producción. De la observación de los subtítulos rehablados en televisión se desprende una tendencia: mientras que en Europa suelen ser más condensados o resumidos, en Norteamérica se prefiere la literalidad. Mientras aquellos defienden la imposibilidad de lectura de los subtítulos literales, estos otros pretenden que condensarlos es censurar el contenido audiovisual.

En el centro de la controversia está la velocidad de lectura ideal de los subtítulos. Una pregunta intrincada dado que el público que los usa no es homogéneo, tampoco

lo son sus habilidades lectoras. A los parámetros que se usan para medir la velocidad ideal de lectura de un texto íntegro, debemos añadir otros: los subtítulos presentan el mensaje de forma parcial y por un periodo de tiempo que el lector no puede controlar. Por ende, mientras se lee un subtítulo también hay que procesar otros estímulos, como las imágenes. Las personas mayores y los niños, aquellas con discapacidad auditiva o intelectual, pero también quienes aprenden la lengua, los que no desean molestar a los demás en el hogar, en una sala de espera, un restaurante, un centro comercial o en el gimnasio, todos somos lectores potenciales de subtítulos accesibles¹.

Esta forma de medir la velocidad ideal de un subtítulo se basa en el número de caracteres y fortalece la idea de que a menos caracteres mayor velocidad de lectura. Sin embargo, otros estudios basados en los parámetros de textualidad apuntan que si se respetan principalmente las dos centradas en el texto (coherencia y cohesión), la velocidad

¹Para que el lenguaje refleje nuestra sociedad inclusiva, preferimos subtítulos accesibles a "subtítulos para sordos" (tampoco su acrónimo SpS).

COCHLEAR IMPLANT SYSTEM

Neuro 2 – Dónde el sonido encuentra el diseño

El procesador de sonido IC más pequeño del mercado

El Neuro 2 es el procesador de sonido IC más pequeño del mercado, además de elegante y discreto. Pero es mucho más que solo su tamaño. Combina gran calidad de sonido, con innovadoras características y prácticos accesorios para todas las edades y estilos de vida, gracias a su diseño ya ha ganado varios premios.

Todos los aspectos del Neuro 2 han sido diseñados para cumplir con los más altos estándares de calidad. Ya sea su cable transparente reforzado con kevlar, o la clasificación IP68 – la mayor protección contra el polvo y agua en la industria de IC – alcanzada sin accesorios adicionales, Neuro 2 está diseñado para brindarle confianza auditiva durante todo el día, los 365 días del año.





de lectura aumenta y no depende tanto del número de caracteres. En el mismo orden de ideas, la longitud y la forma de las palabras también inciden en la velocidad de lectura. En todo caso, se necesitan más estudios con todos los grupos de lectores de subtítulos y con todas las modalidades de subtitulación. Mientras tanto, los rehabladores a los que no se les ha impuesto una política editorial deben usar su propio criterio y sensibilidad. Optar por una mayor literalidad o por una estrategia de reducción total o parcial de un elemento del discurso oral depende también del programa que se subtitle y del nivel de representatividad escrita de los rasgos de la oralidad de dicho elemento que en ocasiones pueden provocar ambigüedad.

¿Cómo es el entrenamiento de un rehablador? ¿Cuál es la formación y perfil requerido?

Cuando se empezó a rehablar en televisión, las propias empresas eran las que formaban al personal. Si eran subtituladores de plantilla, ya poseían ciertas competencias lingüísticas² y un conocimiento extenso del contenido de las emisiones que iban a subtitularse, de la mecánica interna de subtitulación y del público meta.

A medida que el rehablado fue ganando terreno, aumentaban las necesidades de personal haciéndose necesario diseñar el perfil, las habilidades y las técnicas que eran

precisas desarrollar para desempeñar un oficio para el que no existía nombre ni formación académica. Una profesión emergente con similitudes (y diferencias) con la interpretación simultánea. Para rehablar, es necesario escuchar y hablar a la vez poniendo a contribución la memoria operacional (comprender, almacenar y reproducir) y activando una serie de destrezas que van desde el dominio de la pronunciación, el léxico y la gramática hasta conocimientos socioculturales y pragmáticos. En ese mismo instante, habrá que corregir un discurso pasado mientras se sigue trabajando en tándem con el programa de reconocimiento automático del habla, escuchando y rehablando una y otra vez. En todo el proceso, hay que dividir su atención no ya en dos tareas, sino en varias (*multitasking*).

La formación de un rehablador es interdisciplinar y, dependiendo de sus conocimientos previos, puede durar entre unas semanas y varios meses.

En Estados Unidos, donde además la *National Verbatim Reporters Association* (NVRA) emite una certificación oficial, desde el primer momento aparecieron numerosas ofertas privadas de cursos y formaciones de escritura de voz.

En Europa, en 2006, el departamento de lingüística de la Universidad de Bolonia organizó el Primer seminario internacional sobre nuevas tecnologías en subtitulación intralingüística en tiempo real (Forlì). Marcaba el reconocimiento académico de la profesión. A la vez que la accesibilidad se fue integrando a los Estudios de Traducción, comenzaron a impartirse cursos y módulos de rehablado, presenciales y en línea, como parte de la formación en traducción audiovisual.

Desde entonces, la comunidad científica, principalmente europea, trabaja sin tregua por uniformizar el nombre de la profesión, armonizar su perfil profesional y homogeneizarlo dentro de la enseñanza superior.

¿La subtitulación por rehablado se atiende (o debería) igual a la Norma UNE153010:2012, en el caso español?

Hay que aplaudir todos los impulsos por homogeneizar el subtulado, y la estandarización constituye un pilar fundamental. Las normas UNE son, por definición, de carácter voluntario, si bien cuentan con el consenso de todas las partes interesadas, incluidos los usuarios a través de las asociaciones que los representan.

²Uno de los requisitos para ser subtitulador es demostrar el manejo de manera experta de estrategias de comprensión y producción escrita, en todas sus dimensiones: fonética, morfológica, léxica, ortográfica y sintáctica.

La promoción de los estándares constituye otro de los pilares hacia la homogeneización. En ese sentido España es líder, gracias a la excelente labor de sensibilización y difusión del Centro Español del Subtitulado y la Audiodescripción (CESyA).

La homogenización debe rematarse mediante leyes e incentivos que posibiliten su aplicación. Es la Ley General de la Comunicación Audiovisual aprobada en 2010, y actualizada en 2015, la que reconoce el derecho de las personas con discapacidad auditiva a recibir el 75% de la comunicación audiovisual televisiva de cobertura estatal y autonómica mediante subtítulos. Asimismo, estipula que para "garantizar la calidad del servicio y la satisfacción de las personas destinatarias, los prestadores del servicio de comunicación audiovisual deberán atenerse, en la aplicación de las medidas de accesibilidad, a las normas técnicas vigentes" (p. 19).

La Norma UNE153010:2012 es un compendio de códigos profesionales para la buena práctica subtituladora en todos los dispositivos, formatos y escenarios ya sea en diferido o en directo, sea cual sea la técnica que se emplee para elaborar los subtítulos, y el reabladado es una. Recoge una serie de recomendaciones que abarcan los distintos parámetros que intervienen en la elaboración de los subtítulos, incluyendo los aspectos visuales: posición, tamaño; elección de técnicas para identificación de personajes (color, etiquetas, guiones); cómo deben indicarse los efectos sonoros, música y canciones, elementos paralingüísticos y otras informaciones contextuales; uso de abreviaturas, siglas, símbolos y acrónimos; cómo condensar y reducir; trata de aspectos temporales como el sincronismo y el retardo, de la literalidad y dedica una sección para mostrar cómo ha de calcularse la tasa de precisión de los subtítulos en reabladado según el modelo NER (Romero-Fresco).

Es una norma bastante completa. Si bien es cierto que muchas de las pautas de la norma UNE153010:2012 difieren de las que emplean otros países, como también de los propios estándares de un texto escrito, lo importante es insistir en la necesidad de seguir estudiando y analizando el impacto de cada una de las recomendaciones en la percepción del(los) público(s) meta. Cómo reciben los usuarios estos códigos profesionales es tan importante como la voluntad por mejorarlos.

¿Quiere hacer alguna manifestación para nuestros lectores?

La tecnología ha llegado a un grado de madurez que permite cosas imposibles hace tan solo unos años. Hasta podemos escoger el idioma de los subtítulos, incluso de un programa en directo, según nuestras preferencias.

Sin embargo, esta aparente facilidad con que nos llega la información no es sinónimo de accesibilidad, también reposa sobre un entramado frágil. Es muy importante que los grupos de presión sigan velando por los intereses de ciertos sectores de la población y salvaguardando los logros obtenidos. Durante la crisis que ha sacudido el planeta en 2020, pasando por un Estado de Alarma, se ha hecho sentir más que nunca la necesidad de acceso a la información en tiempo real de todos por igual. Quienes trabajamos en accesibilidad a este lado del Atlántico, y reablamos a distancia, hemos asistido con estupefacción, por ejemplo, a la publicación del Decreto-ley 2/2020 "de mejora y simplificación de la regulación para el fomento de la actividad productiva de Andalucía", que atenta directamente contra el derecho a la información que defendemos de las personas con discapacidad auditiva, reduciéndoles el número de horas de subtitulado, audiodescripción y lengua de signos a las que tendrán acceso. Si esto se extiende a otras comunidades, y a otros países, volveremos a situaciones pasadas no deseables, y todos los adelantos alcanzados no tendrán continuación. ■

Fotos: Gabrielle Dubé-rehablado con Dragon Stenomask (transcripción de actas juzgado)

BIBLIOGRAFÍA

- ▶ AENOR (2012): Subtitulado para personas sordas y personas con discapacidad auditiva. UNE 153010. Madrid: AENOR.
- ▶ BOE-Gobierno de ESPAÑA (2010): LGCA-Ley General de la Comunicación Audiovisual. A-2010-5292.
- ▶ MARIANI, Joseph, ed. (2002): Reconnaissance de la parole. Traitement automatique du langage parlé 2 (Vol. 2). Paris: Hermès Science Publications.
- ▶ NÚÑEZ HIDALGO, Javier (2012): La experiencia en el Parlamento español. Revista taquigráfica. Revista internacional de taquigrafía. (5): 11-14.
- ▶ ORERO, Pilar, PEREIRA, Ana María y UTRAY, Francisco (2007): Visión histórica de la accesibilidad en los medios en España. TRANS. Revista de Traductología. (11): 31-43.
- ▶ ORGANIZACIÓN DE LAS NACIONES UNIDAS (1994): Normas Uniformes sobre la igualdad de oportunidades para las personas con discapacidad. Resolución 48/96.
- ▶ RODERO ANTÓN, Emma y CAMPOS PARRA, Gloria (2005): Las voces de los presentadores de informativos en televisión. Comunicar. 25(13): 11.
- ▶ ROMERO-FRESCO, Pablo (2011): Subtitling through speech recognition: respesaking (1a ed.). Manc