

Diferentes técnicas de producción de subtítulos en la TDT

Para muchas personas, especialmente para las personas mayores o con discapacidad auditiva y para las que están aprendiendo un idioma, los subtítulos son un recurso fundamental para ver la televisión.

En la actualidad, se utilizan tres técnicas diferentes para subtitular la programación de las televisiones dependiendo de la naturaleza del programa: grabado, guionizado o en directo, cada una de ellas con sus ventajas e inconvenientes.

GRABADO

Un programa grabado es aquél cuya producción se realiza en un tiempo suficientemente anterior a su emisión. Ejemplos de estos tipos de programas son películas y series.

Por el hecho de tener el producto terminado antes de su emisión, existe tiempo suficiente para producir el subtitulado e incorporarlo. En estos casos la subtitulación se realiza de forma manual o semiautomática y es posible realizar su revisión y corrección de errores

Subtitular un programa grabado no debe suponer ningún problema técnico y por tanto su calidad debería ser máxima.

FALSO DIRECTO

Por programa guionizado queremos indicar aquellos en el que los locutores se someten estrictamente a un guion preparado y ensayado con la suficiente anterioridad y habitualmente aprovechando del teleprompter. En este caso es posible preparar los subtítulos con anterioridad, cuando el guion está finalizado. Para que los subtítulos aparezcan en el momento correcto, un técnico se responsabiliza de “lanzarlos” de forma manual manejando un ordenador. “Si todo va bien”, la operación se reduce a seleccionar el subtítulo siguiente, esperar a que el locutor inicie su frase y mostrarlo.

Dada la intervención humana en directo, el sistema es susceptible de presentar defectos: ya sea por error del operador, ya sea porque el locutor no respete el guion o improvise.

Esta técnica se usa en aquellos programas que comparten parte del guion cerrado y parte en directo. En este caso se usa un combinando la técnica de semidirecto y estricto directo. En los fragmentos del programa donde los diálogos corresponden a un guion previamente elaborado, los subtítulos se preparan con antelación y son emitidos de manera sincronizada y con ayuda de un operador. En los fragmentos donde los diálogos son improvisados, los subtítulos se crean en tiempo real mediante la técnica conocida como rehablado, la más extendida para el subtitulado de programas de TV.

PURO DIRECTO

Por programas de puro directo queremos indicar aquellos que se emiten en directo y de los que no es posible disponer de un guion cerrado previo pues las personas que intervienen básicamente improvisan. Ejemplos típicos de estos son los programas de debate, los reality (en directo).

En este contexto el problema principal es el tiempo necesario para construir los subtítulos al mismo tiempo que se realizan las locuciones.

La técnica que utilizan la mayoría de los canales de televisión es la denominada “rehablador”. Consiste en que una persona (el rehablador) escucha lo que se dice en el escenario y lo repite frente a un micrófono conectado a una máquina que convierte la voz en texto. Esta máquina solemos llamarla ASR, siglas inglesas de Automatic Speech Recognizer (Reconocedor Automático del Habla).

El problema de esta técnica, es que el rehablador no puede empezar a repetir una frase antes de que el locutor haya acabado de pronunciarla, y algo parecido le ocurre al ASR respecto al rehablador; lo cual introduce un retraso desde que el locutor comenzó la frase hasta que el correspondiente subtítulo que por un lado puede ser muy grande, pero que además es muy variable lo que quiere decir que no es predecible. Esto invalida las técnicas de compensación que se han intentado utilizar aplicando un retraso fijo al video, pues en algunos casos los subtítulos se adelantan mucho y otros se retrasan mucho, siendo muy pocos aquellos subtítulos que aparecen en su justo momento.

Una pregunta típica sobre esta técnica, es ¿Por qué no se utiliza el ARS para convertir en texto directamente lo que dicen los locutores? Eso reduciría el retraso. Lo cierto es que los reconocedores funcionan bastante bien cuando han “aprendido” a reconocer a una persona y cuando es alta la calidad del sonido en términos de ruido y dicción; pero no cuando el locutor no es “conocido” por la máquina o cuando la voz se mezcla con otros sonidos u otras voces.

Así que la solución es enseñar a la máquina a reconocer al rehablador; enseñar al rehablador a hablarle a la máquina; y ubicar al rehablador en un estudio acondicionado.

Programas mixtos

Un tipo de programas que generan una nueva problemática son los programas mixtos en los que se reúnen secciones guionizadas y secciones de puro directo. Ejemplos de estos programas son aquellos programas de noticias en los que se realizan conexiones con reporteros en la calle.

Durante los periodos guionizados los subtítulos están saliendo en su momento correcto; cuando se conecta con el reportero, el sistema de rehablado empieza a genera los subtítulos, que emite con un retraso medio de unos 10 segundos. Al volver al estudio todavía no han sido emitidos los últimos subtítulos de la conexión cuando es necesario emitir los nuevos guionizados. Recuperar este desfase lleva a tener que lanzarlos a más velocidad de la normal lo que en la mayoría de los casos impide que se puedan leer.

Desde 2009 en el CESyA estamos trabajando en una solución que permita reducir en los programas en directo la falta de sincronización entre los subtítulos y el video/audio correspondientes, lo que al mismo tiempo eliminaría la necesidad de recuperación en los programas mixtos. Este trabajo de I+D+i se realiza en base a la patente P201030758 'Procedimiento y Dispositivo para Sincronizar Subtítulos con Audio en Subtitulación en Directo' cuya titularidad pertenece a la Universidad Carlos III de Madrid.