

# Mejora en la accesibilidad de contenidos docentes multimedia: subtitulación

## Improvement in accessibility of multimedia teaching content: subtitling

◆ Francisco Cruz Argudo, David Santín Cristobal, Gloria Vega Lunar

### Resumen

Desde el año 2002 la UC3M lleva trabajando en la generación de contenidos multimedia (docentes y de investigación). Esto nos ha llevado a lo largo de los años a dotarnos de espacios docentes con equipamiento informático y audiovisual que nos permita dicha generación de contenidos, además de desarrollar plataformas que nos permitan catalogar, publicar y distribuir.

Con estos antecedentes, desde finales del año 2009 hemos empezado a trabajar en mejorar la accesibilidad de los contenidos educativos y de investigación, en concreto en el tema de la subtitulación. Para ello, hemos estado durante todo el curso académico 2009-2010 trabajando en una titulación que se imparte en modalidad semipresencial el campus de Colmenarejo: Licenciatura en Biblioteconomía y Documentación, en la cual se graban y emiten en directo todas las clases, y es aquí donde hemos comenzado el piloto.

En el tema de la subtitulación entran en juego varios factores: reconocimiento de voz, generación de los archivos de subtítulos, corrección de los mismos y generación del contenido multimedia definitivo. Todo esto ha sido tenido en cuenta y está en distintas fases de implantación. Además en nuestro caso teníamos la necesidad de tener que generar estos contenidos (en alguno de los casos) en tiempo real para personas con discapacidad que asistían a dichas asignaturas en las clases presenciales, además de poder facilitar tanto a todos los alumnos como al profesor una transcripción en texto plano de la clase entera. En este artículo veremos con más detalle el proceso así como todos los factores.

**Palabras clave:** Subtítulos, contenidos multimedia, reconocimiento de voz, ARCAMM, CESYA.

### Summary

UC3M has been working on creating multimedia content (teaching and for research) since 2002. Over the years, this has provided us with teaching spaces with computer and audiovisual equipment enabling us to create content and also to develop platforms in order us to catalogue, publish and distribute.

With this background, we started working at the end of 2009 on improving accessibility of educational and research content, specifically in the field of subtitling. In the academic year 2009-2010 we worked on a degree course to be taught in a mixed (face-to-face and distance) format at the Colmenarejo campus: Degree in Library and Information Science, where all the classes were recorded and broadcast live. This was the start of the pilot.

Various factors come into play in the area of subtitling: voice recognition, generation of subtitle files, correcting the subtitles and generating the final multimedia content. All of this has been taken into account and is now in different phases of implementation. Also, in our case, we were obliged (in some cases) to generate these contents in real time for people with disabilities who were attending these courses in person; in addition, we had to be able to provide all students and the teacher with a plain text transcript of the whole class. This article provides more detail about the process and all the factors involved.

**Keywords:** Subtitles, multimedia contents, voice recognition, ARCAMM, CESYA.

## 1. Situación de Partida

En la UC3M existen distintos espacios con equipamiento informático y audiovisual que permiten dicha generación de contenidos: aulas de grado, aulas de tele-docencia, salones de grado, aula magna, estudios de grabación,... Para poder poner a disposición de todos los usuarios estos contenidos, ha sido necesario el desarrollo de plataformas que nos permitan catalogar, publicar y distribuir los mismos. Ésta plataforma (ARCAMM) compuesta de varios módulo (agenda, gestión de

◆  
Desde el año 2002  
la UC3M lleva  
trabajando en la  
generación de  
contenidos  
multimedia

◆  
Teníamos la  
necesidad de  
tener que generar  
estos contenidos  
(en alguno de los  
casos) en tiempo  
real para personas  
con discapacidad



contenidos,...) también dispone de un módulo de subtitulación vía WWW. Como resultado de todo esto, en la actualidad la UC3M cuenta con un portal de vídeos que dispone de más de 3.000 contenidos multimedia y más de 6.000 horas.

## 2. Objetivos

Cada vez son más el número de contenidos multimedia docentes y de investigación (clases regladas, seminarios, postgrado, conferencias,...) generados y demandados por parte de la Universidad Carlos III de Madrid. Además, la incorporación de nuestra universidad como pionera en España al Espacio Europeo de Educación superior, no hará sino aumentar el uso de dichas tecnologías como apoyo a la docencia e investigación. Nos enfrentamos por tanto, a un aumento del uso de estas tecnologías por parte de profesores, investigadores y alumnos.

El principal objetivo del proyecto es la mejora de la accesibilidad (mediante la subtitulación) de estos contenidos, tanto desde el ámbito académico como desde diferentes servicios de la universidad:

- **Personas con discapacidad auditiva:** este es uno de los principales objetivos de este proyecto, permitir que personas con discapacidad auditiva puedan seguir las clases a través del portal de videos de la universidad.
- **Entorno:** en situaciones donde no se pueda escuchar el audio de un video por distintos motivos, con el subtítulo y la audiodescripción se puede obtener toda la información que quiere transmitir el video, un ejemplo claro en nuestra universidad está en las pantallas de cartelería electrónica que, dependiendo de su ubicación, no es posible que dispongan de audio
- **Multilinguaje:** un subtítulo ayuda a comprender mejor a una persona que no habla el idioma original del video. Además una vez teniendo el subtítulo se puede adaptar a otros idiomas de manera casi automática.
- **Generación de material adicional:** con un subtítulo se puede obtener contenido adicional como la transcripción literal en texto del audio. Esta funcionalidad está originada en profesores con interés en generar una transcripción de las clases o ficheros de audio grabados, para facilitar este material docente a sus alumnos como apuntes.

Todos estos puntos nos llevaron a la creación de un marco tecnológico que nos permita crear “una prueba de concepto” para evaluar la solución propuesta así como ver la viabilidad y coste que supondría la puesta en marcha como servicio extensible a todos los espacios de grabación y resto de servicios de la universidad.

## 3. Piloto de subtitulación

En la realización de este piloto se ha contado con la **inestimable colaboración** del **CESYA** (Centro Español de Subtitulado y Audiodescripción), que con su herramienta APEINTA nos ha aportado una solución a ciertas fases del proyecto. También han jugado un papel importante los **profesores** de la titulación que nos han ayudado tanto en el proceso de entrenamiento de la herramienta de reconocimiento de voz como aportando ideas y comentarios. Dado que se trata de un proyecto piloto, el mismo ha ido sufriendo modificaciones para adaptarnos a ciertos requisitos planteado por los profesores y los propios integrantes del proyecto.

El principal objetivo del proyecto es la mejora de la accesibilidad de estos contenidos

En la realización de este piloto se ha contado con la inestimable colaboración del CESYA

Durante el primer cuatrimestre del curso 2009/2010 se empezó a trabajar en la mejora de la accesibilidad de los contenidos educativos mediante la inclusión de subtítulos en las clases. Para ello se eligió una titulación que se imparte bajo la modalidad de semipresencialidad en el campus de Colmenarejo: Licenciatura en **Biblioteconomía y Documentación**. El motivo de la elección de esta titulación era doble: se graban y emiten en directo todas las clases y además contaba con una alumna que tenía una discapacidad auditiva que le hacía difícil el seguimiento de las asignaturas.

En el proceso de obtener un subtítulo, entran en juego varios factores: **reconocimiento de voz, generación de los archivos de subtítulos, corrección de los mismos y generación del contenido multimedia definitivo** y por último la **publicación y distribución**. A continuación una breve explicación de cada uno de estos factores, que son los pasos a seguir para poder dotar de un subtítulo a un video:

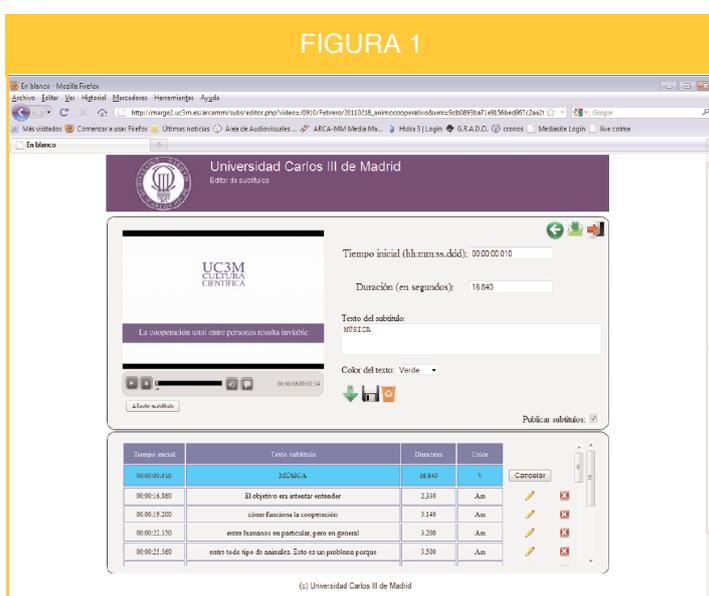
**1. Reconocimiento de voz.** Para una mejor calidad del subtítulo obtenido es necesario realizar previamente un entrenamiento de la herramienta de reconocimiento de voz por parte de los profesores. En la etapa final del proyecto este paso se podía obviar, ya que la herramienta de reconocimiento de voz ofrecía un auto-aprendizaje partiendo de un video con su subtítulo.

**2. Generación de los archivos de subtítulos.** Mediante la herramienta cedida por el CESyA se obtienen los subtítulos correspondientes al video. Esto permite tener una subtitulación en directo (con algunos errores), en nuestro piloto solo hubo subtitulación en tiempo real de dos clases.

**3. Corrección de los mismos.** Esta labor, casi "artesanal", se ha hecho al principio mediante herramientas de libre distribución, al tiempo que se estaba desarrollando una herramienta web a medida para realizar esta tarea de una forma más sencilla, completamente integrada en nuestro portal de videos (ARCAMM). Esta herramienta soporta varios tipos de videos, una vez cargado el vídeo y el archivo, la aplicación realiza el marcado de tiempos de los subtítulos y permite ir corrigiendo mientras el video es visualizado. El sistema, además incluye una gestión de usuarios que permite la asignación de trabajos por usuario. La herramienta es accedida directamente desde la plataforma de catalogación de vídeos y una vez completado la subtitulación el vídeo se puede ver de forma automática a través del portal. Este proceso de corrección lleva aproximadamente 5.2 minutos por cada minuto de video. En el caso que la transcripción se hiciera directamente (sin la generación automática de los archivos de subtítulos) llevaría 12 minutos por cada minuto de video.

En el proceso de obtener un subtítulo, entran en juego varios factores

Para una mejor calidad del subtítulo es necesario realizar previamente un entrenamiento de la herramienta de reconocimiento de voz



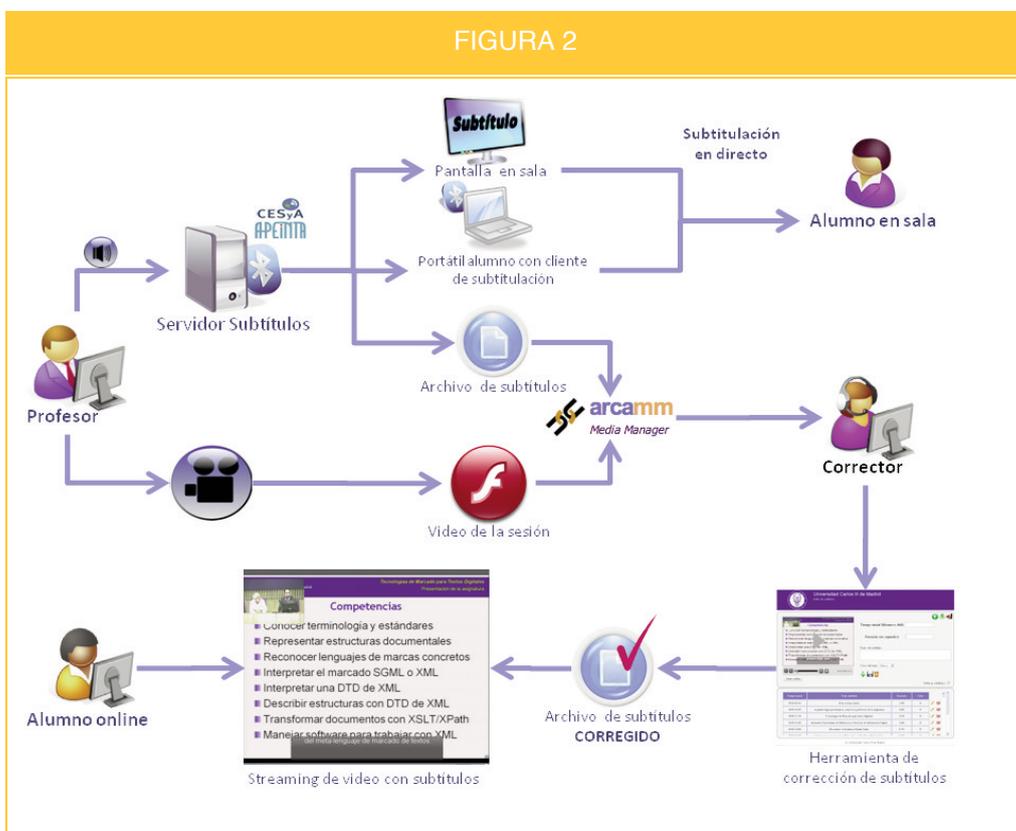


**4. Publicación y distribución** del contenido multimedia generado. Para poder publicar y distribuir el video junto con su subtítulo fue necesario adaptar el portal de videos (arcamm) para que soportara esta funcionalidad. Así que se pasó de que el alumno se tuviera que descargar por separado el video y el subtítulo y verlo en su pc o dispositivo móvil de manera local con un reproductor[1], a poder verlo de manera integrada dentro del propio portal, para esto se desarrolló un reproductor, y se cambio el formato de codificación de los videos a flash[2]. Así, de igual manera los subtítulos generados se pueden incluir en el nuevo sistema de grabación de clases LCS[3].

En la siguiente tabla se muestra de manera gráfica todo el proceso comentando anteriormente:

Para poder publicar y distribuir el video junto con su subtítulo fue necesario adaptar el portal de videos (arcamm)

Para esto se desarrolló un reproductor, y se cambió el formato de codificación



En la siguiente tabla se indica todo el contenido subtulado durante el curso 09/10:

3º Diplomatura Biblioteconomía y Documentación. Curso 09/10					
Asignatura	Nº vistas	Nº Clases	Nº Clases Subtituladas	% clases subtituladas	Horas subtituladas
Estudio de Usuarios	417	7	7	100%	5:15
Catalogación Automatizada	638	7	4	57%	3:00
Catalogación Materiales Especiales	882	6	6	100%	4:30
Automatización de Centros y Servicios de Información	257	4	4	100%	3:00
<b>Totales</b>	<b>2194</b>	<b>24</b>	<b>21</b>	<b>87%</b>	<b>15:45</b>

De tal manera que se llegó casi a tener el 90% de las clases subtítuladas (las que faltaron pertenecían al primer cuatrimestre cuando se inició el piloto), ofreciendo de esta manera el acceso casi completo al curso entero para la alumna con discapacidad auditiva.

Con los datos comentados anteriormente se tuvieron que invertir 82 horas exclusivamente para la corrección de los subtítulos, en el resto del proyecto se tuvo que invertir más tiempo en los apartados correspondientes de modificación del portal de videos y generación de los subtítulos.

### 4. Conclusiones y líneas futuras

El aportar accesibilidad a los contenidos multimedia, será algo casi imprescindible a medio y largo plazo en la UC3M, principalmente para hacer accesibles los contenidos a personas con discapacidad, y partiendo de aquí las ventajas vistas anteriormente (apartado 2, Objetivos).

Siguiendo con este piloto, para poder escalar, será necesario automatizar todo lo posible el proceso de la obtención de los subtítulos, teniendo en cuenta que la corrección de los subtítulos es algo que hay que hacer de manera manual para que el subtítulo siga con toda fidelidad a su vídeo original.

Además estamos trabajando en la evaluación de generación de subtítulos sin necesidad de entrenamiento y evaluando algunas otras herramientas de reconocimiento de voz. Durante todo este año se han extraído datos sobre la carga de trabajo tanto de los docentes como de los servicios implicados y su posible escalabilidad. Por último estamos trabajando en la automatización completa del proceso mediante el desarrollo de una aplicación que nos permita la generación de subtítulos de una forma desatendida.



Será necesario automatizar todo lo posible el proceso de la obtención de los subtítulos

### Referencias

[1] <http://marge2.uc3m.es/arcamm/item.php?mod=vod&id=6953038657dd8ec440632a004379c6a2>

[2] <http://marge2.uc3m.es/arcamm/item.php?mod=vod&id=a7de5a74c0c6418557fc5b9af5119a1f>

[3] <http://marge2.uc3m.es/arcamm/item.php?mod=vod&id=a7de5a74c0c6418557fc5b9af5119a1f>

**Francisco Cruz Argudo**

(paco@di.uc3m.es)

**David Santín Cristobal**

(dsantin@di.uc3m.es)

**Gloria Vega Lunar**

(gvega@di.uc3m.es)

Universidad Carlos III de Madrid

Servicio de Informática y Comunicaciones

Área de Audiovisuales