

Subtitulado Cerrado para la Accesibilidad de Personas con Discapacidad Auditiva en Entornos Educativos

Closed Captioning for Accessibility of Hard of Hearing People in Educational Environments

Pablo Revuelta Sanz

CESyA

prevuelt@ing.uc3m.es

José Manuel Sánchez Pena

Departamento Tecnología Electrónica

jmpena@ing.uc3m.es

Javier Jiménez Dorado

CESyA

jjdorado@ing.uc3m.es

Belén Ruiz Mezcua

Departamento Informática

bruiz@inf.uc3m.es

Universidad Carlos III de Madrid,
Av. Universidad, 30, 28911 Leganés, Madrid, España

Resumen: El objetivo de este proyecto es contribuir a la integración de las personas con discapacidad auditiva en la educación. El sistema está basado en el reconocimiento automático del habla (ASR) y en conversión de texto a voz. Se ha implementado una arquitectura cliente servidor con comunicación inalámbrica que puede funcionar con tres tipos diferentes de dispositivos. El sistema genera dos recursos educativos como salida, uno es la voz del ponente y otro es la transcripción de esta. El proceso ASR está realizado con Dragon NaturallySpeaking y la conversión de texto a voz está realizada con el Speech API de Microsoft.

Palabras clave: subtitulado en directo, reconocimiento del habla automático, subtitulado cerrado, accesibilidad a la educación, aplicación multidispositivo, educación inclusive.

Abstract: The aim of this project is facilitate the integration of hard of hearing people in education. The system is based on automatic speech recognition (ASR) and in text to speech conversion. A client-server architecture was implemented with wireless communication which can run on three different devices. The system provides two output files with the transcription and the audio of the speech. ASR process was done with Dragon NaturallySpeaking and the Text to Speech process with the Microsoft's Speech API.

Keywords: live captioning, automatic speech recognition, closed captioning, educational accessibility, multidevice application, inclusive education.

1 Introducción

Las personas con discapacidad y en particular las personas sordas se tienen que enfrentar a múltiples dificultades en sus vidas. También en la hora de acceder a la educación existen barreras que les impiden disfrutar y participar del proceso educativo plenamente por lo que nos encontramos lejos de lo que llaman *Educación Inclusiva* [1]. Los avances tecnológicos y las investigaciones audiológicas, como los implantes cocleares o bucles de inducción, han supuesto una gran integración de los estudiantes sordos en el proceso educativo y, naturalmente, en su vida diaria. Sin embargo, todavía se encuentran con varios problemas como la comprensión del lenguaje y del habla como consecuencia de la sordera [2] pero también debido al ambiente ruidoso de las clases, la distancia con el profesor, las reverberaciones, etc. [3]. Para solucionar estos problemas queremos proponer una nueva ayuda técnica que facilite la integración de los estudiantes con discapacidad auditiva en el nuevo modelo de educación inclusiva. Para ello utilizaremos los conceptos de subtitulado cerrado y subtitulado en directo [4]. Este sistema permite transcribir el discurso del profesor y lo envía vía radio de forma individual y en forma de subtítulos a los estudiantes sordos. Además, los estudiantes con problemas del habla pueden escribir preguntas y comentarios en sus dispositivos portátiles

para que una voz sintética reproduzca lo que han escrito.

2 Estructura del Sistema

El sistema implementado tiene una arquitectura cliente servidor. El servidor corre en el ordenador del profesor/a que puede seguir siendo usado para otras aplicaciones. Los estudiantes sordos tendrán cada uno, o podrán compartir, un dispositivo portátil donde una aplicación cliente permite visualizar los subtítulos.

2.1 Servidor

El servidor se divide en las siguientes partes:

- Aplicación ASR
- Programa principal
- Comunicaciones
- Aplicación Texto a Voz

El modo en que interactúan cada una de las partes se muestra en la figura 1.

Reconocimiento Automático del Habla. La aplicación de ASR se consigue mediante el reconocedor de voz comercial Dragon NaturallySpeaking (DNS). Se ha creado una conexión interna TCP para mandar de forma automática el texto transcrito al programa principal. Se eligió DNS en un principio por su menor tasa de errores aún sabiendo que el tiempo empleado es mayor que el ViaVoice.



Fig. 1. Interacción de los procesos que corren en el servidor.

Programa Principal. Está implementado con LabVIEW 8.5. Proporciona una sencilla interfaz de usuario al profesor/a. Entre sus funciones se encuentra el manejo de las conexiones TCP y de las interfaces de comunicación. Sincroniza y gestiona las aplicaciones de ASR y de texto a voz. Se encarga de transformar la transcripción para crear subtítulos según la norma de subtítulo para teletexto [5]. Almacena en un fichero de texto (XML) la transcripción de la clase así como los comentarios de los estudiantes sordos recibidos y en un fichero de audio (WAV) la voz del micrófono.

Comunicaciones. Existen dos modos alternativos de comunicación: un enlace serie radio a 433MHz y un enlace Bluetooth. Ambos son necesarios para poder dar cobertura a diferentes dispositivos clientes.

Texto a Voz. La voz sintética se controla desde el programa principal mediante una comunicación interna de ActiveX que usa el Speech API 5.1 de Microsoft. El profesor/a puede controlar el tipo de voz, la velocidad, el volumen y la salida física.

2.2 Clientes

Existen tres tipos diferentes de clientes:

- un circuito electrónico con un pequeño display en forma de gafas y un teclado.
- una PDA
- un ordenador portátil o de sobremesa

El dispositivo microcontrolado se comunica a través del enlace radio serie. Posee un display en unas gafas y un teclado para poder leer los subtítulos y escribir preguntas.

La PDA está disponible para Windows Mobile, Pocket PC y Palm OS. Ha sido desarrollada con LabVIEW 8.5 al igual que la aplicación para el ordenador. La PDA sólo se comunica con Bluetooth mientras que el ordenador puede utilizar ambos sistemas de comunicación.

La interfaz del ordenador portátil o sobremesa es una versión avanzada de la PDA desarrollada para Windows que contiene funcionalidades extra.

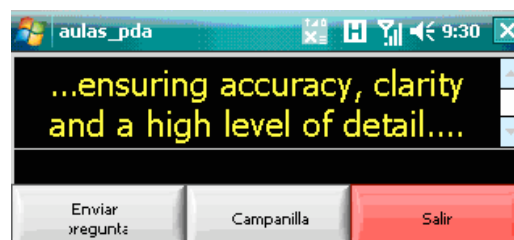


Fig. 2. Interfaz para PDA.

3 Conclusiones

El sistema es compatible con sistemas de FM y bucles magnéticos. Tiene una cobertura de 50 metros para el enlace radio y hasta 25 con algunos dispositivos Bluetooth. Funciona de forma adecuada cuando se entrena el reconocedor consiguiendo muy bajas tasas de error (~5%), sobre todo si se adapta el vocabulario a la materia a impartir. La mayor desventaja que posee el sistema es el retardo (hasta 6 segundos) que existe debido al tiempo que invierte Dragon. Esta latencia se solventa si cambiamos a ViaVoice aunque se necesite mucho más entrenamiento para conseguir las mismas tasas de error. Otro inconveniente es el ruido de fondo y sobre todo de las voces de los estudiantes. Para ello se requiere la utilización de micrófonos que minimicen estos efectos. El siguiente paso es conseguir incorporar los signos de puntuación. Aunque este sistema ha sido diseñado para estudiantes sordos también puede utilizarse para estudiantes extranjeros o con problemas de comprensión o del habla. Además, existe la posibilidad de aplicar este sistema para la accesibilidad en museos con guías, conferencias, congresos, etc [6].

Bibliografía

1. Arnáiz Sánchez, P.: Educación Inclusiva: Una escuela para todos. Ediciones Aljibe, Málaga (2003)
2. Davis, J. M., Effenbein, J., Schum, R. and Bentler R. A.: Effects of Mild and Moderate Hearing Impairments on Language, Educational, and Psychosocial Behavior of Children. *Journal of Speech and Hearing Disorders*, 51. 53-62. (1986)
3. Pekkarinen, E. and Viljancn, V.: Acoustic Conditions for Speech Communication in Classrooms. *International Journal of Audiology*, 20(4). 257-263 (1991)
4. Revuelta P., Jiménez J., Sánchez J. M. y Ruiz B.: Online Captioning System for Educational Resources Accessibility of Hard-of-Hearing people. ICCHP (2008)
5. AENOR, 2003. UNE 153010: Subtitulado para personas sordas y personas con discapacidad auditiva – Subtitulado a través del Teletexto. (2003)
6. Revuelta P., Jiménez J., Sánchez J. M. y Ruiz B.: Multidevice System for educational accessibility of Hearing impaired students. Accepted at CATE (2008)