



POLITÉCNICA



**UNIVERSIDAD POLITECNICA DE MADRID
E.T.S. DE INGENIEROS DE TELECOMUNICACION**

ACTA DE EXAMEN

Asignatura: PROYECTO FIN DE CARRERA

TÍTULO DEL PROYECTO: DESARROLLO DE UN SISTEMA AVANZADO DE AYUDA A LA COMUNICACIÓN ORAL PARA PERSONAS SORDAS.

APELLIDOS Y NOMBRE: MARTÍN MAGANTO, RAQUEL	CALIFICACIÓN 10.00 <i>PARA CALIFICACIÓN DE HONOR</i>
----------------------------------------------------------	----------------------------------------------------------------

Tutor: D. RUBÉN SAN SEGUNDO HERNÁNDEZ

Madrid, a

16 de Julio de 2010.

EL VOCAL PRIMERO

EL PRESIDENTE

EL VOCAL SECRETARIO

MIEMBROS DEL TRIBUNAL

Presidente: D. JAVIER FERREIROS LÓPEZ

Vocal: D. JUAN MANUEL MONTERO MARTÍNEZ

Secretario: D. RUBÉN SAN SEGUNDO HERNÁNDEZ

PROYECTO FIN DE CARRERA

TÍTULO: Desarrollo de un Sistema Avanzado de Ayuda a la Comunicación Oral para Personas Sordas

AUTORA: Raquel Martín Maganto

TUTOR: D. Rubén San Segundo Hernández

DEPARTAMENTO: Departamento de Ingeniería Electrónica (DIE)

GRUPO: Grupo de Tecnología del Habla (GTH)

MIEMBROS DEL TRIBUNAL CALIFICADOR:

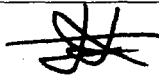

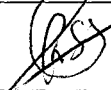
PRESIDENTE: D. Javier Ferreiros López

VOCAL: D. Juan Manuel Montero Martínez

SECRETARIO: D. Rubén San Segundo Hernández

SUPLENTE: D. Luis Fernando D'Haro Enríquez

FIRMA

FECHA DE LECTURA: 16 de Julio de 2010

CALIFICACIÓN: Matrícula de Honor, 10*

UNIVERSIDAD POLITÉCNICA DE MADRID

**ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE
TELECOMUNICACIÓN**



PROYECTO FIN DE CARRERA

**DESARROLLO DE UN SISTEMA AVANZADO DE
AYUDA A LA COMUNICACIÓN ORAL PARA
PERSONAS SORDAS**

Raquel Martín Maganto

Madrid, 2010

RESUMEN

Este Proyecto Fin de Carrera se centra en desarrollar una nueva versión de un sistema software de traducción automática de habla en castellano a Lengua de Signos Española (LSE), y viceversa, que incluye representación gráfica de signos y generación de voz. El objetivo es implementar un sistema de aplicación real que facilite la comunicación entre las personas sordas y los funcionarios de la Administración Pública, en un entorno relacionado con el servicio de solicitud y renovación del DNI, ampliado a las tramitaciones asociadas al permiso de conducir.

El sistema completo consta de dos aplicaciones principales correspondientes a la traducción LSE-voz y voz-LSE, y de una herramienta de edición de signos.

La primera de ellas cuenta con un entorno de selección de signos mediante glosas; un avatar para su representación; un módulo de traducción que los convertirá en texto en castellano, y un conversor texto a voz que permitirá al funcionario escuchar la frase traducida. En este proyecto se ha mejorado su interfaz adaptándola a las necesidades y preferencias de los usuarios finales. Para ello, entre otras modificaciones, se desarrolla una herramienta de creación e incorporación a la aplicación de gifs de signos; se integra un módulo para lenguaje SMS y se actualiza el diseño, obteniendo uno mucho más visual, con imágenes y animaciones sustituyendo a texto.

El segundo subsistema consta de un reconocedor automático de habla que decodifica la voz procedente del micrófono en texto; un módulo de traducción que lo convierte a LSE, y un bloque de representación de signos mediante un agente animado.

Para la traducción, ambos sistemas combinan diferentes estrategias e incorporan un sistema de decisión, dependiente de umbrales, que seleccionará el mecanismo de traducción que proporcionará la salida final. La traducción basada en ejemplos se mantiene en los dos sistemas, mientras que, la basada en reglas se amplía en el sistema LSE-voz y se desactiva en el de voz-LSE tras obtener resultados comparables con la traducción estadística. Además, se añade en ambos sistemas el método de traducción estadístico que proporciona mejor resultado.

El editor empleado para la generación de la base de datos de signos dispone de un módulo de traducción SEA (Sistema de Escritura Alfabética)-HamNoSys (Notación simbólica equivalente a la codificación que precisa el avatar) y la posibilidad de seleccionar los movimientos asociados a un signo. Se desarrolla una versión más manejable y potente, incorporando una herramienta de creación de bitmaps y gifs, y ampliando el registro de gestos no manuales.

Además, se estudian, implementan y evalúan nuevas estrategias de traducción de voz en castellano a LSE basada en ejemplos, las cuales incorporan información del reconocimiento de habla o de la traducción estadística en los costes empleados en el cálculo de la distancia, realizan una segunda categorización, o filtran palabras por confianzas de reconocimiento. Para su evaluación se emplea un corpus bilingüe castellano-LSE de ejemplos y frases salida del reconocedor de voz aplicado sobre archivos de audio con la grabación de los ejemplos en el dominio del DNI. Los mejores resultados se alcanzan al incluir información de la traducción en los costes de edición

utilizados para calcular la distancia de alineamiento, consiguiendo una tasa de traducción correcta de **97,43%**.

PALABRAS CLAVE

Lengua de Signos Española (LSE), Glosa, Interfaz gráfica de usuario, Traducción automática, Traducción basada en ejemplos, Traducción basada en reglas, Traducción estadística basada en subfrases, Traducción estadística basada en transductores de estados finitos, Reconocimiento automático de habla, Agente animado virtual, Tasa de palabras correctas, PIWER, NIST, BLEU.