

# PLAN DE EVALUACIÓN PARA RECONOCIMIENTO AUTOMÁTICO DEL HABLA FRENTE A HABLA PATOLÓGICA

## 1.- Introducción

El presente Plan de Evaluación para Reconocimiento Automático del Habla frente a Habla Patológica proporciona las bases para la evaluación de aplicaciones en este campo de Tecnologías del Habla.

El objetivo de esta evaluación es fomentar los trabajos de investigación en relación con el reconocimiento automático del habla en entornos no colaborativos. Con esta finalidad se plantea un plan de evaluación que consiste en el reconocimiento de palabras aisladas sobre un corpus de voces con o sin patologías. El vocabulario a reconocer lo forman 57 palabras del "Registro Fonológico Inducido" Marc Monfort, Adoración Juárez, Ed. CEPE, ISBN: 8486235987.

01 - arbol	13 - flan	25 - luna	37 - piano	49 - sol
02 - boca	14 - fresa	26 - llave	38 - pie	50 - tambor
03 - bruja	15 - fuma	27 - mariposa	39 - pi~na	51 - taza
04 - cabra	16 - gafa	28 - moto	40 - pistola	52 - telefono
05 - campana	17 - globo	29 - ni~no	41 - platano	53 - toalla
06 - caramelo	18 - gorro	30 - ojo	42 - playa	54 - toro
07 - casa	19 - grifo	31 - pala	43 - preso	55 - tortuga
08 - clavo	20 - indio	32 - palmera	44 - pueblo	56 - tren
09 - cuchara	21 - jarra	33 - pan	45 - puerta	57 - zapato
10 - dedo	22 - jaula	34 - peine	46 - raton	
11 - ducha	23 - lapiz	35 - periodico	47 - semaforo	
12 - escoba	24 - lavadora	36 - pez	48 - silla	

Tanto para el entrenamiento como para la evaluación se utilizarán parte de los corpus de señales de dos bases de datos disponibles en la actualidad:

El corpus que denominaremos "Canarias" ha sido grabado por la Universidad de Las Palmas de Gran Canaria (persona de contacto Juan Luis Navarro Mesa) que contiene 19 locutores sin patologías y 1077 locuciones y 45 locutores con patologías con 2470 locuciones.

El corpus que donominaremos "Alborada" esta siendo grabado por la Universidad de Zaragoza (persona de contacto Eduardo Lleida Solano) que contiene locuciones de estudiantes del colegio de educación especial "Alborada" de Zaragoza con edades comprendidas entre 10 y 21 años.

Las dos bases de datos se partirán en dos subconjuntos, uno para entrenamiento/adaptación y ajuste y otro para la evaluación propiamente dicha.

Los participantes se comprometen a presentar los resultados de la evaluación en una sesión especial que tendrá lugar durante las IV Jornadas en Tecnología del Habla. La participación se realiza a modo individual o en equipo donde representante deberá ser estudiante.

## **2.- Medición de prestaciones**

Al tratarse de una evaluación de reconocimiento de palabras aisladas, con un palabra por fichero, directamente las prestaciones se medirán en términos del

número de palabras correctamente reconocidas/número de palabras a reconocer

Como información adicional se utilizará la matriz de confusión entre palabras.

## **3.- Condiciones de evaluación**

Se definirán dos condiciones de evaluación, una dependiente del locutor y otra independiente.

Para la evaluación dependiente del locutor, el material de entrenamiento contendrá al menos 2 repeticiones de cada palabra de cada locutor a evaluar. Para la evaluación independiente del locutor, el material de entrenamiento contendrá mas de 500 locuciones de habla no patológica y mas de 1000 locuciones de habla patológica.

Todos los archivos contendrán muestras digitalizadas de audio con únicamente una palabra. El formato será el siguiente:

- Frecuencia de muestreo: 16 KHz
- Canales de audio: mono
- Tamaño de muestra: 16 bits
- Alineamiento de octeto: LSB primero (little endian).
- Ficheros PCM sin cabecera ni compresión.

Los archivos de audio tendrán la extensión **.pcm**.

## **4.- Formato de los archivos de etiquetas a generar**

Los resultados del reconocimiento, para cada sistema evaluado, se deberán presentar en un archivo de texto (.res), con un nombre identificador del equipo y del sistema. Por cada línea, y separado por ":" se indicará el nombre del fichero de test, la palabra reconocida y opcionalmente una medida de confianza definida como un decimal entre 0 y 1, siendo 1 la máxima confianza. Un ejemplo, para el sistema 1 del equipo I3A, el fichero de resultados sería

Archivo: I3A\_1.res

B\_00\_23.pcm:zapato:0.789

B\_01\_45.pcm:semáforo:0.123

....

....

## 5.- Datos de evaluación

El material de evaluación se definirá tanto para el reconocimiento dependiente del locutor, donde se intentará disponer de mas de 40 locutores para la evaluación, como para el reconocimiento independiente del locutor donde se intentará disponer de mas de 20 locutores para la evaluación.

El formato de estos archivos de audio será el mismo que el de los archivos de entrenamiento y se nombrarán con la misma extensión (.pcm).

## 6.- Procedimiento para la evaluación

El procedimiento con las fechas para la evaluación es el siguiente:

- El 21 de Julio de 2006 se dispondrá de los planes de evaluación y se abre el periodo de inscripción.
- La fecha límite de inscripción será el 15 de Septiembre de 2006.
- A partir del 16 de Agosto de 2006 se podrá disponer del material de entrenamiento y desarrollo para las distintas evaluaciones. Es necesario estar inscrito en la evaluación para recibir el material.
- El 18 de Octubre de 2006 se liberarán las bases de datos para la evaluación.
- El 27 de Octubre de 2006 a las 24:00 es la fecha límite para recibir los resultados en el formato y método indicados.
- El 3 de Noviembre de 2006 se enviarán los resultados de la evaluación.

## 7.- Envío de segmentaciones

Los resultados de reconocimiento se enviarán por correo electrónico a la organización. Deberán ser completas, conteniendo por tanto todo el conjunto de datos de evaluación.

Los ficheros de texto con los resultados, con la extensión “.res”, deberán enviarse comprimidas en un archivo \*.zip a: **Eduardo Lleida Solano** (lleida@unizar.es)

Los resultados estarán disponibles una vez se hayan enviado dichos resultados a los participantes. Esto permitirá realizar análisis previos a la celebración de las IV Jornadas de Tecnologías del Habla.

Cada participante deberá remitir una descripción del sistema enviado a la evaluación, que deberá incluir:

- Nombre del sistema
- Condiciones de evaluación (base de datos de entrenamiento)
- Descripción de la aproximación algorítmica

Esta descripción se enviará con la misma plantilla utilizada para las comunicaciones regulares de las IV Jornadas en Tecnología del Habla. Las descripciones recibidas se distribuirán como parte del material de análisis de la evaluación.