

SISTEMA DE CÓMPUTO PARA LA COMUNICACIÓN HOMBRE -MÁQUINA

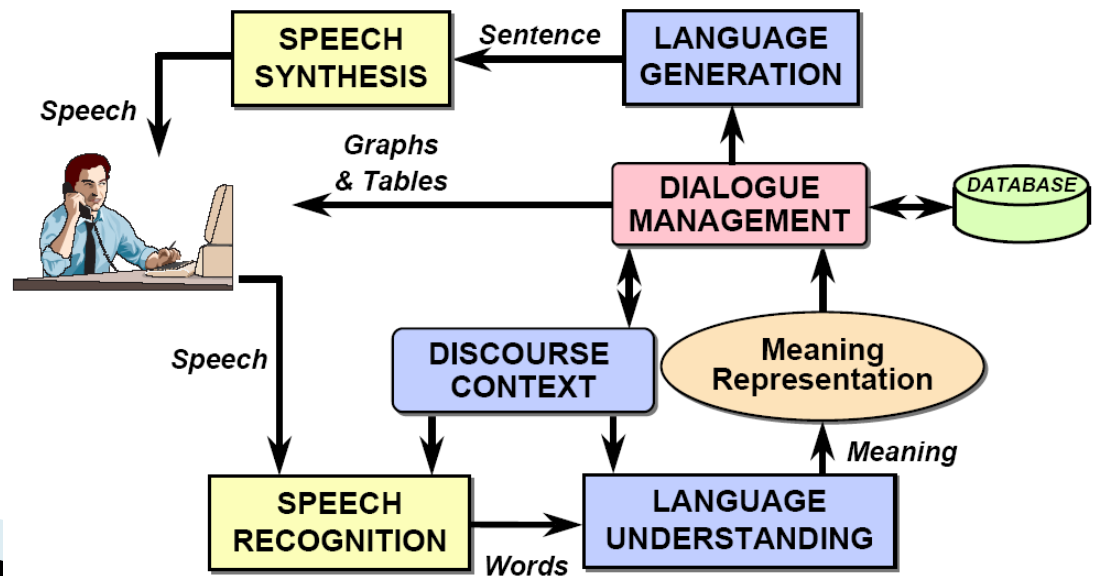
Dr. Sergio Suárez Guerra
Dr. José Luis Oropeza Rodríguez
Dr. Luz Noé Oliva Moreno
Dr. Roberto Herrera Charles
Dr. Alfonso Gutiérrez Aldana

Objetivo principal

Disponer de los modelos y herramientas suficientes para la comunicación hombre – máquina usando aprendizaje automático (machine learning).

RESUMEN

- ▶ *Se presenta un proyecto de investigación y de desarrollo tecnológico, que logre dar respuesta de forma automática, a que las capacidades humanas de producción del habla y su reconocimiento, sean realizadas por un sistema de cómputo.*
- ▶ *En los resultados obtenidos se tendrán los herramientas y módulos necesarios para aprendizaje automático de idiomas y análisis, en tareas de reconocimiento de personas así como de especies animales.*

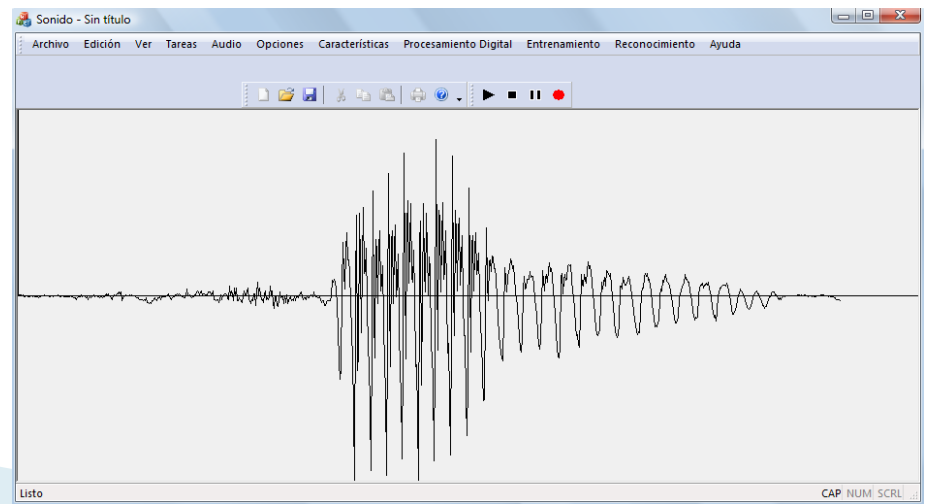


Objetivos específicos

- ▶ *Diseñar un modelo computacional de la adquisición del lenguaje hablado en edad temprana: palabras, sílabas, fonemas.*
- ▶ *Implementar traducción acotada entre los idiomas aprendidos.*
- ▶ *Llevar a cabo separación de locutores (voz forense) y certificación de locutores.*
- ▶ *Realizar la conversión de voz entre usuarios.*
- ▶ *Analizar voces de diferentes especies para realizar su reconocimiento y clasificación.*
- ▶ *Trabajar en la Interfaz de señales EEG (BCI) – Brain Computer Interface) para su conversión a voz.*
- ▶ *Entre otros.*

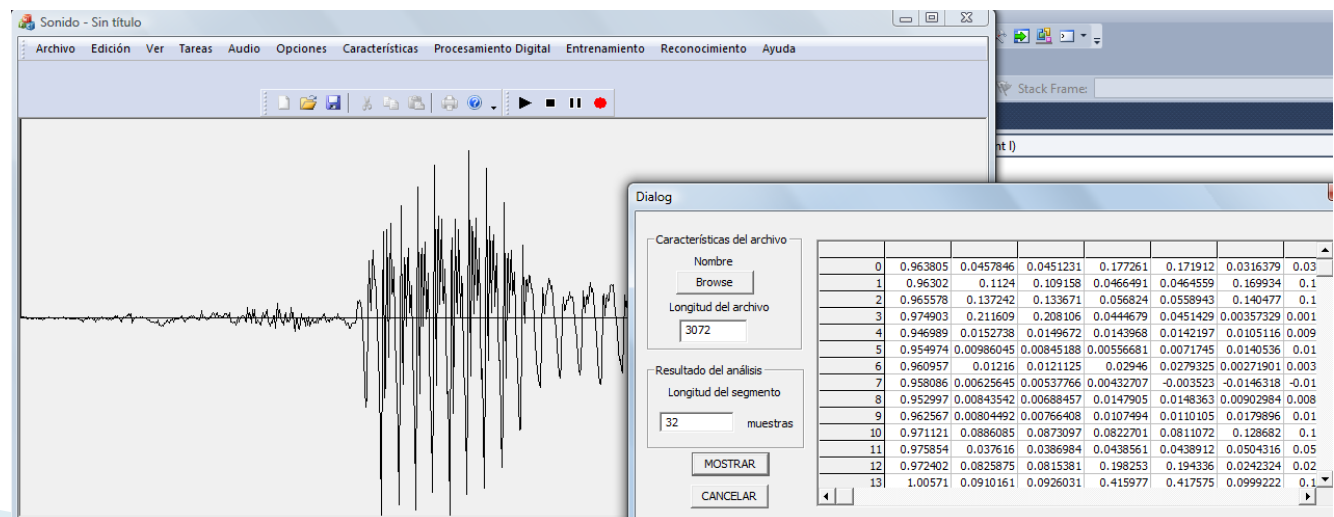
HERRAMIENTAS Y MÓDULO CREADOS

- ▶ *Software para lectura, extracción de características y análisis de señales de voz.*
- ▶ *Módulo para etiquetado fonético.*
- ▶ *Módulos de programas que implementan técnicas de machine learning.*
- ▶ *Módulos para reconocimiento de voz.*
- ▶ *Módulo para separación de fuentes sonoras.*



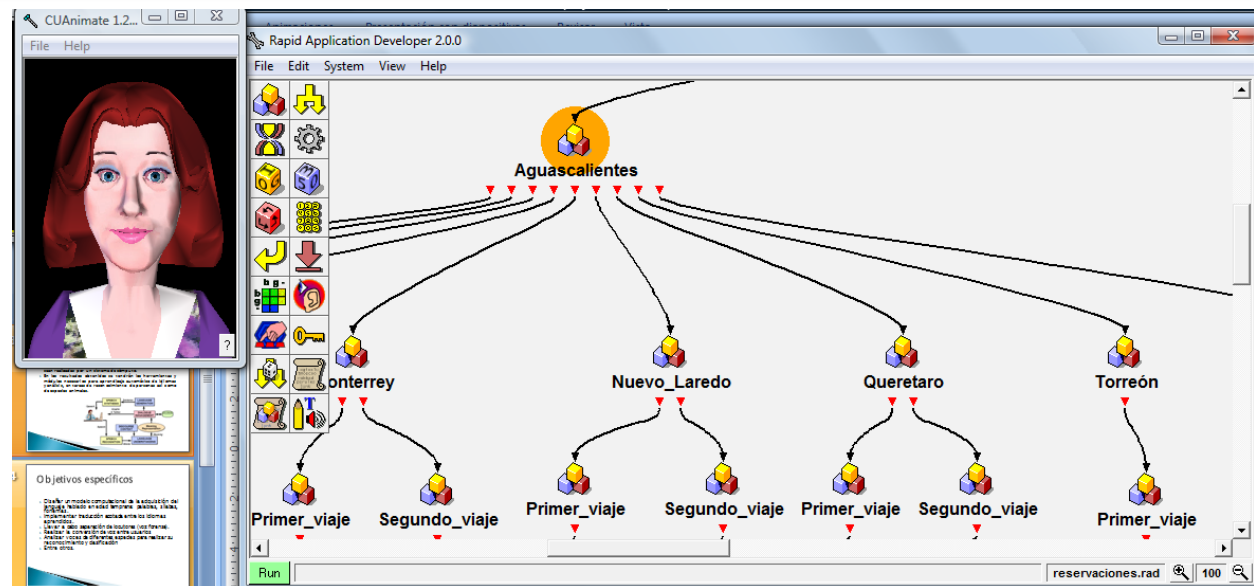
MÓDULOS Y PROPUESTA A DESARROLLAR

- ▶ *Automatizar el entrenamiento del corpus de voz a reconocer.*
- ▶ *Integración de los módulos en prototipos de hardware y software.*
- ▶ *Módulos para análisis y clasificación de especies animales.*
- ▶ *Incorporar los módulos de traducción entre idiomas.*
- ▶ *Incorporar módulos de conversión de voz.*
- ▶ *Diseño y realización de módulos para la certificación de locutores.*
- ▶ *Creación y búsqueda de las bases de datos para las aplicaciones.*
- ▶ *Entre otras propuestas.*



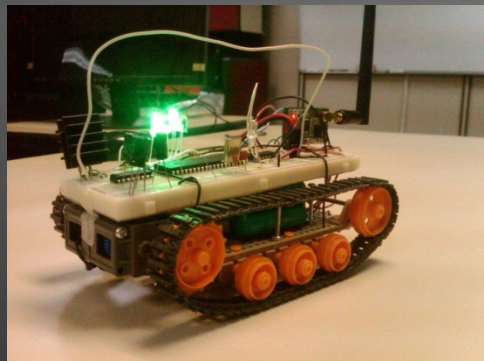
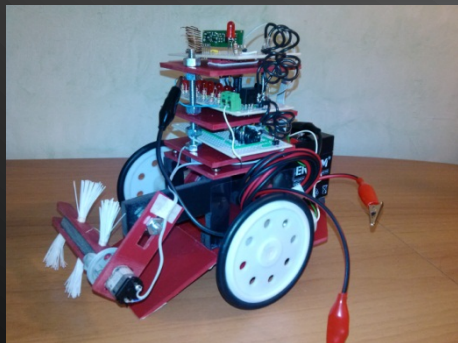
Integración

- ▶ Integrar estudiantes y profesores de ES.
- ▶ Asignación de tareas identificadas en las propuestas a desarrollar para la elaboración de un proyecto integral.
- ▶ Capacitación de los profesores y estudiantes participantes.

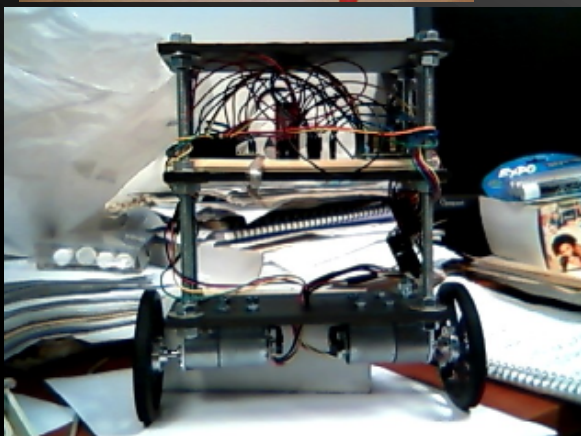


CSLU (Center for Spoken Language Understanding)
Toolkit

MODELOS Y EQUIPOS YA CONSTRUIDOS PARA SER UTILIZADOS PARA LA INTERACCIÓN HOMBRE – MÁQUINA



*CASA
INTELIGENTE*



EXPARAM: sistema para la implantación y rehabilitación del lenguaje

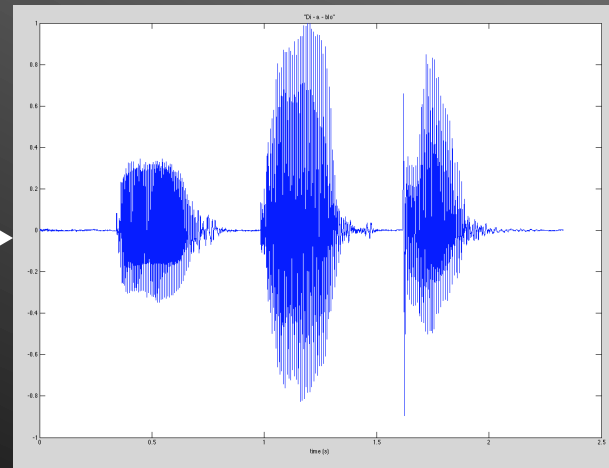
The screenshot displays the EXPARAM software interface. At the top, the title bar reads "Exparam" and the menu bar includes "Programa", "Opciones", "Reproducir", "Grabar", "Parámetros", "Visualización", "Cálculos", and "Ayuda".

The main control area features a "Sonido File" field containing "C:\Exparam\Voces\prueba.wav" and a "Profesor" button. Below these are buttons for "Pausa", "Detener", "Limpiar", and "Repetir" (with an unchecked checkbox). A set of eight red arrow buttons (four left-pointing and four right-pointing) is arranged in two rows, with a "Comparar" button in the center. To the right, there are buttons for "Capturar", "Terminar", and "Salvar_en", along with volume controls labeled "VMedia", "VMax", "VJitter", and "VMin".

Parameter settings include "Frec. Muestreo" (12800) and "Num. Muestras" (25600), both in dropdown menus.

The lower section shows two spectrograms. The top one is labeled "prueba" and "AO" (Analog Output), with a red waveform and a "PRUEBA" label. The bottom one is labeled "Campana" and "AO", with a yellow waveform and a bell icon.

SISTEMA DE INTERACCIÓN HOMBRE – MÁQUINA USANDO ANDROID



RETROALIMENTACIÓN

**Librería del Motor de Reconocimiento de HMM
(HRec)**

Capa de Reconocimiento

*Rutinas de actualización del Modelo
(HMap)*

*Rutinas de
Soporte de
Entrenamiento de
HMM (HFB)*

*Librería de
Adaptación
(HAdapt)*

*Rutinas de
Soporte de Red y
Lattice (HNet)*

*Librería de
Gestión de Lattice
(HLat)*

Capa de Servicios de Modelado Acústico

Capa de Servicios de Modelado del Lenguaje

*Archivos de
parámetros para voz
de E/S
(HParm)*

*Rutinas de
Utilidad para
HMM
(HUtil)*

*Definición del Tipo
de dato del Modelo
HMM (HModel)*

*Rutinas de Soporte
de Entrenamiento de
HMM (HTrain)*

Capa de Servicios de Soporte para Entrenamiento

*Etiquetas de voz
(HLabel)*

*Gestión de archivos
de audio
(HAudio)*

*Gestión de
diccionario (HDict)*

*Modelado del
Lenguaje (HLM)*

Capa de Servicios de Soporte

*Cuantificación
Vectorial (HVQ)*

*Procesamiento Digital
de Señales (HSigP)*

*Modo Gráfico
(HGraf)*

*Gestión de tipos de
ondas (HWave)*

Capa de Servicios Secundarios

OPERACIONES MATEMÁTICAS (HMath)

**GESTIÓN DE
MEMORIA (HMem)**

**(ENTRADA/SALIDA)
(HShell)**

**(DIAGNÓSTICO)
(HShell)**

**(INTERFAZ DE
USUARIO) (HShell)**

Capa de Servicios Básicos