

Nuevos Aportes al entrenamiento de Personas Hipoacúsicas

Garin Juan Manuel
Facultad de Ingeniería
UNMDP
Republica Argentina
jmgarin@fi.mdp.edu.ar

Agüero Pablo Daniel
Facultad de Ingeniería
UNMDP
Republica Argentina
pdaguero@fi.mdp.edu.ar

Tulli Juan Carlos
Facultad de Ingeniería
UNMDP
Republica Argentina
jctulli@fi.mdp.edu.ar

ABSTRACT

This on-going project focuses on the need of software for the assistance of deaf people in Argentina, with a special emphasis in the development of freeware with total support of Argentinian dialects.

Our software employs the latest technological advances in the software area by using JAVA programming language. The main advantage of this approach is the high portability and the avoidance of multiple code development to support different platforms.

The user interacts with the software through the computer screen. The system asks the user to utter some specific sound, word or sentence, which is recorded using a microphone. Then, the speech signal is analyzed and compared with existing reference patterns in a speech database. The results are shown in the screen. They include both the user's mistakes and possible actions to correct them.

All the data generated during the interaction between user and software are sent via Internet to the servers at the university where the development of the project takes place. This information is important for future improvements of the analysis tools and to check the correct performance of the implemented algorithms under different recording conditions.

RESUMEN

Este proyecto en desarrollo analiza la problemática argentina actual sobre la no existencia de programas para computadora destinados a la asistencia de personas hipoacúsicas, de distribución gratuita y en idioma español local.

Debido a esto, se decidió diseñar un programa que emplee los últimos avances tecnológicos en el área informática, utilizando el lenguaje de programación JAVA, que nos permite una amplia portabilidad del programa, eliminando las limitaciones de incompatibilidad de programación entre sistemas operativos.

Su funcionamiento consiste en indicarle en pantalla al usuario lo que se pretende que este pronuncie, capturar su habla a través de un micrófono, analizarla y comparar los resultados con patrones preestablecidos. Luego de esto se le informa nuevamente a través de la pantalla los posibles errores lingüísticos a fin de que este pueda ir mejorándolos con el continuo entrenamiento.

Los datos de este proceso son enviados vía Internet a los servidores del grupo de desarrollo permitiendo realizar análisis posteriores tanto por el profesional a cargo como por los desarrolladores, lo que facilita mantener un registro de sus actividades y progresos.

KEYWORDS

hipoacúsicos, programación java, entrenamiento del habla, distribución gratuita,

INTRODUCCIÓN

En la actualidad existen personas que presentan una disminución en la capacidad auditiva, las cuales se pueden clasificar en dos grupos, los hipoacúsicos y los sordos profundos (cofóticos).

La hipoacusia es la disminución de la percepción auditiva. La misma puede provenir de distintas causas congénitas tanto de origen genético, neurosensorial, como por mal formaciones; o bien adquirirse a través de enfermedades tales como otitis, tímpanosclerosis, colesteatomas, etc.

Las personas que nacieron con alguna de estas anomalías o que las adquirieron de niños, no cuentan con la realimentación auditiva necesaria, lo que les impide percibir con claridad los sonidos del ambiente y los que ellos mismos emiten.

Esta dificultad les produce grandes inconvenientes en el aprendizaje del habla, lentificando el desarrollo de la dimensión del lenguaje, del pensamiento abstracto, y obstaculizando la categorización de los objetos, la estructura y sistematización de la realidad. Dando origen a alteraciones en el ritmo, timbre de voz (ronquera, monotonía), tono, etc.

Garin, J., Agüero, P., Tulli, J. (2008). Nuevos Aportes al entrenamiento de Personas Hipoacúsicas. En J. Sánchez (Ed.): Nuevas Ideas en Informática Educativa, Volumen 4, pp. 60-65, Santiago de Chile.

Los sordos profundos, son personas que no tienen percepción auditiva. Debido a esto, desde niños aprenden a comunicarse a través del lenguaje de señas, lo que lleva a que tengan aún más dificultades que los hipoacúsicos en la comunicación a través del habla.

En ambos grupos a su vez se originan distintas alteraciones en la articulación de fonemas por ausencia o alteración de estos por otros, dependiendo del grado de sordera que tengan. Estas alteraciones se pueden clasificar en tres grupos: dislalias, diglosias y disartrias.

Las dislalias son trastornos producidos por la disminución de la capacidad auditiva, ocasionando que las personas que las padecen no logren discriminar correctamente los fonemas, generándolos de manera equívoca o directamente excluyéndolos. Por ejemplo dicen “mai” en vez de “maíz” o “tre” en vez de “tres”.

Las diglosias son trastornos de origen orgánico provocado por lesiones físicas o malformaciones de los órganos articulatorios y periféricos. Estas últimas, según la parte del aparato fonatorio que afectan, se clasifican en: labiales, mandibulares, dentales, linguales y palatales, para las cuales se pueden citar, como ejemplo, las correspondientes afecciones: labio leporino, prognatismo, agenesias dentarias, glosectomía y paladar hendido.

Por último, las disartrias son anomalías en la articulación de naturaleza motriz, generadas por lesiones en el sistema nervioso central, como así también por enfermedades en los nervios o músculos de la lengua, faringe y laringe. A éstas se las puede clasificar a su vez en cinco grupos: flácida, espástica, atáxica, extrapiramidal y mixta. En las dos primeras la afección está localizada en las neuronas motrices inferior y superior correspondientemente. La atáxica es ocasionada por una afección en el cerebelo generada por lesiones cerebelosas bilaterales o generalizadas. La extrapiramidal es generada por lesiones en el sistema que lleva el mismo nombre, el cual forma parte del sistema paleocéfalo. Y por último las mixtas, que son una conjunción de las mencionadas anteriormente.

Las personas que tienen estas alteraciones son tratadas por los especialistas del habla a través de la utilización de métodos auditivo-visuales, llevando a cabo una reeducación ortofónica y logopédica, siendo ellos los que interpretan el sonido y le informan al paciente lo que dijo a través de señas o dibujos. Esto implica que el profesional no cuenta con más información sobre lo que dijo el paciente que lo que pudo captar e interpretar a través de la experiencia.

Debido al avance tecnológico en el área informática, se dispone de diversos tratamientos asistidos por computadora que complementan el trabajo del profesional y permiten un conocimiento más amplio y preciso de los sonidos del paciente y de sus anomalías.

En la actualidad existen diversos programas para computadoras orientados a la asistencia de personas hipoacúsicas, tales como el SpeechViewer de IBM [1], Laureate Learning Systems [2], Vocaliza [3], etc.

El SpeechViewer de IBM es un visualizador fonético para los profesionales del habla, que cuenta con dos partes principales:

- Terapia clínica que incluye ejercicios para una terapia del habla directa.
- Gestión de terapia que suministra información sobre el rendimiento del paciente y sus progresos.

La compañía Laureate Learning Systems desarrolló diversos paquetes de programas para la asistencia de personas con capacidades diferentes, de los cuales uno está dirigido a niños con problemas en el habla. Dentro de este paquete, se pueden encontrar programas tanto de ejercitación general como específicos, tales como sustantivos, verbos, pronombres personales, etc. Uno de ellos es el “TalkTime with Tucker” que está basado en juegos, en los que el niño va interactuando y avanzando niveles a través del habla. Esto es ampliamente recomendable ya que evita el tedio tan característico de otros programas.

El programa Vocaliza pertenece al proyecto Comunica de la Universidad de Zaragoza, el cual trabaja en tres niveles de lenguaje: fonológico, semántico y sintáctico. El mismo está desarrollado para el tratamiento de niños, ya que se basa principalmente en juegos e interacción con dibujos. También cuenta con un sintetizador de voz, que les permite a las personas con alguna percepción auditiva saber como se debe pronunciar correctamente una palabra. Este último, a diferencia de los mencionados anteriormente, es para el idioma español y su distribución es gratuita lo que permite el acceso a todas las personas.

En la actualidad no existe ningún programa de asistencia de personas con hipoacusia, para el idioma español local. Si bien Vocaliza está desarrollado para el idioma español, las acepciones no se corresponden con las de nuestro país, por lo que no es apto para ejercitar el habla de nuestros niños. Por esta razón hemos decidido desarrollar un programa diseñado en lenguaje de programación JAVA™ (por su portabilidad) y de distribución gratuita, para que sea accesible a todas las personas que padezcan de estas afecciones en nuestro país.

Este artículo describe el trabajo del programa y sus características principales, el cual está dividido en tres subsecciones que explican más detalladamente cada bloque del programa: Sesiones multiusuario (Sección 2.1), Análisis de la señal e interfaz gráfica (Sección 2.2), Historial y Comunicación vía Internet (Sección 2.3). Seguido a esto, se encuentran las conclusiones y la bibliografía.

DESCRIPCIÓN DEL SISTEMA

El objetivo de este proyecto es desarrollar un programa de distribución gratuita de entrenamiento para personas con discapacidades auditivas.

Muchos de los lenguajes de programación permiten a través de los compiladores obtener código ejecutable propio del microprocesador. En el caso de JAVA el proceso de compilación genera un pseudo-código (opcodes) que debe ser interpretado por una máquina virtual JAVA. Esta última es la que ofrece un entorno de ejecución homogéneo para todas las aplicaciones JAVA, independizándolas de los sistemas operativos y de las incompatibilidades del código fuente entre microprocesadores de distintos fabricantes.

La máquina virtual JAVA se ejecuta en segundo plano, recibe las directivas del programa, las convierte al lenguaje del procesador y las ejecuta. Debido a todo este proceso, la ejecución de una aplicación JAVA es levemente más lenta que una aplicación compilada en código nativo del

microprocesador, pero presenta como gran ventaja poder ser ejecutada en la mayoría de los sistemas operativos actuales, eliminando de esta forma el grave inconveniente que posee la programación convencional.

En la figura 1 se presenta el diagrama en bloques del funcionamiento general del programa. En el mismo se pueden ver tres grandes bloques funcionales (encerrados en líneas punteadas) que serán desarrollados en las siguientes secciones: sesiones multiusuario (sección 2.1), análisis de la señal e interfaz gráfica (sección 2.2), e historial y comunicación vía Internet (sección 2.3).

2.1 Sesiones multiusuario

Este programa cuenta con un sistema multiusuarios, permitiendo que las computadoras de uso público de las escuelas puedan ser utilizadas por varias personas, sin eliminar la posibilidad de que una persona en particular pueda utilizar el programa en su domicilio.

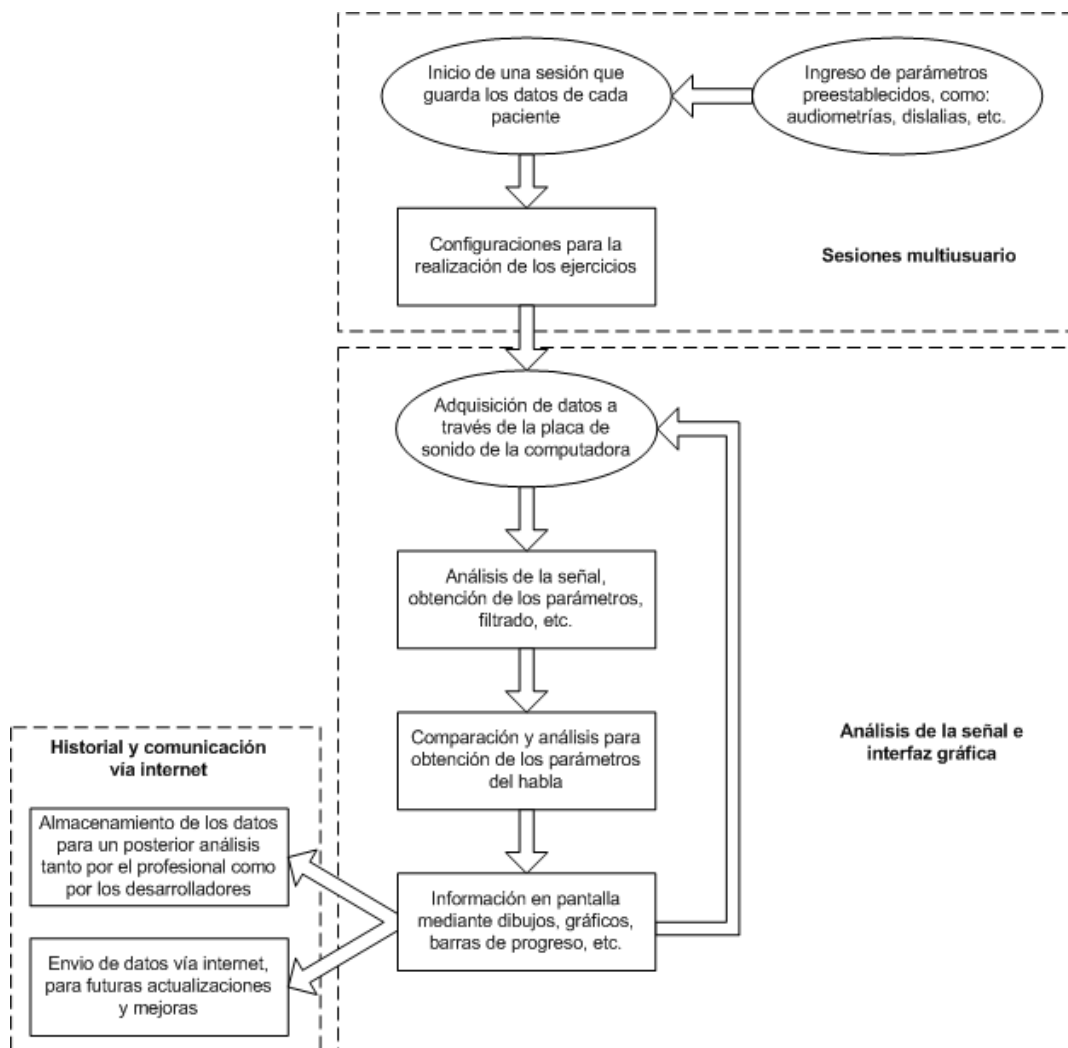


Figura 1. Funcionamiento general del programa

Para cada paciente se deben ingresar sus datos personales y los correspondientes a sus deficiencias conocidas, tales como audiogramas (figura 2), dislalias (mencionadas en la introducción), etc.

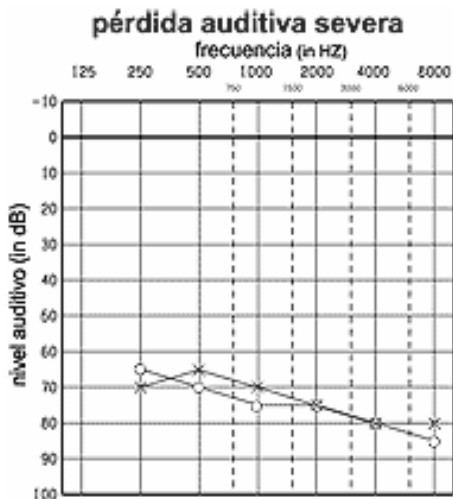


Figura 2. Audiograma de una persona con pérdida auditiva severa

Este programa cuenta también con un sistema de exportación para que el paciente pueda migrar de una computadora a otra sin perder sus datos y su historial de trabajo. Debido a que el sistema posee un servicio de almacenamiento del historial en el servidor de los desarrolladores, el paciente también podrá utilizar múltiples ordenadores sin la necesidad de copiar los datos a medios extraíbles (diskettes, pen drives, etc). Este último aspecto del funcionamiento se explicará con más detalle en la sección 2.3.

La información, provista en el momento de la carga de los datos del paciente, es muy importante, ya que de esta manera el programa cuenta con información previa al momento de tomar decisiones, haciendo que el proceso de análisis sea más corto y efectivo. Por ejemplo: sabiendo que el paciente al momento de decir palabras omite las vocales iniciales, cuando el programa le pida que diga la palabra “ahora” (transcripción fonética SAMPA: “/aora/”), este va a saber que lo primero que dijo es la “o” y no la “a”. Permitiendo de este modo que el análisis continúe con los siguientes fonemas sin detenerse a tratar de interpretar una “a” en lugar de una “o”.

Una vez inicializada la sesión, el profesional debe acceder a un menú para configurar los ejercicios que deberá realizar el paciente durante la práctica.

Existen dos modos principales de configuración de las prácticas. En el primero se utiliza la información sobre las deficiencias del paciente ingresadas con anterioridad, con las cuales el programa filtra los ejercicios correspondientes al tratamiento de estas, lo que lleva a una mayor rapidez a la hora de la configuración. Esto permite también que el

paciente pueda trabajar en su casa sin necesidad de estar con el profesional simplemente siguiendo los pasos que éste le indique. La otra forma de configuración es la manual, en la cual el profesional genera una secuencia de trabajo arbitraria, seleccionando palabras de un listado o directamente ingresando un texto. Este modo presenta como ventaja que el programa pueda ser utilizado para una evaluación específica.

2.2 Análisis de la señal e interfaz gráfica

Las personas hipoacúsicas presentan en su mayoría una disminución muy grande de la capacidad auditiva perdiendo así la realimentación sonora que esta les brinda. Debido a esto, esta última se reemplaza a través de una interfaz gráfica como ser el monitor de la computadora.

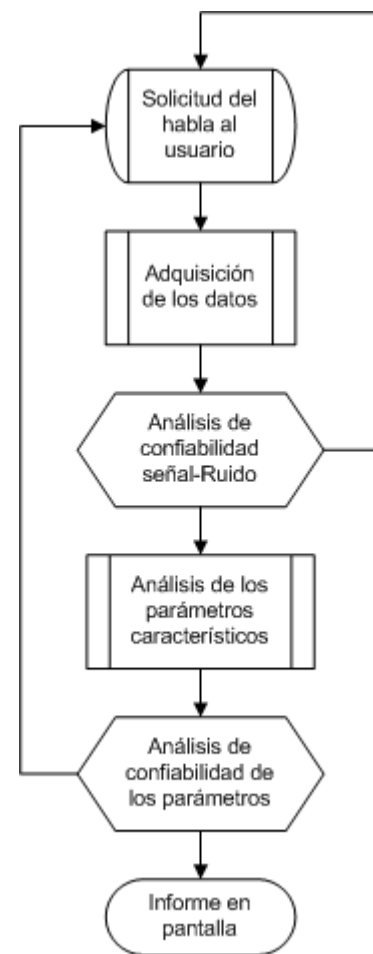


Figura 3. Pasos que recorre la información, desde que es grabada por el micrófono hasta que se informa en pantalla.

La utilización del lenguaje de programación JAVA™, nos permite ofrecerle al usuario una interfaz gráfica amigable y operable tanto por niños, adolescentes o adultos. Para los niños se ofrece una versión con juegos y dibujos, en la cual el paciente va avanzando niveles y ganando puntos a medida que pronuncia correctamente lo que le indica el programa, generando de esta forma un estímulo en el niño para que

continúe su entrenamiento. A su vez estos juegos pueden ser configurados por el profesional para el trabajo de una deficiencia en particular. Para los adolescentes y adultos que tienen más conocimientos sobre los parámetros del habla, se ofrece otra versión con gráficas en tiempo real de los distintos parámetros, presentando en pantalla:

- Barras de progreso para indicar la intensidad de la señal, sonoridad, tono y ritmo.
- Superposición de la grafica de la acentuación utilizando la información del pitch
- Precisión fonemática, indicando con un parpadeo, el fonema que se dijo incorrectamente.
- Duración del habla, superponiendo la grafica temporal con un patrón correcto.

Cuando el paciente pronuncia lo que se le indica en pantalla, su voz es grabada a través de un micrófono, convertida digitalmente por la placa de sonido de la computadora, filtrada para eliminar ruidos y enviada a las secuencias de análisis.

Una vez adquirida la señal, el programa realiza diversos análisis previos antes de iniciar los correspondientes a los parámetros, debido a que el habla de las personas hipoacúsicas tiene en general mucha intensidad de ruido.

Si el ruido es muy grande y se analiza la señal en búsqueda de los parámetros característicos, posiblemente se los encuentre de manera errónea, lo que significaría informar al paciente incorrectamente. En consecuencia, si el programa detecta esta circunstancia, descarta la señal y le solicita al usuario que repita la emisión.

Si la señal pasa la prueba de confiabilidad de señal-ruido, se la envía a las secuencias de análisis de parámetros, las cuales nos retornan información acústica, tal como: el tono, la acentuación a través de la frecuencia fundamental (pitch), el volumen, la intensidad, ritmo, la sonoridad y los fonemas con su respectiva ubicación.

Para una mayor confiabilidad en el análisis de los parámetros de la señal, el sistema procesa los datos a través de varios algoritmos en paralelo que obtienen la misma característica acústica. Estos algoritmos están altamente optimizados para permitir una ejecución más rápida del programa en computadoras de bajo rendimiento, ya que no siempre se cuenta con computadoras nuevas en los establecimientos educativos ni en los domicilios particulares.

Luego de obtener las características acústicas, los resultados son nuevamente analizados para determinar su confiabilidad, efectuando los siguientes estudios: coeficientes de señal ruido (SNR), similitud con el modelo estadístico (HMM), shimmer (medida de la inestabilidad de la amplitud del tono fundamental AF0), jitter (diferencia de amplitud e intensidad entre periodos del sonido), etc; luego de lo cual, si no se cumplen los requisitos, se descartan los datos y se le pide al usuario que repita la emisión. Si la confiabilidad de los parámetros es aceptable, se los compara con los de la base de datos de una persona oyente y se obtienen los resultados

finales del análisis para luego informarle al paciente, las características de su emisión.

2.3 Historial y comunicación vía Internet

Este programa cuenta con un servicio de almacenamiento de información y envío vía Internet a los servidores pertenecientes al grupo de investigación que desarrolla el programa.

El envío de datos del paciente permite que éste pueda trabajar en múltiples computadoras, pudiendo así ejercitar en la escuela junto al profesional y continuar las prácticas en su casa sin la necesidad de recargar sus datos, ya que estos se actualizan vía Internet.

En el momento en que el paciente se registra en nuestra pagina Web para descargar el programa, ingresa sus datos personales y un nombre de usuario que utilizará el programa para identificarlo. Durante el registro, el usuario debe seleccionar su condición de aceptación o no del contrato, en el cual se informa sobre el envío de datos vía Internet y el uso compartido de datos personales. Esto ultimo no es una limitación para el uso del programa, pero el no aceptarlo restringe la utilización del servidor para almacenar los datos personales, lo cual implica que el usuario no puede migrar tan fácilmente de una computadora a otra.

Durante el transcurso de la sesión, un subprograma trabaja en paralelo almacenando datos y enviándolos vía Internet, previa aceptación de las condiciones por el usuario, entre de los cuales podemos distinguir los datos personales y los de la práctica.

Dentro de los datos personales enviados al servidor se encuentran los audiogramas, dislalias conocidas, historial de trabajo, etc; los cuales son descargados del servidor una vez que el paciente inicia la sesión en otra computadora evitando de esta forma que tengan que ser cargados nuevamente.

El segundo grupo de datos enviados al servidor son los correspondientes a la evolución de la práctica, dentro de los cuales se encuentran dos tipos de archivo, los de audio y los de texto. Los archivos de audio enviados contienen la voz del paciente que fue grabada y comprimida durante toda la sesión utilizando un formato sin pérdidas. En los archivos de texto se guardan los datos que corresponden a la configuración de la práctica, incluyendo las palabras y sus respectivos errores, pronunciados por el paciente según lo indicado en cada sesión, permitiendo de este modo que el profesional pueda realizar un análisis posterior y así aplicar las acciones correctivas necesarias.

Los datos son guardados también dentro de la computadora, en carpetas independientes para cada usuario, con el objeto de poder trabajar sin conexión a Internet.

Si el profesional no se encuentra con el paciente, este puede descargar los datos del servidor, para lo cual deberá estar

registrado. De esta manera se mantiene un seguimiento del trabajo del paciente, conociendo los avances y dificultades que presenta a la hora de realizar los ejercicios. De este modo se hace posible el tratamiento en forma particular para personas que no se encuentren en cercanías a los centros médicos o escuelas especiales.

Dentro de este servicio de envío de información vía Internet, se encuentra otro subprograma que realiza informes de errores del programa, lo que les permite a los programadores poder realizar las actualizaciones necesarias.

CONCLUSIONES

En este artículo se ha descrito un programa para la asistencia y entrenamiento de personas hipoacúsicas que aprovecha los avances tecnológicos en el área informática, permitiéndoles la mejora de su habla a través del continuo entrenamiento.

En la actualidad se comprobó que la utilización de este tipo de programas para la asistencia de personas con hipoacusia, acelera de forma notable las mejorías en el habla permitiéndoles una comunicación más fluida y eficiente.

Se está trabajando conjuntamente con el Centro de Fonoaudiología (C.E.F.A) Escuela de sordos e hipoacúsicos de la ciudad de Mar del Plata, con la cual la Universidad Nacional de Mar del Plata tiene un convenio específico de colaboración.

Este proyecto cuenta con una página web en donde se irán colocando los avances del sistema y el programa para su descarga [4].

REFERENCIAS

- [1] <http://www.dif.gob.mx/cta/soluciones/speech.html>
- [2] <http://www.laureatelearning.com>
- [3] <http://www.vocaliza.es>
- [4] <http://elaf1.fi.mdp.edu.ar/pegasus/>
- [5] Carlos Vaquero, Oscar Saz, Eduardo Leida, "Vocaliza: an application for computer-aided speech therapy in Spanish language", Grupo de Tecnología de la Comunicación I3A, Universidad de Zaragoza, IV Jornadas en Tecnología del Habla, noviembre de 2006.
- [6] Elizabeth E. Crawford, "Acoustic signals as visual biofeedback in the speech training of hearing impaired children", Tesis Doctoral, Universidad de Canterbury, 2007
- [7] Emagnet Consultoria SA, "Dislalias", Gabinete Psicopedagógico logopedia, 2007
- [8] José Martínez Ledesma, Gaspar González Rus, Mercedes López Torresilla, "Las NNTT en la rehabilitación logopédica: el visualizador fonético Speechviewer III a la vista del visualizador del habla de la Universidad Politécnica", Rep. Tec., Tecnonet 2000.
- [9] Lawrence R. Rabiner, Fellow IEEE, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", Proceedings of the IEEE, vol. 77, n°. 2, February 1989.
- [10] Paul Christopher Bagshaw, "Automatic prosodic analysis for computer aided pronunciation teaching", Tesis doctoral, Universidad de Edimburgo, 1994
- [11] Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, "Spoken Language Processing" cap 2 y 6, Prentice-Hall 2001