

## **COMUNICA - PLATAFORMA PARA EL DESARROLLO, DISTRIBUCIÓN Y EVALUACIÓN DE HERRAMIENTAS LOGOPÉDICAS ASISTIDAS POR ORDENADOR**

*Oscar Saz, W.-Ricardo Rodríguez, Eduardo Lleida, Carlos Vaquero, Antonio Escartín*

Grupo de Tecnologías de las Comunicaciones (GTC)  
Instituto de Investigación en Ingeniería de Aragón (I3A)  
Universidad de Zaragoza, España  
{oskarsaz,wricardo,lleida,cvaquero}@unizar.es, aescartinv@gmail.com

### **RESUMEN**

Este trabajo presenta *Comunica*, la plataforma desarrollada por el Instituto de Investigación en Ingeniería de Aragón (I3A) en colaboración con diversas entidades educativas para el desarrollo y distribución de diferentes herramientas de logopedia y enseñanza lingüística basadas en Tecnologías del Habla. El conjunto de aplicaciones desarrolladas proporciona entrenamiento desde los niveles prelingüísticos a los niveles pragmáticos del lenguaje utilizando diferentes Tecnologías del Habla para proveer el aprendizaje de una forma rápida, efectiva y no supervisada. *Comunica* propone también un cauce de comunicación directo con la comunidad educativa a través de Internet que permite conocer su valoración de las herramientas, evaluarlas y evolucionarlas para cubrir las necesidades existentes. La viabilidad del proyecto se estudia mediante la valoración de las herramientas por los usuarios y mediante el estudio y desarrollo de nuevos algoritmos para el trabajo con pacientes con alteraciones en el habla y el lenguaje.

### **1. INTRODUCCIÓN**

Existen cada vez un mayor número de aplicaciones orientadas a la enseñanza que buscan aprovechar la potencialidad de interacción que presenta hoy en día la expansión y crecimiento de los ordenadores personales e Internet. En estos casos, se busca aprovechar la facilidad que las nuevas generaciones de alumnos presentan con estos elementos y utilizar su multimodalidad en la interacción para llevar a cabo un aprendizaje de un modo semisupervisado que apoye la labor del educador. Esta interactividad de la enseñanza ha encontrado un especial desarrollo en las aplicaciones de apoyo a la enseñanza del lenguaje, tanto en cuanto se permite el uso de la multimodalidad con imagen y sonido para la interacción y las Tecnologías del Habla que son la base de estas aplicaciones permiten una enseñanza robusta y no supervisada. Los proyectos realizados en este sentido han sido variados y se han orientado a mejorar la capacidad articularia de los alumnos [1], su capacidad de lectura oral y comprensión [2] o

Este trabajo ha sido subvencionado por el proyecto TIN-2005-08660-C04-01 del Ministerio de Educación y Ciencia del Gobierno de España.

el aprendizaje de idiomas extranjeros [3], entre otros.

Este tipo de aplicaciones requieren de avances en las Tecnologías del Habla en que se basan, Reconocimiento Automático del Habla [4] y verificación de pronuncias principalmente, para proveer de una realimentación efectiva y robusta al alumno sobre sus capacidades y necesidades en el proceso de aprendizaje. Actualmente ya se han desarrollado avances en este sentido que ponen estas tecnologías al alcance de la comunidad logopédica.

Es por eso, que el Instituto de Investigación en Ingeniería de Aragón (I3A) ha puesto en marcha *Comunica* para utilizar el conocimiento adquirido en otras tareas de Tecnologías del Habla y aplicarlo al campo de la discapacidad, pudiendo llegar a cubrir el mayor número de áreas posibles. El presente trabajo presenta las circunstancias que rodean la implantación de *Comunica* como plataforma para la distribución y el desarrollo de las aplicaciones logopédicas desarrolladas. En trabajos previos se pueden encontrar resultados científico-técnicos sobre las aplicaciones y algoritmos desarrollados [5, 6, 7, 8].

La organización del artículo es la siguiente: En la Sección 2 se presentan las motivaciones y objetivos principales en *Comunica*. La Sección 3 proporciona una revisión de las herramientas desarrolladas en *Comunica*. Los medios de distribución y evaluación de las herramientas son presentados en las Secciones 4 y 5 respectivamente. Finalmente, las conclusiones de este trabajo se extraen en la Sección 6.

### **2. OBJETIVOS Y REQUERIMIENTOS**

Los objetivos dentro de *Comunica* son el proveer de un marco de desarrollo, evaluación y distribución estable de herramientas informáticas de ayuda a la logopedia. Los requerimientos que se propusieron son los siguientes:

Desde el punto científico-técnico, se busca que las herramientas provean una mejora logopédica a través de su uso de forma fiable. Los sistemas de Tecnologías del Habla utilizados en ellas deben ser robustos y capaces de enfrentarse a cualquier característica personal en el trastorno del usuario. La evaluación de los sistemas debe ser llevada a cabo para garantizar esta situación.

Desde el punto de vista de aplicación, interfaz e interacción, se debe buscar facilitar en todo momento a los

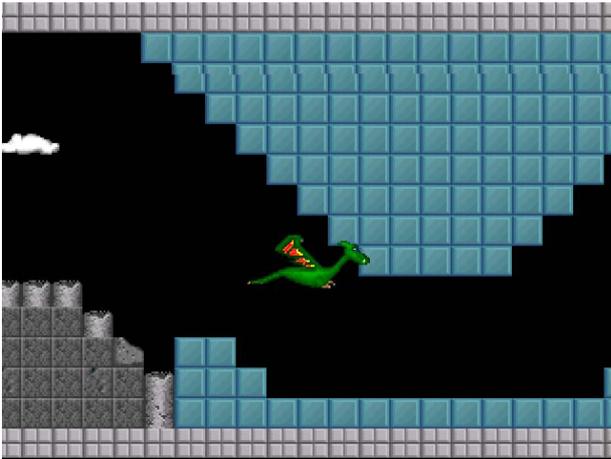


Figura 1. Realimentación visual en PreLingua.

usuarios un manejo sencillo de las herramientas. La complejidad de las Tecnologías del Habla implementadas no debe verse en la superficie y tanto alumnos como educadores deben poder configurar y trabajar fácilmente con las herramientas. El trabajo con la herramienta debe poderse llevar a cabo de forma no supervisada para facilitar las sesiones de logopedia y alcanzar al mayor número de alumnos posible. Por otro lado, se ahonda en la necesidad de trabajar con sistemas de comunicación aumentada y alternativa que permitan la total accesibilidad por parte de los alumnos independientemente de sus capacidades físicas o psíquicas.

### 3. HERRAMIENTAS DESARROLLADAS

En esta Sección se presentan las cuatro herramientas desarrolladas bajo *Comunica* orientadas a apoyar el aprendizaje lingüístico para cuatro grupos de usuarios distintos.

#### 3.1. PreLingua

*PreLingua* es una herramienta orientada al trabajo de las características prelingüísticas por parte de alumnos en fase pre-oral [6]. Consiste en un conjunto de aplicaciones que mediante realimentación visual como la presentada en la Figura 1 buscan hacer consciente al alumno de las características básicas de la producción oral. Los elementos trabajados actualmente por *PreLingua* son la distinción entre voz y no voz (distinguiendo a su vez la producción de sonidos sordos de los sonidos sonoros), el control de la intensidad de la voz, de la frecuencia fundamental y la producción de las vocales del alumno.

La tecnología debajo de las aplicaciones se basa en procesado de señal de voz como detectores de voz-silencio, estimadores de pitch basados en error de predicción de los Coeficientes de Predicción Lineal (LPC) y detección y análisis de formantes. La Figura 1 presenta una actividad consistente en salir de un laberinto en la que el movimiento en el eje horizontal se produce según la presencia o no de voz y el movimiento en el eje vertical se puede controlar por la intensidad de la emisión o por el valor del pitch.

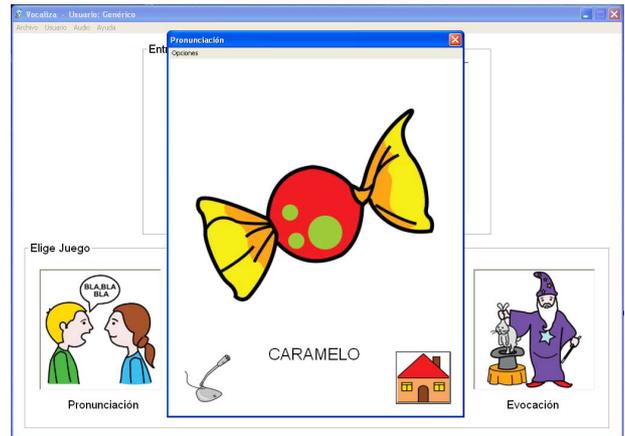


Figura 2. Interfaz visual en Vocaliza y VocalizaL2.

#### 3.2. Vocaliza

*Vocaliza* se dirige al trabajo con alumnos con problemas en el nivel articulatorio del lenguaje [5] mediante el trabajo con la producción de palabras aisladas, frases simples e introducción de la semántica mediante adivinanzas con respuesta en forma también de palabras aisladas. Las actividades con las que se trabaja esta capacidad articulatoria del alumno presentan una interacción basada en texto, imagen y audio como se puede ver en la Figura 2 para facilitar el trabajo del alumno. Cada actividad provee de una forma audio-visual realimentación al mismo sobre su capacidad oral para llevarla a cabo.

La aplicación utiliza Reconocimiento Automático del Habla en cada una de las actividades para decidir cuál ha sido la realización oral del usuario y dar o no la actividad por superada. Para evaluar la capacidad articulatoria del usuario se utiliza un algoritmo de evaluación de pronunciaciões a nivel de palabra [5] que le indica la calidad de su pronunciaciões una vez que la actividad ha sido superada.

#### 3.3. Cuéntame

*Cuéntame* apunta al trabajo en niveles superiores del lenguaje que en *PreLingua* y *Vocaliza*. En este caso, se trabaja con alumnos con una articulación buena que, debido a trastornos del desarrollo presentan problemas en el uso del lenguaje como herramienta de comunicaciões. Este tipo de trastornos pueden suponer que el alumno no sea capaz de crear frases completas o sufra un bloqueo lingüístico a la hora de llevar una conversaciões o responder a preguntas concretas de sus padres o educadores.

La aplicación trabaja todos estos elementos con un interfaz audio-visual similar a *Vocaliza* planteando tres actividades diferentes: Respuesta a preguntas, donde el alumno debe responder con frases sintácticamente correctas a las diferentes preguntas planteadas por la aplicación; descripción de objetos, donde el alumno da la descripción de los objetos presentados de acuerdo a unas cualidades predefinidas que la aplicación le propone; y, por último, navegaciões por un escenario e interacciões virtual con objetos como se presenta en la Figura 3, donde



**Figura 3.** Interfaz visual en *Cuéntame* en la actividad de escenario.

el alumno lleva a cabo un diálogo con la aplicación para conseguir un objetivo planteado por la misma. Esta herramienta prioriza la evaluación de la capacidad de generación del lenguaje del alumno sobre la articularia, evaluando la complejidad de la frase utilizada por el alumno y su aproximación al modelo de lenguaje estimado por la propia aplicación.

#### 3.4. *VocalizaL2*

*VocalizaL2* es una evolución de *Vocaliza* que proporciona una realimentación más precisa al usuario en su pronunciación. *Vocaliza* fue diseñado para proporcionar una evaluación de la pronunciación a nivel de palabra que fuese más fácil para los alumnos de educación especial a los que estaba inicialmente dirigido. *VocalizaL2* busca dar de una evaluación a nivel fonético que permita al usuario conocer con precisión qué partes del conjunto de fonemas del Español le presentan más dificultad.

De esta forma, *VocalizaL2* se dirige también especialmente a usuarios interesados en el aprendizaje del Español como segundo idioma. Desde el punto de vista tecnológico, la aplicación da la evaluación a nivel de fonema a través de una medida de confianza calculada como el resultado de obtener la red de fonemas que mejor se aproxima en términos de verosimilitud mediante el algoritmo de Viterbi a la señal producida por el usuario [8] y compararla a la cadena de fonemas teóricamente producida por el usuario según se le ha requerido por la aplicación como se muestra en la Figura 2.

### 4. DISTRIBUCIÓN

La distribución es uno de los motores que se le ha querido dar a *Comunica*. Ninguna de las aplicaciones desarrolladas tendría sentido si no se buscara hacerlas llegar al mayor número posible de usuarios. En este caso, se utiliza la potencialidad que Internet ofrece para llegar al mayor número de personas. La distribución a través de la página web de *Comunica* [9] permite una distribución sin trabas y en contacto con todos los usuarios, educadores y

logopedas de España y Latino América que han mostrado su interés en las aplicaciones desarrolladas.

Entre los recursos que proporciona la página web está la posibilidad de realizar consultas sobre las aplicaciones, recibir boletines con las novedades que se van realizando, descargar de forma gratuita todas las aplicaciones y sus recursos y acceder a las publicaciones científico-técnicas generadas en *Comunica* entre otras. En la web, también se puede obtener una selección de pictogramas desarrollados por el Centro Aragonés de Tecnologías de la Educación (CATEDU) para su integración en *Vocaliza* y que también se encuentran disponibles en el portal Aragonés de la Comunicación Aumentativa y Alternativa (ARASAAC).

### 5. EVALUACIÓN

La evaluación de los sistemas en *Comunica* se lleva a cabo trabajando en dos direcciones: Por un lado, evaluar la capacidad pedagógica de las herramientas y por otro lado, evaluar los algoritmos de Tecnologías del Habla utilizados en las herramientas y evolucionarlos.

#### 5.1. Evaluación Pedagógica

Para la evaluación de las herramientas, se ha trabajado con diversas instituciones educativas. En el desarrollo tanto de *PreLingua* como de *Vocaliza* y *Cuéntame* se ha trabajado conjuntamente con el Centro de Educación Especial (CPEE) Alborada [10]. Sus educadores han valorado positivamente el uso de la multimodalidad y la interacción oral como forma de motivar al alumnado de la escuela para el trabajo diario en logopedia. También valoran la facilidad que tienen las Tecnologías del Habla para proporcionar una realimentación fiable a los alumnos. El trabajo con estos profesionales ha permitido también valorar el uso de las herramientas desarrolladas para otras tareas, como es el caso de *PreLingua* para tareas de estimulación temprana.

*VocalizaL2* ha contado con la evaluación realizada por personal y alumnado de las clases de Español del Vienna International School, donde jóvenes de diferentes nacionalidades llevan a cabo sus estudios de Español. Una evaluación más completa de todas las aplicaciones se está llevando a cabo actualmente recogiendo la opinión de los usuarios de la herramientas a través de la web; donde se les inquiera sobre la interacción y multimodalidad que ofrece la aplicación, así como sobre la utilidad que proporcionan las Tecnologías del Habla implementadas.

#### 5.2. Evaluación Científica

Para la evaluación de todos los algoritmos requeridos en las aplicaciones se ha adquirido un corpus que contiene 3.192 palabras aisladas de un vocabulario de 57 palabras provenientes de 14 jóvenes locutores con trastornos en el habla y el lenguaje [7]. Para la validación de los resultados obtenidos se ha obtenido también un corpus con 9.576 palabras aisladas de locutores sin discapacidad en el mismo rango de edad como referencia del habla infantil y juvenil; y se ha realizado una anotación manual de los errores en la pronunciación por parte de los locutores con alteraciones lingüísticas con el objetivo de conocer

la afección exacta de cada locutor en su nivel articulatorio. Esta anotación ha demostrado que un 17,61 % de los fonemas en el corpus han sido eliminados o sustituidos, afectando a un 47.71 % de las palabras, dando cuenta de la gravedad del trastorno lingüístico de alguno de los locutores.

Con estos corpora se ha trabajado en varias líneas de trabajo: En términos de Reconocimiento Automático del Habla se ha estudiado el tema de la adaptación al locutor y modelado acústico probando cómo un sistema de Reconocimiento bien adaptado puede descartar con una exactitud del 88 % las palabras con varios fonemas incorrectamente pronunciados [5]. La adaptación al locutor obtiene una reducción en tasa de error del 61 % en la tarea de Reconocimiento, en un trabajo que no sólo aporta conocimiento sobre el funcionamiento de *Vocaliza*, *VocalizaL2* y *Cuéntame* sino que va en la dirección del desarrollo de sistemas de interacción oral con el entorno para discapacitados. La adaptación y modelado léxico también se ha trabajado para intentar introducir la información de cómo desvirtúan los locutores la pronunciación canónica de las palabras en el sistema de Reconocimiento. Por último, se ha trabajado en la evaluación de medidas de confianza para verificación de pronunciaciones, obteniéndose resultados prometedores utilizando algoritmos basados en determinar cuál es la secuencia de fonemas pronunciados más probable y compararla con la pronunciación canónica de la palabra [8] obteniéndose un 18.6 % en el Equal Error Rate.

## 6. CONCLUSIONES Y LÍNEAS FUTURAS

En este trabajo se ha presentado el marco de trabajo de *Comunica*, y los objetivos planteados en su creación. Se ha visto como el conjunto de herramientas desarrolladas abarcan todos los niveles en la adquisición del lenguaje (prelenguaje, articulación y pragmática) mediante el uso de Tecnologías del Habla. Por otro lado, se ha visto como utilizando canales directos de comunicación a través de Internet con los usuarios finales se puede realizar una evaluación y evolución de los sistemas más rápida y directa.

Como líneas de investigación futuras dentro de *Comunica* queda el desarrollo de nuevas herramientas como tutores de lectura, así como el constante estudio para la mejora de las ya existentes. La evaluación final de las aplicaciones mediante estudios y encuestas, así como el desarrollo de nuevos algoritmos para implementar dentro de las mismas es la otra gran línea de trabajo a llevar para conseguir los objetivos finales planteados en *Comunica*.

## 7. AGRADECIMIENTOS

Los autores quieren agradecer la colaboración para este trabajo de José Manuel Marcos y César Canalís del C.P.E.E. Alborada y Victoria Rodríguez del Vienna International School.

## 8. BIBLIOGRAFÍA

- [1] A. Hatzis, P. Green, J. Carmichael, S. Cunningham, R. Palmer, M. Parker, y P. O'Neill, "An integrated toolkit deploying speech technology for computer based speech training with application to dysarthric speakers" in *Proceedings of the 8th Eurospeech-Interspeech*, Geneva, Switzerland, September 2003.
- [2] J. Duchateau, L. Cleuren, H. Van Hamme, y P. Ghesquiere, "Automatic assessment of children's reading level" in *Proceedings of the 10th Eurospeech-Interspeech*, Antwerp, Belgium, September 2007.
- [3] F.-C. Chou, "Ya-Ya language box - A portable device for English pronunciation training with speech recognition technologies" in *Proceedings of the 9th Eurospeech-Interspeech*, Lisbon, Portugal, September 2005, pp. 169-172.
- [4] M.-S. Hawley, P. Green, P. Enderby, S. Cunningham, y R.-K. Moore, "Speech technology for e-inclusion of people with physical disabilities and disordered speech" in *Proceedings of the 9th Eurospeech-Interspeech*, Lisbon, Portugal, September 2005, pp. 445-448.
- [5] C. Vaquero, O. Saz, E. Lleida, y W.-R. Rodríguez, "E-inclusion technologies for the speech handicapped" in *Proceedings of the 2008 ICASSP*, Las Vegas (NV), USA, April 2008, pp. 4509-4512.
- [6] W.-R. Rodríguez, C. Vaquero, O. Saz, y E. Lleida, "Speech technology applied to children with speech disorders" in *Proceedings of the 4th International Conference on Biomedical Engineering*, Kuala Lumpur, Malaysia, June 2008, pp. 247-250.
- [7] O. Saz, W. Rodríguez, E. Lleida, y C. Vaquero, "A novel corpus of children's impaired speech" in *Proceedings of the Workshop on Children, Computer and Interaction*, Chania, Greece, October 2008.
- [8] S.-C. Yin, R. Rose, O. Saz, y E. Lleida, "Verifying pronunciation accuracy from dysarthric speech" in *Proceedings of the 10th ICSLP-Interspeech*, Brisbane, Australia, September 2008.
- [9] A. Escartín, O. Saz, C. Vaquero, W.-R. Rodríguez, y E. Lleida, "*Comunica* framework Web site: <http://www.vocaliza.es>" 2008.
- [10] B. Martínez, P. Peguero, J. Ezpeleta, J. Falcó, E. Lleida, J. Mínguez, y O. Saz, "Universidad y educación especial: Desarrollo y resultados de la colaboración entre el CPS y el CEE "Alborada" in *Proceedings of the III Congreso sobre Universidad y Discapacidad*, Zaragoza, Spain, November 2007.