

Tecnología de voz utilizada en la terapia del lenguaje de niños con deficiencias auditivas.

M. C. Nancy Aguas García

*Universidad del Caribe, Departamento de Ciencias Básicas e Ingenierías
Cancún, Q. Roo, México, <http://www.unicaribe.edu.mx>, nagarcia@unicaribe.edu.mx*

Resumen.

En este documento se presenta la investigación, análisis y diseño de un módulo de enseñanza de lenguaje oral para niños con deficiencias auditivas. En el módulo se propone el método de aprender el lenguaje oral escuchando, leyendo y pronunciando palabras y/o frases comunes del español, apoyado en un tutor animado que muestra el lugar y forma de articulación, imágenes y sonidos digitalizados. La herramienta se desarrolla utilizando estrategias de educación especial y tecnología de reconocimiento de voz. Dicho módulo será el primero en formar parte de un sistema denominado SAEL: Sistema de Apoyo a la Enseñanza del Lenguaje, proyecto que apoya la enseñanza del lenguaje bilingüe (de señas y oral) de los niños del Centro de Atención Múltiple de Cancún, México.

Abstrac.

This paper presents the research, analysis and design of the Language Teaching Module for the SAEL System: Teaching of the Language Support System. This project supports bilingual language teaching (sign and oral) to hearing impaired children using strategies and methods in step with the needs of schools and teachers of Special Education. This module proposes a method designed to teach oral language reading and pronouncing words and / or phrases that are commonly used in Spanish with the help of an enthusiastic tutor who shows how to articulate the words or expressions. The method works by using methods in special educational strategies applying voice recognition and digital image technologies for the teaching and practice of basic language elements.

1. Introducción.

En los países subdesarrollados, la mayoría de los niños con problemas de tipo auditivos son excluidos de todo tipo de educación. Más del 90% de ellos son de padres oyentes, por lo que se requiere de formación especial para atenderlos.

En México existen pocos maestros que se especializan en la enseñanza del lenguaje de señas. La cantidad se reduce para la enseñanza del lenguaje oral, se necesita de mucha paciencia con el alumno pues la mayoría no tiene la habilidad de articular rápidamente y con claridad además, se requiere de mucho tiempo para que el niño pueda llegar a decir algo entendible. Las escuelas de educación especial son en su mayoría públicas y no cuentan con recursos suficientes para adquirir medios que les sirvan de apoyo en terapias de lenguaje.

En esta era del auge de la informática y de la rápida evolución de los sistemas es factible la elaboración de programas que faciliten las tareas cotidianas de oficinas, escuelas y empresas; sin embargo es todavía poco lo que se ha enfocado al desarrollo de tecnología que permita ayudar a personas discapacitadas en países subdesarrollados o en vías de desarrollo, principalmente dentro de éste ámbito. En contraparte, los países primer mundistas han iniciado con el desarrollo de prototipos de programas que representan una herramienta de apoyo para terapeutas y discapacitados, integrándolos total o parcialmente a la población económicamente activa. Por tal motivo, es importante desarrollar en México un proyecto que apoye a la enseñanza del lenguaje oral en niños con necesidades especiales. "La verdadera integración de los alumnos deficientes auditivos es la que permite que se conviertan en adultos competentes y miembros de una sociedad más justa y no transformarlos únicamente en individuos oyentes fracasados". (Vicente, 2003)

1. 1. Situación actual de la educación especial en México.

La siguiente tabla y gráfico, extraídas del sitio en Internet del Instituto Nacional de Estadística Geografía e Informática¹ y de la Secretaría de Educación Pública², muestran los porcentajes de la población mexicana con alguna discapacidad:

Menores con Discapacidad (1995)	Estimación de la Organización Mundial de la Salud	Censo del INEGI (2000)
2 700 000	2 300 000 (10% de la población)	191 541

Tabla 1. Estimaciones de la dimensión de la población con discapacidad en edad escolar

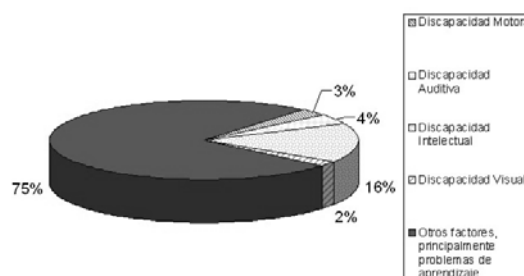


Gráfico 1. Porcentajes de la población que presentan alguna discapacidad en nuestro país

El estado de Quintana Roo está formado por 8 municipios, la población es de 874,963 habitantes. En el 2000, el índice de discapacidad de la población de 0 a 14 años fue 0.7% del total de la población, es decir que aproximadamente 21,000 personas contaban con algún tipo de discapacidad. No se conoce con exactitud el número de niños con problemas de hipoacusia o sordera pero se sabe que evolucionó en un 4.9% del ciclo escolar 2000-2001 al ciclo 2001-2002, es decir que el número de alumnos incremento en 282 personas.³

1.2. Centro de Atención Múltiple de Cancún.

En Quintana Roo existen varias instituciones pero, de las seis escuelas de educación especial que existen, solo una atiende a niños con problemas de audición: el Centro de Atención Múltiple de Cancún (CAM). La matrícula escolar es de 113 alumnos y manejan únicamente dos niveles educativos en grupos de 10 a 14 personas: preescolar y primaria. Su directora, la Profa. Maria Luisa García Lanz afirma que el 98% de los alumnos tienden a expresarse utilizando la lengua de señas, de ese porcentaje el 3% utiliza el lenguaje oral y el de señas, el 1% se expresa utilizando únicamente el lenguaje oral y otro 1% se comunica mediante jalones o pujidos. Este Centro cuenta con 9 docentes y un terapeuta del lenguaje, se reciben y atienden a niños desde los 0 a los 14 años.

1.2.1. La terapia del lenguaje

Los procesos que se llevan a cabo en la terapia son:

- El estudiante realiza diversos ejercicios de movilización de órganos que intervienen en producción del habla, especialmente la lengua, labios y mandíbula.
- El profesor le muestra al alumno un dibujo de la palabra a pronunciar.
- El profesor realiza la pronunciación de la palabra.
- El estudiante observa el movimiento labial.
- El estudiante pronuncia la palabra viéndose en un espejo.
- El profesor escucha la pronunciación realizada por el estudiante.
- El profesor corrige su pronunciación.
- El estudiante realiza el proceso nuevamente.

¹ <http://www.inegi.gob.mx>

² <http://www.sep.gob.mx/wb2/>

³ <http://www.seyc.gob.mx>

2. Recursos informáticos para la intervención logopédica.

Los discapacitados auditivos se han visto beneficiados por una serie de programas que les permiten tener acceso a través de la computadora a una serie de actividades encaminadas a poder comunicarse con los demás. Los programas responden a necesidades concretas como potenciar aspectos de vocabulario, uso de reglas gramaticales, etc.

Para Miguel Puyuelo, los programas cubren cuatro grandes áreas: comprensión oral, expresión oral, comprensión lectora y expresión escrita, favoreciendo estrategias relacionadas con reconocimiento, selección, interpretación, etc. (Puyuelo, 2003). Aunque existen diversos sistemas, no se pueden usar como herramienta de apoyo porque se requiere sea en español mexicano. Algunos ejemplos de software existente son:

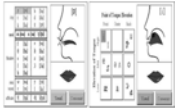
Sistema	Características	Ventajas	Desventajas	Imagen miniatura
Speech Viewer III	Sistema de ejercitación del habla, trabaja la corrección fonética y control de la fonación.	Retroalimentación inmediata, basada en la producción oral, que muestra el tono, intensidad, duración y fonemas aislados.	Diseñado en idioma inglés.	
Globus	Sistema visualizador fonético, cuando un sonido es detectado de muestra un estímulo.	Permite llevar a cabo actividades de imitación y juegos.	Diseñado en idioma español de España.	
Speech Sounds on cue	Software para práctica del lenguaje independiente: fonemas y palabras.	Muestra una filmación del hablante articulando e indica como debe realizar la pronunciación.	Diseñado en idioma inglés.	
Dr. Speech 4	Sistema de juegos infantiles donde se trabaja el tono, volumen, nacimiento articular, pronunciación.	Permite guardar datos de manera individual, el niño recibe retroalimentación	Diseñado en idioma inglés.	
Icatiani	Sistema de apoyo a la terapia de lenguaje que muestra el punto y modo de articulación	Permite almacenar los avances y errores de estudiantes, maneja una cara parlante.	Esta diseñado en español mexicano pero solo trabaja vocales.	

Tabla 2. Tabla comparativa de recursos informáticos para intervención logopédica

3. Metodología para la creación del Módulo.

El presente proyecto trabaja con una metodología basada en el “Ciclo de Vida del Desarrollo de Sistemas en Espiral” (Pressmann, 1998). Los propósitos de la metodología, para este proyecto, se centran en proporcionar un marco de referencia para el análisis,

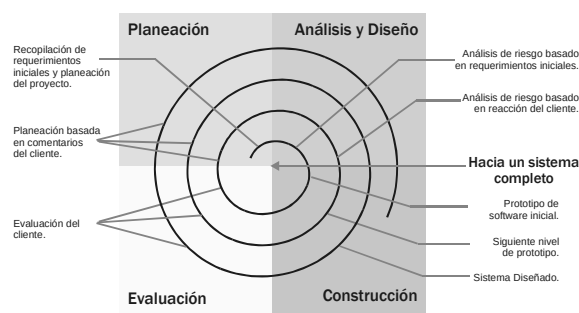


Figura 1. Ciclo de Vida de Sistemas en Espiral

diseño y construcción con una visión integral y holística; suministrar un marco que permita evaluar los resultados obtenidos; ponderar el impacto del módulo en el aprendizaje y desarrollo de los alumnos del CAM Cancún; identificar su contribución a los métodos y estrategias de aprendizaje y poner a disposición un marco metodológico para el diseño o evaluación de futuros sistemas de este tipo.

4. Modulo de Enseñanza del Lenguaje Oral

4.1. Planeación

Se desea desarrollar un módulo de apoyo a la enseñanza del español mexicano para niños con deficiencias auditivas, se debe recopilar información sobre dicho problema.

4.2. Análisis y Diseño

Teniendo conocimiento del medio ambiente, se analiza la problemática, se identifican las necesidades y se crea una estructura. Se propone usar el método de aprender el lenguaje oral escuchando, leyendo y pronunciando palabras o frases comunes del español, se requiere dar retroalimentación de forma inmediata a la pronunciación realizada por el niño. Para diseñarlo es necesario definir algunos componentes de los ambientes de aprendizaje:

a) *Formulación de propósitos y objetivos.* El objetivo es crear herramienta de apoyo basada en las estrategias propuestas por la escuela y el uso de tecnología informática.

b) *Perfil del usuario.* Esta enfocado al apoyo de estudiantes con deficiencias auditivas y los profesores de educación especial y/o terapeutas de lenguaje; el alumno puede practicar y evaluar su pronunciación y el profesor lo utiliza como un medio de apoyo y seguimiento.

c) *Selección del contenido.* Lilian Flores considera que la adquisición del lenguaje hablado debe empezarse con frases completas del contexto real en que vive el niño (Flores, 1991).; por tal consideración, el contenido constará de letras, palabras y frases agrupadas en tópicos que permitirán al niño desarrollar el lenguaje y ampliar su vocabulario.

d) *Selección de estrategias de aprendizaje.* Montserrat Coy afirma que las estrategias a ser consideradas en la programación de terapias del lenguaje son: estimulación visual, auditiva y de expresión (Coy, 2002); por lo que se propone usar una imagen y sonido correspondiente a la letra, palabra o frase, así como una cara animada que muestre la forma de pronunciar y los órganos articulatorios que entran en acción.

e) *Diseño de la estructura.* La arquitectura del módulo para estudiantes se basa en el uso de una interfaz gráfica como medio de interacción con el niño. En la figura 2 se observan los diversos componentes que interactúan: interfaz gráfica, imágenes y sonidos digitalizados, cara animada, base de datos, grabador de voz, reconocedor de voz y un script de verificación para dar retroalimentación sobre la pronunciación realizada.

e) *Diseño de la interfaz de acceso al sistema.* En la figura 3 se muestra un bosquejo del diseño de la interfaz cuando se ha seleccionado un tema. El alumno podrá practicar la



Figura 2. Arquitectura del sistema.

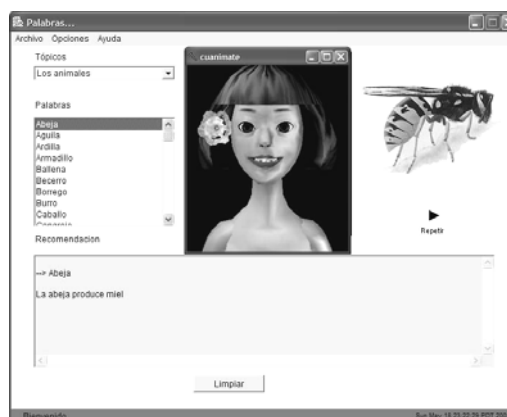


Figura 3. Bosquejo de la interfaz final

pronunciación eligiendo de entre tres temas: vocales, palabras o frases. Después de elegir un tema, debe seleccionar un tópico y luego un elemento específico a practicar, al seleccionarlo se obtendrá de la base de datos la imagen, sonido, punto y modo de articulación, mismos que el alumno observará en pantalla con ayuda del tutor animado. Ahora puede llevar a cabo la grabación de su voz para luego proceder a la evaluación. La señal que produzca será almacenada. Para evaluar, el archivo almacenado pasará por un reconocedor automático de voz que convertirá la información en fonemas, dichos fonemas pasarán por un script de verificación que comparará cada fonema reconocido con el fonema “correcto”. Este proceso se hace generando una gramática específica para el elemento a ser analizado y definiendo las posibles pronunciaciones del mismo. El resultado arroja los errores encontrados y con esta información se puede dar retroalimentación al alumno.



Figura 4. Grabadora de voz

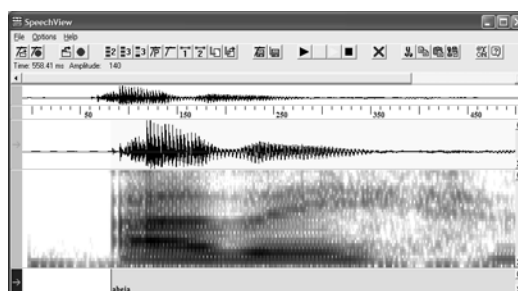


Fig. 5. Espectrograma de la palabra reconocida

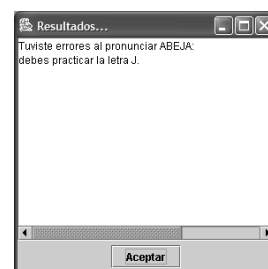


Fig. 6. Retroalimentación

A la par, es necesario diseñar una interfaz para profesores con la cual se tenga acceso a los datos de niños, de tal forma que puede dar de alta perfiles de alumnos, consultar el registro de cada uno (de esta forma dar seguimiento de forma individual), realizar cambios en los datos y borrar del sistema a algún usuario.

4.3. Construcción

Después de estudiar diversas herramientas de desarrollo de sistemas con tecnología de voz, se eligió la caja de herramientas CSLU Toolkit⁴; para construir un reconocedor de voz se debe:

1. *Especificar las categorías fonéticas que la red reconocerá:* conjunto de fonemas del español mexicano.
2. *Recolectar una base de datos de voz:* debe ser grabada por niños de hasta 13 años, ya que deben existir diversos ejemplos de pronunciaciones, de esta forma cuando el reconocedor sea entrenado con los datos no restringirá al niño a un sonido perfecto sino que se dará un margen de error. Se plantea grabar a 50 niños leyendo un vocabulario que contenga los fonemas y combinaciones representativas para la práctica del lenguaje. Las grabaciones se realizarán en un lugar cerrado, apoyados de una computadora personal y un micrófono. Las frases se grabarán a 8000 Hz y serán almacenadas en archivos .wav. Cada archivo será guardado en un directorio.
3. *Transcribir de la base de datos.* Deberá escribirse en cada archivo lo que debió haberse dicho.
4. *Etiquetar de los Datos.* Identificar las porciones de la señal de voz, comenzando a nivel de palabras hasta llegar fonemas, de esta forma se representa el contenido fonético de una pronunciación.
5. *Preparar los archivos para entrenar una red neuronal.*
6. *Entrenar una red neuronal.*

⁴ <http://cslu.cse.ogi.edu/>

7. *Determinar la red con mejor desempeño.*

4.4. Evaluación

Esta fase abarca la planificación y la realización de todas las tareas de verificación, incluyendo pruebas, revisiones, auditorías y todas las tareas de validación. Previamente debe llevarse a cabo la instalación o implantación del sistema para lo cual se requiere dar entrenamiento a los usuarios y documentar el sistema.

Para evaluar el sistema es necesario comprobar que el software funciona de acuerdo con las expectativas del usuario. Es necesario:

- Verificar el comportamiento en entorno real (software, hardware, recursos humanos, procedimientos, bases de datos, etc) y el funcionamiento de las interfaces.
- Adecuar la documentación del usuario.
- Ejecutar el sistema en condiciones límites y de sobrecarga.

También se recomienda probar el desempeño del reconocedor de voz. Este proceso puede realizarse mediante la verificación los fonemas reconocidos por el reconocedor de voz o por el juicio humano. (Franco, 1997), (Bernstein, 1996).

5. Conclusiones

Para construir el sistema requerido es necesario realizar el diseño del mismo tomando en cuenta los componentes de ambientes de aprendizaje enfocados a la terapia de lenguaje.

El utilizar las estrategias de estimulación visual, auditiva y de la expresión ayuda a crear ambientes robustos para el apoyo de enseñanza del lenguaje.

Una red neuronal entrenada con voces de niños es capaz de reconocer errores en las pronunciaciones; utilizar un reconocedor automático de voz permite dar retroalimentación de forma inmediata.

El método de verificar y detectar errores en la pronunciación generando una gramática específica para el elemento a ser analizado y definiendo las posibles pronunciaciones de cada elemento reduce el espacio de búsqueda y el porcentaje de error.

Referencias

- Bernstein, J., Christian, B., "For speech perceptions by human or machines, three senses are better than one", Proceedings of the International Conference on Spoken Language Processing, 1477-1480, 1996.
- Coy Robert, Monserrat, "Consideraciones sobre la programación del lenguaje", Revista Psicología, La adquisición del lenguaje: bases socio afectivas y cognitivas de la adquisición de lenguaje, Editores Mexicanos Unidos, S. A., 2002.
- Flores, Lilián, Berruecos, Pedro, "El niño sordo de edad preescolar", Ed. Trillas, México, 1991.
- Franco, H., Neumeyer, L., Kim, Y., "Automatic pronunciation scoring for language instruction", Proceedings of ICASSP, Abril 1997.
- Pressman, Roger, "Ingeniería del Software: Un enfoque práctico", Ed. McGraw.Hill, 1998.
- Puyuelo, Miguel, San José, Mª Pilar, "Deficiencia auditiva. Aspectos neuropsicológicos y comunicativos. Intervención por medio de software educativo y técnicas de biofeedback", First international Congress on Neuropsychology in Internet, <http://www.uninet.edu/union99/congress/libs/langs/102.html>, 2003.

Vicente, Germán, Pérez, Jorge, Ara, Ignacio, Campillo, M^a Carmen, “Código universal de gestos para las clases de educación física con deficientes auditivos”, <http://www5.ulpgc.es/servidores/webedufis/articul/cusordos/>, 2003.