

Icatiani: un Sistema de Apoyo para la Adquisición del Lenguaje

M.C. Ma. Teresa Toledo C. & Dra. Ingrid Kirschning A.
Grupo de Procesamiento Automático de Voz TLATOA,
CENTIA, Universidad de las Américas, Puebla
teretoledo@yahoo.com , ingrid@mail.udlap.mx

Abstract—El principal objetivo de este trabajo fue el desarrollo de un prototipo que apoye la adquisición del lenguaje que se ajuste a algunas de las necesidades de la escuela de educación especial Jean Piaget. De esta manera un instructor puede dar atención a más niños a la vez. El sistema está desarrollado utilizando el CSLU Toolkit y reconocimiento y síntesis de voz para Español Mexicano desarrollados en TLATOA.

Palabras Clave—Adquisición de Lenguaje, Computadoras en la Educación, Procesamiento de Voz.

I. ADQUISICIÓN DEL LENGUAJE E HIPOACUSIA

La audición ha sido el medio por el cual el humano adquiere el lenguaje verbal [1]. Por medio del lenguaje se adquiere conocimiento del mundo que nos rodea y la comprensión del mismo y por supuesto el desarrollo intelectual. La hipoacusia se refiere a los problemas de percepción auditiva muy baja de lo normal.

La hipoacusia infantil es un gran problema para la adquisición del lenguaje y eso hace la necesidad de buscar otras alternativas para que el infante no crezca sin él [2]. Esto representa un problema difícil porque no sólo los padres tienen que tener el ánimo de ayudar a su hijo a desarrollarse bien en el ambiente que le rodea, sino también los médicos y futuros maestros especialistas en terapia del lenguaje que lo traten.

A. Terapia de lenguaje en la escuela Jean Piaget de Puebla

En México se cuenta con pocas escuelas de enseñanza especial, así como el número de terapeutas es escaso. En la ciudad de Puebla hay una escuela de enseñanza especial. La escuela pública Jean Piaget CAM (Centro de Atención Múltiple) de turno matutino, es una institución de enseñanza especial para estudiantes (desde niños hasta jóvenes de edad secundaria) que padecen desordenes de lenguaje por diversas causas como pueden ser: síndrome de Down, sordera total o hipoacusia, etc.

Cuenta con una población estudiantil de aproximadamente 250 y un personal de 30 personas contando entre ellos maestros de enseñanza especial, psicólogos y terapeutas. El plan de estudios que llevan es el dictado por la SEP (Secretaría de Educación Pública) para la educación especial grupal y han considerado que el niño debe de

desenvolverse en la sociedad por lo que han establecido que el método a enseñar sea el oral, tratando de evitar el lenguaje por señas. La terapia de lenguaje es individual, puesto que cada niño tiene un día y una hora específica, desafortunadamente sólo hay tres terapeutas (servicio social) para una población tan grande de alumnos [3].

Por instrucciones de la SEP la metodología está enfocada al oralismo y la lectura de labios, evitando el lenguaje por señas.

No hay una base científica en la que establezca qué método es el mejor, no obstante, hasta ahora, más escuelas de enseñanza especial para niños sordomudos han adoptado el método oral, porque al sordo le ayuda a no estar aislado de los demás, y poder integrarse a la sociedad. Con éste método el pensamiento del sordo se hace idéntico al normo-oyente aun considerando las imperfecciones articulatorias, expresando sus pensamientos según Perelló [4].

La terapia del lenguaje en la escuela Jean Piaget implica:

- ◆ La enseñanza del punto y modo de articulación de las vocales y consonantes.
- ◆ La entonación, la modulación y el ritmo.
- ◆ Las terapias son individuales.

El material de ayuda que utilizan en la terapia es:

- ◆ Un Cuaderno de trabajo por cada niño, que contiene fonemas y dibujos que enseñan los puntos antes mencionados.
- ◆ Un espejo, confetti y las palmas de la mano para reforzar las lecciones.
- ◆ Láminas donde se expone la entonación de frases completas.
- ◆ Uso de tarjetas de colores para ubicar al niño en el género y el tiempo.

Consideran importante empezar con las vocales, pues son las que tiene una menor dificultad de emitirlas, aunque la /i/ y la /e/ sean algo difíciles de distinguir.

II. SISTEMAS AUTOMATIZADOS PARA LA TERAPIA DEL LENGUAJE

Existen programas y aplicaciones sobre todo en inglés y uno que otro en español de España, que ayudan a los

terapeutas y a médicos a establecer las terapias que ayuden a los niños gradualmente a la adquisición del lenguaje hablado; por supuesto que tales programas se enfocan en una terapia en particular, sea para la articulación, repetición, entonación, ritmo, discriminación, etc. Lo siguientes apartados se mencionan algunos de ellos.

- **El libro electrónico:** En el centro de educación especial “El buen pastor” en España han desarrollado un “libro electrónico” que ayuda al proceso de enseñanza-aprendizaje, su contenido ayuda a intensificar el estímulo de los alumnos. [5]
- **Aphasia Tutor 0: Sights ‘n Sounds:** Es un sistema de terapia de lenguaje que mejora la articulación y recuperación de la palabra, manejando la discriminación auditiva y su repetición [6].
- **Speech Sounds On Cue:** Este sistema se enfoca a las prácticas de lenguaje independiente: fonemas individuales (sonidos), así como palabras completas. Es apropiado para alumnos que necesiten oír y ver un sonido hecho (niños con dificultad de articulación) [6].
- **Terapia del lenguaje con el programa Dr. Speech 4 de IBM:** Es un sistema que cuenta con varios juegos interactivos, donde el niño recibe retroalimentación del cambio de tono, volumen, fonación expresada y no expresada, nacimiento articular y medir la fonación máxima, permitiendo guardar datos estadísticos para ser consultados por los clínicos o terapeutas sobre las articulaciones emitidas por el niño [7].
- **El tutor de vocabulario:** La escuela oral “Tucker-Maxon” en Portlan Oregon hace uso de un sistema desarrollado por el Center for Spoken Language Understanding (CSLU), llamado CSLU Toolkit [8,9]. Este es un software que contiene herramientas integradas de reconocimiento y síntesis de voz y una aplicación de desarrollo (RAD); el software es muy interactivo y fácil de usar. Los terapeutas construyen sus propias lecciones con el uso de dichas herramientas [10], asimismo han y siguen cooperando con los desarrolladores del Toolkit para el diseño y elaboración de software de entrenamiento de lenguaje [11].

Existen más sistemas, sin embargo, en Español Mexicano para terapias de lenguaje existen muy pocos, a la vez que son caros.

La principal inquietud de las instructoras de la escuela es el poder brindar una mejor atención a los problemas particulares de cada estudiante, lo cual se dificulta enormemente por la falta de personal.

Por ello se desarrolló una herramienta que apoyara la terapia de lenguaje individual. Esta herramienta llamada ICATIANI esta basada en tecnología de voz e imágenes para hacer que los niños practiquen la pronunciación [12].

Gracias a TELETON la escuela Jean Piaget ya cuenta con 11 computadoras PC con procesador Pentium III, equipadas con micrófono, bocinas, tarjeta de sonido, pantalla a color de 15”, suficientes para correr el sistema que describiremos a continuación.

III. DISEÑO DE LAS LECCIONES

En base a las entrevistas con las maestras y terapeutas de la escuela, se determinó trabajar inicialmente con pares de vocales, indicando el punto y modo de articulación. El sistema esta dirigido a niños entre 3 y 8 años de edad y de manera personalizada se registrarán los avances y errores de los estudiantes al tratar de pronunciar las vocales. Cabe mencionar que el sistema en ningún momento sustituirá al terapeuta, solo será un instrumento de apoyo para él en sus terapias.

A. La Fonética Articulatoria

La fonética articulatoria estudia la producción física del sonido del habla y los órganos articuladores (boca, labios, lengua, cuerdas vocales y el velo paladar) implicados al producir el sonido de vocales o consonantes; es una rama de la fonética¹, ésta a su vez forma parte de la lingüística que es una disciplina que trata sobre el estudio de las formas de la lengua humana de acuerdo al idioma tratado y las bases de su formación [13].

1) Punto de articulación:

El término punto de articulación se refiere a la zona donde se realiza el contacto entre la lengua y los dientes, paladar y los labios, para de este modo interrumpir y modificar la salida del aire. Se cuenta con siete zonas de articulación que son: Bilabial, Labiodental, Dental, Alveolar, Alveopalatal, Palatal y Velar.

2) Modo de articulación:

En este caso los sonidos de clasifican de acuerdo a la manera de cómo se producen. En el caso de las vocales el aire pasa libremente desde los pulmones a la laringe hasta llegar a la boca sin obstrucción audible. Para las consonantes se presenta una obstrucción del aire y por ello se tiene una clasificación para cada caso de acuerdo a las zonas de articulación: Nasal, Oral, Lateral, Vibrante, Oclusivo, Fricativo y Africativo.

En la figura 1 se muestra el triángulo vocálico con una clasificación de la posición de la lengua y los labios, así como la producción del sonido para cada vocal del Español.

¹ Trata del estudio puramente físico del sonido.

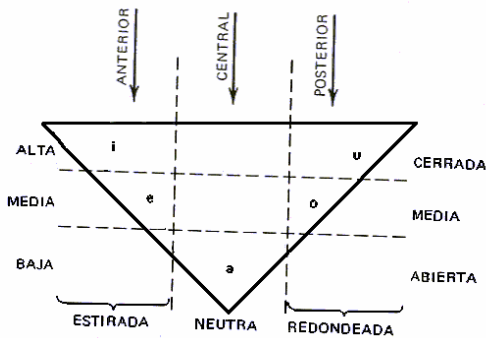


Fig. 1: El cuadro vocálico [13]

Las etiquetas laterales (Alta, Cerrada, Media etc.) indican la posición de la lengua con respecto al paladar; las etiquetas arriba del triángulo (Anterior, Central y Posterior) indican la articulación de las vocales según la posición de la lengua dentro de la boca; y por último las etiquetas debajo del triángulo (Estirada, Neutra y redondeada) señala la posición de los labios al pronunciar las vocales. La figura 2 muestra la relación entre la posición de la lengua y el cuadro vocálico [13]. En las imágenes se ilustran las vibraciones de las cuerdas vocales y el círculo da a entender la redondez de los labios.

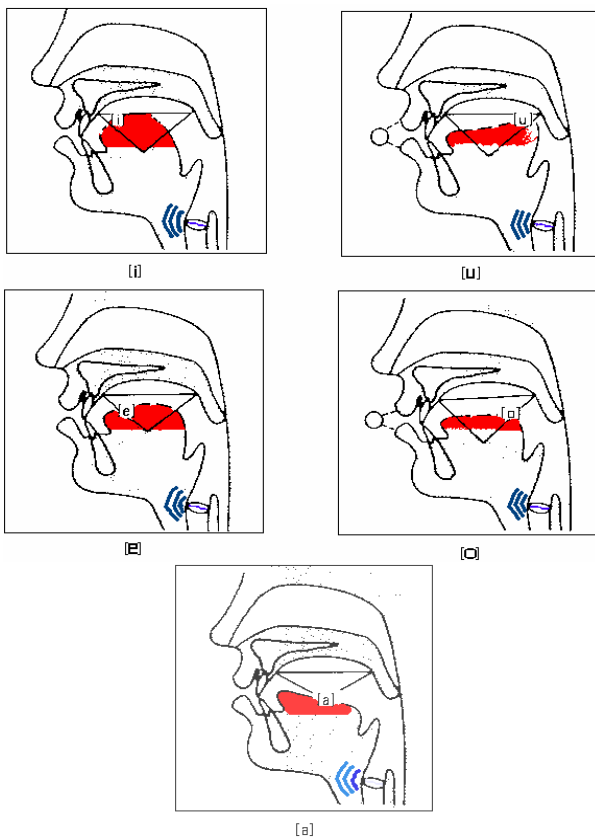


Fig. 2: La articulación de las vocales en Español: i, u, e, o, a [13].

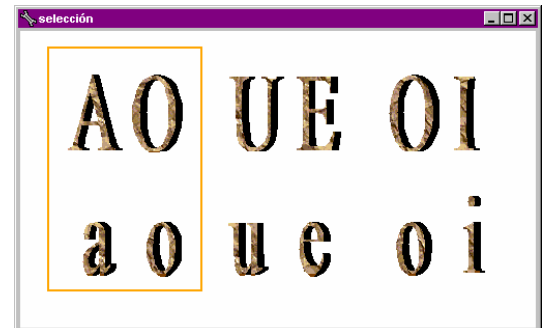


Fig. 3: Aspecto de las vocales por pareja

B. Contenido de las Lecciones

Se ha especificado con anterioridad de que el método a usar es el oral para la adquisición del lenguaje hablado. Las metodologías implementadas serán el punto y modo de articulación y la lectura labio-facial que son las dos metodologías para ayudar a la articulación; las lecciones serán de las vocales ya que son más fáciles de diferenciar y marcan el inicio de la terapia.

El orden a enseñar las vocales sea mejor en parejas: /a/ con /o/, /u/ con /e/ y /o/ con /i/ tanto en mayúsculas como en minúsculas como se muestra en las imágenes de la figura 3 [3].

Algunos autores [1] consideran que para la adquisición del lenguaje hablado debe de empezarse desde las frases completas en el contexto real en que vive el niño, es decir, la rutina diaria que realiza el niño (se levanta, se baña, se viste, se desayuna, etc.) haciendo que los padres le hable de frente como si su hijo escuchara, sin embargo, otros dicen que primeramente hay que empezar por la desmutización², enseñarle primeramente a articular los sonidos empezando con las vocales [13].

En [4] se considera que el orden apropiado de enseñar las vocales es: /a/, /o/, /u/, /e/, /i/ puesto que representa un orden lógico para aprenderlas. La primera pareja de vocales, /a/ y /o/ son distintas en la forma de pronunciarlas debido a la posición de los labios, no obstante, son casi idénticas en la posición de la lengua (ver figura 1); de la misma forma se agrupan /u/ y /e/ y en la tercera pareja /o/ e /i/ para distinguir el contraste de la posición de la lengua y los labios. Sin embargo, el autor comenta que realmente el terapeuta es el que decide el orden, porque es el experimentador y sabe qué es lo que le ha ayudado a obtener los resultados deseados.

C. Aspectos sobre la retroalimentación

El sistema hace uso de imágenes que indican al estudiante cómo pronunciar las vocales y le retroalimentan sobre su desempeño.

² Se refiere "a la enseñanza de la producción de los sonidos" [4].

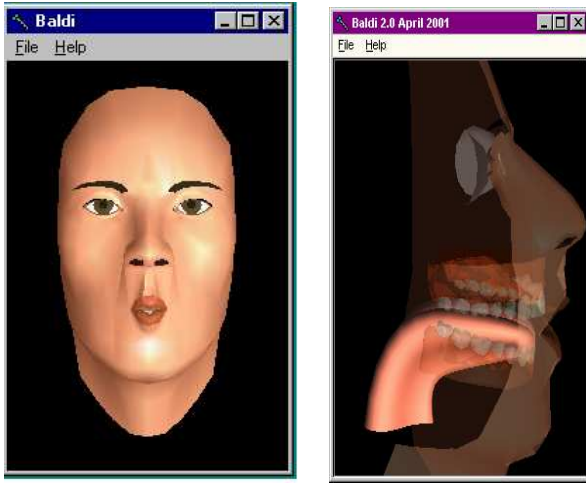


Fig.4: Baldi pronunciando la /u/, tanto de frente como de perfil con la piel transparente.

Al inicio de la lección Baldi pronunciará varias veces la misma vocal de frente, de perfil y con la cara transparente para que se vean las posiciones de lengua, labios y dientes (ver fig.4.)

Debido a que los niños fijan su atención en la cara animada (para leer sus labios) no se les debe distraer con demasiadas formas diferentes de retroalimentación. Esta le debe dar la misma cara. De manera que si el niño pronunció correctamente la vocal indicada, Baldi sonríe y de lo contrario pone una cara triste (fig 5.).

D. Diagrama de Contexto

El diagrama de contexto de la figura 6 muestra que el programa será utilizado por el estudiante y el instructor. El instructor puede dar de alta nuevos estudiantes, verificar el avance de cada uno y modificar las lecciones.

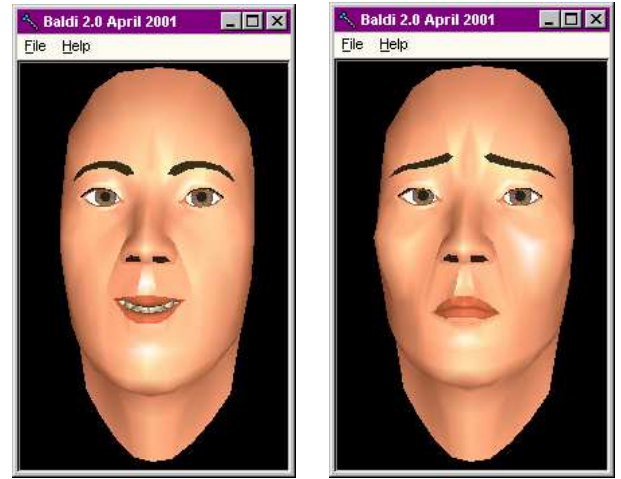


Fig. 5: Uso de expresiones faciales para retroalimentar al estudiante.

El estudiante por su parte necesita dar su nombre para que el sistema lo identifique. Luego seleccionará la lección y Baldi, la cara animada le dará las instrucciones. Baldi también le da retroalimentación sobre su desempeño.

E. El CSLU Toolkit

El CSLU ha desarrollado una serie de herramientas y tecnologías que ayudan a la investigación y desarrollo de aplicaciones de la ciencia del habla y del lenguaje [8]. Todas ellas están integradas en un software denominado CSLU Toolkit; consta del reconocimiento de voz, síntesis de voz y una aplicación para el desarrollo fácil de sistemas de voz. Los reconocedores y las voces del sintetizador en Español Mexicano fueron desarrollados por el grupo TLATOA [14].

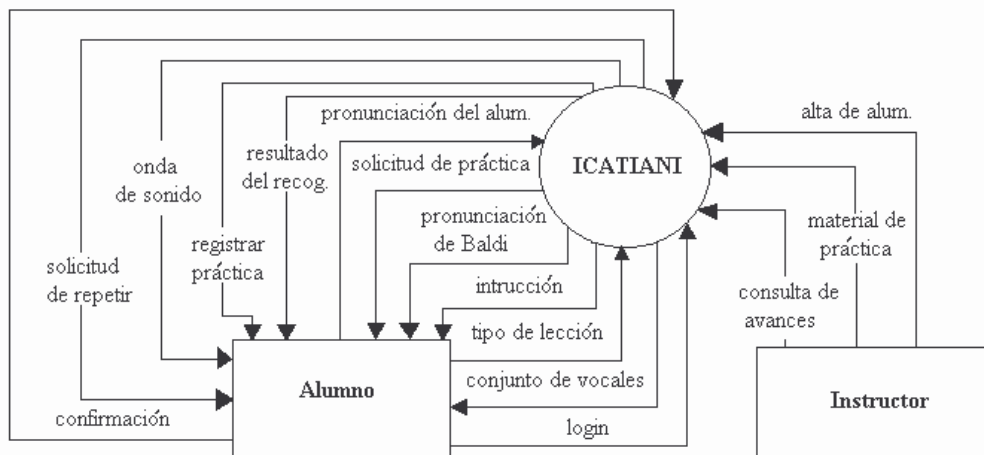


Fig. 6: Diagrama de contexto de ICATIANI.

Algunos de los componentes del CSLU Toolkit que se utilizaron son:

- ♦ **RAD (Rapid Application Developer):** es un ambiente gráfico para diseñar aplicaciones e interfaces de voz, donde se puede construir diálogos con una cara animada que habla (ver Fig. 7)
- ♦ **Animación facial:** Baldi es una cara animada en 3D³, en la cual se aprecia la coordinación conjunta de la articulación de las palabras, el movimiento de los labios y la expresión facial [15].
- ♦ **Reconocimiento de voz:** el Toolkit tiene incorporado el reconocedor de voz en Español Mexicano para adultos desarrollado por el grupo Tlatoa (www.udlap.mx/~sistemas/tlatoa).
- ♦ **Ambiente de programación:** el sistema cuenta con un ambiente de programación para C y Tcl con librerías de software y una serie de Api's que son la base de la programación.

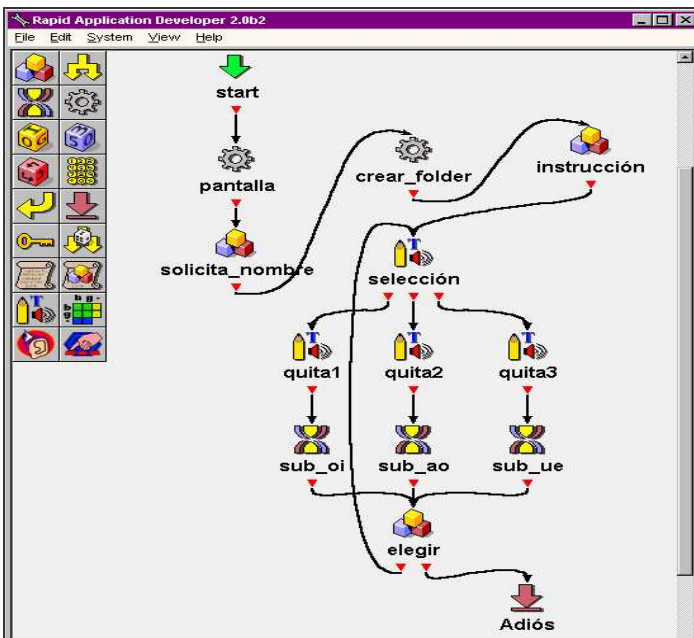


Fig. 7: El programa principal (canvas del RAD)

IV. EL SISTEMA ICATIANI⁴

Utilizando el RAD del CSLU Toolkit se programó la secuencia lógica de las lecciones. Esta consta de los siguientes pasos: primero el estudiante entra al sistema en donde se le pide su nombre. Así Baldi lo saluda con su nombre y puede registrar todo lo que haga en un archivo de texto. Este archivo contendrá la fecha de entrada al sistema, la vocal seleccionada, la vocal que pronunció, número de aciertos y rechazos.

Inicialmente Baldi le pide al estudiante que escoja un par de vocales a practicar. Dependiendo del par escogido ahora puede escoger cual de las dos vocales desea practicar primero. Baldi le pronuncia la vocal deseada y luego le pide que la pronuncie el niño. En este momento un reconocedor tratará de reconocer lo que diga el niño, si es la vocal solicitada la cara sonríe, de lo contrario se muestra triste y le pide intentarlo de nuevo.

A. Reconocimiento de Voz

Inicialmente parecía que no era recomendable utilizar reconocimiento de voz para este sistema, ya que el tono y fuerza con la que los niños hipoacúsicos hablan varía de manera extraordinaria, demasiado para un reconocedor convencional. Sin embargo se hicieron algunas pruebas y con la ayuda del terapeuta el niño empezó a entender por qué Baldi no entendía lo que decía. El niño tenía que controlar mucho su voz.

Para fines de esta aplicación se simplificó la tarea del reconocedor para que tomara como vocabulario solamente las dos vocales en cuestión

B. Caption

Como un apoyo adicional para el entendimiento de las instrucciones, se agregó la opción de utilizar “captioning” o subtítulos. Con esta herramienta el alumno puede leer las instrucciones cuando no puede leer los labios de Baldi. (esto es para aquellos que ya saben leer).

C. Baldi

La cara animada de Baldi habla moviendo los labios de manera acorde a lo que dice, además de que habla utilizando síntesis de voz de concatenación de difonemas (Festival).

El RAD permite colocar los diferentes módulos genéricos y unirlos para representar la secuencia del proceso. Existen módulos para reconocimiento, para iniciar sub-rutinas, etc. (Figuras 7, 8 y 10). Sin embargo, cada uno de ellas debe programarse para que se presenten las caras y las diferentes imágenes y pantallas en el momento adecuado, se realice el reconocimiento de voz y las acciones de acuerdo al resultado del reconocedor. El lenguaje de programación utilizado fue Tcl.

⁴ ICATIANI es una palabra náhuatl que quiere decir “el que habla”.

³ Programada en la Universidad de Santa Cruz, California

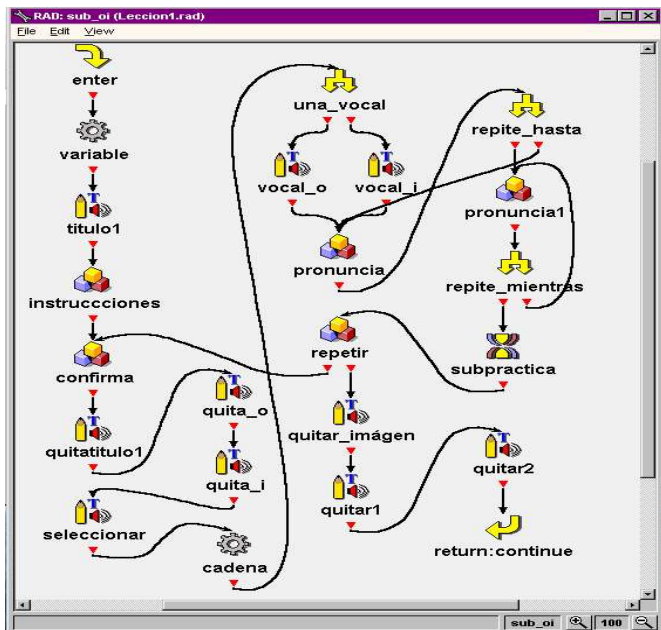


Fig. 8: el subprograma de la lección articulación autiva-visual

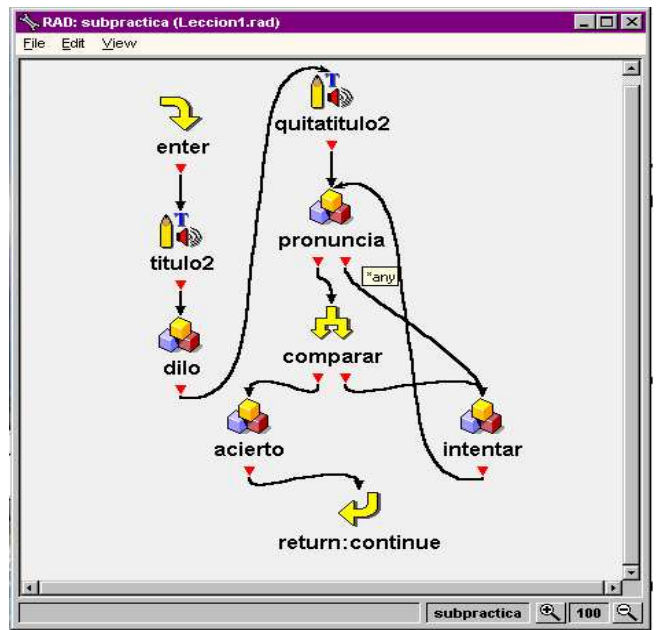


Fig. 9: La práctica de la lección (vocal /i/)

A continuación se presentan algunas de las pantallas del sistema durante las lecciones. En la figura 10 se puede observar a Baldi pidiendo al estudiante que escoja una vocal. Esto es una vez que este ya escogió uno de los 3 pares de vocales de la fig. 3.

Cuando Baldi explica la pronunciación de una vocal lo hace primero con un diagrama que trata de explicar el flujo de aire y después a través de las imágenes animadas con la cara de frente y de perfil y una vez con la cara transparente (Fig. 11).

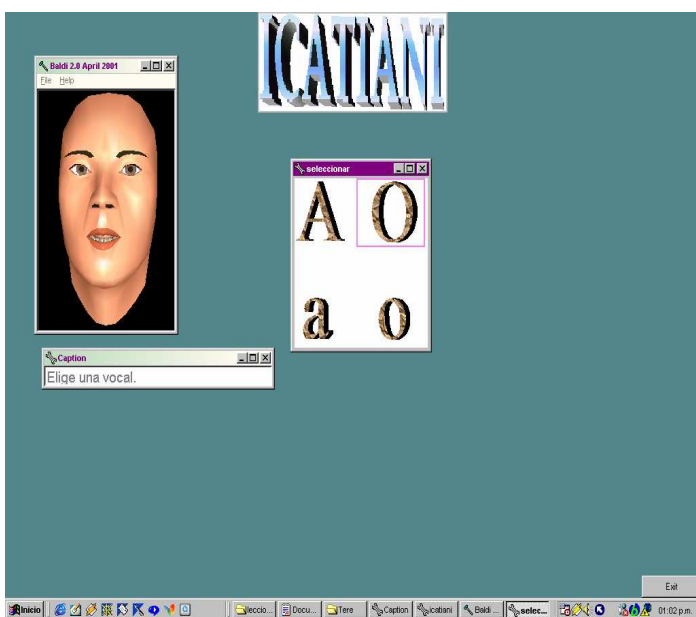


Fig. 10: Baldi pide al estudiante que elija una vocal del par de vocales seleccionadas previamente.

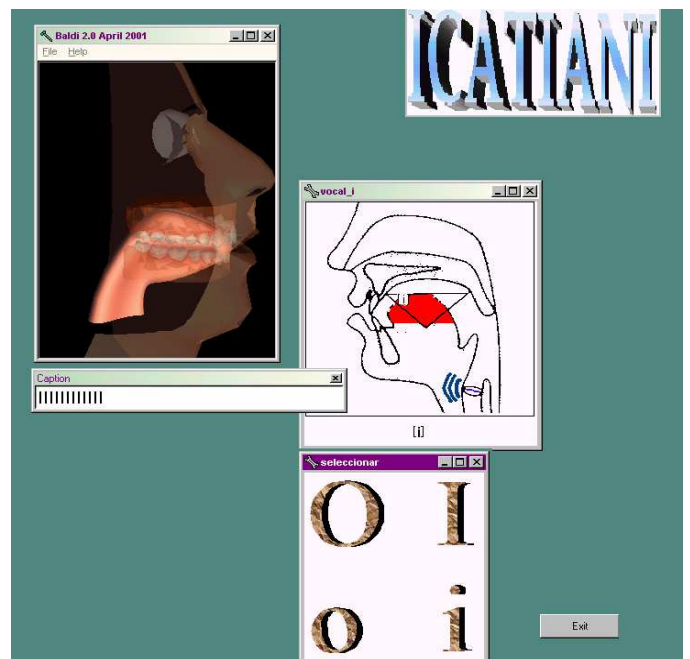


Fig. 11: Baldi pronunciando la /i/ de perfil con la piel transparente

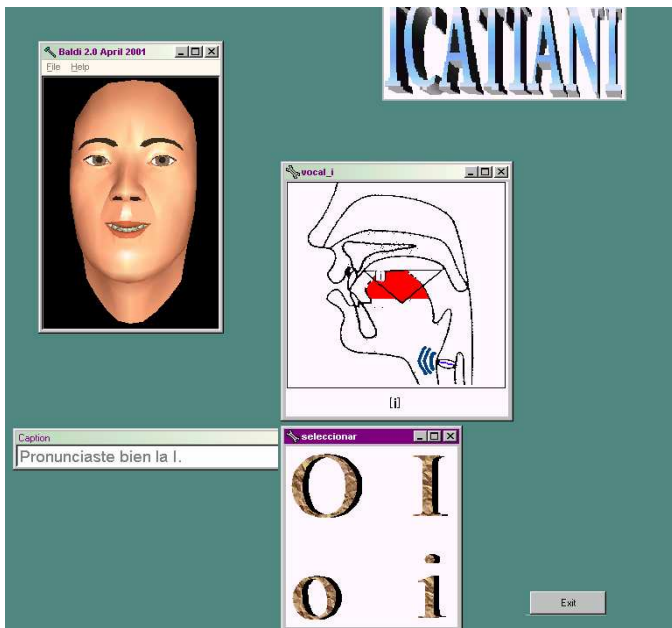


Fig. 12: La expresión de alegría de Baldi, una muestra de que el reconocedor aceptó la pronunciación del alumno

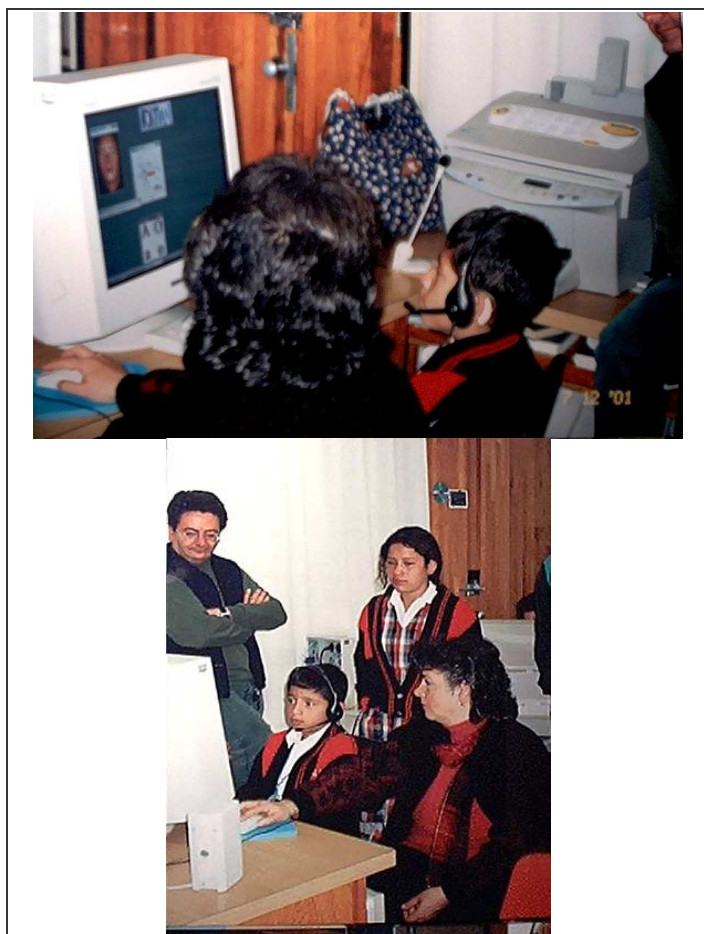


Fig. 13: Estudiantes de la Escuela Jean Piaget probando el sistema Icatiani.

La figura 12 muestra un ejemplo del caso cuando el reconocedor acepta la expresión como correctamente pronunciada y le sonríe al estudiante diciendo “Pronunciaste bien la i”. También se observa el cuadro de texto en el que se despliega lo que dice Baldi marcando palabra por palabra al momento que las dice.

Finalmente la figura 13 muestra una de las sesiones de prueba con estudiantes de las Escuela Jean Piaget, una instructora y profesores que supervisaron este trabajo como sinodales de tesis.

V. CONCLUSIONES

La desventaja enorme que tiene la escuela de enseñanza especial en la actualidad es que no cuenta con terapeutas de planta, sino con uno dos estudiantes que están prestando servicio social, lo cual hace que las terapias carezcan de continuidad.

Gracias al apoyo e interés de las maestras y encargadas de los estudiantes en la escuela se logró un sistema que ayuda a los niños practicar su pronunciación. Pueden trabajar 11 niños al mismo tiempo y la instructora puede atender a cada uno de los problemas individualmente sin hacer esperar a otros su turno.

El uso de reconocimiento de voz genera un elevado número de rechazos de expresiones que al oído humano le parecerían prácticamente correctas. Una fuerte preocupación fue que esto desmotivara a los chicos. Sin embargo, la experiencia mostró que al estar jugando con la computadora ellos quieren permanecer el mayor tiempo posible, mejorando su pronunciación. En algunos casos la maestra tuvo que intervenir y alentar al estudiante dándole a entender que lo estaba haciendo muy bien, sólo faltaba un poco más para que Baldi lo entendiera.

Actualmente se esta desarrollando una segunda versión en la que cada vocal se relaciona con palabras que comiencen con esa letra y sus respectivas imágenes.

Es importante señalar que en ningún momento se pretende sustituir al terapeuta o al maestro, sino que se le provee con una herramienta para poder dedicarle más tiempo a los niños que lo necesitan.

Este primer sistema Icatiani ha demostrado que diversas metodologías para la enseñanza del lenguaje, como técnicas de entonación, modulación, del comienzo de la articulación, del punto y modo de articulación, entre otras, pueden aprovechar muy bien el uso de las tecnologías de voz. Se han desarrollado algunos sistemas comerciales pero son escasos y bastante caros. Con el uso del CSLU Toolkit resulta relativamente fácil partiendo de la base que ya hemos programado. El CSLU Toolkit mismo es gratuito y accesible via internet.

Se realizaron varias etapas de prueba durante las cuales se fueron depurando pequeños errores en la interacción con los niños. Un detalle que causaba problemas a los niños que utilizaban el sistema por primera vez, era que debían esperar un poco (esperar el “beep”) para empezar a hablar al reconocedor. Esto no se ha podido modificar ya que el sistema

no cuenta aún con tecnología “barge-in”, pero los niños se acordaron y al usarlo por segunda vez contaban con los dedos de la mano un cierto tiempo (2 segundos) antes de hablar.

Actualmente Icatiani es utilizado por un hombre adulto en Santiago, Chile y estamos obteniendo valiosa retroalimentación sobre el funcionamiento y posibles mejoras.

Es enorme la satisfacción de ver como los niños disfrutaban el sistema, poniendo todo de su parte y que no se desaniman por el hecho de que no pronunciaron bien. Cuando vuelven a intentarlo y lograrlo se reflejaban en sus rostros una sonrisa, porque Baldi los felicitaba muy feliz por el fruto de su esfuerzo.

AGRADECIMIENTO

Deseamos expresar nuestro agradecimiento a la directora de la escuela Jean Piaget, Psic. Ma. del Rosario Espinosa Franco y a la Lic. Lydia Estela Valadez Cedillo profesora especialista en audición y lenguaje. Su cooperación e interés han sido muy valiosos en el desarrollo de ICATIANI cuyo objetivo es cubrir una de tantas necesidades en cuanto a herramientas de apoyo para reforzar la enseñanza.

BIBLIOGRAFÍA

- [1] L. Flores & P. Berruecos, “El niño sordo de edad preescolar”, Ed. Trillas, México, 1991.
- [2] J.J. Delgado, “Hipoacusia infantil”, PrevInfad, 14 junio, 1999.
- [3] L. Valadez, & M.R. Espinosa, Entrevista personal, Septiembre, 2001.
- [4] J. Perelló, F. Tortosa, “Sordomudez”, Ed. Científico-Médica, Audiofoniatría y Logopedia, Barcelona, España, 3ª. Ed. 1978. <http://www.medynet.com/usuarios/PrevInfad/Audicion.htm>
- [5] M. Gómez, “Las nuevas tecnologías de la información y la comunicación y su aplicación a las necesidades educativas especiales. Cómo generar recursos didácticos para el aula”, revista Almeceas (nº 5 Junio 99) del CPR de Cieza y en Magisterio Español (nº 41; 9 Junio 99). <http://paidos.rediris.es/needirectorio/tema21.htm#inicio>
- [6] Bungalow Software, “Speech Sounds on cue and AphasiaTutor 0: Sights ‘n sounds”, agosto, 2000. <http://www.bungalowsoftware.com/sights.htm>
- [7] H.D. Zaoming, “Speech Skill Builder for Children”, Tiger DRS inc. 1990. <http://www.drspeech.com/Paper.html>
- [8] R. Cole, “[Tools for research and education in speech science](#)”, Proceedings of the 14th International Congress of Phonetic Sciences; San Francisco, agosto 1999. http://www.tmos.org/tech/papers/ron_icphs1.html
- [9] Tucker-Maxon, “Tutor de vocabulario”, 1998-2001. http://www.tmos.org/tech/vocabulary_tutor/vt1.html
- [10] P. Stone, “Revolutionizing language use in oral deaf education”. Proceedings of the 14th International Congress of Phonetic Sciences; San Francisco, agosto, 1999. <http://www.tmos.org/tech/papers/S0837.html>

- [11] P. Connors, A. Davis, G. Fortier, K. Gilley, B. Rundle, C. Soland & A. Tarachow, “Participatory design: Classroom applications and experiences”. Proceedings of the 14th International Congress of Phonetic Sciences; San Francisco, agosto 1999. <http://www.tmos.org/tech/papers/S0836.html>
- [12] M.T. Toledo, “Cómo enseñar a niños hipoacúsicos aprender a hablar utilizando reconocimiento de voz”, Tesis de Maestría en Ciencias con Especialidad en Sistemas Computacionales, Universidad de las Américas, Puebla, mayo 2002.
- [13] R. Barrutia, T. Terrell, T. David, “Fonética y fonología españolas”, Ed. John Wiley & Sons, U.S.A., 1982.
- [14] I. Kirschning, I., “Research and Development of Speech Technology & Applications for Mexican Spanish at the Tlatoa Group”. Development Consortium at CHI 2001, Seattle, WA, abril, 2001, pp.49-50.
- [15] R. Cole, T. Carmell, P. Connors, M. Macon, J. Wouters, J. deVillers, A. Tarachow, D. Massaro, M. Cohen, J. Beskow, J. Yang, U. Meier, A. Waibel, P. Stone, G. Fortier, A. Davis, C. Soland, “[Intelligent Animated Agents for Interactive Language Training](#)”. STiLL: ESCA Workshop on Speech Technology in Language Learning, Stockholm, Sweden, mayo 1998. <http://www.tmos.org/tech/papers/ilt.html>