

Aplicaciones sociales de las tecnologías de la lengua

Social applications for language technology

Santiago Aguilera, J. Ignacio Godino, Sira E. Palazuelos y J. Luis Martín

Las tecnologías de la lengua tienen gran cantidad de aplicaciones sociales, considerando como tales las ayudas a discapacitados, y las aplicaciones médicas. En algunos casos, los sistemas clásicos de síntesis o reconocimiento de voz son de uso inmediato pero en otros se requieren aplicaciones específicas, como pueden ser las prótesis auditivas, audífonos e implantes cocleares. Este artículo es una revisión de las distintas discapacidades o enfermedades, que pueden ser paliadas o diagnósticas aplicando técnicas de proceso de voz o texto.

Language technologies can provide many social applications, such as tools for the handicapped as well as medical applications. In some cases, classical voice synthesis and immediate recognition systems are perfectly suitable to fulfill these needs but in other occasions specific applications are involved, such as in hearing prostheses, audiophones and cochlear implants. The article reviews various disabilities and illnesses, which can be helped or diagnosed applying voice or text processes.

Hay personas que tan sólo tienen problemas de fonación, manteniendo intactas el resto de funcionalidades. Para ellas un simple sintetizador de voz es suficiente para resolver su problema de comunicación oral. Pero en muchos casos, la incapacidad para hablar se debe a un problema de movilidad, que afecta al aparato fonador y a muchas otras partes del cuerpo, como es el caso de las personas que padecen parálisis cerebral, esclerosis lateral amiotrófica o tetraplejia.¹ En estos casos, el problema fundamental no es el sintetizador, sino el mecanismo por el que el usuario puede controlarlo. Existen fundamentalmente dos métodos de control de los sistemas de comunicación: *directo* y *por barrido*.

Método de control directo. El usuario directamente o con la ayuda de algún instrumento unido a su cuerpo maneja el sistema. Las personas con un buen control de cabeza suelen utilizar un licornio (fig. 1), una varilla sujeta a la cabeza mediante una diadema, con la que se accede directamente a una imagen fija, a una pantalla de ordenador sensible al tacto o a un teclado envuelto en una carcasa que evita la activación de teclas adyacentes.

También se emplea un tablero silábico, en el que sobre una cartulina aparecen todas las sílabas posibles para el español, con las que el usuario compone sus mensajes ayudado, en muchas ocasiones, ayudado por licornio. Ésta ha sido, y sigue siendo hoy en día, una ayuda a la comunicación ampliamente utilizada por muchas personas con problemas de fonación.

Métodos de control por barrido. En estos métodos suelen emplearse sistemas tecnológicamente más avanzados como puede ser un ordenador, donde se van realizando secuencialmente distintas opciones, y cuando llegamos a la deseada se pulsa un interruptor para seleccionarla. Es un interruptor que debe adaptarse al movimiento o acción mejor controlada por parte del usuario. Existen muchos modelos de pulsadores controlados por diversos movimientos, como los que pueden ser accionados con la mano o con el pie, con la barbilla o neumáticos. Uno de los sistemas de escritura por barrido más empleados es aquel en el que se colocan sobre una matriz las letras del abecedario y los signos de puntuación más frecuentes. En un paso posterior se va señalando secuencialmente cada uno de los símbolos, y el usuario pulsa el interruptor cuando está sobre el símbolo deseado (fig. 2).

Una velocidad típica de escritura con este procedimiento es unas seis palabras por minuto, con lo que resulta muy útil dotar al sistema de mecanismos de aceleración del proceso. Los más utilizados son los denominados métodos de predicción de palabras, por los que un sistema con mayor o menor inteligencia, basándose en las últimas palabras escritas y en las letras que el usuario lleva escritas de la presente palabra, intenta «adivinarla» para reducir el tiempo y el esfuerzo de escritura. Los sistemas de predicción de palabras se basan en una de las importantes áreas de la tecnología de la lengua: el procesamiento de lenguaje natural.²

Ayudas a personas ciegas o con deficiencias visuales

Para las personas con estos problemas existen múltiples sistemas que utilizan el tacto, como el código Braille o el oído, con síntesis de voz, para transmitir información escrita a las personas ciegas.³

Información táctil

El código Braille realiza una representación táctil de cada una de las letras, números y signos de puntuación utilizados en los textos escritos. Fue inventado por Louis Braille, que nació en 1809 en Coupvray, cerca de París

, y ciego desde los tres años. Cada símbolo está formado por seis puntos colocados sobre una matriz de dos columnas y tres filas. Para representar los distintos símbolos utilizados en la escritura se realzan distintos puntos de dicha matriz (fig. 3). Existen prefijos que actúan sobre todos los símbolos que les siguen para indicar propiedades como subrayado, mayúsculas o negrilla. Aunque en sus orígenes fue pensado para soporte papel, en la actualidad existen periféricos informáticos que funcionan con este código. Existen teclados Braille, configurados por seis teclas (una para cada punto), celdas Braille como salida de ordenadores, compuestas por seis pequeños cilindros que mediante un dispositivo electromecánico se activan unos u otros para indicar los distintos caracteres. También existen impresoras que por procedimientos de impacto o térmicos producen textos en este código.

Información auditiva

La información se suele transmitir en forma de síntesis de voz, y debemos decir que para aumentar la velocidad de trabajo, las personas ciegas suelen configurar los sintetizadores para que hablen a velocidades muy superiores de la habitual. Para las personas no habituadas suele resultar imposible el entendimiento. En la actualidad, hay toda una línea de investigación abierta para incrementar la calidad de los sintetizadores de voz cuando funcionan a altas velocidades.

Las principales ayudas técnicas, basadas en tecnología del habla, que suelen utilizar las personas ciegas son:

- *Lectores de pantallas de ordenador.* Son programas que capturan la información existente en la pantalla de un ordenador y se las transmiten a las personas ciegas bien mediante síntesis de voz o códigos Braille, seleccionable por el usuario, leyendo los textos y describiendo los gráficos. Están especialmente diseñados para transmitir información de Internet, suministrando toda la información sobre los enlaces que aparecen en la página activa, y también describiendo los gráficos, si tienen un texto vinculado. Existe toda una metodología para el diseño de páginas de Internet accesibles (<http://www.w3.org>). Uno de los programas más utilizados es el JAWS (Job Access With Speech).⁴

- *Agendas personales.* Son pequeños equipos portátiles con un teclado Braille (seis teclas, cada una asociada a un punto de la matriz Braille) que permiten a las personas ciegas introducir texto para ser consultado o editado posteriormente. Disponen de un navegador que se va desplazando por el texto, indicando cambios de línea, párrafo o página, que nos permite acceder por síntesis de voz o por línea Braille a la información allí contenida.

- *Lectores de documentos.* Unos dispositivos informáticos compuestos por un escáner un *reconocedor óptico de caracteres* (OCR en su siglas en inglés) y un sintetizador de voz, que con un manejo muy simple va leyendo un libro o documento.

Las tecnologías en la rehabilitación de la voz

Existen numerosos equipos para ayuda a la rehabilitación de la voz, empleados en gabinetes de logopedia o foniatría, que permiten entrenar diferentes aspectos como la prosodia o la articulación de fonemas. Sus formas de trabajo son dos, bien mediante la imitación de patrones establecidos por el

rehabilitador con su voz o almacenados en el sistema y mediante el control de videojuegos mediante voz, con el fin de estimular a los pacientes más pequeños.

Los dos equipos más utilizados en España son el Visualizador Fonético de IBM,⁴ y el sistema VISHA (VISualizador del HAbra), desarrollado por la Universidad Politécnica de Madrid (UPM)⁵ (fig. 4).

Respecto a las ayudas para el diagnóstico de patologías de la voz, los últimos productos aparecidos en el ámbito de la otorrinolaringología empiezan a incorporar ayudas al diagnóstico basadas en técnicas de proceso de voz; a modo de ejemplo vamos a describir dos de los ofrecidos por la firma Kay Elemetrics.⁶ En el fonetograma se realiza una representación sobre unos ejes cartesianos del tono fundamental y energía producidos por un paciente al emitir un fonema aislado. Durante la prueba el paciente debe producir las máximas variaciones posibles de ambos parámetros, con el fin de ver su margen de emisión que posteriormente será valorado por el médico especialista.

Con el *Multidimensional voice análisis* se realiza una representación en forma polar de hasta 25 parámetros extraídos de la señal de voz. Entre ellos se encuentran tono fundamental, energía, Jitter (variaciones de tono fundamental), Shimmer (variaciones de energía) o distintas medidas del ruido contenido en la señal. En él se representa los márgenes de normalidad con anillo verde y los parámetros del paciente en el sólido rojo.

Las ayudas auditivas

Las personas sordas o con elevadas pérdidas de audición suelen utilizar prótesis auditivas para mejorar la percepción de los sonidos que se producen a su alrededor, estas prótesis son fundamentalmente de dos tipos: audífonos e implantes cocleares. Hasta hace unos pocos años, los audífonos utilizaban técnicas analógicas de procesamiento del sonido y, básicamente, se limitaban a realizar una amplificación de la señal selectiva en frecuencia. Con el advenimiento de los audífonos digitales donde se hace un auténtico procesado digital de la señal de voz, la capacidad de manipulación de la señal, y de adaptación a las pérdidas del paciente son mayores. Para más información sobre audífonos se pueden consultar las páginas de Internet de los distintos fabricantes, en concreto la de la firma OTICON⁷ está estructurada de una forma muy didáctica.

Los implantes cocleares pretenden emular, mediante unos pequeños electrodos, la función de excitación del nervio auditivo que realizan las células ciliadas de la cóclea. Está formado por dos elementos totalmente independientes: el *transmisor* y el *receptor*.

El *transmisor de colocación externa* al paciente está formado por un micrófono, un procesador de señal y una pequeña bobina que transmite la información, a través de la piel del usuario, al receptor.

El receptor está alojado en el oído del paciente mediante una intervención quirúrgica.^{8,9,10} Se compone de una antena que recibe la información del transmisor y de unos electrodos que se alojan en la cóclea y son los encargados de excitar el nervio auditivo.

Conclusiones

Aunque la tecnología de la lengua, entendida de una forma amplia, a día de hoy es una gran ayuda a las personas discapacitadas. Debemos decir que tan sólo se está aprovechando una pequeña parte del *know how* existente en esta tecnología y que una investigación aplicada, seria y rigurosa, tomando como punto de partida el conocimiento existente, podría producir incrementos de calidad de vida espectaculares en muchos colectivos de discapacitados, así como progresos importantes en el diagnóstico de determinadas patologías del aparato fonador.

Bibliografía

¹ En <http://www.ceapat.org> se puede consultar un amplio catálogo de ayudas técnicas para personas discapacitadas.

² SIRA E. PALAZUELOS: «Aportación a la predicción de palabras en castellano y su integración en sistemas de ayuda a personas con discapacidad física», Tesis Doctoral, Universidad Politécnica de Madrid, Mayo de 2001.

³ <http://cidat.once.es>

⁴ <http://www-3.ibm.com/able/spv3.htm>

⁵ <http://lir.die.upm.es>

⁶ <http://www.kayelemetrics.com/>

⁷ <http://www.oticon.com/>

⁸ <http://www.utdallas.edu/>

⁹ <http://www.cochlearimplant.com/>

¹⁰ <http://www.cochlear.com/>

Nota: Las figuras que incluye este artículo han sido extraídas de las direcciones de Internet referenciadas en este apartado.

Santiago Aguilera,¹ J. Ignacio Godino,¹ Sira E. Palazuelos² y J. Luis Martín²

¹ Departamento de Ingeniería Electrónica de la ETSI de Telecomunicación de la Universidad Politécnica de Madrid, Madrid.

² Escuela Politécnica, Universidad de Alcalá, Alcalá de Henares, Madrid.

Figura 1 Sistema de escritura por barrido, que incorpora predicción de palabras



Figura 2 Representación Braille del abecedario

a b c d e f g h i j k



l m n ñ o p q r s t u



v w x y z á é í ó ú ü

