

# El poder de la voz sube decibelios

Los sistemas de reconocimiento de voz mejoran la productividad de las empresas

Escribir sin teclear es posible gracias a las tecnologías de reconocimiento de voz que permiten transcribir la voz humana, convirtiéndola automáticamente en texto escrito. Las ventajas que puede reportar en términos de mejora de eficiencia y, por tanto, de reducción de costes operativos no están pasando desapercibidas por las empresas.

**M<sup>a</sup> Luisa Melo.-**

Y es que el teclado cumplió en 2005, 137 años, una edad considerable, pese a la que el diseño del mismo no haya variado en exceso, en comparación con los vertiginosos avances de otros dispositivos hardware. El siguiente nivel no pasa tanto por los teclados enrollables e impermeables de reciente lanzamiento, que también, sino más bien por las tecnologías de reconocimiento vocal, un mercado en creciente expansión, pero con diferentes perfiles

evolutivos en función del país de despliegue.

En este sentido, Jesús María Boccio, responsable de Speechware, reconoce que “el mercado de reconocimiento de voz es un éxito en Estados Unidos, si bien en España aún es incipiente. Básicamente se debe a que la adopción de tecnologías de reconocimiento de voz es más amplia en el sector sanitario, siendo éste de carácter privado en Estados Unidos. De ahí que en Europa, todavía está costando un poco porque la sanidad pública aún no está del todo preparada”.

**Natural e intuitivo**

Las ventajas de estos sistemas son muchas, desde corregir, editar y formatear dicho texto con la voz, hasta interaccionar y controlar tanto el sistema operativo como las demás aplicaciones con “comandos vocales”. El objetivo no es otro que complementar, que no sustituir, el teclado y el ratón por un “interfaz” mucho más natural e intuitivo como es la propia voz del usuario.

Un sistema de reconocimiento vocal profesional permite generar texto en diversas lenguas y sin necesidad de teclear, dictando al equipo, ya sea un ordenador personal, una grabadora digital portátil o incluso a una agenda electrónica personal (PDA), siendo la velocidad media de cualquier usuario de entre 160 a 180 palabras por minuto.

Por otro lado, también es posible obtener una gran precisión en el reconocimiento de la voz, dictando incluso texto libre. En este punto, vale la pena advertir que esta velocidad varía en función del producto de que se trate.



“El mercado de reconocimiento de voz es un éxito en Estados Unidos, si bien en España aún es incipiente, porque la sanidad pública aún no está del todo preparada”,  
Speechware

En este sentido, Boccio reconoce “las ventajas que genera estos sistemas que permiten optimizar el uso e integración de Dragon Naturally Speaking, siendo posible lograr un porcentaje muy elevado de precisión con un mínimo de entrenamiento o adaptación de la aplicación a la voz de cada usuario”. Dicho porcentaje aumenta cuando se “enseña” a la aplicación, corrigiéndose los errores de transcripción, lo que agiliza notablemente las funciones administrativas”, según Boccio.

Por su parte, Juan Prieto Madoño, responsable de desarrollo de negocio de Ydilo, reconoce que entre las principales

ventajas de estas tecnologías para el ámbito corporativo destaca la “posibilidad de desplegar servicios avanzados de atención telefónica, que van mucho más allá de ser un simple filtro o “separador” de llamadas, para proporcionar una serie de funcionalidades similares a las que podría ofrecer un operador que estuviera al otro lado de hilo telefónico”.

Así, es posible hacer un uso más eficiente de los recursos humanos destinados a la atención al cliente, que quedan liberados para realizar tareas de un mayor valor añadido, lo que puede repercutir en una mayor satisfacción del cliente, a su vez.

Otro punto de vista interesante a la hora de revisar los beneficios que pueden reportar este tipo de tecnologías de reconocimiento de voz es su utilidad en el aprendizaje de idiomas. Domingos Bruno, director comercial de Latioamérica y España de Auralog, empresa especializada en la aplicación de la tecnología del reconocimiento de voz al aprendizaje de idiomas, explica que “genera una interactividad enriquecedora”, especialmente gracias a la detección automática del error de pronunciación, una tecnología que permite al alumno loca-

## El Supremo agiliza legajos a golpe de voz



La profusión de casos y juicios es una realidad en las mesas del Tribunal Supremo español. Las tecnologías de reconocimiento de voz podrían aminorar la lentitud en la gestión de los procesos de la justicia española. Con este propósito, el Tribunal Supremo ha iniciado un proyecto

piloto con un grupo de 10 personas, Magistrados del Tribunal Supremo y miembros de la Secretaría de Gobierno y del Gabinete Técnico de implantación de la tecnología de Speechware. En función de los resultados se definirían las necesidades para la implantación progresiva de la nueva tecnología en el resto del organismo. Fernando Román García, jefe del gabinete del presidente del Tribunal Supremo de España, reconoce los beneficios de su utilización apuntando la “racionalización de los métodos de trabajo en las distintas fases del proceso de elaboración de resoluciones del Tribunal. Además, permitiría recuperar la abundante información de todo tipo (jurisprudencial y gubernativa), que, por su antigüedad, se encuentra manuscrita, mediante su conversión a soporte informático, con la ventaja adicional de poder conservarla, ya transformada, en soporte tecnológico seguro y de duración ilimitada, sin los riesgos inherentes al manejo físico del papel.

Sin embargo, el responsable también alude a “cierta resistencia inicial” a la hora de implementar una nueva herramienta, así como la “necesidad de adaptación a una tecnología con la que no se está familiarizado y cuyo uso precisa para obtener la máxima rentabilidad, además, de un período de formación”.

lizar la palabra en que hizo el mayor error de pronunciación en toda la frase.

Más allá de enfoques verticales, este tipo de tecnologías pueden llegar a optimizar las infraestructuras de TI, ya que



puede ser utilizada como apoyo para el sistema de comercio por voz obteniendo así mayor rendimiento de las inversiones previas realizadas en software y hardware.

#### Campos de desarrollo

El sector sanitario, de la abogacía, magistratura o entornos sociales para discapacitados son algunos de las áreas en las que se prevé un mayor despliegue y desarrollo de este tipo de tecnologías. Un punto en el que está de acuerdo Boccio, quien, además, apunta "la necesidad de desarrollar un vocabulario específico para el mundo de la abogacía, que consiga facilitar las funciones de dictado", ya que en España existen alrededor de 170.000 letrados censados.

En definitiva, cada sector de actividad presenta unas necesidades de comunicación específicas: "algunos sectores centran su atención telefónica en la atención al cliente, otros como canal comercial y de desarrollo de negocio y algunos como parte de sus procesos productivos internos", expone Prieto. Cualquier empresas que utilice el teléfono de forma intensiva en sus procesos de negocio, ya sea con clientes, proveedores o en el propio proceso de producción, puede advertir los beneficios de estas tecnologías,

## Iberdrola aligera tareas administrativas

Iberdrola concluía durante el primer trimestre de 2005 su apuesta por esta tecnología. "Consistió en una prueba de evaluación del producto Dragon Naturally Speaking v7 (motor de reconocimiento de voz y transcripción a texto) complementado con el producto Transcription Aid (permite la utilización del anterior en un entorno de grupo de trabajo con asistentes) en varios departamentos de Iberdrola, cuyo resultado fue satisfactorio", resume Francisco Javier Morrás, director de sistemas de Iberdrola. Y es que los beneficios que genera a los departamentos que pueden crear documentos textuales (áreas corporativas, en general), gestionar correo electrónico o realizar capturas de datos en formato numérico son considerables. En el capítulo de ventajas, Morrás reconoce que "la impresión de uso en cuanto a amigabilidad fue favorable, la precisión media de acierto se situó en torno al 98% y, aunque depende mucho del escenario probado, se llegó a lograr un ahorro de tiempo del 60%".



con lo que destacan sectores como el financiero, telecomunicaciones, transporte y logística, distribución y gran consumo, administraciones públicas, entre otros.

El campo de aplicación de estas soluciones es muy amplio, desde compañías de telefonía, financieras, de venta de entradas, transporte y logística o la administración pública.

En lo que respecta a las aplicaciones, se pueden desplegar tanto para servicios básicos de información hasta venta de entradas, gestión de programas de fidelización, captación de nuevos clientes, banca telefónica o portales del inversor. En cuanto a las oportunidades en el sector

público, el responsable de Ydilo destaca que "también puede desplegar servicios de atención telefónica automatizada como información sobre horarios de atención al público, direcciones y rutas de acceso, servicios de petición de cita, ofertas de empleo, becas y subvenciones a empresas, pago de multas e impuestos varios o servicio de notificación personalizada, entre otros".

Fuentes de Microsoft, han apuntado como sectores claves de implantación "la banca por teléfono, los centros de atención al cliente". Asimismo, hay que atender la expectativas de desarrollo de cualquier tipo de servicios para deficientes visuales, ya que permite dictar en vez de teclear en un ordenador, así como "asistencia y servicio de información telefónica, para sustituir las interfaces de tonalidad pesadas que pueden ser fuente de confusión, en el sector servicios para hospitales, servicios administrativos públicos o seguros", según Bruno.

Según IDC, los sistemas de reconocimiento de voz ayudan a optimizar la gestión de llamadas de cualquier volumen; ofrecer nuevas aplicaciones a los centros de atención al cliente; generar bases de datos bajo captura de la información,

que se genere en cada contacto cliente-empresa y viceversa para realizar un auténtico marketing *one-to-one* en tiempo real o cumplimentar formularios, encuestar a clientes permitiendo a los operadores humanos resolver otros problemas.

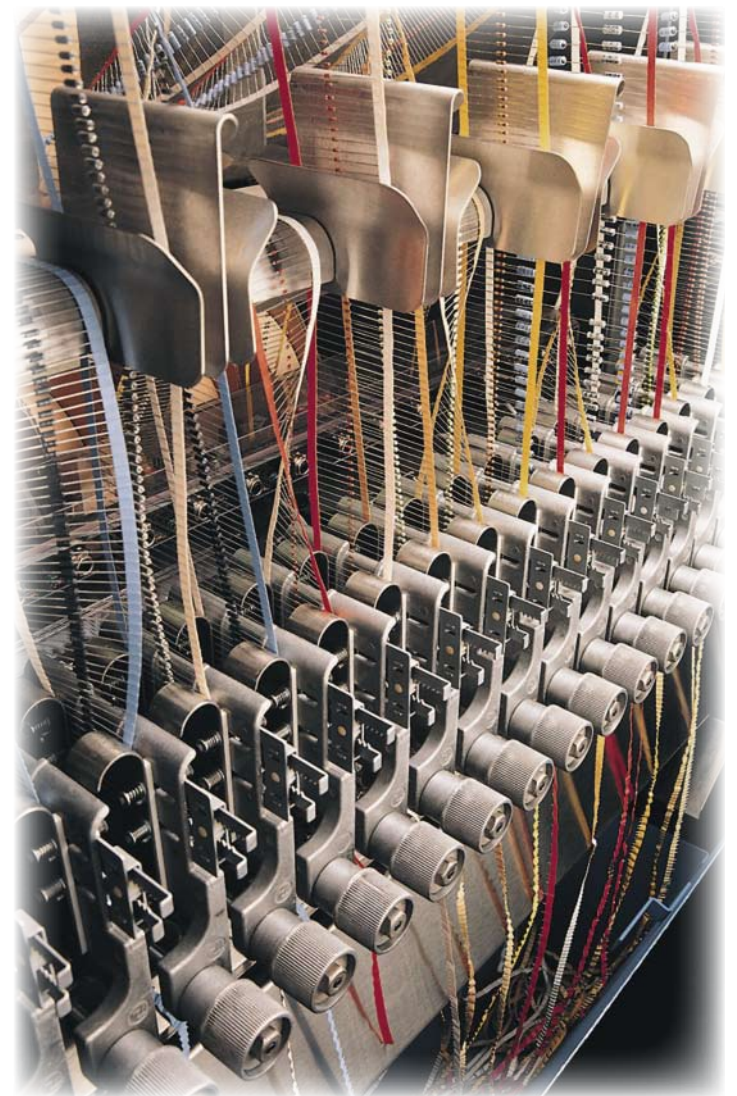
#### Ondas de futuro

Si revisamos las perspectivas de los sistemas de reconocimiento de voz podemos apreciar que están emergiendo considerablemente, ya que "cada vez existe una demanda mayor de sistemas interactivos con posibilidades de conversación", según fuentes de IDC. Asimismo, los ingresos provenientes del negocio de reconocimiento de voz, integra-

dos con otras aplicaciones multimedia se verán incrementados, debido a que los servicios de voz de última generación ofrecen multitud de opciones con nuevos beneficios para las empresas como la venta de productos, información en tiempo real 24x7, recepción de pedidos, reservas de viajes o redención de concursos y promociones.

De "prometedoras" califica Prieto las perspectivas de futuro, ya que "caminan hacia la convergencia del reconocimiento de voz con las tecnologías multimedia en lo que llamamos multimodalidad, aprovechando las capacidades que ofrece la telefonía de tercera generación". La multimodalidad consiste en la posibilidad de que la interacción entre usuario-proveedor de servicio pueda llevarse a cabo en cualquiera de los formatos multimedia disponibles en la actualidad y en un futuro próximo. El modo final en que el intercambio de información se realice dependerá de la naturaleza de esa información, de las capacidades del dispositivo de recepción o de las preferencias del propio usuario.

**"La tecnología del reconocimiento de voz aplicada al aprendizaje de idiomas genera una interactividad enriquecedora", Auralog**



**"Las perspectivas de futuro son prometedoras, ya que caminan hacia la convergencia del reconocimiento de voz con las tecnologías multimedia", Ydilo**