



## La traducción cibernética del español

### Descripción

Vivimos en un entorno económico y social en continuo cambio. La globalización creciente gracias a la utilización de las nuevas tecnologías de la información es un hecho que plantea nuevos retos y nuevas oportunidades, también para los idiomas.

El fenómeno más paradigmático de este nuevo entorno global es Internet. No ha existido a lo largo de la historia otro sistema de información que haya crecido y evolucionado tanto y tan deprisa como la red. Cada tres meses progresa en madurez lo que a cualquier otro medio le cuesta un año. Por ejemplo, durante el periodo 1996-2000, el número de usuarios de Internet ha crecido a un ritmo de más del 31% anual en Europa y del 18% en los Estados Unidos, con cifras de más de 70 y 85 millones de usuarios respectivamente.

Las nuevas tecnologías en general, e Internet en particular, han hecho que surjan nuevas necesidades de expresión, nuevos conceptos que nuestros idiomas no pueden describir adecuadamente por carecer de términos o palabras precisas. La mayor parte de esta nueva terminología proviene del inglés. No es extraño. Algunos datos estadísticos recientes revelan que casi el 70% de las páginas de la red están escritas en ese idioma. El español, contrariamente a lo que pudiera pensarse dado su enorme potencial de usuarios, está pobremente representado. Tan sólo el 2,42% de las páginas *web* están en castellano, lo que sitúa a nuestro idioma en un modesto sexto puesto por detrás del inglés, el japonés, el alemán, el chino y el francés.

En cualquier caso, cabría esperar que la progresiva incorporación de nuevos usuarios hispanohablantes a Internet —un fenómeno que se producirá de forma creciente en los próximos años— hará aumentar la necesidad de prestar servicios y ofrecer información en castellano. No obstante, esta tendencia no debe hacernos pensar que los internautas o usuarios de tecnologías estarán únicamente interesados en información disponible originalmente en su propio idioma. Al contrario, también es sensato prever que van a cobrar importancia aquellos mecanismos que permitan a los usuarios acceder rápida y eficientemente a información y servicios disponibles en otros idiomas. En este sentido, estamos asistiendo, por ejemplo, a una auténtica eclosión de los sistemas de traducción automáticos. Así, la consultora IDC estima que el mercado de la traducción automática alcanzará un volumen de negocio de 378 millones de dólares en el año 2003.

Ya existen programas informáticos, como el anunciado recientemente por IBM, capaces de traducir páginas de Internet a múltiples idiomas —entre ellos el español— en tiempo real. Este sistema proporciona traducción de texto bidireccional, entre inglés y español, alemán, francés e italiano, así

---

como de inglés a chino, japonés y coreano. Ello supone un paso adelante en la comunicación global y en el comercio electrónico al ayudar a eliminar las barreras del lenguaje.

Otro aspecto que cobra una relevancia especial es el desarrollo de tecnologías que permitan una comunicación más natural con el ordenador. En este sentido, IBM lleva trabajando desde hace más de veinticinco años en tecnologías de reconocimiento del habla para varios idiomas, entre ellos el castellano. Con la tecnología actual, los usuarios pueden no sólo dictar textos al ordenador, sino también realizar funciones avanzadas (navegación por Internet, manejo del correo electrónico y del sistema operativo) utilizando la voz, en lugar del teclado o el ratón tradicionales.

El sistema de reconocimiento del habla está basado en complejos métodos probabilísticos y modelos lingüísticos. La conversión de la palabra hablada en texto se realiza a través de algoritmos que aíslan, identifican e interpretan los componentes fonéticos individuales del habla humana. Este proceso resulta altamente eficaz y permite alcanzar una tasa media de aciertos del 96%. El ordenador es capaz, por ejemplo, de elegir correctamente entre palabras homófonas, como «a» y «ha», y diferenciar la palabra «coma» del signo de puntuación. Esta tecnología tiene múltiples aplicaciones, como el control y activación de ordenadores mediante la voz, recuperación de información en bases de datos de lenguaje natural o aplicaciones telefónicas.

Desde 1983, la Unidad de Lenguaje de IBM España ha trabajado tanto en el área de tratamiento de textos como en el de tratamiento del habla. Esta unidad de ingeniería lingüística en castellano ha desarrollado recursos básicos (léxicos, corpus y analizadores) y productos como correctores, traductores y sistemas de dictado automático. Es conveniente reseñar que esta unidad estudia las necesidades específicas del castellano sin olvidar las características particulares de los hablantes en función de su origen (diferentes regiones españolas y latinoamericanas).

En la actualidad se están realizando esfuerzos importantes en proyectos basados en la tecnología de comprensión del lenguaje natural, que va mucho más allá del simple reconocimiento de la voz y permitirá a las personas *dialogar* con los dispositivos informáticos.

Debemos ser conscientes de que la herramienta de comunicación humana fundamental, la voz, se perfila como el instrumento de vanguardia en las tecnologías de la información en este nuevo milenio. En este sentido, el estudio concienzudo de los diferentes idiomas y la adaptación de las tecnologías a sus particularidades es una prioridad de las empresas que trabajamos en el sector. Pensamos que es imprescindible dedicar más esfuerzos, no sólo desde la perspectiva tecnológica, sino también de la lingüística, al español hablado, y no con afán utilitarista, sino porque a través de la voz va a ser posible un enriquecimiento cultural y lingüístico si ponemos en ello el empeño y los recursos necesarios.

Es cierto que a través del lenguaje escrito hemos logrado afianzar todo aquello que nos une a los que hablamos el idioma español, a pesar de las notorias diferencias fonéticas. No obstante, la tecnología va a seguir potenciando la comunicación hablada, más rápida y natural.

El desarrollo de este tipo de tecnologías en castellano está garantizado por la relevancia intrínseca —económica y social— que tiene nuestro idioma en el contexto mundial. Las empresas de nuestro país debemos esforzarnos en potenciar la utilización de la lengua española, hasta situarla en una posición de vanguardia en el ámbito de las tecnologías de la información.

#### **Fecha de creación**

30/05/2001

**Autor**

Juan Pi Llorens

*Nuevarevista.net*