

# Informática y humanidades

POR M.R.J.

Los humanistas y estudiosos de las ciencias sociales que deseen sacar el máximo partido de su ordenador, sin necesidad de ser especialistas en Informática, son el objetivo declarado del autor de este tratado (1), que les ofrece la posibilidad de llegar al manejo de recursos informáticos complejos, de tipo externo, partiendo de sistemas y procedimientos sencillos con los que normalmente contarán en su propia casa.

F. A. Marcos Marín, catedrático de Lingüística de la Universidad Autónoma de Madrid, cuyos numerosos trabajos le han labrado una sólida reputación como investigador dentro y fuera de España, muestra desde hace casi un cuarto de siglo especial inquietud por aprovechar para estas disciplinas el inmenso caudal de eficacia que la informática les aporta. Con el presente libro persigue poner al alcance de los humanistas que trabajan en lengua española y portuguesa su amplia experiencia práctica en el acopio y tratamiento de la información, bien se trate de búsqueda de fuentes, manejo de sistemas, programas y procedimientos que durante años ha utilizado y, como ocurre en varios casos, contribuido a crear.

Como él mismo reconoce en la introducción, la gran diversidad y hondura del campo que abarca, en el que se hallan comprendidos estudiosos, investigadores de la lingüística, educadores y gerentes de empresa, ha exigido del autor un planteamiento eminentemente práctico que abre posibilidades orientativas. Acorde con ello, la obra presente está jalonada de numerosos ejemplos que allanan el terreno a los neófitos y brindan nuevas perspectivas a los humanistas que necesitan cada vez más consultar, recoger y transformar información, procesos ya imposibles de abordar con éxito sin recurrir a la potencia de los ordenadores.

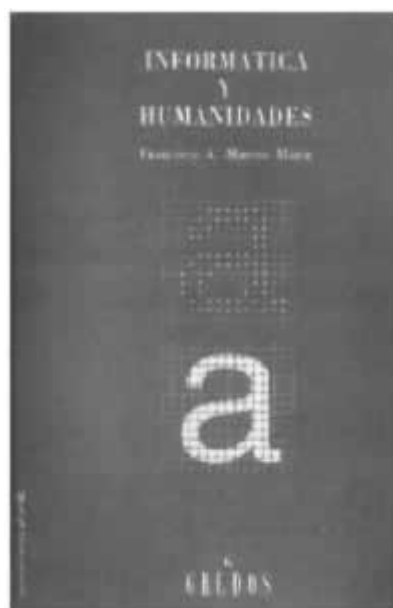
La obra se divide en apartados que, conservando coherencia entre sí, tienen valor por sí mismos y se pueden leer con provecho como unidades independientes. Así ocurre con la presentación de instrumentos físicos de trabajo: disco duro y disco compacto, y con los instrumentos lógicos, en los que desde los sistemas operativos se hace una incursión en profundidad hacia los sistemas de escritura, tanto editores como procesadores, que han sufrido un vuelco conceptual con la aparición de nuevos dispositivos periféricos que han dado lugar al multimedia -técnica en la que el ordenador organiza y trata la información que aportan distintos medios- y al hipertexto, que permite la construcción de un acceso personalizado a la información, prácticamente a medida.

## POLÍTICA E INDUSTRIA



Las industrias de la lengua, concepto que engloba los productos de la intervención de medios técnicos sobre el lenguaje natural, escrito o hablado, constituyen un atractivo núcleo dentro de esta obra. La importancia que cobran estas industrias es tal que en el cuarto programa marco (1994-1998), la Comisión de las Comunidades Europeas ha asignado una cifra cercana a los 800 millones de ecus a su promoción. No es extraño, pues sólo en Europa hay más personas trabajando sobre tratamientos de texto que en labores agrícolas. Y aún más: dos tercios de la mano de obra activa trabaja sobre símbolos de información. Se ofrece aquí un bosquejo realista del entorno lingüístico europeo, la internacionalidad económica, las relaciones de lengua y tecnología en el mercado internacional y la perspectiva y situación desde España e Iberoamérica. El autor expone también su convencimiento de la necesidad de movilizar la voluntad política para que los negocios basados en el idioma español, tan prometedores por la gran implantación de esta lengua y que se revelan como de la máxima importancia económica y cultural, no sean, paradójicamente, promovidos y explotados en su mayor parte por personas y empresas extranjeras. Para ello, se propone como muy necesaria la creación, en el ámbito nacional, de un organismo coordinador de industrias de la lengua. Estas industrias ofrecen, en lo que se refiere a productos, la posibilidad de coordinar esfuerzos en principio dirigidos a fines distintos, aprovechando así mejor los recursos invertidos. El Dr. Marcos Marín señala la necesidad de defender el idioma apoyándose en que, sin duda, una lengua no es una masa de hablantes, sino base para el desarrollo de una industria o de unas industrias y nadie que se ocupa de una industria o la crea espera que se defienda sola. Como él mismo ha dicho y escrito en repetidas ocasiones, el supuesto contrario sólo puede ser sostenido por aquellos «que no sepan nada, no de lengua, sino de industria».

La participación desde la Universidad Autónoma de Madrid en programas de gran ambición, como el comunitario Eurotra, y la experiencia recogida en su etapa como director del Área de industrias de la lengua en la Sociedad Estatal Quinto Centenario, permiten a F.A. Marcos describir el panorama que ofrecen la traducción automática y el desarrollo de las industrias del idioma en España desde el comienzo de este decenio, con especial mención al desarrollo de archivos digitales, programas educativos y medios informáticos de enseñanza.



## HACIA UN CORPUS DE REFERENCIA

En el apartado correspondiente a archivos digitales, se dedica cumplido espacio a los corpus. Se trata archivos digitales textuales uno de cuyos fines -entre otros de no menor importancia- es servir como punto de partida para el tratamiento automático de la lengua. Aquí se urge a acometer el corpus de referencia del español contemporáneo, que ha de incluir el español de América, además del peninsular e insular de España. Esto responde a la necesidad de establecer nexos entre los resultados de la investigación universitaria y las aplicaciones industriales. Aunque no hay un corpus de referencia completo del español general, en dos años y para la Sociedad Estatal Quinto Centenario, el autor dirigió la realización de un corpus oral peninsular y dos escritos, uno del español de Argentina y otro del de Chile, demostrando así la viabilidad del proyecto general, que sigue abierto.

En la sección dedicada a archivos de texto e imagen, se presenta y describe el Admyte (Archivo digital de manuscritos y textos españoles) que es el corpus de mayor amplitud de esta naturaleza que, como proyecto único, existía en julio de 1994. El esfuerzo desplegado para su realización no tiene parangón en ninguna otra lengua. Esta combinación única de textos y facsímiles electrónicos fue asimismo auspiciada desde 1990 por la Sociedad Estatal Quinto Centenario y el autor de este libro fue su director científico.

Los archivos bibliográficos son, por sus especiales características, objeto de exposición aparte en esta obra. Constituyen una revolución en el método bibliotecario y el autor nos ofrece instrucciones para lograr entrar en ellas a través de las diversas redes. Amplios ejemplos de los principales recursos ilustran esta parte. La dedicada a filología textual e informática nos lleva, desde el análisis conceptual de los procesos de escritura, composición de textos, transmisión, copia y lectura a los procedimientos y programas de crítica textual informatizada y edición unificada (Libro de Alexandre). Los sistemas de recuperación de información léxica (análisis de archivos digitales y extracción de la información que contienen) constituyen otro capítulo y el dedicado a terminología, aparte de aclarar y establecer las relaciones de esta disciplina con la lexicografía y otras afines, analiza las necesidades nuevas de creación, mantenimiento, información y traducción que, sobrepasando en mucho las capacidades físicas de los terminólogos, exigen de estos el uso de la potencia informática.

Por último, los que deseen navegar por el ya inmenso acervo de la información automatizada apreciarán la parte dedicada a acceso remoto y comunicaciones, en la que el lector hallará ejemplos y aclaraciones sobre correo electrónico, protocolos, y navegación por Internet, la autovía de información que está revolucionando la comunicación electrónica. Una amplia documentación (en todo tipo de soportes) permite al lector ahondar en los temas tratados y un glosario ad hoc aclara sus dudas sobre conceptos fundamentales y acrónimos más utilizados.

Como reflexión final, sabedores de que la limitación de estas líneas imposibilita un comentario más acorde con la dimensión de la obra, hemos de decir que la consideramos un valioso instrumento para los que precisen de una introducción actual y práctica a las soluciones que la informática y las modernas comunicaciones aportan a las tareas relacionadas con la lengua. Entre sus posibles destinatarios -aparte de los lectores por defecto ya mencionados- vemos con claridad a los estudiantes de las nuevas licenciaturas de traducción, quienes precisan

manejarse entre los medios informáticos que hoy existen para recopilación de documentación y traducción asistida y automática. Necesario es decir, como el propio autor manifiesta en su introducción, que no se trata de un manual de ordenadores, ni de lingüística informática, traducción automática, inteligencia artificial, programación, tratamiento del habla, programación, lingüística aplicada u otros temas que no se abordan directamente, aunque algunos de los ellos aparecen como consecuencia de su desarrollo, de suerte que el lector entrará en contacto con estas disciplinas de la mejor forma: encarándolas a través de otras que quizá le son más familiares.

(1) Francisco A. MARCOS MARIN, Informática y Humanidades. Manuales Gredos. Madrid, 1994. 805 páginas.

