

FRANCISCOA. MARCOS MARÍN, *Informática y humanidades*, Madrid, Gredos, 1994, 816 pp.

Con una claridad expositiva fuera de lo común en este tipo de textos, Marcos Marín nos acerca, con una mezcla de filología entendida en el sentido más clásico del término e informática en los niveles más avanzados, a lo que es, en muchos casos gracias a su dedicación, una realidad ya en la investigación en Humanidades dentro de la Universidad española.

Reseñar esta obra es, no obstante, una tarea difícil porque en ella se analizan cuáles son los últimos programas que pueden aplicarse en campos tan diversos como la crítica textual, la lexicografía informática, el almacenamiento de textos orales y escritos o las industrias de la lengua, y se describen todos aquellos medios que las modernas tecnologías han puesto al servicio del humanista y del investigador en general para la consulta bibliográfica a través del ordenador, las intercomunicaciones y los bancos de datos (“No es un libro de lingüística computacional o de traducción automática, ni de inteligencia artificial. Tampoco es un libro sobre tecnología del habla, ni de programación, ni un manual de uso del ordenador, ni una introducción a la lingüística aplicada, en general, o un manual de enseñanza con ayuda del ordenador” –p. 9–, “Es un libro con un fuerte componente informativo, incluso general, que recoge sistemas, programas y procedimientos, desde los sistemas operativos, el hardware o los paquetes de programación, experimentados por el autor” –p. 8–).

Debido a la informática, nuestros esquemas como investigadores han cambiado en la última década y, como muy bien se indica en la introducción, la incorporación a esa tecnología se traduce, además, en un orden distinto, en una “situación social en la que no cuentan tanto las ventajas de rapidez y eficacia –con ser muy importantes– como la apertura de una nueva perspectiva intelectual”. Aunque los instrumentos de la lingüística aplicada no son sino eso, instrumentos, su aplicación es tal que ahora parecen asequibles empresas y nuevas investigaciones impensables hace solamente unos cuantos años.

A través de diez capítulos, más una introducción de conceptos generales, y un apartado final dedicado a “Documentación” (entendida ésta en un sentido más amplio que la tradicional bibliografía), el autor nos va describiendo, generalmente con ejemplos prácticos ya realizados con esos mismos programas y técnicas, y con un lenguaje claro y sencillo, desde los hechos más elementales en informática como la explicación del hardware necesario hasta la descripción del software actualmente disponible para los análisis filológicos. Al mismo tiempo, se combina todo ello con la historia (corta para nuestros habituales modelos, pero ya relativa-

mente larga comparada con la rapidez con que se suceden técnicas y programas informáticos en estos últimos años) de las llamadas “Industrias de la lengua”, “ingeniería lingüística” o “tecnología lingüística”, especialmente en el entorno europeo, así como se presenta un análisis real de la situación del español como lengua internacional y de la tecnología de la lengua en nuestro ámbito. Las posibilidades abiertas son múltiples: “Las nuevas tecnologías de la información constituyen un apoyo fundamental del idioma, en los dos sentidos históricos: hacia el pasado, porque medios de almacenamiento electrónico y recuperación de la información como los discos láser y los archivos de texto e imagen salvaguardan nuestros tesoros escritos; hacia el futuro, porque, tal vez en una sociedad sin papeles, podemos crear instrumentos gigantescos de almacenamiento lingüístico, como las grandes bases textuales de los corpus de referencia, que nos permiten hoy lograr sistemas de interrogación de bases de datos en lengua natural y nos llevarán en el futuro hacia sistemas expertos con diálogos en lengua natural y el sueño del hombre desde los primeros escribas asirios que compilaban diccionarios: la regularización de la terminología y la traducción automática” (p. 57).

Dos son los capítulos que más llaman la atención, quizá porque en ellos se han logrado ya trabajos que demuestran fehacientemente que ese futuro abierto es, sin duda, un presente realizable dentro de la investigación española. La creación y la publicación de los dos primeros números de archivos digitales textuales contenidos en el proyecto ADMYTE y la obtención por ordenador de ediciones críticas “unificadas”, gracias al programa UNITE, que permite el tratamiento filológico de textos en versos.

Con respecto al primero, el *Hispanic Seminary of Medieval Studies* de Madison (Wisconsin) hacía algunos años que venía coordinando un trabajo de recogida y análisis de un corpus de textos españoles que tendrían como fin la elaboración del *Diccionario medieval*, y también a través de Oxford (*Oxford Text Archive*) era posible adquirir digitalizados algunos textos en diversas lenguas, una parte de ellos españoles. Con ADMYTE se ha pretendido iniciar una “biblioteca del descubrimiento”, en la que el CD-ROM posibilita, por un lado, la conservación de obras, manuscritos e incunables, y, por otro, disponer en cualquier ordenador de un corpus de textos e imágenes que se puede reproducir o que se puede utilizar como base para cualquier tratamiento posterior. Siguiendo las normas de transcripción adoptadas por el equipo de Wisconsin en Madison, se han publicado ya dos volúmenes. El primero (*Volumen 0*), considerado como disco instrumental, incluye la Base bibliográfica (que comprende la *Bibliografía Española de Textos Antiguos*, la *Bibliografía de Textos Antiguos Portugueses* y la *Bibliografía de Textos Catalanes Antiguos*), el programa TACT de recuperación de información textual, las transcripciones de la colección de *Textos Medievales del Hispanic Seminary of Medieval Studies* y de otras obras realizadas con las mismas normas y el programa UNITE para la construcción de ediciones críticas automatizadas. El segundo (*Volumen I...N*) ofrece imágenes facsimilares de textos y de miniaturas de algunos

códices, transcripciones de esas obras, programas de recuperación de la información en texto e imagen y un diccionario interno de lemas y formas de las palabras contenidas en el corpus digitalizado. Con un ordenador personal y el entorno Windows 3.0 ó 3.1 es posible acceder fácilmente a diccionarios como el de Alfonso de Palencia o los de Antonio de Nebrija, o a textos legales, científicos, poéticos, crónicas, biografías, libros de viajes, etc. “La materialidad de las cifras es a veces fría y, a veces, muy reveladora. Si consideramos en cifras ADMYTE-I hemos de pensar que en los trece gramos de peso del CD-ROM se incluyen sesenta y un títulos (sesenta y cuatro en ADMYTE-0), lo que supone más de ocho mil páginas facsímiles de extraordinaria calidad y casi cuarenta millones de caracteres, que corresponden a su transcripción en formato ASCII” (p. 183).

El sistema UNITE, otra de las novedades importantes comentada muy detalladamente a lo largo de este libro, permite la comparación de hasta diez textos en verso distintos y realizar ediciones que el prof. Marcos Marín llama “unificadas”, basadas principalmente en la *coherencia intratextual*: “Son, en suma, ediciones críticas, pero distintas, porque en este caso no se ha partido de qué requiere el filólogo de una edición crítica y qué puede hacer el ordenador para satisfacer esa necesidad, sino de la perspectiva inversa, a saber, qué puede hacer el ordenador en el terreno de la comparación de textos, del registro de las variantes resultado de esa comparación, del control de esas variantes sobre el conjunto de la comparación y de la aplicación de los datos globales para establecer el texto básico, que no pretende ser el arquetipo, sino la mejor copia más antigua posible, según los manuscritos conservados, es decir, una variante del antecesor común más próximo del original” (p. 376). El ordenador ofrece, además, multitud de datos que, recogidos en el aparato crítico, permitirá a cualquier lector reconstruir o discrepar de la versión ofrecida.

Son éstos dos de los muchos ejemplos de las posibilidades que comenta el libro ya que, como versa en uno de los títulos de uno de los capítulos, el ordenador es, hoy por hoy, “una ventana al mundo” que nos permite con un MODEM, sin movernos, acceder a los archivos bibliográficos a través de redes de telecomunicación como REDIRIS, que facilita la interconexión, en el ámbito español, de más de cien organizaciones, la mayor parte de ellas universidades y centros públicos de investigación, y, en el ámbito internacional, la comunicación con redes como IXI de COSINE (red europea).

Otras muchas aplicaciones puede tener la informática en humanidades, todas las que cada investigador vaya necesitando y adaptando a sus propios análisis. El proyecto ADMYTE, por ejemplo, contempla un índice de formas que contiene las palabras recogidas en los textos y desde hace algunos años son conocidos y utilizados programas como *WordCruncher* o *MicroOCP* que facilitan la realización de concordancias de textos. Marcos Marín presenta de forma muy detallada, además, el programa TACT (pp. 583 y ss.), desarrollado en Canadá y que permite el análisis de textos codificados siguiendo las pautas marcadas por Madison

(Wisconsin). Dentro de este apartado hubiese resultado interesante incluir otros programas de análisis del léxico y que hoy son utilizados en la elaboración de diccionarios. No debemos olvidar tampoco las últimas técnicas introducidas en la elaboración de atlas lingüísticos sonoros, tanto en lo que a informatización de proyectos se refiere como al tratamiento de los datos recogidos que, en el futuro, gracias al CD-ROM, formarán bancos de datos acústicos y podrán sumarse a los archivos digitales orales.

De todo un panorama de posibilidades, en definitiva, dispone el filólogo como medios auxiliares de su quehacer investigador. Manuales como éste actualizan y ofrecen una información a la que accedemos, en la mayoría de los casos, muy fragmentariamente. La recopilación, en este sentido, se muestra fundamental, no sólo para simplificar los pasos que debemos recorrer y no volver a caminar sobre lo ya realizado sino también para aunar esfuerzos y, en cierta medida, avanzar bajo las mismas coordenadas.

Dolores Corbella

Universidad de La Laguna