

REAL ACADEMIA ESPAÑOLA, *Diccionario de la lengua española*. 21ª edición, versión en CD-ROM, Madrid, Espasa-Calpe, 1995.

Abierta la posibilidad de aplicar las nuevas tecnologías desarrolladas en el campo de la informática a la consulta de la obra lexicográfica, se ha editado en forma de disco compacto la última edición del diccionario de la Academia. Imagen exacta de la obra impresa, la versión digitalizada reproduce fielmente no sólo la información léxica del diccionario en papel sino también su disposición formal¹. No pocas resultan, sin embargo, las posibilidades que el nuevo soporte ofrece tanto al profano de la lengua como al curioso o investigador de la misma. Posibilidades que, en última instancia, conducen a los cimientos de la obra, al quehacer lexicográfico desarrollado en la misma.

La disposición tradicional de los artículos de definición permitía acceder a la información sólo a partir de la búsqueda alfabética de los lemas. A esa búsqueda práctica, pero limitada, la versión electrónica suma la opción del *diccionario inverso* en la que los lemas se ordenan de derecha a izquierda, es decir, desde el final de la palabra al principio de la misma. Se contempla, además, la posibilidad de realizar búsquedas léxicas a partir de una serie de comodines que, en caso de duda ortográfica o de interés por la variación formal de una parte de la voz, sustituyen un número variable de caracteres.

Esta conjunción de posibilidades en la búsqueda a través de la macroestructura ha supuesto una mejora en la calidad de la consulta, en el acceso a la información por parte del usuario de la obra tradicional. Así, el investigador se ha visto asistido en su tarea con la disposición de las distintas informaciones del artículo de diccionario en una base de datos: estructura informática a modo de ficha en la que cada una de las partes en que se divide el artículo se ha convertido en un campo de acceso a la información o cuerpo del lema. Con ello desaparece el hermetismo del artículo lexicográfico, bloque compacto e inaccesible desde todo punto que no fuere el lema o entrada, para permitir búsquedas de carácter más o menos complejo a través del texto de la microestructura.

El diccionario electrónico contiene, en forma de posibilidades de búsqueda, una organizada infraestructura que dirige el acceso a la información léxica. Además de las opciones ya presentadas que parten de la macroestructura (*diccionario* y *diccionario inverso*), la consulta de la microestructura puede hacerse o bien a partir de una búsqueda de carácter formal o textual, opción que se presenta bajo la forma de diferentes ín-

¹ Las únicas diferencias entre el texto de la versión impresa y el de la versión electrónica atienden a cuestiones de formato con el objeto de mejorar la consulta en la pantalla del ordenador:

El texto que presenta la pantalla se atiene escrupulosamente al del Diccionario en su versión impresa; se han mantenido las variantes tipográficas: negritas, cursivas... Los colores, cuando aparecen, son elementos adicionales de la versión electrónica, y se superponen a las características tipográficas originales. La única diferencia es que las definiciones y formas complejas están separadas por punto y aparte, en vez de aparecer unas a continuación de otras, separadas por una doble raya. Además, en esta versión las definiciones únicas se han numerado también. (*Manual de instrucciones* de la edición en CD-ROM, p. 36).

dices, o bien eligiendo una búsqueda de carácter temático a partir de las informaciones contenidas en los artículos, opción que se presenta bajo la forma de distintos árboles de materia (*árbol de lenguas, árbol de categoría gramatical, árbol de usos: ámbito geográfico, árbol de usos: materia y nivel, etc.*).

Los índices, enumeración alfabética exhaustiva de los elementos que aparecen en los artículos de definición, permiten acceder al contenido de la obra a través de la forma léxica o gramatical elegida por el usuario. De acuerdo con las diferentes informaciones que comprende el artículo (abreviaturas, texto y formas complejas), tres son los índices que presenta el diccionario electrónico:

a) El *índice de abreviaturas y marcas* permite conocer las voces que recogen una determinada abreviatura (gr., psicol., frec., etc.) o una determinada marca de uso (*afectivo, antiguo, apócope, prefijo, etc.*) en su artículo de definición.

Desde esta opción de búsqueda es posible realizar no sólo un estudio léxico que relacione distintas voces con una abreviatura común (abreviatura que, de alguna manera, las califica e identifica), sino el acercamiento a la organización interna de la obra, al sistema de marcas que subyace a la concepción de la misma.

b) El *índice de todas las palabras* permite hacer búsquedas dentro del texto de la definición y de la etimología (en ambos campos a la vez o sólo en uno de ellos).

Esta opción sirve, entre otras posibilidades, para llegar a conocer una parte importante de la teoría lexicográfica de la obra: las relaciones de hiperonimia o relaciones de inclusión semántica (*instrumento*, por ejemplo, aparece 988 veces en 852 artículos, *utensilio* aparece 110 veces en 107 artículos y *aparato* aparece 511 veces en 471 artículos).

c) La búsqueda de formas complejas permite también acceder a las expresiones o frases en las que entra una voz. La voz *hombre*, por ejemplo, aparece en 102 expresiones: [*hombre*] *de bien*, [*hombre*] *de capa negra*, [*hombre*] *de dos caras*, [*hombre*] *de edad*, [*hombre*] *rana*, [*hombre*] *como un solo hombre, etc.*

Frente a los índices, cuyas búsquedas cubren en su totalidad el texto del artículo de definición, los árboles ordenan por materias la búsqueda léxica (el uso dialectal de la voz, el valor diastrático, la lengua etimológica, la información gramatical, etc.). Cinco son los árboles del diccionario electrónico:

a) El árbol de lenguas recoge, agrupadas por estratos, las lenguas etimológicas que aparecen tanto en el texto de la definición como en el texto de la etimología: *lenguas originarias, castellano o español y sus dialectos, otras lenguas peninsulares, árabe, latín, griego, romances, indígenas americanas y otras*, son las ramas principales del árbol que a su vez se expanden en sucesivas ramas de menor extensión.

b) El *árbol de categoría gramatical* recoge por categorías (*adjetivo, adverbio, artículo, conjunción, interjección, preposición, pronombre, sustantivo y verbo*) la información sintáctico-gramatical que aparece en las acepciones de los lemas. Éstas, a su vez, estructuran en forma de nuevas ramas la información que se incluye en cada una de ellas.

c) El árbol de ámbito geográfico organiza la información de uso diatópico: *España, América, África y Asia* son las ramas principales que a su vez se subdividen en grupos de menor extensión geográfica.

d) El árbol de materia recoge, por campos de uso técnico y niveles de uso diastrático, las especificaciones que sobre estos usos aparecen en el diccionario; tres son las ramas principales: *profesiones y disciplinas, variantes socialmente marcadas y variantes desprestigiadas*.

e) El árbol de opciones avanzadas recoge información de carácter heterogéneo (morfológica, pragmática, semántica, etc.) que se estructura en cuatro grandes ramas: *usos restantes, cuestiones etimológicas, formación de palabras y cuestiones semánticas*.

Otra de las opciones de consulta que ofrece el diccionario electrónico es la de la búsqueda múltiple, forma de acceso filtrado o restringido a la información de la obra. Cuatro operadores lógicos (Y, O, NO, paréntesis) se combinan con un determinado número de elementos léxicos para construir expresiones complejas, o suma de condiciones, que delimitan la búsqueda en función de los intereses de quien consulta la obra. La operación permite acceder desde la microestructura a la macroestructura, desde la noción de concepto al nombre que lo designa. Un inconveniente, sin embargo, se deriva de la fidelidad que esta forma de búsqueda manifiesta con respecto al texto escrito de la obra. Con la expresión compleja "*animal Y alas*", por ejemplo, no accederemos al conjunto de voces que los hablantes, conceptualmente, considerarían dentro del grupo de seres animados que comparten dichas condiciones morfológicas. La causa está en que no todos esos lemas contienen la voz *animal* y la voz *alas* en su artículo lexicográfico. Es, pues, la redacción de los artículos de definición, y no su contenido semántico, lo que dirige o condiciona la selección léxica, convirtiendo en una búsqueda de carácter textual lo que prometía ser una compleja búsqueda de carácter semántico.

Queda sólo por presentar el salto hipertextual, la opción más automática de las hasta ahora descritas, la que permite navegar libremente por los recodos del diccionario. Inherente a la naturaleza de toda obra lexicográfica es la utilización de elementos léxicos (ser, elemento, idea, persona, que, etc.) para definir los lemas o entradas, elementos léxicos a su vez. La consulta de la obra, continua glosa o paráfrasis de sí misma, puede convertirse, en el diccionario en papel, en un continuo ir y venir de páginas en busca de las voces poco conocidas del artículo lexicográfico. En el diccionario electrónico, sin embargo, el salto hipertextual permite desplazarse de forma inmediata al artículo de la voz seleccionada en el texto de pantalla. La hipertextualidad posibilita, pues, el acceso automático a los numerosos segmentos, los artículos, en que se divide el diccionario. La obra se rige, por consiguiente, por un nuevo orden de búsqueda léxica, más acorde con las necesidades del usuario y con la propia naturaleza de la obra lexicográfica. En el diccionario electrónico de la Academia, sin embargo, este recurso adolece de las imperfecciones de un cerebro artificial no programado para distinguir homógrafos creados por la flexión verbal: *como*, por ejemplo, aparece en el texto del lema *comer*; el salto hiper-

textual reconoce los homónimos con el valor de adverbio (*como* 2) y de nombre (*como* 1 y *como* 3), pero no la forma conjugada del propio verbo comer, primera persona del presente de indicativo; no reconoce, pues, las formas verbales flexionadas que aparecen en el texto de la definición (*dícese, tratándose*, etc.).

Por último, queda por enumerar otra serie de ventajas y posibilidades que la versión en CD-ROM ofrece frente a la obra impresa. El nuevo soporte permite elegir la forma de presentación en pantalla (esquema de acepciones/artículo lexicográfico completo, ocultación/presentación de las etimologías y de las formas complejas, presentación del texto justificado o no justificado en su margen derecho), además de proporcionar otra serie de facilidades para el acceso a las voces o incluso para la investigación dentro de la propia obra, como son la información léxico-cuantitativa que acompaña a una consulta, la recuperación de las búsquedas anteriormente realizadas, la creación de un archivo externo a la obra en el que recoger información propia o personal (bloc de notas) o la posibilidad de exportar información, a través del portapapeles², a cualquier programa del entorno *windows* (opción editar).

La versión electrónica del diccionario de la Academia supone, pues, una metamorfosis en la explotación de la obra. Libera y agiliza la búsqueda léxica al mismo tiempo que permite no sólo penetrar en las estructuras que rigen y articulan los criterios seguidos en la confección de la obra sino penetrar en el conocimiento léxico. Cabe señalar, sin embargo, que con la digitalización del diccionario, punta del iceberg, se inaugura un vasto campo de trabajo lexicográfico. A pesar de las muchas ventajas que tanto al investigador como al usuario ofrece esta nueva versión electrónica, sus limitaciones y carencias son todavía obstáculo importante al conocimiento. Si bien el programa informático intenta acercarnos a la estructura de pensamiento del hombre (banco de datos léxicos, asociaciones entre voces, etc.), sus posibilidades reales se ven reducidas a la simple búsqueda textual o formal, siendo obligado un segundo nivel de análisis, un análisis, en definitiva, más lingüístico. A la obra digitalizada le falta, en primer lugar, un analizador morfológico con el que deshacer las ambigüedades y limitaciones del salto hipertextual. Asimismo, la manifiesta fidelidad que presenta con respecto a la redacción de la obra impresa la convierte en cómplice de unas mismas incoherencias lexicográficas. Poso de una larga tradición de tres siglos de lexicografía académica, de revisiones lentas e irregulares, el diccionario electrónico carece, al igual que su predecesor, de una teoría lexicográfica aplicada de manera sistemática en la obra. Son, precisamente, las posibilidades del diccionario en CD-ROM (la liberación en el acceso a la información de la obra unido al automatismo en la consulta de la misma) las que delatan esa falta de cura lexicográfica en la composición y revisión del diccionario académico. De la digitalización se des-

² Área de almacenamiento que se utiliza para transferir datos de un documento a otro o de una aplicación a otra. (*Microsoft Windows. Manual del usuario*. Microsoft Corporation, 1990-1992).

prende, pues, una información valiosísima para la enmienda de errores lexicográficos acumulados a lo largo de las veintiuna ediciones. La informática ofrece, además, la precisión y pulcritud necesarias para la corrección exhaustiva de esos mismos errores. De esta primera versión electrónica debiera derivarse, pues, una nueva edición actualizada y revisada del diccionario académico.

Carmen Morales (Universidad Autónoma de Barcelona)