

DICCIONARIOS EN SOPORTES INFORMÁTICOS

O. SANTANA, Z. HERNÁNDEZ, J. PÉREZ, G. RODRÍGUEZ, F. CARRERAS *

En las páginas que siguen pretendemos anticipar de qué manera los diccionarios experimentarán en los próximos años una transformación de la mano de los avances tecnológicos en informática. En cómo se presenten y en cómo se usen —lo formal— seguramente se cumplirá un notable avance; de hecho es innegable que ya están empezando a seguir ese camino. En la concepción de lo que es y para qué es un diccionario —lo sustancial— el cambio no será tan acusado puesto que la informática no persigue el propósito de crear ideas o elementos, sino el de facilitar la plasmación y utilización de las buenas ideas sólidamente contrastadas.

No obstante, a medida que la utilización de los recursos se hace más sencilla surgen otros planteamientos que dan solución a nuevas necesidades; normalmente, el proceso se acelera de forma importante y cuando ello ocurre se implantan nuevas formas de representación y de utilización del conocimiento, lo que daría lugar a productos que estarían mucho más allá de lo que podríamos considerar simplemente como diccionarios mejorados.

Las dos vías principales a las que las nuevas tecnologías prestarían las transformaciones más relevantes a los diccionarios son :

a) el aspecto físico: en el sentido de aquel con el que estamos acostumbrados a identificar un diccionario —el papel— cambiará al publicarse en soportes diferentes del clásico;

b) la flexibilidad de consulta: que se verá beneficiada por la utilización de algoritmos desarrollados para búsquedas en bases de datos textuales, así como de otros basados en el conocimiento del lenguaje a partir del desarrollo de las técnicas de procesamiento del lenguaje natural; permitirá la realización de búsquedas complejas e insuficientemente especificadas, muy distintas de las que estamos acostumbrados.

Ambos factores implican un cambio del modelo de interacción entre los usuarios y los diccionarios; el diálogo se verá enormemente facilitado por la uti-

lización de *interfaces* amigables cuidadosamente desarrolladas; el uso se verá revalorizado por el mayor alcance de los resultados obtenidos al aprovechar las capacidades multimedia de los sistemas empleados.

SOPORTES INFORMÁTICOS

Uno de los cambios más rápidamente evidenciables que están experimentando los diccionarios se refiere a su presentación en un medio directamente procesable con el que se hacen accesibles al público que desea utilizarlos. Aunque ya existían otras ofertas en CD-ROM, no hace aún mucho que la propia Real Academia Española ha editado el suyo y lo cierto es que ahora es cuando se ha incorporado a los ordenadores personales como estándar un lector de CD-ROM.

Con la tecnología actual, un CD-ROM puede almacenar una cantidad de información digitalizada de unos 640 Mbytes, lo que resulta más que suficiente para trasladar a este medio cualquier diccionario de uso corriente; por ejemplo, el *Diccionario de la Real Academia Española* ocupa menos de 40 Mbytes, y eso incluye el *software* específico necesario para su manipulación. No obstante, ya se estudian los estándares para implantar a corto plazo soportes con capacidad para 3,5 gigabytes o más

(un gigabyte = 1.024 Mbytes). Además, hay que tener en cuenta que la capacidad de almacenamiento puede incrementarse sensiblemente con la aplicación de técnicas de compresión de datos que se hallan ampliamente desarrolladas.

Otro soporte que en el futuro puede resultar altamente interesante lo constituyen las tarjetas de memoria: dispositivos sin partes móviles y actualmente con un tamaño aproximado al de una tarjeta de crédito. Estas características las hacen especialmente aptas para su utilización en equipos de consulta portátiles, cuyo peso y dimensiones no superan los de un diccionario de bolsillo aunque sí sus prestaciones. Se espera que este tipo de medio haya alcanzado para el año 2000 una capacidad de almacenamiento de gigabytes —un orden de magnitud por encima de los actuales CD-ROM.

Y naturalmente, no se puede dejar de lado la omnipresente Internet, **la red**, que se ha convertido en un tópico habitual en telediarios y tertulias con una inusitada explosión de entusiasmo más propia de una revelación metafísica. Está produciendo una revolución social en la que ya hay quien apunta a una redefinición futura del término analfabetismo —se deja

entrever que todo aquel que *no existe en internet* simplemente *no existe*—. Internet no es más que **una red de redes** que proporciona la posibilidad de comunicarse entre sí a millones de usuarios en todo el mundo, permitiéndoles ofrecer y obtener todo tipo de servicios entre los que el acceso a diccionarios en línea sería uno más.

En el momento en que redactábamos este artículo realizamos una búsqueda en Internet del término “*on line dictionary*”, eligiendo al azar uno de los varios *ingenios de búsqueda* que se ofrecen; el resultado fueron 469 referencias. Esto no significa que encontráramos 469 diccionarios en línea; de hecho, una sola de esas referencias, titulada “*on-line dictionaries*”, ofrecía enlaces a 106 diccionarios organizados en 25 categorías: “diccionarios especiales

y tesauros”, “diccionarios multilingües”, “lenguas indoeuropeas clásicas”, “lenguas romances”... y 17 enlaces a otras listas de diccionarios.

Lamentablemente, una búsqueda del término “diccionario” en servidores de búsqueda en español —¡*Olé!* y ¡*Ozú!*— no produjo ningún resultado. En compensación, resulta grata la noticia de que la Real Academia Española de la Lengua se encuentra imbuida en la actualidad en un proyecto para la generación de un *corpus* de Referencia del Español Actual al que se podrá acceder a través de Internet; no cabe duda de que se aprovecharán las posibilidades que la informática ofrece a la lexicografía y la filología para intentar abarcar todas las modalidades del uso de cualquier palabra.

La mayoría de los diccionarios que se encuentran en Internet son para la traducción entre dos o más lenguas; predominan los que una de sus lenguas es el inglés, lo cual resulta lógico, puesto que la traducción se convierte en una necesidad de primer orden en un medio destinado a la libre comunicación mundial. Existe un llamado Internet Dictionary Project cuyo objetivo es crear un conjunto de diccionarios de traducción entre múltiples lenguas en base al trabajo voluntario de los propios usuarios de la red. También hay algunos diccionarios de significados, pero en mucha menor medida.

En general, tienen en común una *interfaz* bastante mediocre; se limita a una ventana donde el usuario puede escribir la palabra que se desea buscar y obtener como resultado una pequeña página de formato sencillo, tanto que desde ella no siempre es posible lanzar una nueva búsqueda y menos aún refinar la realizada. Esta deficiencia de los diccionarios actualmente disponibles en línea puede deberse a su juventud; tampoco hay que olvidar que la mayoría han sido desarrollados por voluntarios —organiza-



ción entre múltiples lenguas en base al trabajo voluntario de los propios usuarios de la red. También hay algunos diccionarios de significados, pero en mucha menor medida.

En general, tienen en común una *interfaz* bastante mediocre; se limita a una ventana donde el usuario puede escribir la palabra que se desea buscar y obtener como resultado una pequeña página de formato sencillo, tanto que desde ella no siempre es posible lanzar una nueva búsqueda y menos aún refinar la realizada. Esta deficiencia de los diccionarios actualmente disponibles en línea puede deberse a su juventud; tampoco hay que olvidar que la mayoría han sido desarrollados por voluntarios —organiza-



ESPASA CALPE

REAL ACADEMIA ESPAÑOLA, DICCIONARIO DE LA LENGUA ESPAÑOLA,
EDICIÓN EN CD-ROM, MADRID, ESPASA CALPE, 1996

ciones o particulares— sin ánimo de lucro, mientras que los diccionarios clásicos han sido sustentados por editoriales con carácter comercial. Sin duda, estas editoriales no tardarán en poner sus diccionarios a disposición de los usuarios de Internet, pero antes tendrán que resolver la forma de asegurar su justo beneficio: por suscripción o mediante pago en línea por tiempo, tipo o volumen de consulta.

Los diccionarios en Internet pueden también verse beneficiados del carácter participativo de la misma, pues es raro el recurso que no cuenta con un medio para hacer llegar sugerencias a su administrador. Esta filosofía puede redundar en que los propios usuarios contribuyan a la actualización del diccionario desde una doble vertiente: las propuestas para corrección de posibles erratas y de crecimiento merced a nuevas acepciones.

En relación con los equipos portátiles de consulta es menester señalar que tales equipos pueden basarse simplemente en el empleo de las redes de comunicación inalámbricas con acceso a Internet; en cuanto esa tecnología esté suficientemente extendida, los soportes físicos en poder del usuario —CD-ROM, tarjetas de memoria o cualquier otro dispositivo— tenderán a desaparecer. Lo cierto es que en la actualidad hay ocasiones y lugares en los que puede dar la impresión de que en vez de movernos por una *autopista de la información* lo hacemos por una *carretera de montaña*, con atas-

cos y cuellos de botella que provocan una lentitud exacerbante de las comunicaciones en horas punta. Es necesario que el ancho de banda de las vías que soportan físicamente las comunicaciones se vea incrementado de forma generalizada para obtener el máximo partido de la inmensa cantidad de información y servicios disponibles.

TÉCNICAS DE ACCESO A LA INFORMACIÓN

Por razones obvias de limitación física los diccionarios en soporte de papel están enfocados a una forma de búsqueda muy concreta. Su organización los hace aparecer divididos en artículos, y se accede por medio de una entrada por la que se encuentran ordenados alfabéticamente. Salvo excepciones significativas, tales términos se hallan, en lo que podríamos denominar *forma canónica*, sin flexiones ni derivaciones, lo que constituye una pequeña parte de la lengua real. Esta organización condiciona al usuario a buscar las palabras tal como aparecen en la lista de entradas, o sea, como una expresión restringida de lo que se denomina búsqueda exacta.

Los diccionarios en soporte informático se pueden beneficiar de la aplicación de los algoritmos desarrollados durante años para la localización de documentos en bases de datos textuales que permiten abarcar un amplio abanico de tipos de búsqueda:

- **Exacta:** consiste en determinar todas las localizaciones de una palabra.

- **Más similares:** se localizan palabras que guarden un cierto parecido estructural con la de búsqueda, de acuerdo con alguna métrica que mida las diferencias entre cadenas de caracteres. Por ejemplo: rida - {ría, vida, rica, risa, sida, oída, brida...}. Esta técnica permite *escribir derecho con renglones torcidos*, en el sentido de aplicarse no ya para encontrar términos parecidos a uno dado, sino para encontrarlos no sabiendo su correcta ortografía, pudiendo soslayar posibles errores gráficos.

- **Máscara fija:** se recuperan todos los artículos en los que se encuentra una palabra de longitud igual a la de búsqueda y que difiere en ciertos caracteres no especificados. Por ejemplo: t?m?r - {tumor, temor, tomar...}

- **Máscara flexible:** igual que la anterior, salvo que las posiciones no especificadas pueden ser ocupadas por cualquier número de caracteres, incluido ninguno. Por ejemplo: t*m*r - {tumor, temor, tomar, tambor, temblor, terminar...}

• **Truncamientos:** permiten recuperar artículos que contienen palabras que empiezan, terminan o presentan alguna coincidencia central con la secuencia de caracteres explicitada en la búsqueda. Por ejemplo: tosl - {tos, tose, tosa, tostada...}, ltipo - {subtipo, fenotipo, genotipo...}.

Los tipos de búsqueda enumerados hasta aquí pueden centrarse tanto sobre la entrada en el diccionario como sobre el cuerpo del artículo; los dos siguientes son más adecuados para localizar ocurrencias en las definiciones.

• **Cercanía y antecendencia:** se caracterizan por limitar el número de palabras —radio de búsqueda— que pueden aparecer entre dos especificadas; con una restricción adicional en la segunda para establecer el orden de aparición en que deben hallarse. Por ejemplo, en: “*aerobio*— Organismo que precisa de *oxígeno* en su normal *metabolismo*”, *oxígeno* aparece antes y a una distancia de cuatro palabras de *metabolismo*.

• **Frases:** tiene como objetivo la localización de artículos en los que se encuentra una frase especificada. Por ejemplo, si buscamos *moliendo el pimiento* podemos hallar “*pimentón*.— Polvo que se obtiene *moliendo el pimiento* colorado después de seco...”

Es posible componer expresiones de búsqueda mediante operadores lógicos —Y, O, NO, (...)— en lo que se podría denominar una *búsqueda compleja*. Por ejemplo, la expresión: *enfermedad Y sangre Y hepatl*, localiza “*ictericia*— *Enfermedad* que consiste en ciertos trastornos *hepáticos* que producen aumento de bilis en la *sangre*...”; o mediante: *manchal Y (ojo O córnea)* se puede obtener “*albugo*— Nombre aplicado a ciertas *manchas* blancas de la *córnea* debidas a gránulos de grasa depositados en ella...”

Por otro lado, no menos útiles resultarán las técnicas derivadas de la aplicación de los avances obtenidos en el procesamiento del lenguaje natural, cuyos algoritmos y métodos pueden tanto facilitar la consulta como mejorar la respuesta.

Disponer de un reconocedor y generador morfológico posibilita la realización de búsquedas con independencia de la flexión o derivación que afecta a la palabra; de esta manera se pueden llevar a cabo localizaciones mucho más complejas y cercanas al lenguaje que los usuarios tienden a utilizar de manera espontánea. Con estas herramientas se puede acceder a un diccionario ya no sólo a través de la palabra introducida, sino extendiéndose a cualquier forma flexionada o derivada de ella; no cabe duda que de esta manera se enriquecen y facilitan las búsquedas. En el ejemplo de *pimentón*, en vez de “moliendo el pimiento”, podríamos no encontrar una respuesta, ya que no

se halla en esa forma en el diccionario, pero si disponemos de un operador morfológico que permita relajar la consulta hasta otras palabras morfológicamente relacionadas, entonces hallaríamos las ocurrencias de “moliendo el pimiento”, aun no habiendo formulado exactamente esa consulta. También es de utilidad cuando se pretende localizar entradas en un diccionario y la forma en la que se hace la petición está afecta-

da de flexión o derivación: acceder a *andar* con *anduve*, a *comer* con *cómetelo*, a *fuerte* con *fortísimas*, a *joven* con *jovencitos*, a *pronto* con *prontito*... o por prefijación: acceder a *encaminar* con *reencaminar*, a *mundo* con *submundo*... En la línea del procesamiento morfológico se puede avanzar hasta considerar relaciones de derivación o de composición más amplias, incluso las que conlleven cambio de categoría gramatical; así por ejemplo, de una búsqueda con expansión morfológica de *jugada* y *fullería* conjuntamente en un mismo significado se podría obtener “*donillero*.— *Fullero* que agasaja y convida a aquellos a quienes quiere inducir a *jugar*” y “*tahúr*.— *Jugador fullero*”,



gracias a que *jugador* y *jugar* son formas relacionadas con *jugada* y *fullero* con *fullería*.

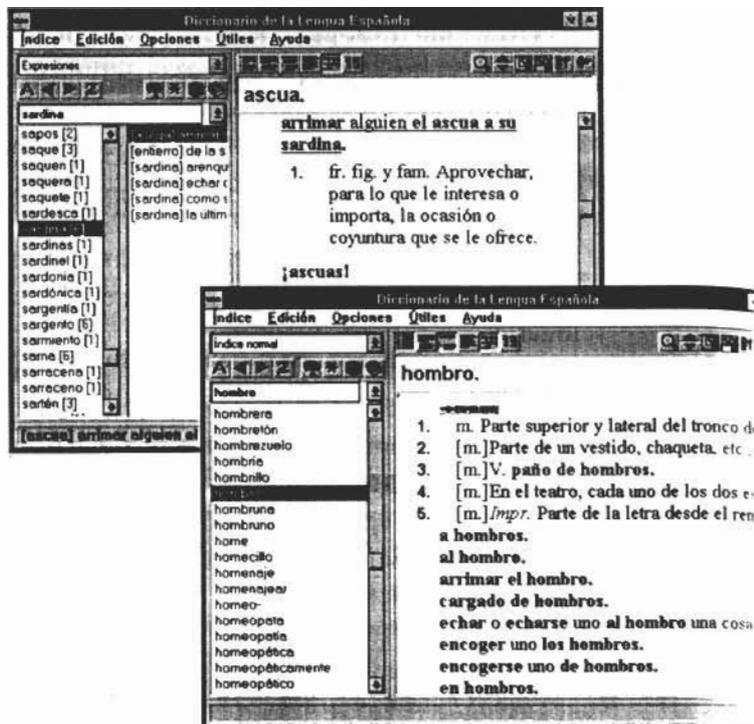
Mas allá de la búsqueda apoyada en el tratamiento morfológico de las palabras se encuentra la posibilidad de expandir el proceso a un nivel superior: la búsqueda semántica. Obviamente, las palabras obtenidas por el generador morfológico están semánticamente relacionadas, ya que son formas flexionadas o derivadas de la misma palabra; conviene habilitar correspondencias entre palabras morfológicamente distantes como las que se establecen en un diccionario de sinónimos —relaciona palabras por significados— o en un diccionario ideológico —conecta términos que tienen alguna relación conceptual—. Por ejemplo se podría obtener *tahúr* de *jugador tramposo* a través de la relación semántica entre *tramposo* y *fullero*.

Una adecuada combinación de todas estas posibilidades de búsqueda proporcionaría una extraordinaria pontencialidad para la localización de información en diccionarios.

DICCIONARIOS EN CD-ROM

Ya se encuentran disponibles algunos diccionarios en CD-ROM que incorporan determinadas técnicas avanzadas de acceso a su información. Expondremos a continuación dos de los diccionarios de definiciones más conocidos.

La edición electrónica del *Diccionario de la Lengua Española de la Real Academia* admite búsquedas alfabéticas directas o inversas en la lista de entradas —incluye las formas masculina y femenina—. Permite acceder a las formas complejas; por ejemplo, entre las que contienen la palabra *sastre* se encuentran “buen hombre, pero mal sastre” bajo la entrada **hombre**, “jabón de sastre” bajo la entrada **jabón** y “buen sastre” o “corto sastre” bajo la entrada **sastre** entre otras. Mediante el índice de todas las palabras se pueden localizar términos en los cuerpos pudiendo matizar la búsqueda según definiciones, etimologías y apariencia con la que aparecen escritas las palabras en el cuerpo; éstas se pueden combinar mediante operadores lógicos para llevar a cabo la denominada búsqueda múltiple. Así, por ejemplo, existen 3.897 entradas en las que aparece la palabra *parte* y 15 en las que figura *invariable*, la búsqueda múltiple *parte Y invariable* arroja las entradas: **adverbio**, **conjunción**, **gerundio** y **preposición**. No permite otros tipos de búsquedas tales como



ALGUNAS PANTALLAS DEL DICCIONARIO DE LA LENGUA ESPAÑOLA, EDICIÓN EN CD-ROM, MADRID, ESPASA CALPE, 1996

frases, cercanía...; todas las búsquedas se basan en el concepto de búsqueda exacta y la incorporación de tratamiento morfológico es nula. Resulta una aportación interesante la clasificación de las entradas, a través de las marcas, mediante distintos tipos de árboles: lenguas, categoría gramatical, usos...

La Editorial Gredos ha publicado en CD-ROM el *Diccionario de María Moliner*, que posee cierto tratamiento morfológico además de distintos tipos de búsqueda. Se puede acceder a la lista de entradas alfabéticamente —despreciando las tildes—, situándose en la forma canónica y en caso de no encontrarla en la que le sigue alfabéticamente. La denominada búsqueda en las entradas, a partir de una petición, ofrece la lista de formas canónicas a partir de las cuales puede obtenerse la palabra solicitada mediante un tratamiento morfológico. Las alteraciones morfológicas tenidas en cuenta son cambio de género y de número, conjugación verbal y grado superlativo en los adjetivos; por ejemplo, ante la consulta *amigas* responde como posibles entradas **amigar**, **amigarse** y **amigo**. En caso de no obtener respuesta se intenta ofrecer una lista de sugerencias por similitud de caracteres; por ejemplo, ante *amigor* sugiere **amigo**, **amigar** y **amagar**. La conjugación incluye tratamiento de pronombres enclíticos, por ejemplo ante *comértela* responde **comer** y **comerse**, aunque en ocasiones se comporta de forma excesivamente permisiva; por ejemplo,

responde —no como sugerencia— **comerse** ante *comersese* y **comer** ante *comedos* o *comeos*. El grado superlativo, a pesar de ser una incorporación interesante, adolece de algunos defectos; por ejemplo, no reconoce *riquísimo* ni *gagúísimo* y devuelve **rico** ante *ricísimo* y **gago** con *gagísimo*. Adicionalmente permite máscaras fijas y flexibles mediante el uso de caracteres comodín, por ejemplo con **scom?r* responde **descomer** y **escomar**. La búsqueda en las definiciones permite localizar las entradas que contienen la expresión de búsqueda; las expresiones pueden ser palabras, frases o combinaciones, mediante operadores lógicos, de palabras que pueden llevar comodines. Un amplio conjunto de filtros posibilita la aplicación de transformaciones morfológicas, la definición de la proximidad entre las palabras, la distinción entre mayúsculas y minúsculas y la definición de los bloques de búsqueda.

UN PROCESADOR MORFOLÓGICO COMO UTILIDAD DE UN DICCIONARIO

El Grupo de Investigación en Estructuras de Datos de la Universidad de Las Palmas de Gran Canaria ha desarrollado un procesador morfológico. Las flexiones consideradas son género y número para sustantivos y adjetivos, grado superlativo para los adjetivos y conjugación para los verbos con inclusión de pronombres enclíticos. Como derivación de sustantivos y adjetivos se consideran los sufijos apreciativos —aumentativos, diminutivos y despectivos—. El procesador faculta la obtención de formas derivadas que conllevan cambio de categoría gramatical—sustantivación, adjetivación, verbalización y adverbializa-



CARÁTULA Y PANTALLA DE LA EDICIÓN EN CD-ROM DEL DICCIONARIO DE MARÍA MOLINER, MADRID, GREDOS, 1996

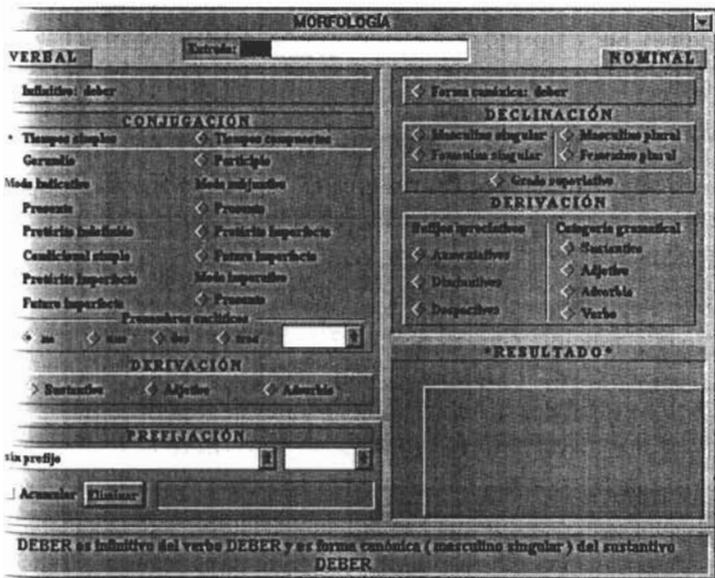
ción—; incluye un módulo que permite tratar palabras con prefijos. Una *interfaz* amigable permite formas de exploración imaginativas y versátiles.

A partir del reconocimiento de la palabra deduce su forma canónica y dirige el trabajo hacia la parte verbal, nominal o ambas. El diseño en dos grandes bloques *VERBAL* y *NOMINAL* permite: a) visualizar todas las opciones morfológicas simultáneamente de forma simple y clara; y b) usar en paralelo la flexión verbal y la nominal para aquellas palabras susceptibles de admitir las dos formas como *deber* o *ser*.

Del infinitivo reconocido, las acciones correspondientes a la morfología verbal permiten: a) conjugar los diferentes tiempos simples o compuestos para cada uno de los modos; b) incorporar enclíticamente uno, dos o tres pronombres que pueden elegirse de entre las posibilidades válidas para cada caso, y c) presentar sustantivaciones, adjetivaciones o adverbializaciones del verbo en proceso.

En la parte *NOMINAL* la selección de un género y número permite elegir derivaciones mediante sufijos apreciativos —*Aumentativos*, *Diminutivos* y *Despectivos*—, y si la entrada es adjetivo, su grado superlativo. Para algunos adverbios permite el superlativo y el diminutivo. La derivación por cambio de *categoría gramatical* de las entradas sustantivas proporciona sus adjetivaciones, de los adjeti-

PANTALLA DEL PROCESADOR MORFOLÓGICO DESARROLLADO POR EL GRUPO DE INVESTIGACIÓN EN ESTRUCTURAS DE DATOS DE LA UNIVERSIDAD LAS PALMAS DE GRAN CANARIA



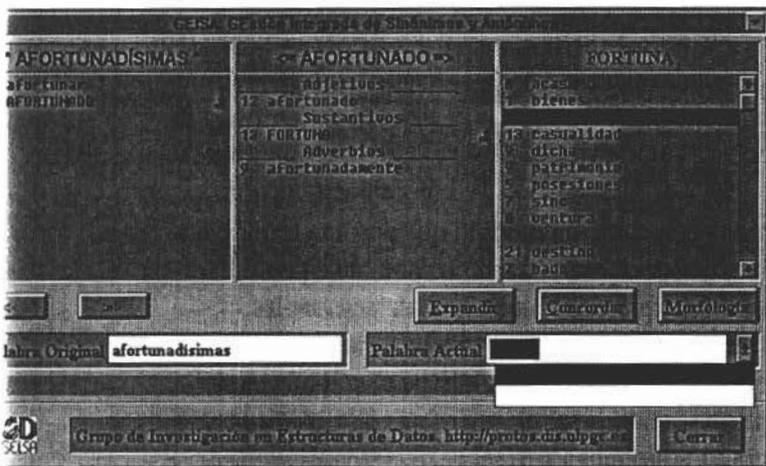
vos sus sustantivaciones y adverbializaciones y de los adverbios sus adjetivaciones; en cualquier caso, ofrece los verbos relacionados morfológicamente.

EL PROCESADOR MORFOLÓGICO APLICADO A LA GESTIÓN DE UN DICCIONARIO DE SINÓNIMOS Y ANTÓNIMOS

El objetivo perseguido fue obtener una versión informatizada con las siguientes características: a) uso fácil y amigable; b) acceso tanto gracias a sus entradas exactas como a partir de sus formas flexionadas o derivadas; c) flexibilidad en la exploración merced a navegaciones alternativas basadas en palabras relacionadas morfológicamente; d) recu-

La primera ventana presenta el resultado del reconocimiento bajo una cabecera donde muestra *Palabra Original* —inscrita entre asteriscos para distinguirla del resto de las ventanas—. Mantiene un vínculo entre cada una de las respuestas del reconocedor morfológico con *Palabra Original*: al recorrer las diferentes opciones aparecen en el área de mensajes sus categorías y la relación morfológica de flexión o derivación con *Palabra Original*.

Seleccionar una forma canónica en la ventana inicial conlleva: a) fijar la categoría y flexión de *Palabra Original*, y b) usarla como entrada al diccionario de sinónimos para comenzar la navegación; implica además: c) copiarla en *Palabra Actual*, y d) habilitar los botones *Concordar* y *Morfología*. Al resaltar una palabra en cualquier otra ventana aporta su categoría gramatical como información general —de particular interés para la concordancia—. Dado que el proceso puede extenderse más allá de las tres ventanas que están visibles simultáneamente habilita los botones << o >> para visualizar las listas que han quedado fuera.



OTRO EJEMPLO DEL PROCESADOR MORFOLÓGICO DESARROLLADO POR EL GRUPO DE INVESTIGACIÓN EN ESTRUCTURAS DE DATOS DE LA UNIVERSIDAD LAS PALMAS DE GRAN CANARIA

peración simple de la historia de una navegación apoyada en hasta tres entradas simultáneamente visibles; e) obtención de la respuesta morfológicamente concordante con la palabra de partida, y f) facultad para dar salida a cualquier inconsistencia en la concordancia mediante un abundamiento en la flexión y derivación de la respuesta haciendo uso del procesador morfológico.

La forma de funcionamiento de la *interfaz* es la siguiente:

Al escribir en *Palabra Original* se inicia el proceso —activa automáticamente el botón *Iniciar*. Actuar sobre *Iniciar* implica realizar el análisis morfológico de *Palabra Original* —en caso de que no sea reconocida aparece un mensaje indicándolo—; admite así no sólo formas canónicas o infinitivos que sean cabeceras en el diccionario, sino cualquiera de sus formas flexionadas o derivadas.

LOS DICCIONARIOS I

La próxima generación de diccionarios sobre soporte electrónico va a suponer toda una revolución comercial y de uso. Los diccionarios son herramientas para trabajar con palabras, y lo esperable es que acaben unidos a los grandes programas de procesamiento de textos, dentro de una panoplia de herramientas destinadas a ayudar al usuario, como correctores de estilo, programas de búsqueda de información, resúmenes automáticos de documentos y demás. La pérdida de importancia de los soportes locales (disco duro, CD-ROM) a favor de las conexiones por línea va a provocar una situación nueva: el usuario obtendrá de la red solamente la información que necesite, y por supuesto sólo pagará la que use.

Podemos suponer la revolución editorial que supondrá no vender *a priori* toda la obra (como es ahora el caso), sino sólo una licencia de consulta. En vez de depender de la distribución en librerías, el editor dependerá probablemente de acuerdos con los fabricantes de procesadores de textos, que,

Actuar sobre el botón *Concordar* permite flexionar o derivar *Palabra Actual* en concordancia con *Palabra Original*: aplica a *Palabra Actual* la misma flexión o derivación que representa *Palabra Original* respecto a la opción elegida en la ventana producto de su reconocimiento. Ya que tal operación puede dar lugar a más de una posibilidad, presenta una lista con las formas concordantes; al elegir una, la coloca en *Palabra Actual* para su transferencia al generador morfológico o a otras aplicaciones.

Si existe alguna dificultad o se produce una violación gramatical y no es factible conseguir la concordancia, muestra dos tipos de mensaje: a) *Atención*, indica cómo ha procedido —normalmente por ajuste de la flexión a la categoría gramatical o por cambio de género— y b) *Fallo*, señala el error y aconseja usar el procesador morfológico — botón *Morfología*— para dar con su forma flexionada, derivada o relacionada que se considere más idónea —faculta una mayor riqueza de matices en la forma final.

Siempre que exista una palabra resaltada en cualquiera de las ventanas permanece activado el botón *Expandir*; su acción despliega un abanico de entradas del diccionario relacionadas morfológicamente que sugieren novedosas posibilidades de navegación. Permite mayor flexibilidad en las opciones de exploración que el diccionario impreso; pues mientras la solución manual sólo admite ojear las palabras del entorno alfabéticamente cercanas, esta propuesta ofrece una mejor perspectiva por su carácter morfológico: las relaciones sufijales eliminan opciones molestas alfabéticamente próximas y las prefijales proponen alternativas interesantes que aparecerían en páginas muy distantes.

INTERACCIÓN USUARIO-DICCIONARIO

Con lo expuesto anteriormente es indudable que la interacción entre los diccionarios y sus usuarios se verá radicalmente alterada respecto a su situación actual; por ejemplo, cuando se desea utilizar en un documento la información —el significado o la palabra— encontrada en un diccionario de papel tendrá que copiarse a mano; un diccionario en soporte informático ofrece la posibilidad de realizar esa operación de forma automática. La transformación tecnológica redundará en una mayor comodidad, flexibilidad y facilidad de uso; tales características se verán favorecidas por el alto grado de conectividad con otras herramientas y tipos de datos. Es importante señalar la enorme potencialidad de aplicaciones que integren a varios diccionarios de propósitos diferentes proporcionándose apoyo mutuo en las búsquedas.

En un sistema de almacenamiento de información impresa lo que prima es la interrogación; su secuencia de interacción consiste en que el usuario proporciona el término a buscar y obtiene una respuesta. Una noción fundamental en la interacción con los nuevos diccionarios consiste en la navegación; es corriente en Internet, donde se ha acuñado la denominación de *internautas* para los usuarios que buscan información a través de sus múltiples enlaces. Cuando se accede a una determinada información, como resultado de una interrogación directa o procediendo desde otra, se encuentran enlaces que conectan con otros elementos de información relacionados que el usuario puede alcanzar navegando a través de ellos. Sobre un diccionario en soporte

L SIGLO XXI

JOSÉ ANTONIO MILLÁN

facilitarán la *interfaz* de consulta con determinados diccionarios, y quizá no con otros. Por último, cabe también la posibilidad de que los grandes productores de programas textuales desarrollen sus propios bancos de datos lexicográficos, o bien compren su uso a los editores tradicionales (y ambas cosas ya están ocurriendo). Ello significaría la pérdida de una importante cuota de mercado para los editores en papel, aunque no necesariamente un peor servicio para el usuario final.

Pero uno de los aspectos más apasionantes del futuro inmediato es la relación que establecerá el usuario con el diccionario. Por supuesto, la consulta de una palabra será instantánea a partir de cualquiera de sus formas: desde *condujiste* se tendrá acceso a la entrada *conducir* (ahora muchos extranjeros y no pocos españoles se preguntan por qué en el diccionario no viene *conduje*). Los homógrafos y las distintas acepciones serán aclaradas sobre la marcha por programas de análisis del contexto: consultar *planta* desde un artículo sobre vegetales nos llevará a la acepción botánica, y en un contexto de edificación a su acepción arquitectónica. *(Pasa a la página siguiente)*

informático, este tipo de acción resulta cómoda por su mecanización y potente porque puede además dirigirse a otros diccionarios. Claro que también esta forma de actuar funciona en los diccionarios de papel

(viene de la página anterior) Las consultas se sinónimos o antónimos no arrojarán una lista ciega entre la que el lector ha de escoger, sino que también responderán contextualmente: “Te *condujiste* con corrección” llevará a “Te *comportaste* con corrección”, mientras que “*Condujiste* con temeridad” llevará a : “*Guiaste* con temeridad” (por supuesto, el programa diccionario ajustará las concordancias de la frase para cualquier sustitución o paráfrasis).

Nuevas *interfaces* ayudarán en la toma de decisiones, haciendo uso de todos los recursos necesarios: indicadores, colores... Podemos imaginar a alguien que necesita un equivalente para *molestar*: al pedirlo surgirá una “barra de aceptabilidad” dotada de un cursor. Éste aparecerá situado en un punto intermedio (indicando que molestar pertenece a un registro neutro). Si se mueve el cursor hacia la parte de arriba, rotulada como FORMAL, aparecerá *importunar*. Si se desplaza hacia abajo, hacia VULGAR, irá apareciendo *fastidiar *chinchar (de color rosa, para indicar que es familiar) *jorobar *hacer la cusqui (en estas dos últimas el color rosa se hace más fuerte, para indicar que son vulgares) *joder (que, en rojo, nos indicaría su inaceptabilidad).

Las ayudas para buscar un término apropiado harán uso del conocimiento activo del análisis componencial, las solidaridades léxicas y fraseologías. Para localizar lo que necesita, el usuario actuará con naturalidad: podrá utilizar la paráfrasis o el término general que tenga más a mano, o bien dejar indicado su deseo de completar una expresión. Podrá dejar escrito “hacer entrar” (porque no atina con la palabra exacta) y obtendrá *introducir*. Si escribe “Crece el número de delitos que se *hacen* en las grandes ciudades”, y pide ayuda se le ofrecerá una lista: **cometen* **perpetran*. Podrá escribir: “Era la primera vez que “ba una conferencia” (porque no sabe qué se hace con las conferencias, si se *dicen* o *cuentan* o qué exactamente...), y el programa le sustituirá el comodín: *pronunciaba*. La información de régimen preposicional estará igualmente disponible, y escribir “pienso de que” despertará alguna alarma de programa.

Pero no pensemos en una presencia intromisoria y regañona: la *interfaz* será lo suficientemente ergonómica para que el programa de apoyo tenga la presencia que desee el usuario, ya sea la de un censor constante que

cuando en la definición de una palabra aparece alguna otra que nos es desconocida; entonces podemos buscar esta última para entender la definición de la primera —siempre que el diccionario sea autocon-

mira por encima del hombro, o la de un consejero agazapado a la espera de que se le llame.

Habrán rutinas que funcionen constantemente en el transfondo, que examinen cada palabra y cada combinación de palabras y que además tengan una visión global del texto (por ejemplo, calibrando el registro o nivel global de un escrito para poder señalar una intromisión formal en un contexto coloquial, o viceversa). Pero todo ello, insisto, de forma tan patente o tan subterránea como desee el usuario.

Muchas cosas cambiarán en la forma y en el contenido de los diccionarios. Habrá presentaciones gráficas e interactivas de campos semánticos, que darán acceso a palabras, giros y expresiones, y que utilizarán descendientes avanzados de las técnicas hipertextuales para permitir al usuario afinar conceptos y deslindar matices. Pienso en el nuevo tratamiento de las metáforas que se abre en contextos electrónicos, y me imagino el campo semántico de la lucha desplegado, y cómo podremos acotar sus usos metafóricos. Al elegir LA DISCUSIÓN ES UNA LUCHA se destacarán “*atacar*”, “*defenderse*”, un “argumento *acorazado*”, “*no abandonar las posiciones*”, “*atrincherarse* en las opiniones de uno” ... Al escoger EL AMOR ES UNA LUCHA aparecerían las “*armas* de mujer” “*asedio* amoroso”, la “*resistencia*”, el “*rendirse* a los encantos de alguien” ... Y de esta forma nuestra comprensión de la lengua, y las posibilidades de usarla, se desplegarán de forma asombrosa.

Como es evidente, no sólo los hablantes nativos se beneficiarán de estas ayudas: quienes aprenden una lengua extranjera tendrán al alcance de la mano una información preciosa que activará de forma explícita muchos mecanismos interiores de los hablantes.

¿Cuándo veremos todo esto? Aunque estos diccionarios futuros suponen un reto informático y de diseño, sobre todo tienen que partir de un conocimiento de nuestra lengua que hoy estamos lejos de poseer. Dicho de otra manera: si (según recordaba hace poco José Antonio Pascual) nos faltan los equivalentes españoles de obras como *The BBI Combinatory Dictionary of English*, *el Collins Cobuild Dictionary* o *el Longman Language Activator*, ¿cómo vamos a emprender sus equivalente electrónicos? Mientras las instituciones (públicas o privadas) que deberían crear estas obras no cumplan con su deber, este futuro portentoso no comenzará en nuestra lengua.

nido y no incluya en sus definiciones palabras que luego no figuren definidas, a veces ocurre.

Los soportes informáticos de almacenamiento de información tienen la capacidad de incluir datos multimedia —otro capítulo que está cada vez más en boga—; de tal manera que cualquier documento comprenderá no sólo texto, sino además sonidos, gráficos, imágenes y secuencias de animación. Un diccionario español-inglés presentará con la palabra inglesa correspondiente a la española su transcripción fonética y será posible oír su pronunciación ideal e incluso visionar una imagen del objeto al que identifica. La definición de la palabra *caballo* no se limitará a “mamífero del orden de los perisodáctilos, solípedo, de cuello y cola poblados de cerdas largas y...”, sino que podrá incluir una secuencia de 20 ó 30 segundos que muestre, por ejemplo, una manada de caballos pastando. El universo sensorial en el que se transformarán los textos hará que la consulta de diccionarios deje de ser una tarea poco atractiva y restringida a especialistas o estudiosos para convertirse en una actividad fascinante.

Con la incorporación de los avances tecnológicos a todos los hogares, el futuro se abre como una puerta esperanzadora hacia un mundo en el que la información sea manipulada con facilidad y agrado por una población cada vez más amplia, donde las distancias tanto geográficas como idiomáticas y culturales pierdan el carácter de frontera y el conocimiento se universalice en beneficio de la humanidad y de la Tierra en general.

REFERENCIAS

Diccionario de la Lengua Española. Edición electrónica. Versión 21.1.0. Real Academia Española de la Lengua y Espasa Calpe S.A. 1995.

Diccionario de María Moliner. Edición en CD-ROM. Editorial Gredos. 1996.

Pérez Aguiar, José R.: *Reconocimiento y Generación Integrada de la Morfología del Español. Una aplicación a la gestión de un diccionario de sinónimos y antónimos*. Tesis Doctoral. Universidad de Las Palmas de Gran Canaria. 1996.



REALE

Revista de estudios de adquisición
de la lengua española

DIRECTOR

Pedro Benítez Pérez

Suscripciones

Suscripción anual: España 1.800 ptas.
(2 números) Europa 2.000 ptas.
Resto del mundo 2.000 ptas.
Precio del nº suelto: 1.500 ptas

Dirección

Secretaría de REALE
Departamento de Filología
Universidad e Alcalá
C/Colegios, 2
28801 Alcalá de Henares (Madrid)
Teléfono: 91-885 44 23 Fax: 91-885 44 13

UNIVERSIDAD DE ALCALÁ

