

GEISA: Un diccionario de sinónimos en formato electrónico.

*Santana, O.; Pérez, J.; Carreras, F.;
Santos, S.; Rodríguez, G.; Hernández, Z.
Grupo de Investigación en Estructuras de Datos
Departamento de Informática y Sistemas
Universidad de Las Palmas de Gran Canaria
<http://protos.dis.ulpgc.es>*

0. INTRODUCCIÓN

Con la popularización y el uso generalizado de los ordenadores cada vez se hace más patente la necesidad de facilitar la comunicación entre las personas y las máquinas; paralelamente, la lingüística ha experimentado un desarrollo importante en los últimos años y ha interesado a la informática hasta constituir una disciplina que se ha dado en llamar lingüística computacional. Es menester resaltar que de lo que trata este encuentro no es precisamente de acercar el lenguaje humano al de la máquina, sino más bien de todo lo contrario, lograr que el ajeno y frío lenguaje del ordenador y su *modus operandi* se aproximen cada vez más a las necesidades de la persona. Bajo este enfoque evolucionan las técnicas de procesamiento de lenguaje natural y la tendencia a que la interfaz hombre-máquina sea cada vez más agradable: es el reto de las industrias de la lengua.

En esa línea de facilitar la tarea humana figuran:

a) el desarrollo de herramientas de ayuda para la generación de textos escritos (diccionarios de significados, de sinonimia, ideológicos, tesauros,...)

b) el análisis de textos para la detección de vicios y depuración de estilos (frecuencias de aparición de palabras, empleo de formas verbales y no verbales, aliteraciones,...)

c) la mecanización de la corrección ortográfica, etcétera.

Los objetivos perseguidos son:

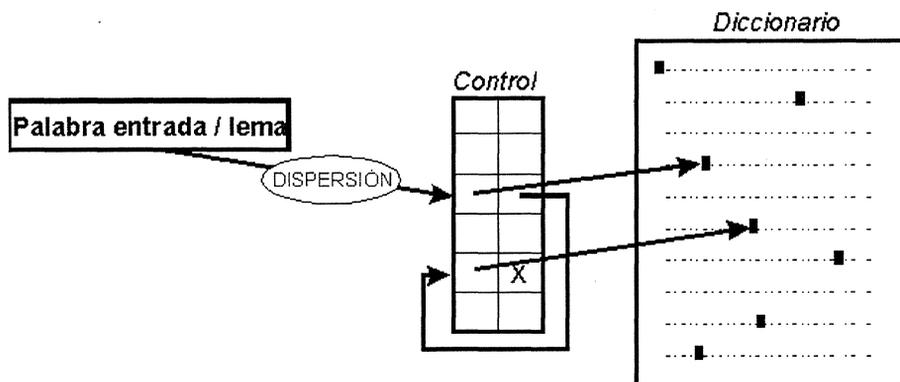
- 1.- Situar en un sistema informático un diccionario de sinónimos y antónimos conforme con un equilibrio adecuado entre ocupación y tiempo de respuesta.
- 2.- Considerar como entradas las voces consignadas en el diccionario que tomamos como base y cualquiera de sus formas flexionadas o derivadas.
- 3.- Navegación clásica que adelanta la lista de sinónimos-antónimos correspondiente a un elemento de la anterior.
- 4.- Proporcionar formas alternativas de navegación por medio de palabras morfológicamente relacionadas.
- 5.- Ofrecer el sinónimo o antónimo elegido con sus accidentes gramaticales en concordancia con la palabra de entrada.
- 6.- Interacción con el diccionario en un entorno cómodo.

1.- EL DICCIONARIO DE SINÓNIMOS Y ANTÓNIMOS EN UN SISTEMA INFORMÁTICO

Para comprobar la eficacia del sistema se toma el *Gran Diccionario de Sinónimos y Antónimos* de Espasa-Calpe. Tal diccionario dispone de unas 35.000 entradas: cada una posee una lista de sinónimos y antónimos —incluye frases— con un tamaño promedio de 10 elementos por lista. El universo de palabras de este diccionario está compuesto de unos 20.000 sustantivos, 10.000 adjetivos, 6.000 verbos, 400 adverbios de modo y otras 400 formas invariantes tales como preposiciones, conjunciones, interjecciones, locuciones o frases y palabras de otros idiomas.

Como el manejo de un diccionario de sinónimos no requiere una extrema velocidad de acceso a sus listas, queda almacenado en memoria externa, lo cual facilita su consulta a usuarios que no disponen de ordenadores muy potentes. El acceso a los artículos se lleva a cabo mediante una función de dispersión por división con tratamiento para convertir su carácter alfabético en valor numérico; las colisiones se resuelven por cadenas separadas. Cada registro contiene una entrada y su lista de sinónimos, antónimos y frases sinónimas y antónimas. Se permite la redundancia que implica el que cada registro contenga los sinónimos en forma explícita ya

que no es significativa la reducción de almacenamiento externo que se conseguiría sustituyéndolos por un código numérico o empleando algoritmos de compresión.



La solución llevada a cabo usa dos ficheros externos, uno de *Control* (384Kb) y el propio *Diccionario* (2,18Mb). El disponer de *Control* —registros de tamaño fijo— que permite direccionar a *Diccionario* —registros de tamaño variable— donde se almacenan las listas, implica un ahorro de espacio debido a que evita la existencia de huecos en el fichero. Se necesitan hacer tantos accesos a *Control* como colisiones haya y uno a *Diccionario* por cada una de sus resoluciones.

A costa de hipotecar memoria principal, una opción que disminuiría los requerimientos de almacenamiento secundario y el número de accesos consistiría en disponer de *Control* en memoria interna ya que el proceso trataría con memoria secundaria una sola vez en lugar de dos por colisión. Dado el bajo número de colisiones (1,3 por entrada), no se ha optado por este arreglo a pesar de que con un costo de memoria interna no demasiado alto podría mejorar la velocidad.

Es posible reducir la ocupación del *Diccionario* reemplazando las palabras de las listas de sinónimos por direcciones a posiciones donde se alojarían todas ellas sin repetición alguna; en este caso, *Diccionario* ocuparía en torno a 1,5Mb y el vector de palabras menos de 0,5Mb. Esta alternativa tendría sentido si se localizaran en memoria principal *Control*, *Diccionario* y vector de palabras (aproximadamente 2,3Mb), en tanto que los accesos a memoria secundaria serían nulos y se alcanzaría una velocidad de navegación mucho mayor. No se ha elegido esta solución para la aplicación porque el aumento de velocidad no justificaría la memoria necesaria.

2.- REDUCCIÓN MORFOLÓGICA PARA ACCEDER A LAS ENTRADAS DEL DICCIONARIO

A fin de permitir el acceso mediante formas flexionadas o derivadas de las voces que hay en el diccionario, se realiza un análisis morfológico de la palabra entrada del que resultan las formas canónicas de las que proviene y su relación de flexión o derivación con cada una de ellas. Para una mejor realización, el sistema que soporta la morfología se orienta más a datos que a reglas gracias a una estructuración adecuada que recoge las características morfológicas del español y cierta información sintáctica.

Debido a que la cardinalidad del universo compuesto por las formas flexionadas y derivadas representaría algo más de sesenta veces el número de formas canónicas, este trabajo aprovecha la morfología del español como lengua flexiva para la división de las palabras en raíces y terminaciones y obtiene una reducción considerable de datos: la cantidad de raíces no duplica el número de formas canónicas y las terminaciones necesarias son apenas unos cientos. La eficacia del proceso de reconocimiento se mantiene alta.

El análisis de una palabra se realiza a través de: a) segmentación en posibles pares raíz-terminación¹, b) localización de cada una de las raíces, c) decodificación de la terminación y d) construcción de la forma principal de la que proviene. A partir de la palabra de entrada se forman sus posibles pares raíz-terminación por medio de un catálogo de terminaciones adecuado a las flexiones y derivaciones consideradas. Para cada una de las raíces que resultan, se lleva a cabo su búsqueda, mediante una función de dispersión, en la estructura que organiza por raíces el universo de palabras. Tal estructura posee información de qué terminaciones se pueden concatenar con cada una de las raíces y qué flexión o derivación representa cada una de ellas; las características flexivas del español permiten representar sintéticamente los diferentes grupos de terminaciones. Por estar cada raíz vinculada a la de la forma canónica de la que proviene, el analizador puede reconstruir las voces que hay en el diccionario, con sus respectivas categorías gramaticales y las flexiones que las relacionan con la palabra original.

Las flexiones y derivaciones consideradas para los sustantivos son las de:

género,

pedagogo ⇒ **pedagogo, pedagoga**

autor ⇒ **autor, autora**

alcahuete ⇒ **alcahuete, alcahueta**

¹ Los términos raíz y terminación se usan en sentido estrictamente gráfico y no exactamente gramatical.

número

pedagogo ⇔ **pedagogos, pedagogas**

y sufijación apreciativa:

aumentativos,

casa ⇔ **casón, caserón**², **casarón**² (masculino singular)³

↳ **casaza, casona, casota, casacha** (femenino singular)

↳ **casones, caserones**², **casarones**² (masculino plural)³

↳ **casazas, casonas, casotas, casachas** (femenino plural)

diminutivos

casa ⇔ **casita, casilla, casina, casica, casuela, casuca**² (femenino singular)

↳ **casitas, casillas, casinas, casicas, casuelas, casucas**² (fem. plural)

y peyorativos.

casa ⇔ **casucho, caserón**² (masculino singular)³

↳ **caseja, casucha, casuca**² (femenino singular)

↳ **casuchos, caserones**² (masculino plural)³

↳ **casejas, casuchas, casucas**² (femenino plural)

Para los adjetivos se considera:

el género,

bueno ⇔ **bueno, buena**

abrasador ⇔ **abrasador, abrasadora**

grandote ⇔ **grandote, grandota**

el número,

bueno ⇔ **buenos, buenas**

² Además de los sufijos discretivos (-**azo**, -**ón**, -**ote** y -**acho** para aumentativos; -**ito**, -**illo**, -**ico**, -**ín** y -**uelo** para diminutivos; -**ucho** y -**ejo** para peyorativos) se tienen en cuenta otros en casos particulares.

³ Un sustantivo puede cambiar su género con la apreciación.

☒ la sufijación apreciativa:

☒ aumentativos,

bueno ⇔ **buenazo, bonazo**⁴, **buenón, buenote, buenacho, bonachón**
(masculino singular)

↳ **buenaza, bonaza**⁴, **buenona, buenota, buenacha, bonachona**⁵
(femenino singular)

↳ **buenazos, bonazos**⁴, **buenones, buenotes, buenachos, bonachones**⁵
(masculino plural)

↳ **buenazas, bonazas**⁴, **buenonas, buenotas, buenachas, bonachonas**⁵
(femenino plural)

☒ diminutivos

bueno ⇔ **buenecito, buenecillo, buenecico, buenecín, buenezuelo, bonito**⁴,
bonillo⁴, **bonico**⁴ (masculino singular)

↳ **buenecita, buenecilla, buenecica, buenecina, buenezuela, bonita**⁴,
bonilla⁴, **bonica**⁴ (femenino singular)

↳ **buenecitos, buenecillos, buenecicos, buenecines, buenezuelos,**
bonitos⁴, **bonillos**⁴, **bonicos**⁴ (masculino plural)

↳ **buenecitas, buenecillas, buenecicas, buenecinas, buenezuelas,**
bonitas⁴, **bonillas**⁴, **bonicas**⁴ (femenino plural)

☒ y peyorativos;

bueno ⇔ **buenejo, buenucho** (masculino singular)

↳ **bueneja, buenucha** (femenino singular)

↳ **buenejos, buenuchos** (masculino plural)

↳ **buenejas, buenuchas** (femenino plural)

☒ el grado superlativo (con sus irregularidades),

bueno ⇔ **buenísimo, bonísimo**⁶, **óptimo**⁶ (masculino singular)

↳ **buenísima, bonísima**⁶, **óptima**⁶ (femenino singular)

↳ **buenísimos, bonísimos**⁶, **óptimos**⁶ (masculino plural)

↳ **buenísimas, bonísimas**⁶, **óptimas**⁶ (femenino plural)

⁴ Como consecuencia de las operaciones de flexión y derivación pueden aparecer alteraciones en las raíces e incluso raíces completamente diferentes, al menos desde el punto de vista gráfico y formal.

⁵ Puede combinarse una raíz alterada con una terminación atípica.

⁶ En el grado superlativo pueden aparecer alteraciones en las raíces e incluso raíces completamente diferentes desde el punto de vista sincrónico, como en el caso de las irregularidades: **pobre** ⇔ **paupérrimo**.

- la adverbialización

bueno ⇨ **buenamente**

- y la adverbialización del superlativo.

bueno ⇨ **buenísimamente, bonísimamente** ⁶ **óptimamente** ⁶

Para los adverbios que lo admiten se consideran sus formas:

- superlativas,

cerca ⇨ **cerquísima** ⁷

lejos ⇨ **lejísimos**

- y apreciativas:

- aumentativas,

arriba ⇨ **arribota**

lejos ⇨ **lejotes**

- diminutivas

ahora ⇨ **ahorita, ahoritita** ⁸, **ahoritica** ⁸, **horitica** ⁹

arriba ⇨ **arribita**

lejos ⇨ **lejitos, lejillos, lejuelos**

- y peyorativas.

arriba ⇨ **arribota**

Para los verbos se considera la conjugación.

Al lema verbal **callar** se accede con sus formas conjugadas:

⁷ **Cerca** como sustantivo femenino tiene los apreciativos correspondientes y su función adverbial es compatible con la nominal como en el caso del adverbio **afuera** con el sustantivo **afueras**.

⁸ Puede aparecer cualquier sufijo.

⁹ Admite modificaciones en la raíz (usado en Colombia según el diccionario de dudas y dificultades de la lengua española de M. Seco).

modo indicativo				
presente	pretérito	condicional indefinido	pretérito imperfecto	futuro imperfecto
<i>callo callas ...</i>	<i>callé callaste ...</i>	<i>callaría callarías ...</i>	<i>callaba callabas ...</i>	<i>callaré callarás ...</i>

modo subjuntivo			modo imperativo
presente imperfecto	pretérito imperfecto	futuro	presente
<i>calles calles ...</i>	<i>callara o callase callaras o callases ...</i>	<i>callare callares ...</i>	<i>calla ...</i>

Se supera a la palabra como elemento de entrada y se identifica como conjugación de un verbo la pareja que representa su tiempo compuesto.

De esta forma conducen al lema verbal **callar** sus tiempos compuestos

modo indicativo				
pretérito perfecto	pretérito anterior	condicional perfecto	pretérito pluscuap.	futuro perfecto
<i>he callado has callado ...</i>	<i>hubo callado hubiste callado ...</i>	<i>habría callado habrías callado ...</i>	<i>había callado habías callado ...</i>	<i>habré callado ...</i>

modo subjuntivo		
pretérito perfecto	pretérito pluscuap.	futuro perfecto
<i>haya callado haya callado ...</i>	<i>hubiera o hubiese callado hubieras o hubieses callado ...</i>	<i>hubiere callado hubieres callado ...</i>

Las formas no personales también son admitidas tanto en sus formas simples como compuestas:

formas simples	formas compuestas
infinitivo ⇒ <i>callar</i> gerundio ⇒ <i>callando</i> participio ⇒ <i>callado</i>	infinitivo ⇒ <i>haber callado</i> gerundio ⇒ <i>habiendo callado</i> participio ⇒ <i>habido callado</i>

En los escasos casos que procede se reconoce el **diminutivo** del **gerundio** ⇒ **callandito**

Y el participio admite su flexión como adjetivo verbal:

masculino singular ⇒ *callado*

masculino plural ⇒ *callados*

femenino singular ⇒ *callada*

femenino plural ⇒ *calladas*

grado superlativo ⇒ **calladísimo** (masculino singular)
 ⇒ **calladísima** (femenino singular)
 ⇒ **calladísimos** (masculino plural)
 ⇒ **calladísimas** (femenino plural)

adverbialización ⇒ **calladamente**

adverbialización del superlativo ⇒ **calladísimamente**

En la conjugación de los verbos se incluye la posibilidad de hasta tres pronombres enclíticos tanto en las formas simples como en las compuestas; así se accede al lema *callar* desde: **callándose**¹⁰, **callate**¹⁰, **callaos**¹¹, **callémonos**¹¹; **húboselo callado...**

Se consideran las irregularidades propias de la conjugación verbal, sean formales o gráficas:

a) que afectan a la raíz

— vocálicas

concibo ⇒ **concebir** (cambiar *e* por *i*)

cuelgo ⇒ **colgar** (cambiar *o* por *ue*)

— consonánticas

conozco ⇒ **conocer** (añadir una *z*)

hago ⇒ **hacer** (cambiar *c* por *g*)

— mixtas

sepa ⇒ **saber** (cambiar *ab* por *ep*)

caigo ⇒ **caer** (añadir *ig*)

¹⁰ Se producen alteraciones en la forma gráfica, pues han de llevar la tilde como consecuencia de la incorporación de los pronombres enclíticos.

¹¹ Aparecen alteraciones que afectan al último carácter del verbo.

b) que afectan al tema

di ⇒ **decir**

haz ⇒ **hacer**

doy ⇒ **dar**

voy ⇒ **ir**

c) otras

— supletivismo

fueron, era ⇒ **ser**

— participios y gerundios irregulares

escrito ⇒ **escribir**

durmiendo ⇒ **dormir**

d) se consideran también los cambios ortográficos

sigo ⇒ **seguir** (cambiar **gu** por **g** delante de **e** o **i**)

delinca ⇒ **delinquir** (cambiar **qu** por **c** delante de **a** u **o**)

quepo ⇒ **caber** (cambiar **c** por **qu** delante de **e**)

leyendo ⇒ **leer** (cambiar el carácter que representa el fonema)

La entrada puede llevar prefijos: así, de **submundos** o de cualquier otra flexión o derivación se accede a **mundo** y **preprocesadas** da lugar a **procesado** y **procesar**. Puede incorporarse más de un prefijo: **subprecondición** conduce a **condición**.

Son consideradas las reglas gramaticales para la unión de los prefijos a las formas.

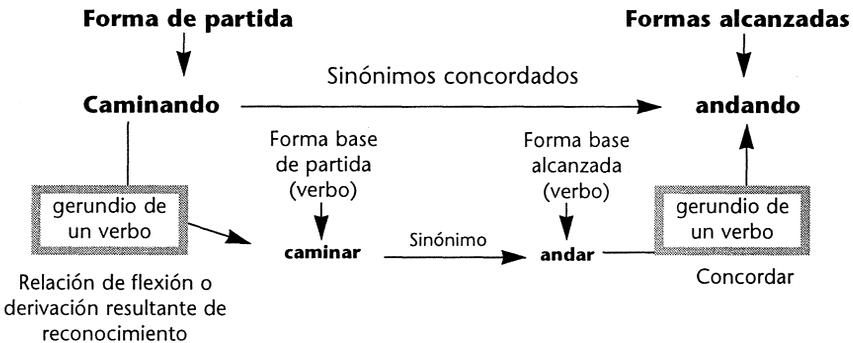
Palabra de entrada	Se reconoce como		
	Prefijo	Forma canónica	Relación Gramatical
contrarrevolucionaria	contra	revolucionario	femenino singular
transiberianos	trans	siberiano	masculino plural
irrealidades	in	realidad	femenino plural
combenediciaditos	con	beneficiado	diminutivo masculino plural

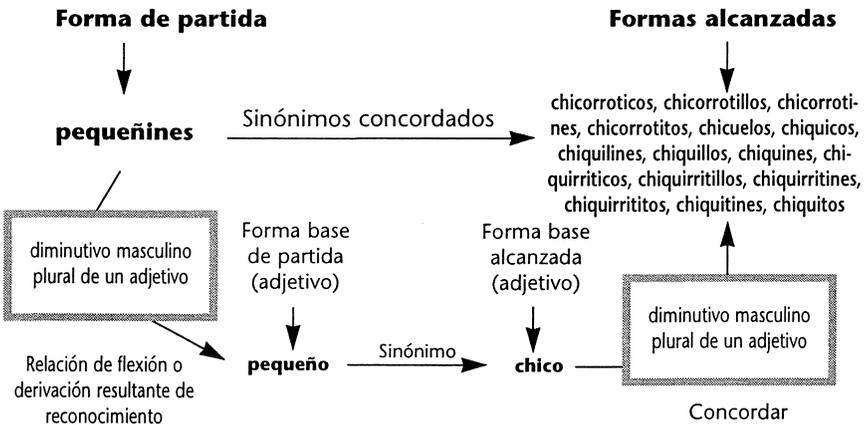
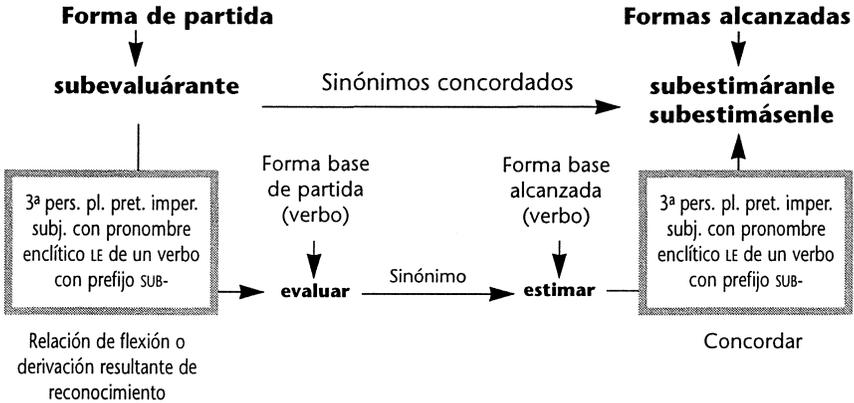
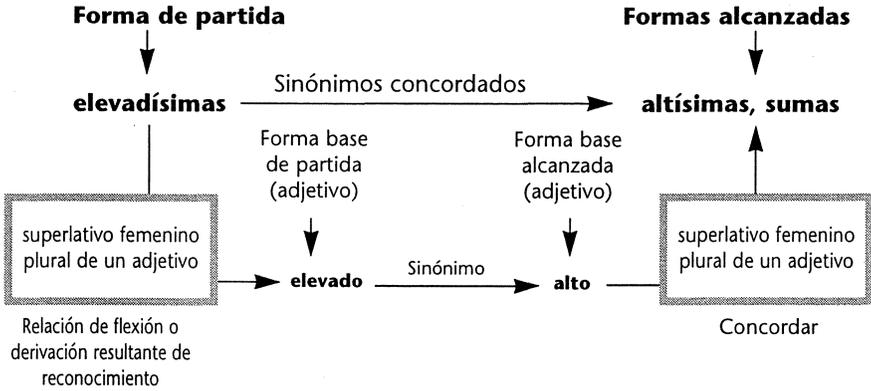
No se flexionan ni derivan, como es lógico, preposiciones, conjunciones, interjecciones, locuciones o frases, ni palabras de otros idiomas.

3. FLEXIÓN O DERIVACIÓN DEL SINÓNIMO O ANTÓNIMO ELEGIDO

Al objeto de obtener el sinónimo o antónimo elegido afectado por los mismos accidentes gramaticales que interesan a la forma introducida, se hace uso de un generador morfológico. La propia estructura de datos que usa el análisis morfológico concede la generación de flexiones y derivaciones de las formas canónicas. El procedimiento consiste en: a) detectar la raíz de la forma canónica y b) buscar la terminación que corresponda a la forma que se desea generar. Se extrae la raíz gramatical de la forma canónica y se busca en la estructura que organiza por raíces el universo de palabras mediante una función de dispersión; para obtener la flexión o derivación buscada se concatena la raíz con la terminación adecuada. A veces es necesario recorrer raíces alternativas que representan alteraciones tipográficas en la raíz producidas al flexionar o derivar como consecuencia de reglas fonológicas, ortográficas o de cualquier irregularidad.

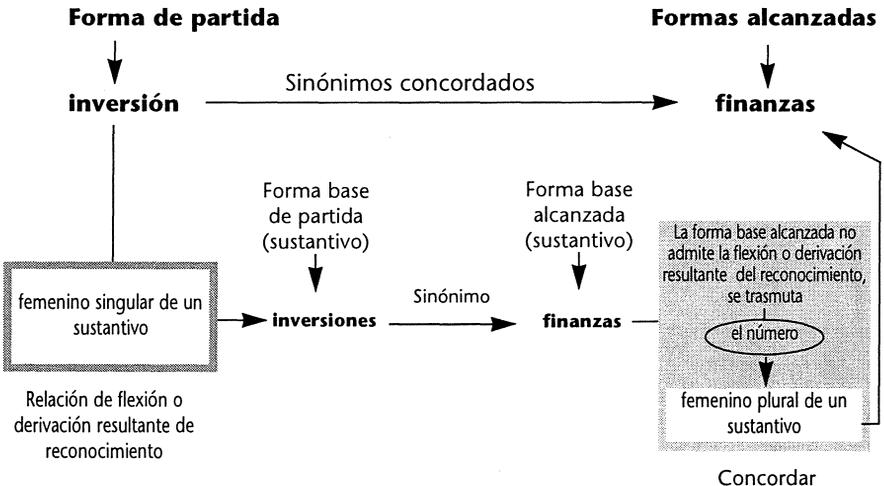
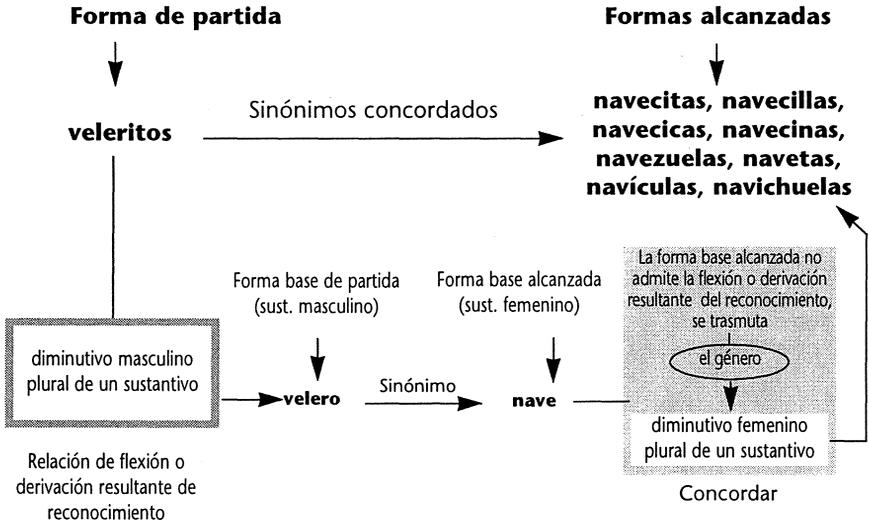
Como resultado del análisis morfológico se obtiene para cada forma canónica su relación de flexión o derivación con la palabra original; ello permite aplicar al sinónimo-antónimo elegido la misma flexión o derivación que relaciona la palabra original con la forma canónica donde comenzó la navegación, lo que en ocasiones da lugar a más de una palabra.



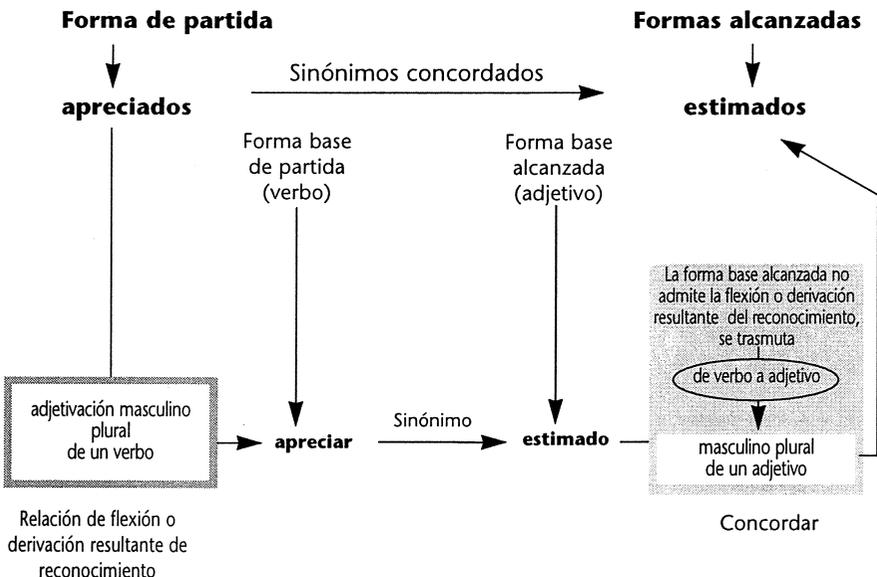
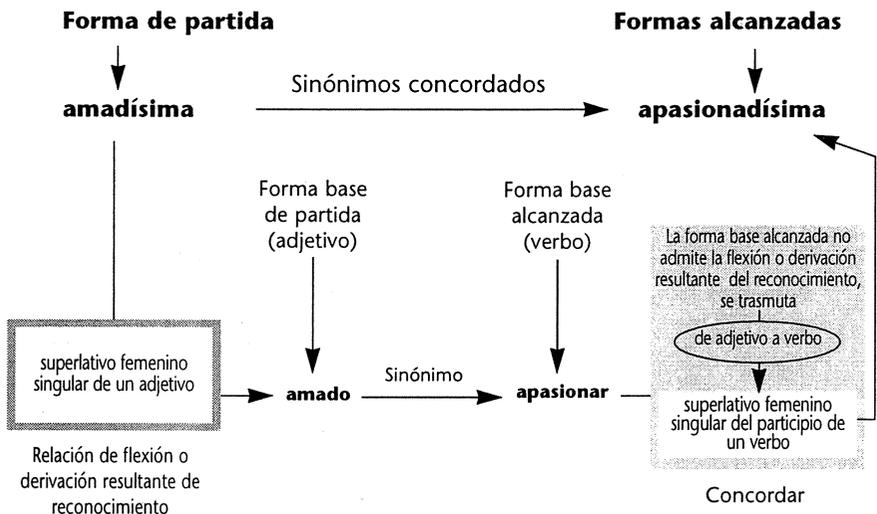


Para poder concordar formas con diferente categoría gramatical se dispone de procedimientos que proporcionan flexiones y derivaciones alternativas:

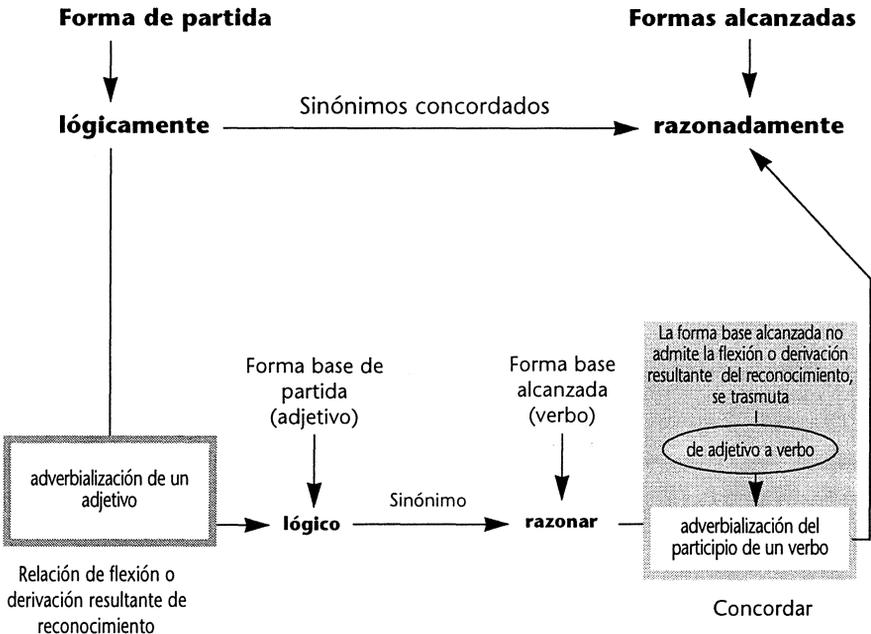
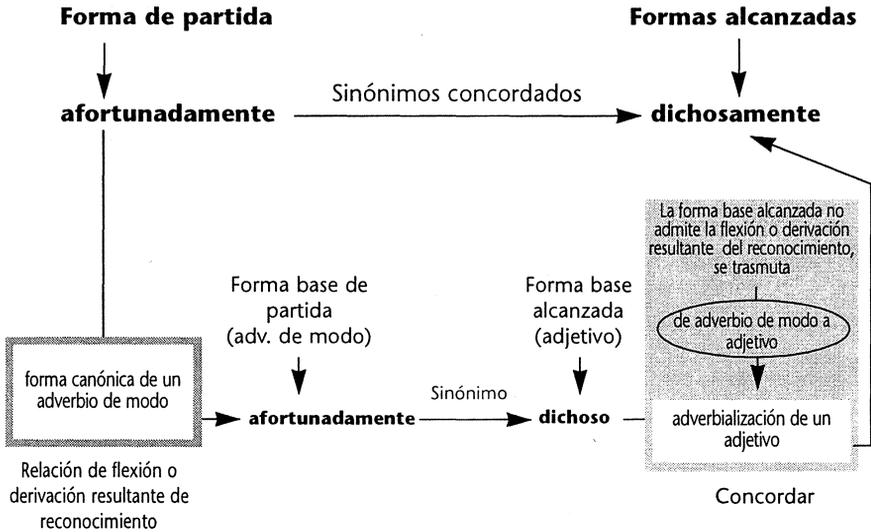
- a) al generar cualquier forma flexionada o derivada puede ser necesario un cambio de género —es frecuente en sustantivos— o de número;

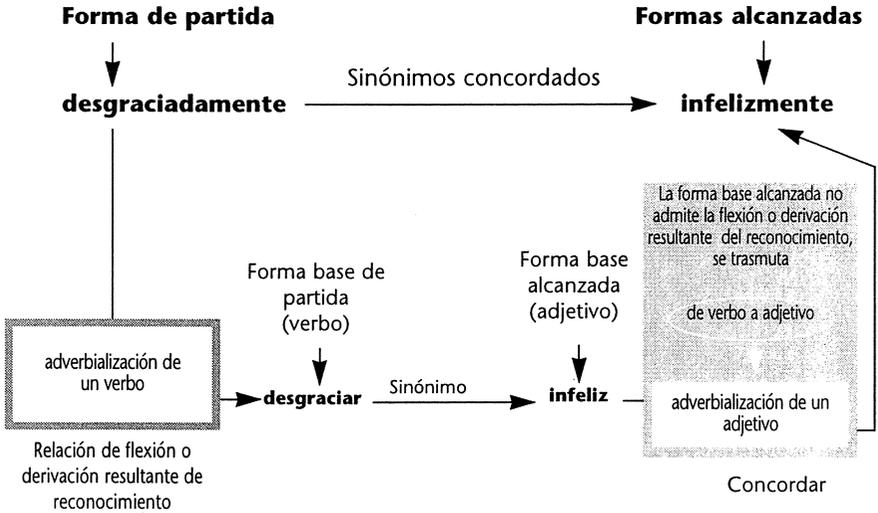


b) toda declinación de un adjetivo se concuerda con un verbo a través de su participio y viceversa;

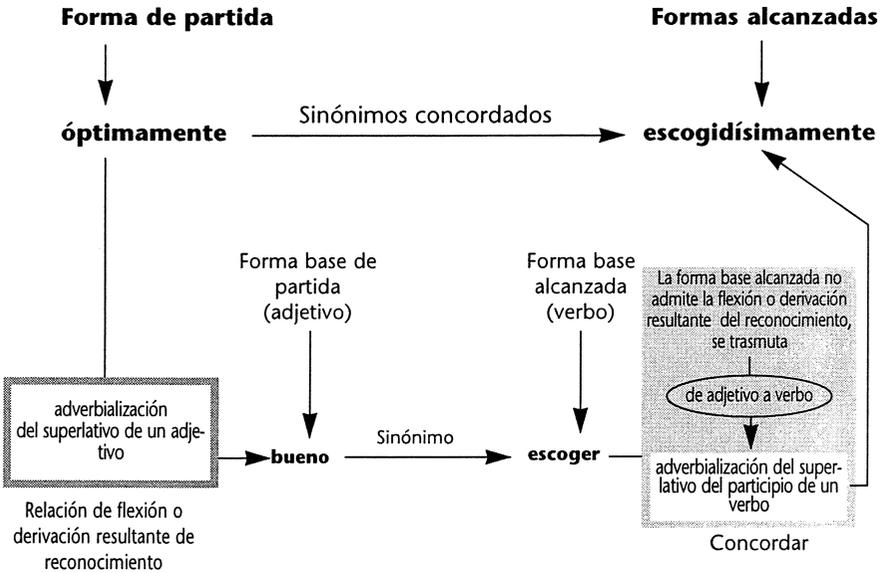


- c) los adverbios de modo, las adverbializaciones de los adjetivos y la adverbialización del participio verbal son alternativas entre sí;





d) la adverbialización del superlativo de un adjetivo lo hace con la adverbialización del superlativo del participio de un verbo.



4. NAVEGACIÓN A TRAVÉS DE PALABRAS MORFOLÓGICAMENTE RELACIONADAS

Se dispone de relaciones morfológicas entre diferentes formas canónicas que facultan la obtención de un conjunto de entradas del diccionario. Permite mayores posibilidades de acceso a las opciones de exploración que la que ofrece el diccionario impreso; mientras la solución manual sólo admite ojear palabras alfabéticamente cercanas, esta propuesta sugiere novedosas posibilidades para recorrerlo: las relaciones sufijales eliminan cabeceras molestas alfabéticamente próximas, y las prefijales proponen alternativas interesantes que aparecen en páginas distantes.

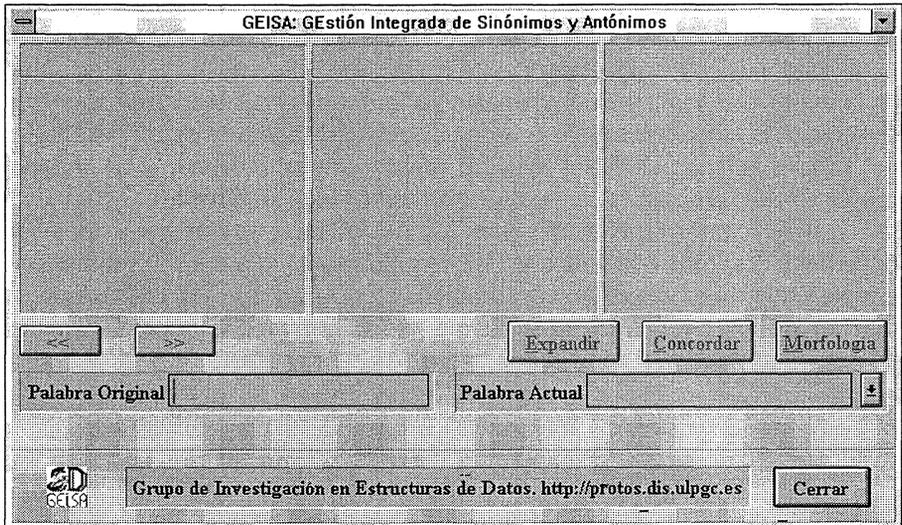
Las relaciones establecidas para los verbos son: sustantivos y adjetivos deverbales y verbos relacionados por cambio prefijal. Para los sustantivos se tienen los verbos derivados. Los adjetivos se enlazan con los verbos que se derivan de ellos y con el adverbio de modo correspondiente. Y para los adverbios de modo se tiene el adjetivo del que provienen.

Cabecera	<i>Lista de alternativas morfológicas</i>
<i>cariño</i> <i>emisión</i>	cariñoso, encariñado, cariñito, cariñosamente, encariñar emisor, emisora, transmisor, emisión, transmisión, retransmisión, emitir, transmitir, retransmitir
<i>hombre</i>	hombruna, hombruno, hombrada, hombracho, hombrillo, hombría, hombrón, hombrar
<i>matar</i>	matador, rematado, matachín, matadero, matadura, matanza, matón, remate, matonismo, rematadamente, rematar
<i>saber</i>	sabedor, sabido, sabio, sapiente, resabido, sabiduría, sapiencia, resabio, resabiar
<i>total</i> <i>ver</i>	totalidad, totalmente, totalizar vidente, visible, visto, vistoso, previsto, visionario, revisable, revisor, supervisor, visión, videncia, vista, vistazo, viso, visor, visual, previsión, visa, visado, visualización, revisión, televisión, supervisión, visiblemente, entrever, rever, prever, visar, revisar, divisar, televisar, supervisar

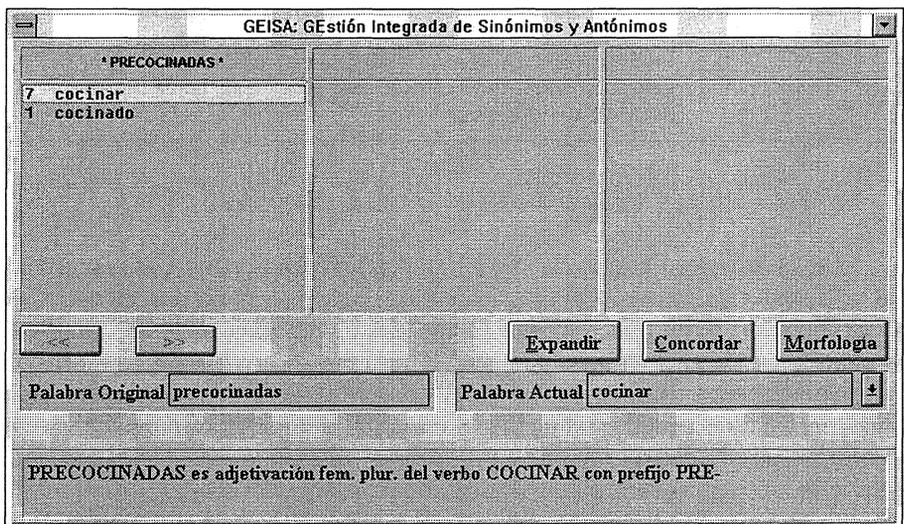
5.- INTERACCIÓN CON EL DICCIONARIO EN UN ENTORNO CÓMODO.

La versión informatizada del diccionario de sinónimos y antónimos presenta las siguientes características: a) uso fácil y claro, b) acceso tanto por 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) recuperación 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 respuesta a cualquier inconsistencia en la concordancia mediante un abundamiento en la flexión y derivación.

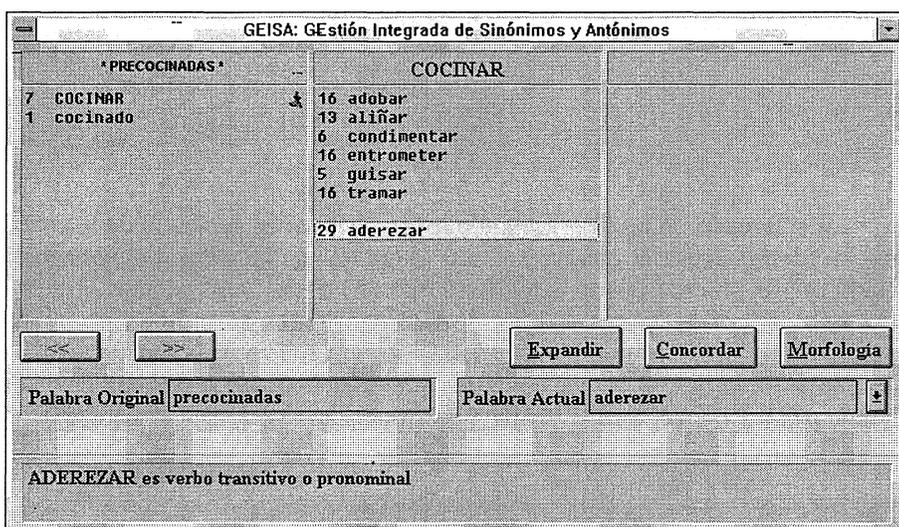
A continuación se expone la forma de funcionamiento de la interfaz utilizada.



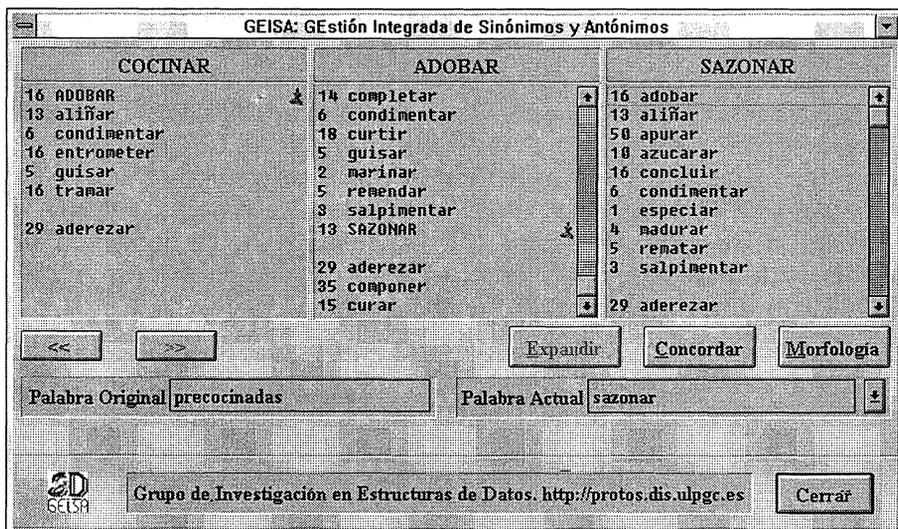
En esta ventana aparecen inhabilitados todos los botones salvo *Cerrar*: la aplicación los habilita en cuanto tenga sentido su uso. A medida que se escribe en *Palabra Original* se realiza simultáneamente el análisis morfológico, que admite no sólo las entradas en el diccionario, sino cualquiera de sus formas flexionadas o derivadas.



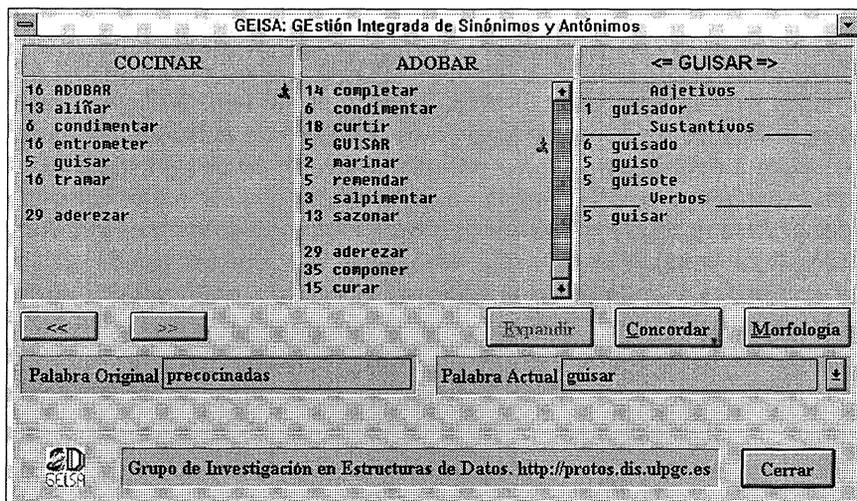
La primera ventana presenta el resultado del reconocimiento bajo un título donde muestra *Palabra Original*, inscrita entre asteriscos para distinguirla del resto de las ventanas. Mantiene un vínculo entre cada una de las respuestas del reconocedor morfológico con *Palabra Original*: al recorrer las 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 conlleva: a) fijar la categoría gramatical y la 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 *Expandir*, *Concordar* y *Morfología*.



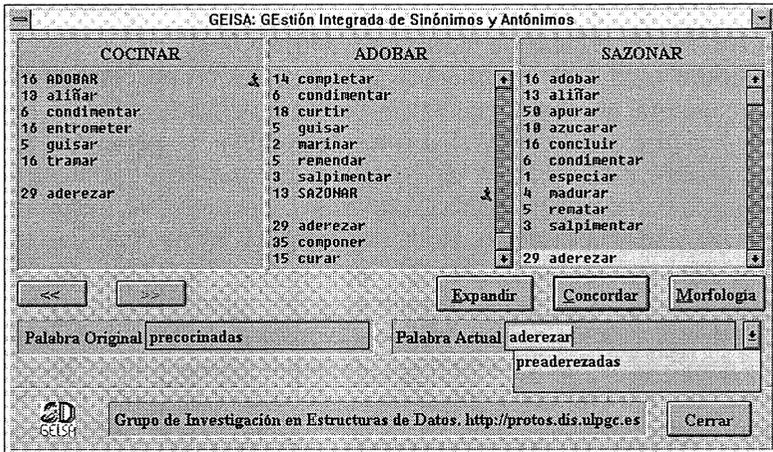
A la izquierda de cada palabra, aparece el número de elementos que posee su lista de sinónimos y antónimos al objeto de anticipar información que ayude a decidir por dónde continuar la investigación. Un *doble clic* o la *tecla de entrada* hace progresar la navegación por el diccionario: consiste en exponer en la siguiente ventana la lista de sinónimos y antónimos que corresponde. Para destacar el recorrido, la palabra seleccionada queda señalada con una marca y aparece en el título de la ventana que contiene su lista. Resaltar una palabra en cualquier ventana que no sea la inicial conlleva que sea expuesta su categoría gramatical como información general —de particular interés para la concordancia— además de situarla en *Palabra Actual*.



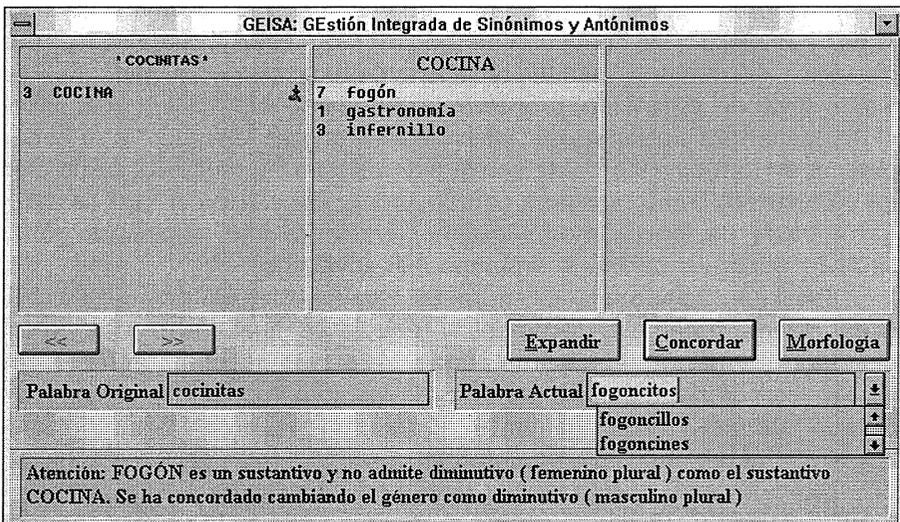
Si el tamaño de las listas de sinónimos y antónimos supera al de la ventana, la interfaz activa el correspondiente desplazamiento vertical u horizontal. Dado que el proceso puede extenderse más allá de las tres ventanas que permanecen simultáneamente visibles, quedan habilitados los botones << o >> para visualizar las listas que han quedado fuera. La marca de la tercera ventana apunta al título que se oculta por la derecha y el título de la primera coincide con la palabra marcada en la ventana que ha desaparecido por la izquierda.



El botón Expandir permanece activado siempre que exista una palabra resaltada en cualquiera de las ventanas y su acción despliega el abanico de entradas del diccionario relacionadas morfológicamente, clasificadas según su categoría gramatical.



Actuar sobre Concordar permite flexionar o derivar *Palabra Actual* en concordancia con *Palabra Original*: significa aplicar a *Palabra Actual* la misma flexión o derivación que representa *Palabra Original* respecto a la opción elegida en la ventana que muestra su reconocimiento. Ya que tal operación puede dar lugar a más de una posibilidad, presenta una lista con las formas concordantes; elegir una implica colocarla en *Palabra Actual* para su transferencia a otras aplicaciones.



Si existe alguna dificultad o se produce una violación gramatical y no es factible conseguir la concordancia, la interfaz juega con dos tipos de mensaje: a) *Atención*, indica cómo se ha producido la flexión o derivación y b) *Fallo*, señala el error y aconseja pasar a la ventana de *Morfología* para dar con una forma flexionada, derivada o relacionada que se considere idónea, lo que faculta una mayor riqueza de matices en la forma final.

El botón *Morfología* transfiere *Palabra Actual* a un procesador morfológico para generar libremente cualquier flexión o derivación de ella. Posibilita tanto la flexión y derivación verbal como la nominal. La parte verbal permite conjugación, incorporación de hasta tres pronombres enclíticos, flexión del participio como adjetivo verbal, diminutivo del gerundio, sustantivación, adjetivación y adverbialización. La sección nominal incluye para sustantivos y adjetivos género, número, sufijación apreciativa y verbalización; para los adjetivos, además se tienen en cuenta el grado superlativo, la adverbialización, la adverbialización del superlativo y la sustantivación, y para los sustantivos, la adjetivación; en los adverbios se consideran sus formas superlativas, aumentativas, diminutivas y peyorativas. Permite incorporar prefijos a las palabras. Evidentemente, como ya se dijo, no flexiona aquellas piezas que carecen de flexión como preposiciones, conjunciones, interjecciones, locuciones o frases, ni palabras de otros idiomas, pues está concebido para el español.

6. CONCLUSIONES Y PERSPECTIVAS

Se aporta una solución a la gestión integral de un diccionario de sinónimos y antónimos de notable dimensión: a) tanto en la entrada como en la obtención del resultado admite prácticamente todas las posibilidades flexivas y derivativas del español, b) proporciona abundante información gramatical, c) una de las principales aportaciones consiste en que la respuesta se puede obtener en forma concordante con la entrada para su inmediata sustitución en un texto, d) como novedosa contribución incluye la capacidad de navegar a través de formas morfológicamente relacionadas, sugiriendo pistas alternativas para encontrar el sinónimo más adecuado.

Se ha conseguido una aplicación en español diseñada para ser de utilidad a quienes tratan con documentos: lexicólogos, analistas de estilo, recuperadores de información textual, traductores automáticos,... Permite una interacción hombre-máquina intensa y muy sencilla, fundada en ventanas de diálogo, botones y demás herramientas que caracterizan a una interfaz gráfica. Necesita de pocos recursos informáticos.

Entre otras posibilidades, la actual propuesta evolucionará hacia la implantación sobre otros soportes como Unix o entorno de red.*

RESUMEN

Se desarrolla una aplicación para la gestión de un diccionario de sinónimos y antónimos con una interfaz clara que: a) admite como entradas los lemas del diccionario y cualquiera de sus formas flexionadas o derivadas, b) permite la navegación clásica en estos diccionarios, c) ofrece formas de exploración alternativas por medio de entradas relacionadas morfológicamente y d) considera la salida con los accidentes gramaticales que afectan a la entrada.

BIBLIOGRAFÍA

- AÏT-MOKHTAR, S.; RODRIGO MATEOS, J. L. (1995): "Segmentación y análisis morfológico de textos en español utilizando el sistema SMORPH", *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, 17, pp. 29-41.
- ALARCOS LLORACH, E. (1994): *Gramática de la Lengua Española*, Madrid, Espasa-Calpe.
- ALVAR EZQUERRA, M. (1993): *La formación de palabras en español*, Madrid, Arco/Libros.
- ARONOFF, M. (1976): *Word Formation in Generative Grammar*, Cambridge, M.I.T. Press.
- BEAR, J. (1986): "A Morphological Recogniser with Syntactic and Phonological Rules", *11th International Conference on Computational Linguistics (COLING-86)*, Bonn.
- BELLO, A. (1984): *Gramática de la Lengua Castellana*, Madrid, EDAF.
- CALDER, J. (1989): "Paradigmatic Morphology", *Proceedings of 4th Conference of European Chapter of ACL*, Manchester, pp. 58-65.
- ESPASA (1991): *Gran Diccionario de Sinónimos y Antónimos*, Madrid, Espasa-Calpe, 4ª ed.
- GAZDAR, G.: "Review article: Finite State Morphology", *Linguistics*, Berlín, 23, pp. 597-607.
- GILI GAYA, S. (1985): *Curso superior de sintaxis española*, VOX, Barcelona, Bibliograf.

* Queremos agradecer al profesor Dr. Manuel Alvar Ezquerro y al Dr. Juan Manuel García Platero del Departamento de Filología Española I de la Universidad de Málaga su colaboración en cuantas consultas le hemos formulado a lo largo del desarrollo del presente trabajo.

- GONZÁLEZ COLLAR, A. L.; GOÑI MENOYO, J. M.; GONZÁLEZ CRISTÓBAL, J. C. (1995): "Un Analizador Morfológico para el castellano basado en Chart", *Actas de la VI Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA'95)*, Alicante.
- HALLE, M. (1973): "Prolegomena to a Theory of Word Formation", *Linguistic Inquiry*, 4, pp. 3-16.
- KAY, M. (1973): "Nonconcatenative Finite-State Morphology", *Proceedings of the Third Conference of the European Chapter of the ACL*, Copenhagen, pp. 2-10.
- MARTÍ, M. A. (1986): "Un sistema de análisis morfológico por ordenador", *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, 4, pp. 104-110.
- MATTHEWS, P. H. (1980): *Morfología: Introducción a la Teoría de la Estructura de la Palabra*, Madrid, Paraninfo.
- MEYA, M. (1986): "Análisis morfológico como ayuda a la recuperación de información", *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, 4, pp. 91-103.
- MORENO SANDOVAL, A. (1992): *Un Modelo Computacional basado en la Unificación para el Análisis y la Generación de la Morfología del Español*, Universidad Autónoma de Madrid, Tesis Doctoral.
- REAL ACADEMIA ESPAÑOLA (1973): *Esbozo de una Nueva Gramática de la Lengua Española*, Madrid, Espasa-Calpe.
- RODRÍGUEZ, G.; HERNÁNDEZ, Z.; SANTANA, O. (1993): "Agrupaciones de Tiempos Verbales en un Texto", *Anales de las II Jornadas de Ingeniería de Sistemas Informáticos y de Computación*, Quito (Ecuador), pp. 132-137.
- RODRÍGUEZ MAGRO, C.; SOPEÑA, L.; VALLADARES, C.; VILLAR, C. (1990): "Clasificación morfológica del léxico castellano para un analizador en ordenador", *Actas del VII Congreso Nacional de Lingüística Aplicada*, Sevilla, pp. 491-503.
- RODRÍGUEZ MAGRO, C.; SOPEÑA, L.; VILLAR, C. (1990): "Confeción de un diccionario de sinónimos en ordenador: Teoría, metodología y resultados", *Lingüística Española Actual*, 12, pp. 81/101.
- SANTANA, O.; HERNÁNDEZ, Z. J.; RODRÍGUEZ, G. (1993): "Conjugaciones Verbales", *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, 13, pp. 443-450.
- SANTANA, O.; RODRÍGUEZ, J. C.; GONZÁLEZ, J. D. (1993): "FRECTEXT: Una Aplicación de Ayuda a la Elaboración de Documentos", *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*, 13, pp. 451-462.
- SCALISE, S. (1987): *Morfología Generativa*, Madrid, Alianza.
- SECO, M. (1991): *Diccionario de dudas y dificultades de la lengua española*, Madrid, Espasa-Calpe, 9ª ed.
- VOX (1990): *Diccionario Actual de la Lengua Española*, Barcelona, Bibliograf.