

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS



TESIS DOCTORAL

**RECONOCIMIENTO Y GENERACIÓN INTEGRADA DE LA
MORFOLOGÍA DEL ESPAÑOL : UNA APLICACIÓN A LA GESTIÓN
DE UN DICCIONARIO DE SINÓNIMOS Y ANTÓNIMOS**

JOSÉ RAFAEL PÉREZ AGUIAR

Las Palmas de Gran Canaria, Mayo de 1996

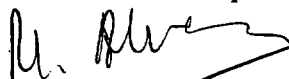
61/1995-96

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
UNIDAD DE TERCER CICLO Y POSTGRADO**

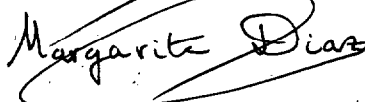
Reunido el día de la fecha, el Tribunal nombrado por el Excmo. Sr. Rector Magfco. de esta Universidad, el/a aspirante expuso esta **TESIS DOCTORAL**.

Terminada la lectura y contestadas por el/a Doctorando/a las objeciones formuladas por los señores miembros del Tribunal, éste calificó dicho trabajo con la nota de apto cum laude Las Palmas de Gran Canaria a 6 de julio de 1996.

El/a Presidente/a: Dr. D. Manuel Alvar Ezquerro,



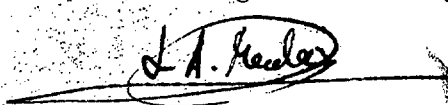
El/a Secretario/a: Dra. Dña. Margarita Díaz Roca,



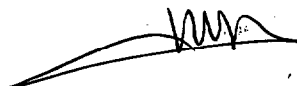
El/a Vocal: Dr. D. Antonio Núñez Ordóñez,



El/a Vocal: Dr. D. Juan Angel Méndez Rodríguez,



El/a Vocal: Dra. Dña. María Victoria Rodríguez Uría,



El/a Doctorando/a: D. José Rafael Pérez Aguiar,



BIBLIOTECA UNIVERSITARIA
LAS PALMAS DE G. CANARIA
Nº Documento.....405644
Nº Copia.....405651

**UNIVERSIDAD DE
LAS PALMAS DE GRAN CANARIA**

FACULTAD DE INFORMÁTICA



TESIS DOCTORAL

**Reconocimiento y generación integrada
de la morfología del español: Una
aplicación a la gestión de un diccionario
de sinónimos y antónimos**

Autor: D. José Rafael Pérez Aguiar

Director: Dr. D. Octavio Santana Suárez

Departamento de Informática y Sistemas

Mayo 1996

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

DOCTORADO EN INFORMÁTICA

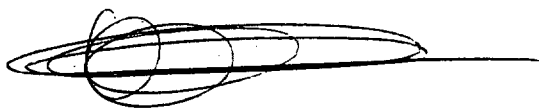
DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS

Programa de Percepción Artificial y Aplicaciones

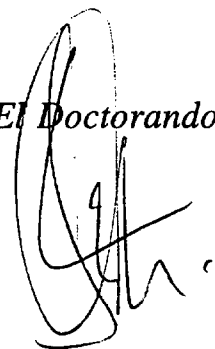
**RECONOCIMIENTO Y GENERACIÓN INTEGRADA
DE LA MORFOLOGÍA DEL ESPAÑOL: UNA
APLICACIÓN A LA GESTIÓN DE UN DICCIONARIO
DE SINÓNIMOS Y ANTÓNIMOS**

*Tesis Doctoral Presentada por D. José Rafael Pérez Aguiar
Dirigida por el Dr. D. Octavio Santana Suárez*

El Director,

A stylized, horizontal signature consisting of several overlapping loops and a long trailing line.

El Doctorando,

A stylized, vertical signature with a large loop at the top and a long trailing line.

Las Palmas de Gran Canaria. Mayo, 1996

A Elsa

RESUMEN

Esta tesis aporta una solución al problema del reconocimiento y la generación de la morfología del español; aprovecha la estructura flexiva de la lengua para orientar su realización a datos más que a reglas —mejora la velocidad. Un pormenorizado estudio de las características morfológicas da lugar a un preprocesador que actúa sobre un diccionario de formas básicas y genera el universo de palabras flexionadas y derivadas que se estructura en forma de raíces y terminaciones —minimiza la ocupación. Se construye un procesador que es el encargado de reconocer y generar.

El análisis de una palabra se lleva a cabo mediante: a) segmentación en posibles pares raíz-terminación, b) localización de cada una de las raíces, c) decodificación de la terminación y d) construcción de la forma principal de la que proviene; el resultado también aporta su categoría gramatical. La generación de una forma derivada o flexionada de una forma principal se logra gracias a: a) la detección de la raíz y b) la búsqueda de la terminación que corresponda a la forma que se desee conseguir.

Las flexiones consideradas son género y número para sustantivos y adjetivos, grado superlativo para los adjetivos y conjugación para los verbos que incluye pronombres enclíticos. Como derivación de sustantivos y adjetivos se consideran los sufijos apreciativos —aumentativos, diminutivos y despectivos. Se dispone además de un conjunto de relaciones morfológicas entre diferentes formas principales que faculta la obtención de formas derivadas que conllevan cambio de categoría gramatical —sustantivación, adjetivación, verbalización y adverbialización. El trabajo incluye un módulo de prefijación que permite incorporar prefijos a las palabras.

Se desarrolla una aplicación a la gestión de un diccionario de sinónimos y antónimos mediante una interfaz amigable que permite formas de exploración imaginativas y versátiles además de ofrecer la respuesta en forma concordante con la entrada.

AGRADECIMIENTOS

Quiero expresar mi profundo agradecimiento al profesor Dr. D. Octavio Santana Suárez, Director de esta Tesis y del Grupo de Investigación en Estructuras de Datos en el seno del cual se ha desarrollado este trabajo, por haber desempeñado un papel esencial en mi formación tanto científica y profesional como humana y personal; sin su tenaz esfuerzo, su certero consejo y su pródiga ayuda no habría sido posible llevar a término este trabajo.

A todos mis compañeros del Grupo de Investigación en Estructuras de Datos mi especial gratitud por su desinteresada e incondicional colaboración y su amable disponibilidad con las que he podido contar siempre.

ÍNDICE

INTRODUCCIÓN.	1
1.- LA MORFOLOGÍA DEL ESPAÑOL.	8
1.1.- DE LA FORMACIÓN DEL PLURAL.	10
1.2.- DE LA FORMACIÓN DEL FEMENINO.	14
1.3.- GRADO SUPERLATIVO DEL ADJETIVO.	18
1.4.- SUFIJOS: CONSIDERACIONES GENERALES Y REGLAS DE DERIVACIÓN.	19
1.5.- LA FORMACIÓN DE AUMENTATIVOS.	21
1.6.- LA FORMACIÓN DE DIMINUTIVOS.	22
1.7.- LA FORMACIÓN DE DESPECTIVOS.	25
1.8.- VERBOS: CONJUGACIÓN Y PRONOMBRES ENCLÍTICOS.	27
1.8.1.- Morfología del verbo.	27
1.8.2.- Grupos de tiempos irregulares.	30
1.8.3.- Modelos de conjugación.	30
1.8.4.- Pronombres enclíticos.	31
1.8.4.1.- El acento ortográfico.	32
1.9.- SUSTANTIVACIÓN, ADJETIVACIÓN Y ADVERBIALIZACIÓN.	33

1.10.- LA PREFIJACIÓN.	35
2.- CREACIÓN DEL LÉXICO.	40
2.1.- DE LA SEPARACIÓN SILÁBICA Y LA ACENTUACIÓN.	43
2.1.1.- Localización de la primera vocal.	44
2.1.2.- Tratamiento del resto de la palabra.	46
2.1.2.1.- Consonantes entre vocales.	46
2.1.2.2.- Grupos vocálicos.	47
2.1.3.- Consonantes que requieren un tratamiento especial.	49
2.1.4.- Determinación de la sílaba tónica.	50
2.1.5.- Actualización de la tilde.	51
2.2.- TRATAMIENTO DE LAS FORMAS PRINCIPALES NO VERBALES.	53
2.2.1.- Raíces y terminaciones. Género y número.	53
2.2.2.- Sufijos apreciativos.	56
2.2.3.- Grado superlativo.	57
2.2.4.- Categorías gramaticales.	59
2.3.- TRATAMIENTO DE LOS VERBOS.	60
2.3.1.- Reglas de cambio de raíz.	62

	iii
2.3.2.- Reglas de excepción.	64
2.3.3.- Reglas de unión.	72
2.3.4.- Reglas de postexcepción.	76
2.3.5.- Participios irregulares.	78
2.3.6.- Reglas de acentuación.	80
2.3.7.- Terminaciones.	81
2.3.8.- Modelos de conjugación utilizados.	84
2.3.9.- Conjugación.	84
2.3.10.- Categorías gramaticales.	88
2.4.- ESTABLECIMIENTO DE RELACIONES MORFOLÓGICAS ENTRE DIFERENTES FORMAS PRINCIPALES.	89
3.- SOLUCIÓN AUTOMATIZADA DEL PROBLEMA.	93
3.1.- DISEÑO E IMPLEMENTACIÓN DEL RECONOCEDOR.	93
3.2.- DISEÑO E IMPLEMENTACIÓN DEL GENERADOR.	100
3.3.- MÓDULO DE ÍNDICES.	102
3.4.- MÓDULO DE ACCESOS EXTERNOS.	104
3.5.- MEDIDAS DE REALIZACIÓN.	107
3.6.- INTERFAZ DEL GENERADOR MORFOLÓGICO. ...	108

4.- DEL RECONOCIMIENTO Y LA GENERACIÓN EN UN DICCIONARIO DE SINÓNIMOS Y ANTÓNIMOS.	126
5.- ORGANIZACIÓN DEL DICCIONARIO DE SINÓNIMOS. . .	134
6.- CONCLUSIONES.	137
BIBLIOGRAFÍA.	140

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 —ciencia específica del lenguaje— ha experimentado un desarrollo importante en los últimos años y se ha ido interesando por 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 aproxime cada vez más a las necesidades de la persona. Bajo este enfoque evolucionan las técnicas de procesamiento de lenguaje natural y la marcada tendencia a que la interfaz hombre-máquina sea amigable.

Lo que inicialmente se planteó sobre bases textuales fueron problemas sencillos tales como la interrogación exacta; luego, la localización por truncamientos y máscaras se fue complicando en la dirección de la búsqueda por similitud tipográfica; ahora se pretende que las búsquedas se realicen con independencia de la flexión o derivación que haya sufrido la palabra y llegar a interrogar en lenguaje natural. En la línea de facilitar la

tarea humana están: el desarrollo de herramientas de ayuda para la generación de textos escritos —diccionarios de significados, de sinonimia, ideológicos, tesauros...—, 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,...—, la mecanización de la corrección ortográfica, etc.

Pero, ¿por qué situar un primer centro de gravedad en la morfología? El análisis morfológico constituye no sólo una base sólida por sí misma sino también el fundamento sobre el que es posible aproximarse con mayor rigor al sintáctico e incluso al semántico. El problema principal que presenta el morfema como elemento básico de la morfología es que el español contiene palabras que sólo pueden segmentarse arbitrariamente, a pesar de que estas palabras pertenecen a las mismas clases gramaticales a las que pertenecen otras que sí son segmentables en morfemas; ni siquiera haciendo la distinción de los estructuralistas entre morfema y morfo se resuelve completamente el problema en las lenguas flexivas. El trabajo de Aronoff aportó un cambio radical en la morfología: su base principal la constituyen las palabras y no los morfemas; reconoce que existen morfemas pero que carecen de significado autónomo fuera de las palabras en que aparecen y la unidad mínima con significado propio pasa del morfema a la palabra. Los

argumentos principales son que existen morfemas sin significado constante y palabras que no pueden dividirse en morfemas.

Las características principales de las lenguas flexivas son: a) dificultad para segmentar neta y coherentemente las palabras en morfos, b) falta de correspondencia entre los segmentos de palabra y los morfemas —las raíces y las terminaciones no son siempre las mismas y las terminaciones son multifuncionales—, c) las palabras llevan terminaciones con información sintáctica y semántica, d) la formación de palabras admite movilidad en sus elementos y e) el orden de las palabras en la oración es relativamente libre. Son varios los autores que distinguen tres enfoques generales sobre la estructura lingüística: a) Palabras y Paradigma, b) Elementos y Colocación y c) Elementos y Proceso. El modelo Palabras y Paradigma es el que sigue el enfoque más tradicional, supone una ventaja sobre los otros modelos y es que los morfemas no están organizados secuencialmente sino que son propiedades de las palabras como un todo; en este modelo la determinación de los formantes de la palabra se reduce a raíces o lexemas y terminaciones o accidentes.

La relación entre morfología y sintaxis es una cuestión abierta que se reforma continuamente y que influye en cualquier trabajo sobre morfología. La historia de la lingüística muestra que tal relación ha fluctuado desde la

completa inclusión de la morfología en la sintaxis hasta la total autonomía entre ambas; no obstante, los elementos flexivos tienen un comportamiento fundamentalmente sintáctico al ser elementos de concordancia. Vale la pena resaltar una diferencia fundamental entre morfología y sintaxis: la noción clave en morfología es el concepto de *palabra posible pero inexistente*, frente a la noción sintáctica de *oración posible o gramatical* frente a *oración imposible o agramatical*. La morfología generativa caracteriza las reglas flexivas por aplicarse a la base (sea palabra o tema) por la derecha y por no cambiar la categoría de la base. La flexión por sustitución es la más adecuada para las lenguas de carácter flexivo aunque existen casos en español de flexión por adición.

Durante mucho tiempo los procesadores morfológicos se han orientado fundamentalmente al análisis o reconocimiento y dejaron de lado los problemas de síntesis o generación de palabras; han tratado más extensamente la flexión que la derivación y la composición. Se aduce que, por ser la lingüística computacional una ciencia aplicada, sus actividades informáticas deben basarse exclusivamente en conocimientos sólidamente contrastados por la teoría; más bien todo lo contrario, por ser una ciencia aplicada no vale que se contente con llevar a la práctica modelos que hayan sido postulados completamente, sino que puede —y de hecho, debe—

investigar la implantación de soluciones que partiendo de casos concretos vayan completándose hasta alcanzar un nivel de manejabilidad que pueda ser de aplicación general. No nos cabe ninguna duda de la complejidad de los fenómenos morfológicos de derivación y de composición pero no hemos querido renunciar a abordarlos por carecer de una propuesta teórica global. Proponemos una solución bidireccional en el sentido de no restringirnos a un análisis de la morfología, puesto que podemos alcanzar la generación de formas morfológicas flexionadas o derivadas a partir de una forma primitiva con la misma información utilizada convenientemente.

Trabajamos exclusivamente con la forma escrita de las palabras —cosa bastante común en los sistemas de procesamiento de lenguaje natural. Nuestro concepto de palabra está perfectamente delimitado: es una cadena de caracteres entre espacios en blanco; la palabra fonológica no es considerada y por tanto los criterios de morfología teórica basados en la prosodia no son aplicables directamente sólo se usan en tanto en cuanto se reflejan en la propia escritura. Por ejemplo, a la hora de explicar las irregularidades del paradigma verbal los textos teóricos dicen: algunas aparentes anomalías que obedecen a principios generales fonológicos del sistema español no se consideran irregularidades:

cre-er $\Rightarrow$ *cre-yendo*

atac-ar $\Rightarrow$ *ataqu-é*

carg-ar $\Rightarrow$ *cargu-e*

unc-ir $\Rightarrow$ *unz-o*

por seguir un criterio estrictamente tipográfico tenemos en cuenta estas irregularidades en el modelo y consideramos como modificación toda cadena que presente alguna variante en sus caracteres —se incluye la aparición o desaparición de una tilde.

El análisis de una palabra se lleva a cabo mediante: 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. La generación de una forma derivada o flexionada de una forma principal se efectúa mediante: a) detección de la raíz y b) búsqueda de la terminación que corresponda a la forma que se desea generar —puede ser necesario recorrer las raíces alternativas si las hay. Se dispone de un conjunto de relaciones morfológicas —prefijales y sufijales— entre diferentes formas principales que faculta la obtención de formas derivadas secundarias o de orden superior e incluso sugiere maneras imaginativas de navegación, amén de dar respuesta a operaciones no inmediatas de derivación tales como la sustantivación, la adjetivación, la

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

verbalización y la adverbialización.

Para una mejor realización, orientamos más a datos que a reglas el sistema sobre el que soportamos la morfología. Donde sí abunda el uso de reglas es en la fase de preproceso que genera el universo; se mantiene almacenado en una estructura de datos manejable —representa las características morfológicas e incluso sintácticas de muchos aspectos del español. La cardinalidad del universo compuesto de formas principales flexionadas y derivadas representa algo más de sesenta veces la de las formas principales; por tanto, aprovechamos la morfología del español como lengua flexiva para la división en raíces y terminaciones, ya que se obtiene una reducción considerable de memoria —el número de raíces no duplica el de las formas básicas y las terminaciones necesarias son apenas unos cientos— y la eficacia de los procesos de reconocimiento y generación se mantiene alta.

1.- LA MORFOLOGÍA DEL ESPAÑOL.

Para el estudio morfológico del español conviene tener en cuenta las diferentes categorías gramaticales y los accidentes o flexiones que pueden sufrir cada una de ellas. Se consideran las siguientes categorías gramaticales: sustantivos, adjetivos, adverbios, verbos y otras expresiones que carecen de flexión —preposiciones, conjunciones, interjecciones, palabras de otros idiomas y algunas locuciones. Aunque sus características no difieren esencialmente de las tenidas en cuenta, otras categorías gramaticales como artículos, pronombres y otras locuciones no han sido consideradas por el momento por no tener cabida en la aplicación que nos proponemos llevar a cabo.

Los sustantivos poseen las desinencias flexivas de género y número y pueden derivarse de ellos, mediante sufijos apreciativos, aumentativos, diminutivos o despectivos. En general y con respecto al género se clasifican en masculinos o femeninos; algunos pueden usarse en los dos géneros: varían su morfología sólo en la desinencia o en toda su forma —heteronimia— o con una misma forma —comunes—, los menos tienen ambigüedad respecto al género.

Los adjetivos tienen desinencias flexivas de género, número y grado

superlativo —ciertas irregularidades afectan al lexema. Admiten derivación apreciativa mediante sufijos aumentativos, diminutivos o despectivos; en general pueden ser adverbializados añadiendo a la forma femenina la terminación *-mente*. Hay adjetivos de dos terminaciones —una para cada género— y de una terminación —no cambian su morfología con el género.

Los adverbios que mayor interés suponen son los caracterizados por su terminación en *-mente*; provienen de un adjetivo y generalmente son de modo, no tienen flexión con respecto al género ni al número.

Los verbos admiten como flexión la conjugación —determinada por el tiempo, la persona y el número—; también aceptan la incorporación de pronombres enclíticos. Existe una importante cantidad de irregularidades en la conjugación de muchos verbos y algunos son defectivos por carecer de alguna de sus formas.

En el apartado de invariantes se han clasificado las frases o locuciones y todas aquellas palabras que no presentan cambios morfológicos: palabras de otros idiomas, preposiciones, conjunciones e interjecciones.

En cuanto a la derivación de palabras en español es de resaltar que, aparte del uso de sufijos apreciativos y de la adverbialización de los adjetivos, existen los procesos de sustantivación —tanto de adjetivos como

de verbos—, de adjetivación —en sustantivos y en verbos— y la verbalización —de sustantivos y adjetivos. La prefijación es una operación de derivación que matiza, corrige o modifica el significado de la palabra sin cambiar su categoría gramatical y con independencia de la flexión.

1.1.- DE LA FORMACIÓN DEL PLURAL.

Son de uso genérico para los sustantivos y los adjetivos las siguientes reglas básicas:

1) Palabras terminadas en consonante:

1.1) Distinta de *n*, *s*, *x*, *z* añaden *es*.

árbol ⇒ *árboles*

1.2) *n* o *s* y agudas añaden *es*,

clan ⇒ *clanes*

gas ⇒ *gases*

si llevan tilde la pierden

camión ⇒ *camiones*

japonés ⇒ *japoneses*

1.3) *n* y no agudas añaden *es*, las llanas ganan una tilde al pasar a esdrújulas.

velamen ⇒ *velámenes*

1.4) *s* y no agudas permanecen invariables.

el guardagujas ⇒ los guardagujas

1.5) *x* permanecen invariables.

el clímax ⇒ los clímax

1.6) *z* la sustituyen por *ces*.

matriz ⇒ matrices

2) Palabras terminadas en vocal:

2.1) No acentuada añaden *s*,

copa ⇒ copas

excepto los monosílabos que añaden *es*,

el yo ⇒ los yoes

la a ⇒ las aes

la i ⇒ las íes

salvo las notas musicales que añaden *s*.

do, re, mi, fa, la, si ⇒ dos, res, mis, fas, las, sis

2.2) *á* u *ó* añaden *es* y pierden la tilde,

abacá ⇒ abacaes

paletó ⇒ paletoes

con excepciones que añaden *s*.

dominó ⇒ dominós

2.3) *é* añaden *s*.

café $\Rightarrow$ *cafés*

2.4) *í* o *ú* añaden *es*.

baladí $\Rightarrow$ *baladíes*

bambú $\Rightarrow$ *bambúes*

Además de las implícitas en las reglas, existe un número importante de otras excepciones que pueden concretarse en:

1) Excepciones a la formación del plural:

1.1) Palabras que sólo se usan en singular:

cariz, cenit, grey, salud, sed,...

1.2) Palabras que sólo se usan en plural:

albricias, ambages, anales, andas, añicos, arras,

enseres, exequias, finanzas, gafas, víveres,...

1.3) Palabras invariantes —igual morfología en singular y plural:

el caos $\Rightarrow$ *los caos*

el déficit $\Rightarrow$ *los déficit*

el quórum $\Rightarrow$ *los quórum*

el superávit $\Rightarrow$ *los superávit ...*

1.4) La desinencia del plural no es un sufijo:

cualquiera ⇒ *cualesquiera*

quienquiera ⇒ *quienesquiera* ...

1.5) Cambian la sílaba tónica al formar el plural:

régimen ⇒ *regímenes*

espécimen ⇒ *especímenes*

carácter ⇒ *caracteres* ...

1.6) Otras irregularidades:

réquiem ⇒ *réquiems*

vermut ⇒ *vermús* ...

1.7) Palabras que tienen dos formas para el plural:

accésit ⇒ *accésits* / *accessis*

currículum ⇒ *currícula* / *currículum* / *currículos* ...

2) Conviene tener en cuenta que hay palabras que pueden cambiar su

significado con el número:

anal ⇒ *anales*

celo ⇒ *celos*

prez ⇒ *preces*

seso ⇒ *sesos* ...

1.2.- DE LA FORMACIÓN DEL FEMENINO.

Son de uso genérico para los sustantivos las siguientes reglas básicas:

- 1.1) Los terminados en *o* cambian la *o* por *a*.

niño $\Rightarrow$ *niña*

- 1.2) Los terminados en consonante añaden una *a*.

horticultor $\Rightarrow$ *horticultora*

con excepciones

el juez $\Rightarrow$ *la juez*

- 1.3) Los terminados en *e* permanecen invariables

el conferenciante $\Rightarrow$ *la conferenciante*

o cambian *e* por *a*.

franchute $\Rightarrow$ *franchuta*

- 1.4) Los terminados en *a* permanecen invariables.

el malabarista $\Rightarrow$ *la malabarista*

Y para los adjetivos se han de tener en cuenta las siguientes reglas:

- 2.1) Los gentilicios que acaban en consonante añaden una *a*.

andaluz $\Rightarrow$ *andaluza*

2.2) Los que acaban en *o*, *ote* o *ete* cambian la vocal final por *a*.

bueno ⇒ *buen**a*

grandote ⇒ *grandot**a*

vejete ⇒ *vej**et**a*

2.3) Los que acaban en *an*, *on* u *or* añaden una *a*.

albarrán ⇒ *albarran**a*

bribón ⇒ *bribon**a*

agresor ⇒ *agresor**a*

salvo los comparativos latinos: *exterior*, *interior*, *superior*,

inferior, *ulterior*, *mayor*, *menor*, *mejor*, *peor*,...

2.4) El resto permanecen invariantes

amable, *azul*, *decente*, *verde*,...

salvo excepciones: *el sirviente* ⇒ *la sirvient**a*,...

Además de las implícitas en las reglas anteriores existen otras excepcionalidades que se pueden concretar en:

1) Excepciones a la formación del femenino:

1.1) Muchos sustantivos son sólo masculinos:

libro, *objeto*,...

1.2) Muchos sustantivos son sólo femeninos:

casa, *legaña*,...

1.3) Hay sustantivos femeninos que acaban en *o*

la mano,...

1.4) Hay sustantivos masculinos que acaban en *a*

el califa,...

1.5) Hay sustantivos de género común y adjetivos de una terminación —tienen la misma morfología cuando se usan en masculino y en femenino—:

el cónyuge ⇒ la cónyuge

el estudiante ⇒ la estudiante

el mártir ⇒ la mártir

el testigo ⇒ la testigo

el tratable ⇒ la tratable ...

1.6) Hay sustantivos de género ambiguo que no se sabe bien a qué género pertenecen; normalmente pueden ser usados en ambos:

el mar / la mar ...

1.7) En casos especiales pueden aparecer otras terminaciones

como *esa, isa, ina* o *triz*:

conde ⇒ *condesa*

papa ⇒ *papisa*

héroe ⇒ *heroína*

actor ⇒ *actriz* ...

1.8) Algunos sustantivos tienen dos formas para el femenino:

actor ⇒ *actora* / *actriz*

generador ⇒ *generadora* / *generatriz*

diablo ⇒ *diabla* / *diabla* ...

tigre ⇒ *tigra* / *tigresa*

1.9) Algunos sustantivos poseen heteronimia —cambian la

forma de la palabra y no sólo su desinencia—:

toro ⇒ *vaca*

carnero ⇒ *oveja* ...

2) Conviene tener en cuenta que existen voces que cambian su

significado con el género:

el cólera / *la cólera*

el corte / *la corte*

el frente / *la frente*

el guardia / la guardia

o según la forma del femenino:

la actriz / la actora

y que el género no es igual al sexo:

la foca macho / el cocodrilo hembra

1.3.- GRADO SUPERLATIVO DEL ADJETIVO.

En los adjetivos se puede considerar el grado superlativo como una flexión con los morfemas flexivos: *-ísimo* para el masculino singular, *-ísima* para el femenino singular, *-ísimos* para el masculino plural e *-ísimas* para el femenino plural.

grande ⇒ *grandísimo / grandísima / grandísimos / grandísimas*

blanco ⇒ *blanquísimo / blanquísima / blanquísimos / blanquísimas*

Hay que tener en cuenta que los acabados en *-ble* lo forman con *-bilísimo/a/os/as*

amable ⇒ *amabilísimo / amabilísima / amabilísimos / amabilísimas*

y que algunos adjetivos poseen superlativo irregular.

pobre ⇒ *paupérrimo / paupérrima / paupérrimos / paupérrimas*

sabio ⇒ *sapientísimo / sapientísima / sapientísimos / sapientísimas*

1.4.- SUFIJOS: CONSIDERACIONES GENERALES Y REGLAS DE DERIVACIÓN.

Una manera usual de formar palabras nuevas es añadiendo sufijos en uso a palabras ya existentes. La palabra original se denomina **primitiva** y la compuesta **derivada**; cuando a un vocablo ya derivado se le añade otro sufijo resulta un **derivado secundario**.

— Si el vocablo primitivo es un verbo: el sufijo se agrega al radical

si éste termina en consonante

am-ar $\Rightarrow$ *am-atorio*

tem-er $\Rightarrow$ *tem-ible*

insist-ir $\Rightarrow$ *insist-encia*

si acaba en vocal el radical puede perder la *e* o la *i* finales.

cambi-ar $\Rightarrow$ *cambi-o*

cambi-ar $\Rightarrow$ *camb-alache*

fre-ír $\Rightarrow$ *fr-itada*

— Si el vocablo primitivo no es verbo y termina en vocal: pierde las

letras finales *a*, *e*, *o*,

cervez-a $\Rightarrow$ *cervec-ero*

verd-e $\Rightarrow$ *ved-or*

blanc-o $\Rightarrow$ *blanc-ura*

y en algunas ocasiones pierde todo el diptongo final.

rab-ia $\Rightarrow$ *ráb-ico*

Si termina en consonante: normalmente se añade el sufijo sin sufrir modificaciones,

verbal $\Rightarrow$ *verbal-ismo*

las terminadas en *-dad* que pierden la *-ad* final.

humed-ad $\Rightarrow$ *humed-ecer*

— Al unir el sufijo si la palabra —o lo que quede de ella— acaba en *z*, ante *e* o *i* se convierte en *c*.

pez $\Rightarrow$ *pec-era*

cabez-a $\Rightarrow$ *cabec-ear*

moz-a $\Rightarrow$ *moc-ita*

— Si acaba en *c* con sonido /k/ (ca, co, cu), antes de *e* o *i* se convierte en *qu*.

tabac-o $\Rightarrow$ *tabaqu-ería*

— Si acaba en *qu* antes de *a*, de *o* o de *u* se convierte en *c*.

miriñaqu-e $\Rightarrow$ *miriñac-ote*

— Si acaba en *g* con sonido /x/ (ge, gi), antes de *a*, de *o* o de *u* se convierte en *j*.

esfing-e $\Rightarrow$ *esfinj-ucha*

- Si acaba en *g* seguida por *a* o por *o* y el sufijo comienza por *e* o por *i* se debe añadir una *u* para mantener el sonido /g/ (ga, go)

pag-o ⇒ *pagu-ito*

y el sufijo comienza por *u* seguida de vocal débil se debe añadir la diéresis para conservar el diptongo.

llag-a ⇒ *llag-üela*

- Si acaba en *g* seguida de *u* más *a* u *o* y el sufijo comienza por *e* o *i* se ha de añadir la diéresis correspondiente.

agu-a ⇒ *agü-eja*

Una clase importante de sufijos la constituyen los **sufijos apreciativos**

— **aumentativos, diminutivos y despectivos.**

1.5.- LA FORMACIÓN DE AUMENTATIVOS.

Pueden formar aumentativos los sustantivos, adjetivos y algunos adverbios.

Los sufijos principalmente utilizados para la apreciación aumentativa son **-ón** y **-azo** que poseen sus formas femeninas **-ona** y **-aza** y los plurales correspondientes **-ones**, **-onas** y **-azos**, **-azas**; también se utilizan **-ote** y **-acho** con sus femeninos y plurales. Son posibles los aumentativos de aumentativos combinando parejas de los anteriores: **-azón**, **-achón**, **-otón**,

-achote, -onazo,... Además de éstos existe un importante número de sufijos que se utilizan con menor frecuencia para la formación de aumentativos: *-ajón, -allón, -anchón, -ancón, -ardo, -arrón, -azón, -ejón, -ellón, -erón, -etón, -orrón, -ucón, -udo, -ujón, -ullón, -urrón*,...

— Algunos monosílabos que terminan en *e*, añaden el interfijo *-ez-* para la formación de sus aumentativos.

t-é $\Rightarrow$ *t-ez-azo*

— Existen algunas voces que pierden la vocal débil de un diptongo intermedio al formar aumentativos.

valient-e $\Rightarrow$ *valent-ón*

1.6.- LA FORMACIÓN DE DIMINUTIVOS.

Los sustantivos, adjetivos, gerundios y algunos adverbios pueden tener diminutivos.

Los sufijos principalmente utilizados para la apreciación diminutiva son *-ito* e *-illo* con sus formas femeninas *-ita* e *-illa* y los plurales correspondientes *-itos, -itas* e *-illos, -illas*; también se utiliza *-ico* aunque más bien es local (Aragón y Andalucía); o con frecuencia en Asturias *-ín* o *-ino* y de forma restringida *-uelo* —admiten femeninos y plurales. Los mencionados sufijos se emplean con carácter general para la formación de

diminutivos, aunque existen casos particulares en los que se añade algún interfijo.

— Añaden el interfijo *-c-* las palabras agudas de dos o más sílabas terminadas en *n* o en *r*

camión ⇒ *camion-c-ito*

sabor ⇒ *sabor-c-illo*

o las acabadas en vocal tónica que la mantienen

café ⇒ *cafe-c-ito*

rondó ⇒ *rondo-c-ito*

y las llanas acabadas en *n*

dictamen ⇒ *dictamen-c-illo*

-z- en el caso de *-uelo*.

ladrón ⇒ *ladron-z-uelo*

— Añaden el interfijo *-ec-* los monosílabos acabados en consonante

son ⇒ *son-ec-illo*

los bisílabos con la primera sílaba en *ue*, *eu*, *ie*, *ei*

cuent-o ⇒ *cuent-ec-ito*

los bisílabos con la última sílaba en *ia*, *io* o *ua*

sabi-a ⇒ *sabi-ec-illa*

las voces terminadas en *io*

cenobi-o $\Rightarrow$ *cenobi-ec-ito*

y los bisílabos terminados en *e*.

cort-e $\Rightarrow$ *cort-ec-ito*

-ez- en el caso de *-uelo/a*.

lot-e $\Rightarrow$ *lot-ez-uelo*

Las palabras terminadas en *z* la cambian por *c*:

pez $\Rightarrow$ *pec-ec-illo*

— Añaden el interfijo *-ecec-* los monosílabos acabados en vocal

té $\Rightarrow$ *t-ecec-ito*

-ecez- en el caso de *-uelo*.

pie $\Rightarrow$ *pi-ecez-uelo*

En derivados de sustantivos y verbos se utiliza *-ín*;

faj-a $\Rightarrow$ *faj-ín*

tremol-ar $\Rightarrow$ *tremol-ín*

en los derivados nominales también combina con *-ete* (*-etín*), *-ello* (*-ellín*),

-ito (*-itín*) y *-orre*, *-ote* (*-orrotín*, *-irritín*)

caf-é $\Rightarrow$ *caf-etín*

fald-a $\Rightarrow$ *fald-ellín*

chic-o $\Rightarrow$ *chiqu-itín* / *chic-orrotín* / *chiqu-irritín*

y en los verbales con *-acho* (*-achín*), *-ancho* (*-anchín*) y *-ario* (*-arín*).

mat-ar $\Rightarrow$ *mat-achín*

parl-ar $\Rightarrow$ *parl-anchín*

danz-ar $\Rightarrow$ *danz-arín*

Con menor frecuencia se usan: *-aco*, *-ajo*, *-ancho*, *-anchuelo*, *-arajo*, *-ato*, *-cín*, *-ececín*, *-ecín*, *-ejo*, *-encho*, *-ete*, *-eto*, *-ezno*, *-ijo*, *-ino*, *-iño*, *-oncho*, *-ote*, *-ucho*, *-uco*, *-ugo*, *-ujo*, *-ullo*,...

Existen algunas voces que pierden un diptongo intermedio al formar diminutivos.

calient-e $\Rightarrow$ *calent-ito*

1.7.- LA FORMACIÓN DE DESPECTIVOS.

Los despectivos se producen con sustantivos y adjetivos.

Los sufijos más utilizados para la formación de despectivos son: *-ejo* y *-ucho*, con sus femeninos *-eja* y *-ucha*, y sus plurales *-ejos*, *-ejas* y *-uchos*, *-uchas*. También se utilizan:

Otros sufijos para la formación de despectivos					
Con alguna frecuencia		Con poca frecuencia		Ocasionalmente	
-acho	-ete	-achón	-ellón	-ango	-aruzco
-aco	-orrio	-achuelo	-encho	-arasco	-enco
-ajo	-ote	-aino	-erón	-arazco	-ete
-anco	-ucho	-ajón	-etón	-aresco	-iche
-aracho	-uco	-alla	-istrajo	-arezco	-il
-ardo	-ujo	-allón	-onazo	-ario	-illo
-arraco	-ujón	-ancho	-oncho	-arisco	-ina
-aruco	-urrón	-anchón	-orrón	-arizco	-oche
-astre	-usco	-ancón	-otazo	-arracho	-ón
-astro	-uzco	-arajo	-ote	-arro	-orio
-azo	-uzo	-arrajo	-ucón	-arusco	-orro
-ejo		-arrón	-ullo		
		-azón	-ullón		
		-ejón	-urrón		

Algunos monosílabos y bisílabos que poseen diptongos o terminan en *e*, añaden el interfijo *-ec-* para la formación de sus despectivos

pi-e $\Rightarrow$ *pi-ec-ejo*

o *-ez-* en su caso

pi-e $\Rightarrow$ *pi-ez-ucho*

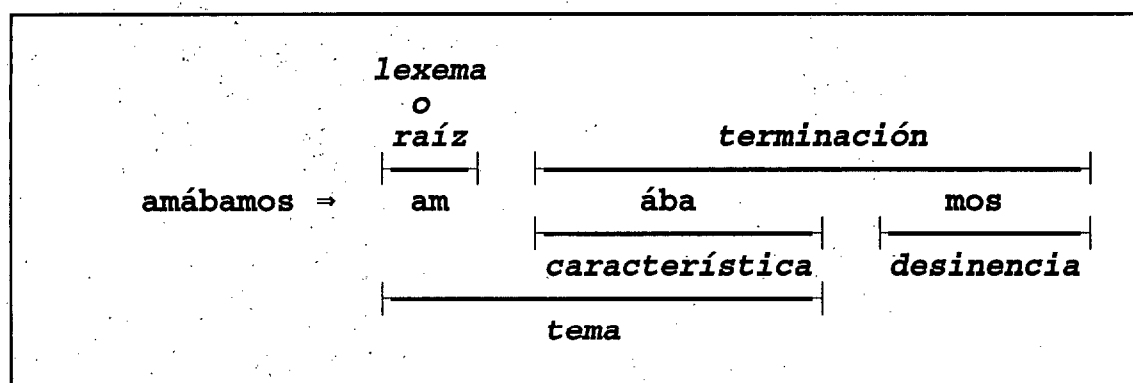
Existen algunas voces que pierden un diptongo intermedio al formar despectivos.

ardient-e $\Rightarrow$ *ardent-ejo*

1.8.- VERBOS: CONJUGACIÓN Y PRONOMBRES ENCLÍTICOS.

1.8.1.- Morfología del verbo.

Por sus características formales, el verbo es aquella parte de la oración que contiene *morfemas* flexivos de número, persona, tiempo y modo. Suele aplicarse la denominación de *desinencias* a los morfemas de número y persona, y la de *características* a los de modo y tiempo. Si se suprime de una forma verbal su desinencia y característica, queda el *lexema*, *raíz* o *radical*. La unión de la raíz con la característica se denomina *tema*. Cualquier forma verbal se construye básicamente con una raíz y una *terminación* —constituida por los morfemas.



Existen dos formas de conjugación denominadas *voz activa* y *voz pasiva* —se forma con el tiempo coincidente del verbo auxiliar *ser* y el participio del verbo. A su vez existen *tiempos simples* y *compuestos* —se forman con el tiempo correspondiente del verbo auxiliar *haber* y el participio del verbo.

CONJUGACIÓN VERBAL (voz activa)			
MODOS	TIEMPOS		NÚMEROS Y PERSONAS
	SIMPLES	COMPUESTOS	
Indicativo	presente	pretérito perfecto compuesto	1ª, 2ª, 3ª: singular y plural
	pretérito imperfecto	pretérito pluscuamperfecto	1ª, 2ª, 3ª: singular y plural
	pretérito perfecto simple (o pretérito indefinido)	pretérito anterior	1ª, 2ª, 3ª: singular y plural
	futuro (o futuro imperfecto)	futuro perfecto	1ª, 2ª, 3ª: singular y plural
	condicional (o potencial simple)	condicional perfecto (o potencial compuesto)	1ª, 2ª, 3ª: singular y plural
Subjuntivo	presente	pretérito perfecto	1ª, 2ª, 3ª: singular y plural
	pretérito imperfecto	pretérito pluscuamperfecto	1ª, 2ª, 3ª: singular y plural
	futuro (o futuro imperfecto)	futuro perfecto	1ª, 2ª, 3ª: singular y plural
Imperativo	presente		1ª sólo plural y 2ª, 3ª: singular y plural
Formas impersonales (o infinitas)	infinitivo	infinitivo compuesto	sin persona ni número (sustantivo verbal, a veces tiene plural)
	gerundio	gerundio compuesto	sin persona ni número (adverbio verbal)
	participio	participio compuesto	sin persona; con género y número como adjetivo verbal

Atendiendo a la forma de su conjugación se pueden clasificar en *primera* (la terminación del infinitivo es *-ar*), *segunda* (*-er*) o *tercera* (*-ir*). Para el tratamiento automatizado de la conjugación no conviene olvidar que la mayor parte de los verbos son regulares y mantienen la raíz invariable en su conjugación, si se exceptúa la posición del *acento de intensidad*. Los verbos irregulares presentan variaciones en su raíz —vocálicas, consonánticas o mixtas— además de otros tipos de anomalías.

— La IRREGULARIDAD VOCÁLICA consiste en el cambio de una vocal por otra u otras:

e por i **concebir** ⇒ **concibo**

o por ue **colgar** ⇒ **cuelgo**

— La IRREGULARIDAD CONSONÁNTICA tiene lugar si se reemplaza una consonante por otra:

c por qu delante de e **caber** ⇒ **quepo**

o se añade una consonante a la consonante final de la raíz del infinitivo:

añadir z **conocer** ⇒ **conozco**

— La IRREGULARIDAD MIXTA aparece cuando se sustituyen una vocal y una consonante por otra vocal y otra consonante:

ab por ep **saber** ⇒ **sepa**

o en la adición del grupo **ig** a la última vocal de la raíz:

añadir **ig** **caer** ⇒ **caigo**

Existen irregularidades que afectan al tema —**di** de **decir**, **haz** de **hacer**, **doy** de **dar**, **voy** de **ir**, etc.— y otras de más difícil sistematización:

— contracciones: **ver** ⇒ **ves** en vez de **vees**.

— verbos con más de una raíz (*supletivismo*): **ser** ⇒ **fueron**, **era**.

— participios y gerundios irregulares: **escribir** ⇒ **escrito**,
dormir ⇒ **durmiendo**

No se consideran propiamente irregularidades los simples cambios ortográficos:

— g por gu delante de e o i **sigo** ⇒ **sigue**

— qu por c delante de a u o **delinquir** ⇒ **delinca**

En estos casos no ha variado el *fonema*, sino el carácter que lo representa. Otras aparentes anomalías obedecen a principios generales

del sistema fonológico español y tampoco se consideran irregularidades —leyendo de leer.

1.8.2.- Grupos de tiempos irregulares.

La irregularidad afecta siempre a más de un tiempo. Son tres los grupos de tiempos que comparten la misma irregularidad:

— Si es irregular el presente de indicativo, también lo son los otros presentes.

— Si es irregular el pretérito indefinido, poseen la misma irregularidad el pretérito imperfecto de subjuntivo y el futuro imperfecto de subjuntivo.

— Si es irregular el futuro imperfecto de indicativo, tiene la misma irregularidad el condicional simple.

El pretérito imperfecto de indicativo carece de irregularidades, salvo raras excepciones que se reducen casi exclusivamente a los imperfectos heredados del latín *era* e *iba*, de los verbos de raíz múltiple *ser* e *ir*.

Los tiempos compuestos y la voz pasiva no presentan más irregularidades que las de los auxiliares *haber* y *ser* y las de algunos participios como *escrito*, *impreso*,...

1.8.3.- Modelos de conjugación.

Tradicionalmente las irregularidades que se dan en la conjugación se sistematizan a través de unos modelos de conjugación, así por ejemplo el Diccionario Actual de la Lengua Española de VOX establece los 92 que se exponen en la tabla siguiente.

Modelos de conjugación					
sacar	descafeinar	dormir	regar	argüir	poner
mecer	rehusar	servir	trocar	andar	predecir
zurcir	reunir	hervir	forzar	asir	querer
realizar	amohinar	ceñir	avergonzar	caber	raer
proteger	prohibir	reír	colgar	caer	roer
dirigir	averiguar	tañer	jugar	dar	saber
llegar	ahincarse	empeller	cocer	decir	salir
distinguir	enraizar	muñir	elegir	erguir	satisfacer
delinquir	cabrahigar	mullir	seguir	estar	ser
adecuar	homogeneizar	nacer	errar	haber	tener
actuar	acertar	agradecer	agorar	hacer	traer
cambiar	entender	conocer	desosar	ir	valer
desviar	discernir	lucir	oler	oír	venir
auxiliar	adquirir	conducir	leer	placer	ver
aislar	contar	empezar	huir	poder	yacer
aunar	mover				

1.8.4.- Pronombres enclíticos.

El empleo de los pronombres enclíticos ha ido reduciéndose con el tiempo, y ahora se usan casi exclusivamente cuando el verbo está en infinitivo, imperativo o gerundio.

Un verbo puede llevar simultáneamente hasta tres pronombres átonos al final; en tal caso, la partícula **se** debe preceder al resto, el de segunda persona —**te, os**— adelanta siempre al de primera —**me, nos**— y cualquiera de estos dos antecede al de tercera —**le, les, la, las, lo, los**.

Al unirse los pronombres enclíticos con el verbo, se producen algunas alteraciones que afectan al último carácter del verbo:

— delante del enclítico **nos** se pierde la **s** de la primera persona del plural del subjuntivo e imperativo

comamos + nos ⇒ comámonos

esta pérdida se produce también en otros tiempos del verbo, pero el

pronombre enclítico es muy poco usual en ellos. En la segunda persona del plural seguida del enclítico se desaparece una s

comamos + selo $\Rightarrow$ comámoselo

— la forma reflexiva de la segunda persona del plural del imperativo pierde la d

comed + os $\Rightarrow$ comeos

1.8.4.1.- El acento ortográfico.

El acento ortográfico de las formaciones con enclíticos está siempre de acuerdo con las reglas generales cuando dicha formación es esdrújula o sobresdrújula. Cuando una forma verbal llana o esdrújula se agrupa con uno o más enclíticos, la vocal prosódicamente acentuada del verbo lleva siempre tilde, lo exija o no cuando se emplea sin enclíticos: **decía-me-lo, veía-la, oía-lo** —de acuerdo con las formas verbales **decía, veía, oía**—, **hablába-se, mirándo-os, quisiéra-lo** —en contraste con **hablaba, mirando, quisiera**. Y cuando una forma verbal aguda —incluyendo las monosilábicas— se agrupa con dos enclíticos, la vocal prosódicamente acentuada del verbo se escribe siempre con tilde, lo requiera o no cuando se emplea sola: **partió-se-le, oír-se-lo, dé-se-la** —de acuerdo con **partió, oír, dé**—, **dá-se-lo, dí-me-lo, decid-nos-lo, pedír-me-la** —en contraste con **da, di, decid, pedir**.

El acento ortográfico de las formaciones con enclíticos deja de estar de acuerdo con las reglas generales en algunos casos en que la formación resulta con acentuación llana. En las formas verbales agudas —incluyendo las monosilábicas— seguidas de un solo enclítico: **da-le, fui-me, decid-me, reír-se, oír-lo; dé-le, salí-me, partió-se** el verbo conserva su acento ortográfico originario; solamente en los cinco primeros ejemplos la formación se atiene a las reglas generales del uso ortográfico (**dale** como

sale, fuime como **fuiste**,...) y en los tres últimos ejemplos diverge de las reglas generales, ya que una palabra llana terminada en vocal normalmente no lleva tilde.

Aparecen también sometidos a un régimen ortográfico especial los imperativos plurales de los verbos reflexivos, o en construcción reflexiva, tras la pérdida de la desinencia **d**. Formas como **marcha-os**, **detene-os** han de emplearse sin tilde, a pesar de que la forma verbal es aguda y de que se agrupa con sólo un enclítico. Sin embargo los verbos de la tercera conjugación llevan tilde a causa del hiato como **partí-os**, salvo la forma **id-os** del verbo **ir** en el que **í-os** es un arcaísmo.

1.9.- SUSTANTIVACIÓN, ADJETIVACIÓN Y ADVERBIALIZACIÓN.

Los adjetivos que acaban en **-ble** pueden formar sustantivos femeninos en **-bilidad**. Análogamente todos los sustantivos terminados en **-bilidad** se relacionan con un adjetivo en **-ble**.

probable ⇔ *probabilidad*

Añadiendo la terminación **-mente** a la forma femenina de un adjetivo se forma en general un adverbio de modo.

bueno ⇒ *buenamente*

afable ⇒ *afablemente*

El participio es el adjetivo verbal (salvo en algunos verbos entre los que se encuentran los auxiliares: *ser*, *haber* o *estar*) y por ello admite todas las flexiones propias de los adjetivos de dos terminaciones.

amar ⇒ *amado*, *amada*, *amados*, *amadísimo*, *amadamente*, ...

El infinitivo de un verbo generalmente se puede considerar como el sustantivo verbal —masculino— y en consecuencia admitiría plural e incluso otras flexiones propias de los sustantivos.

deber ⇒ *el deber, los deberes, ...*

Además existe un buen número de sufijos que permiten a partir de la raíz de un verbo formar sustantivos y adjetivos: *-able, -ación, -adero, -adiza, -adizo, -ador, -adora, -adura, -aje, -amento, -amiento, -ancia, -ante, -antez, -anza, -aria, -arina, -ario, -ativa, -ativo, -atoria, -atorio, -azgo, -azo, -edero, -ediza, -edizo, -edor, -edora, -edura, -encia, -ente, -etivo, -etoria, -etorio, -ible, -ición, -icionario, -idero, -ía, -ón,...*

amar ⇒ *amable*

Algunas terminaciones propias de la verbalización de sustantivos y adjetivos: *-ecer, -izar, -ear, ...*

verde ⇒ *verdecer*

formal ⇒ *formalizar*

baba ⇒ *babear*

y otras que permiten formar sustantivos a partir de otros sustantivos o de adjetivos: *-ismo, ...*

laico ⇒ *laicismo*

o adjetivos a partir de otros adjetivos o de sustantivos: *-ista,*

social ⇒ *socialista*

1.10.- LA PREFIJACIÓN.

Existe otra posibilidad de formación de palabras mediante la introducción de prefijos y aunque puede haber un paralelismo con la sufijación, suele considerarse como composición —cuestión muy debatida. Los prefijos generalmente no producen cambios de categoría gramatical y son independientes de la flexión de la palabra. Normalmente matizan, corrigen o modifican el significado de la palabra, en particular algunos de ellos tienen significación de negación y son de utilidad en la relación de antonimia.

Se pueden agrupar en categorías según su significación en:

- A ambos lados o alrededor de: *anfi-*, *circun-*.
- A distancia o lejos: *tele-*.
- A través de, cambio o al otro lado: *tras-*, *trans-*, *ultra-*, *meta-*.
- Acción secundaria: *so-*, *sub-*, *za-*, *zam-*, *sus-*.
- Arriba, en alto o sobre: *ana-*, *epi-*, *supra-*.
- Atenuación del significado: *so-*, *sub-*, *za-*, *zam-*, *sus-*, *entre-*.
- Aumento, encarecimiento o refuerzo del significado: *re-*, *rete-*, *requete-*, *archi-*, *super-*, *sobre-*, *hiper-*, *ultra-*, *extra-*.
- Bajo o debajo: *so-*, *sub-*, *za-*, *zam-*, *sus-*, *sota-*, *soto-*.
- Compañía: *co-*, *con-*.
- Conforme a: *ana-*.
- Contra, contrariedad, contrario, oposición, opuesto o rechazo: *contra-*, *anti-*, *cata-*, *di-*, *dis-*, *ob-*, *para-*, *re-*, *rete-*, *requete-*.
- Delante, anterioridad en tiempo y espacio o prioritario: *pro-*, *pre-*, *ante-*, *proto-*.
- Dentro: *intro-*, *endo-*.
- Después de en el sentido de detrás: *pos-*, *post-*, *meta-*.

- Doble o dos: *bi-*, *bis-*, *anfi-*, *di-*.
- En vez de o por substitución: *pro-*.
- Entre o en medio: *inter-*, *dia-*.
- Extensión o dilatación: *di-*, *cata-*.
- Exterior, junto o próximo: *para-*, *epi-*, *yuxta-*.
- Fuera, más allá o externo: *extra-*, *ex-*, *es-*, *des-*, *ecto-*, *meta-*,
supra-, *ultra-*, *exo-*.
- Hacia atrás, de nuevo, tiempo anterior o inversión de la acción:
retro-, *des-*, *re-*, *ob-*, *ana-*.
- Igual: *equi-*, *iso-*.
- Inferioridad: *hipo-*.
- Lejos de o separado de: *apo-*.
- Medio, casi o mitad: *semi-*, *hemi-*.
- Negación: *re-*, *a-*, *des-*, *in-*.
- Orden posterior: *so-*, *sub-*, *za-*, *zam-*, *sus-*.
- Origen o procedencia: *di-*, *abs-*, *ab-*.
- Parte de acá: *cis-*, *citra-*.
- Parte de atrás: *opisto-*.
- Por causa de: *ob-*.
- Preeminencia, primacía, superioridad o prioridad: *archi-*, *arce-*,
arci-, *arc-*, *arz-*, *arqui-*, *proto-*, *pro-*.
- Preposiciones: *a-*, *ad-*, *de-*, *en-*, *in-*, *sin-*, *so-*.
- Privación, negación o ausencia: *a-*, *des-*, *ab-*, *abs-*, *ex-*, *ana-*, *in-*,
hipo-.
- Progreso, continuidad de acción o hacia adelante: *pro-*, *para-*.
- Que suple o que hace las veces de: *vice-*, *viz-*, *vi-*.
- Repetición o reiteración de la acción de un verbo: *re-*.

- Segundo lugar en categoría: *vice-*, *viz-*, *vi-*.
- Separación: *ab-*, *abs-*, *dia-*, *dis-*, *di-*.
- Simultaneidad: *sin-*.
- Situación o calidad intermedia: *entre-*, *entro-*.
- Subalterno: *sota-*, *soto-*, *sub-*, *vice-*, *viz-*, *vi-*.
- Unión: *co-*, *con-*, *sin-*, *yuxta-*.

Deben tenerse en cuenta las siguientes reglas:

- + Cuando se añade un prefijo a un monosílabo sin tilde acabado en vocal, *n* o *s* debe tildarse.

anti- + *son* $\Rightarrow$ *antisón*

- + Cuando se añade un prefijo terminado en vocal una palabra que comienza por *r* debe duplicarse la *r* para mantener el sonido fuerte.

contra- + *revolucionario* $\Rightarrow$ *contrarrevolucionario*

- + Cuando se añade un prefijo terminado en vocal fuerte a una palabra que comienza por vocal débil tónica no tildada, aunque vaya precedida de *h*, debe tildarse para romper el diptongo.

entre- + *ido* $\Rightarrow$ *entreído*

re- + *hice* $\Rightarrow$ *rehíce*

- + El prefijo *a-* se convierte en *an-* ante vocal.

a- + *estesia* $\Rightarrow$ *anestesia*

- + El prefijo *ana-* no se usa ante palabras que comienzan por *a*.
- + El prefijo *anfi-* se convierte en *anfí-* ante palabras bisilábicas llanas.

anfi- + *podo* $\Rightarrow$ *anfípodo*

- + Los prefijos *arc-* y *arz-* no se usan ante consonantes.

- + El prefijo *circun-* toma la forma *circum-* ante *b*, *p* o *m* y en ocasiones ante *n*.

circun- + *polar* $\Rightarrow$ *circumpolar*

circun- + *nutación* $\Rightarrow$ *circumnutación*

circun- + *navegar* $\Rightarrow$ *circunnavegar*

- + Los prefijos *con-*, *en-* y *sin-* cambian la *n* por *m* ante *b* o *p*.

con- + *beneficiado* $\Rightarrow$ *combeneficiado*

en- + *paquete* $\Rightarrow$ *empaquetar*

sin- + *pétalo* $\Rightarrow$ *simpétalo*

- + El prefijo *exo-* reduce la *o* ante palabras que comienzan por esta vocal, si está tildada puede mantenerse o no la tilde

exo- + *ósmosis* $\Rightarrow$ *exosmosis* / *exósmosis*

Ante bisílabas llanas en ocasiones se convierte en *exó-*

exo- + *geno* $\Rightarrow$ *exógeno*

exo- + *dermo* $\Rightarrow$ *exodermo*

- + El prefijo *in-* se convierte en *im-* ante *b* o *p*

in- + *borrable* $\Rightarrow$ *imborrable*

in- + *palpable* $\Rightarrow$ *impalpable*

y en *i-* ante *l* o *r*

in- + *lógico* $\Rightarrow$ *ilógico*

en el caso de *r* ésta se dobla para mantener el mismo sonido.

in- + *real* $\Rightarrow$ *irreal*

- + Los prefijos *trans-* y *tras-* reducen la *s* ante palabras que comienzan por esta letra y pueden hacer desaparecer la inicial en palabras que comienzan por *e*.

trans- + *sexual* $\Rightarrow$ *transexual*

trans- + *emitir* $\Rightarrow$ *transmitir*

+ El prefijo *para-* puede convertirse en *pará-* sin regla conocida.

para- + *metro* $\Rightarrow$ *parámetro*

para- + *poco* $\Rightarrow$ *parapoco*

+ El prefijo *tele-* puede convertirse en *telé-* ante palabras bisilábicas llanas.

tele- + *fono* $\Rightarrow$ *teléfono*

tele- + *grama* $\Rightarrow$ *telegrama*

+ En contra de las normas de la Real Academia Española de la Lengua, existen prefijos que no se sienten unidos a la base y se escriben con guión.

anti- + *vertedero* $\Rightarrow$ *anti-vertedero*

2.- CREACIÓN DEL LÉXICO.

Dada la complejidad inherente a la morfología del español se ha optado por producir las formas flexionadas, conjugadas y derivadas, mediante la aplicación de las reglas básicas discutidas en la sección anterior; una segunda fase exige la corrección por diferentes medios —automatizados y manuales— de las formas obtenidas. Para mayor eficacia se adopta un sistema orientado a datos que tiene en cuenta raíces y terminaciones; la aplicación de reglas produciría un sistema poco fiable por complejo y tendría una realización peor. De esta forma se reconoce y se genera todo aquello y sólo aquello que ha sido previamente declarado universo.

Cada palabra, se denominará forma principal —el infinitivo en los verbos—, consta de: a) una categoría gramatical, b) una serie de formas flexionadas y derivadas y c) una lista de palabras de otras categorías gramaticales relacionadas morfológicamente. Cada elemento de este conjunto se segmenta en una raíz y una terminación. El universo de terminaciones está constituido por todas las necesarias para obtener de la forma principal las formas flexionadas y derivadas. Se utiliza un criterio de corte tendente a minimizar el número de raíces alternativas que pueden aparecer junto a la raíz de la forma principal. Cada raíz se acompaña de un indicativo de las terminaciones que puede llevar y con qué función se utilizan. Toda esta estructura constituye el léxico que se considera; se almacena en memoria secundaria de forma que permita un acceso rápido a cada elemento.

Las posibilidades de flexión y derivación en los diferentes elementos del léxico dependen fundamentalmente de la categoría gramatical de la

forma principal. A continuación se expone un cuadro ilustrativo de las posibilidades existentes en cada uno de los casos.

Definición de los elementos del léxico							
Categoría Gramatical	Flexiones				Derivación: Sufijos apreciativos		
	Género	Número	Conjugación	Grado superlativo	Aumentativo	Diminutivo	Despectivo
Verbo			SI			Gerundio	
Sustantivo	SI	SI			SI	SI	SI
Adjetivo	SI	SI		SI	SI	SI	SI
Adverbio							
Invariante							

Existen además relaciones entre los diferentes elementos del léxico que permiten obtener otras formas relacionadas morfológicamente con una dada; en particular posibilitan: sustantivaciones, adjetivaciones, adverbializaciones o verbalizaciones. Las relaciones establecidas se concretan para los verbos en: sustantivos relacionados, adjetivos relacionados y verbos relacionados mediante cambios prefijales; para cada sustantivo existe la lista de verbos relacionados; los adjetivos además de su lista de verbos relacionados poseen un adverbio de modo que les corresponde; y los adverbios de modo tienen asociado el adjetivo del que provienen, figura 1.

Las mencionadas relaciones permiten realizar derivaciones entre las

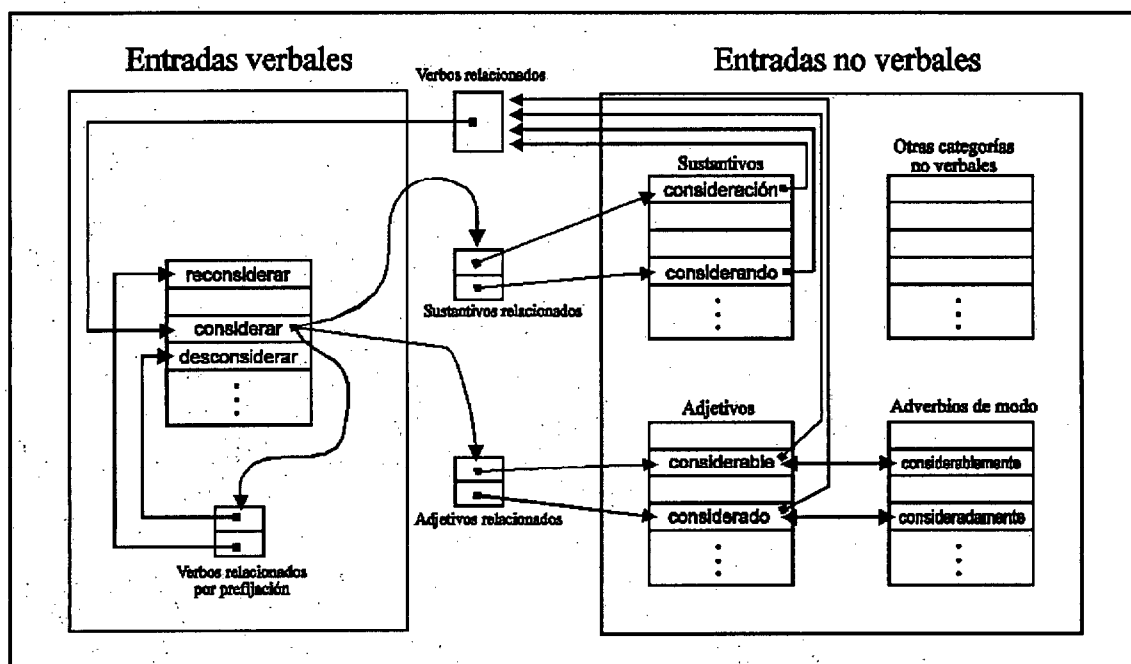


Figura 1: Relaciones entre elementos del léxico

diferentes categorías gramaticales. La siguiente tabla muestra la forma en que se deriva cuando existe un cambio de categoría gramatical —sin tener en cuenta las relaciones prefijales entre verbos. Se incluye la generación de adverbios a partir de adjetivos mediante la terminación *-mente* aunque es más propiamente una forma de composición.

Derivaciones con cambio de categoría gramatical				
Categoría Gramatical	Cambio a categoría gramatical			
	Sustantivo	Adjetivo	Adverbio de modo	Verbo
Verbo	verbo → sustantivo	verbo → adjetivo	verbo → adjetivo → adverbio	
Sustantivo		sustantivo → verbo → adjetivo	sustantivo → verbo → adjetivo → adverbio	sustantivo → verbo
Adjetivo	adjetivo → verbo → sustantivo		adjetivo → adverbio	adjetivo → verbo
Adverbio de modo	adverbio → adjetivo → verbo → sustantivo	adverbio → adjetivo		adverbio → adjetivo → verbo
Invariante				

Las posibilidades se enriquecen cuando en lugar de un verbo se usan todos aquellos que se encuentran morfológicamente relacionados mediante modificación prefijal; se obtiene así un conjunto de formas principales más amplio que incluye las palabras morfológicamente relacionadas a través de cambios prefijales y sufijales.

2.1.- DE LA SEPARACIÓN SILÁBICA Y LA ACENTUACIÓN.

El necesario tratamiento de las tildes como parte de la morfología de las palabras aconseja el diseño de un delimitador silábico y la ulterior aplicación de las reglas de acentuación tanto para la determinación de la sílaba tónica de una palabra como para el proceso de actualización de su tilde.

El proceso empleado para la separación silábica se basa exclusivamente en la grafía de la palabra.

Se lleva a cabo en los siguientes pasos:

- a) Localización de la primera vocal de la palabra.
- b) Tratamiento del resto de la palabra.

Ambos puntos se apoyan en que cada sílaba necesita de un núcleo que contiene una vocal al menos.

2.1.1.- Localización de la primera vocal.

Este paso lo demanda la existencia de ciertos grupos consonánticos que sólo se manifiestan al principio de la palabra. En palabras como *psicología* y *gnomo* las dos consonantes iniciales forman parte de la misma sílaba (*psi-co-lo-gí-a* y *gnó-mo*); en cambio, en casos como los de *snob* y *slogan* la lengua española añade *e* delante de *s* a la hora de pronunciarlas, la primera sílaba será /es/ (*s-nób* y *s-ló-gan*) —la letra *s* no se integra en la misma sílaba que *n* o *l*.

La forma de actuar en esta parte del proceso consiste en examinar carácter a carácter hasta encontrar la primera vocal; respecto a las consonantes que la anteceden se tiene en cuenta:

- 0. Ninguna, la palabra comienza con una vocal. Ejemplo: *a-la*.

1. Una, puede ser cualquiera. Se cuida no confundir el carácter *u* de las sílabas *que*, *qui*, *gue*, *gui* con la vocal *u*. Ejemplo: *pa-la*, *que-so*, *gui-so*.

2. Dos:

2.1.- La primera consonante pertenece a $\{p, b, f, t, c, g, d\}$ o a $\{p, b, f, c, g, l\}$ y la segunda es *r* o *l* respectivamente —grupos tautosilábicos. Ejemplos: *pra-do*, *bra-zo*, *gra-pa*, *pla-ta*, *blan-co*, *fla-co*, *lla-no*.

2.2.- El primer carácter es la letra *c* y el siguiente es la *h*. Ejemplos: *chi-no*, *chu-pa*.

2.3.- La pareja se corresponde con el grupo $\{gn, mn, ps, cn, cz, pt, tm\}$; ejemplos: *gn̄o-mo*, *mne-mo-tec-nia*, *psi-co-lo-gí-a*. O con el grupo $\{sl, sn\}$; ejemplos: *s-l̄o-gán*, *s-n̄ob*.

Las palabras que no cumplan con alguna de las condiciones expuestas se consideran extrañas.

Se traslada la posibilidad de obtener el prefijo *sub-* a partir de una distribución silábica morfológica (*sub-lu-nar*, *sub-a-cuá-ti-co*) y no fonológica (*su-bli-mi-nal*, *su-bir*) a un nivel de conocimiento que trascienda del puramente lexicográfico.

2.1.2.- Tratamiento del resto de la palabra.

El examen de caracteres que progresa hasta la siguiente vocal o el final de la palabra puede encontrarse con una de las dos circunstancias:

1.- Consonantes entre vocales.

2.- Grupos vocálicos.

Sus soluciones fijan las fronteras silábicas en el interior de una palabra.

2.1.2.1.- Consonantes entre vocales.

Atendiendo a su número:

a) Una consonante.

Se considera que forma parte de la segunda sílaba. Ejemplos:

a-la, ho-la, so-lo.

b) Dos consonantes.

Se sitúa la frontera silábica entre ambas consonantes (ai-s-lar, ar-bol) a menos que formen grupo tautosilábico en cuyo caso forman parte de la segunda sílaba (a-tra-er, ha-blar).

c) Tres consonantes.

1. Se considera que la primera pertenece a la primera sílaba y las dos últimas a la segunda si forman grupo tautosilábico. Ejemplos: *as-tro*, *es-tre-lla*.

2. Se sitúa la frontera silábica entre la segunda y la tercera cuando las dos últimas no forman grupo tautosilábico. En tal caso las dos primeras han de ser cualquiera de $\{b, d, l, n, r\}$ seguida de *s* (*subs-tan-cia*, *ads-cri-bir*, *sols-ti-cio*, *trans-for-mar*, *pers-pi-caz*) o *s* seguida de *t* (*post-da-ta*, *ist-mo*) —coda compuesta.

d) Cuatro consonantes.

1. Se fija la frontera silábica entre la segunda y la tercera si las dos últimas forman grupo tautosilábico; en tal caso las dos primeras han de constituir una coda compuesta (*subs-tra-er*).

2. Se considera que sólo la última forma parte de la segunda sílaba y por tanto la primera sílaba acaba en una coda compuesta de tres consonantes extraña al español (*tungs-te-no*).

2.1.2.2.- Grupos vocálicos.

a) Una vocal débil tildada $\{í, ú\}$ seguida o precedida de una fuerte $\{a, e, o\}$ siempre se reparten en dos sílabas —hiato (*dí-a*, *a-ún*).

b) Una vocal fuerte tildada o no seguida de una débil se agrupan en una misma sílaba —diptongo— ya que de existir hiato sería casi exclusivamente un recurso métrico (*cáu-ca-so*, *feu-do*, *vai-vén*).

c) Una vocal débil seguida de una fuerte tildada o no generalmente forman diptongo y para determinar hiato es necesario recurrir a un nivel de conocimiento que trasciende del estrictamente lexicográfico (*vién-do-la*, *ri-ó*, *rue-da*, *ri-a-da*).

d) Dos vocales fuertes distintas se consideran hiato ya que la posibilidad de diptongo es poco frecuente (*co-ar-tar*, *subs-tra-er*).

e) Dos vocales débiles diferentes con una de ellas tildada se considera hiato (*a-cu-í-fe-ro*).

f) Dos vocales débiles diferentes generalmente forman diptongo y para determinar hiato es necesario recurrir a un nivel de conocimiento que trasciende del estrictamente lexicográfico (*rui-do*, *ru-i-na*).

g) Cualquier vocal duplicada —*uu* no figura en español— hace necesario recurrir a un nivel de conocimiento distinto del lexicográfico, aunque frecuentemente pertenecen a la misma sílaba (*vehe-men-te*, *coo-pe-ra-ción*, *le-er*, *pro-hom-bres*).

2.1.3.- Consonantes que requieren un tratamiento especial.

— **Q**: * Comienzo de sílaba en caso de seguirle el carácter *u* y a continuación *e* o *i* (*por-ques*).

* Último carácter de palabra precedido de vocal (*fa-ruq*).

— **Y**: * Comienzo de sílaba y opera como consonante si le sigue una vocal (*pla-ya*).

* A final de palabra actúa como vocal (*ca-ray*).

— **G**: * *gu* constituye una única consonante si el siguiente carácter es *e* o *i* (*gue-rra*).

* *gu* se trata como una consonante más una vocal si el siguiente carácter es una consonante, una *a* o una *o* (*dis-gusto*, *a-gua*, *an-ti-guo*).

* *gü* debe ir seguido de una *e* o una *i*. Ejemplos: *ci-güe-ña*, *pin-güi-no*.

— **H**: * Se obvia si está flanqueada por vocales —es muda (*a-hí-to*).

* Comienzo de sílaba si le antecede *s* y le sigue una *i* o una *u* (*des-hie-lo*, *des-hue-sar*).

* La sílaba comienza en la consonante que precede a la *h* si es distinta de *s* (*a-dhe-ri-do*, *i-nhi-bir*).

2.1.4.- Determinación de la sílaba tónica.

1.- Es aquella sílaba que contiene a la vocal con tilde si la hubiere.

Ejemplos: las agudas *es-pan-té*, *la-drón*, *sa-brás*; las llanas *ca-dá-ver*, *bí-ceps*; las esdrújulas *pé-ta-lo*, *co-mún-men-te*; las sobreesdrújulas *có-me-te-lo*, *fá-cil-men-te*.

2.- Si la palabra no tiene tilde:

2.1.- No tienen sílaba tónica las palabras átonas, indeterminadas a un nivel lexicográfico y casi exclusivamente monosílabos.

2.2.- No átona:

a) Es aguda si termina en consonante distinta de *n* o *s*, acaba en *n* o *s* precedida de otra consonante o se trata de un monosílabo.

Ejemplos: *ca-paz*, *I-sern*, *Ca-nals*, *pie*.

b) Es llana cuando tiene al menos dos sílabas y acaba en vocal o vocal seguida de *n* o *s*. Ejemplos: *ni-ño*, *ger-men*, *es-te-las*.

c) El sufijo *-mente* —identificado en un nivel de conocimiento distinto del lexicográfico— no es tenido en cuenta en la determinación de la tónica —estas palabras tienen dos vocales tónicas la del adjetivo y la de *-mente*. Ejemplos: *su-til-men-te*, *he-roi-ca-men-te*.

2.1.5.- Actualización de la tilde.

Dada una palabra dividida en sílabas se pretende actualizar la tilde —ponerla si no la tiene, quitarla si la tiene, cambiarla de posición si la tiene o dejar la palabra como está tanto si la tiene como si no— al señalar una sílaba como tónica.

Poner la tilde implica: colocarla sin más sobre la única vocal de la sílaba tónica, sobre la vocal fuerte —mantiene diptongo— o sobre la débil —rompe diptongo—; al romper diptongo se retoca la separación silábica.

Quitar la tilde implica: eliminarla y retocar la separación silábica si se forma un nuevo diptongo.

Cambiar de posición la tilde implica las actuaciones correspondientes a quitar y poner una tilde.

1.- Debe llevar tilde si la tónica es anterior a la penúltima sílaba —dejar, poner o cambiar.

2.- La tónica es la penúltima sílaba:

2.1.- Lleva tilde si termina en consonante distinta de **n** o **s**, acaba en **n** o **s** precedida de otra consonante —dejar, poner o cambiar.

2.2.- No lleva tilde si acaba en vocal o vocal seguida de *n* o *s* —dejar o quitar— salvo que precise el acento diacrítico con el fin de diferenciarlos de otros de igual grafía (*cuánto*, *cúyo*, *dónde*, *cuándo*, *cómo*).

3.- La tónica es la última sílaba:

3.1.- Lleva tilde si acaba en vocal o vocal seguida de *n* o *s* —dejar, poner o cambiar.

3.2.- No lleva tilde si termina en consonante distinta de *n* o *s*, acaba en *n* o *s* precedida de otra consonante —dejar o quitar.

4.- Los monosílabos no llevan acento ortográfico, ya que no es preciso señalar que sílaba tiene mayor intensidad (*pan*, *doy*, *fe*, *pie*, *vio*), aunque abundan los que utilizan la tilde diacrítica (*mí*, *tú*, *él*, *sí*, *sé*, *té*, *dé*, *más*, *aún*, *qué*, *quién*, *cuál*).

5.- Los compuestos con el sufijo átono *-mente* llevan tilde si sus primeros componentes lo llevan por sí solos y no en caso contrario.

2.2.- TRATAMIENTO DE LAS FORMAS PRINCIPALES NO VERBALES.

En esta tesis se ha colectado información procedente de diferentes fuentes dotándola de la normalización necesaria para su tratamiento automático. Una misma forma puede proceder de etimologías diferentes y categorías distintas; a efectos morfológicos se considera un único elemento del léxico y por tanto se reúnen las categorías y consecuentemente sus posibilidades de flexión. A efectos de tratamiento morfológico, en la unificación se ha tenido en cuenta la categoría con mayores alternativas de cambio morfológico: primero adjetivo, luego sustantivo y en último lugar adverbio de modo —el resto son invariantes o asimilables morfológicamente a alguna de las anteriores.

2.2.1.- Raíces y terminaciones. Género y número.

A los efectos de generación de formas flexionadas y derivadas, dada una palabra se tomará como raíz la parte del morfema que permanece invariante a las diferentes operaciones de flexión y derivación que pueda sufrir —género, número, aumentativo, diminutivo, despectivo y grado superlativo.

No se consideran invariantes las apariciones o eliminaciones de tildes

im/agen $\Rightarrow$ *im/ágenes*

anfitri/ón $\Rightarrow$ *anfitri/ones*

ni los cambios tipográficos (z por c, g por j, c por qu, etc.).

soe/z $\Rightarrow$ *soe/ces*

falan/ge $\Rightarrow$ *falan/jucha*

ar/co $\Rightarrow$ *ar/quito*

Además de la raíz existe una terminación principal que da lugar a la forma principal. Si la palabra es invariante ante flexiones o no las admite, la terminación principal será vacía. La raíz se considerará vacía si la flexión de la palabra produce alteraciones en su primer carácter —aparición de una tilde.

orden $\Rightarrow$ *órdenes*

Las formas nominales disponen de una lista de cuatro terminaciones principales que proporciona el masculino singular, el femenino singular, el

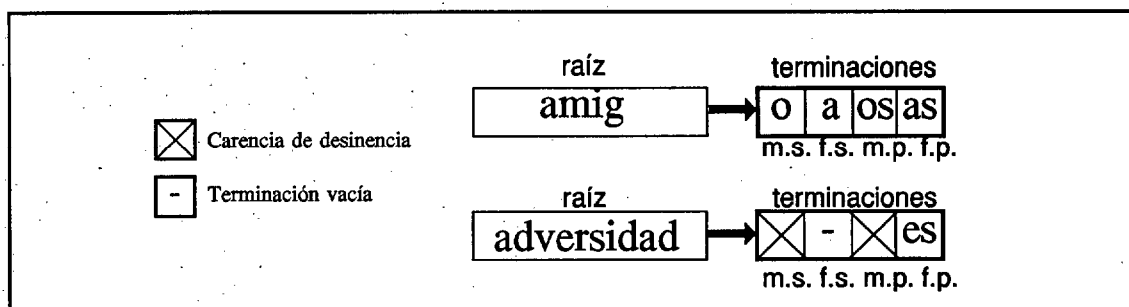


Figura 2: Raíces y terminaciones principales

masculino plural y el femenino plural, figura 2; carecer de alguna de las desinencias se marca con ☒ y puede existir como terminación la ristra nula ☐ indicando que en tal caso no se debe añadir nada a la raíz.

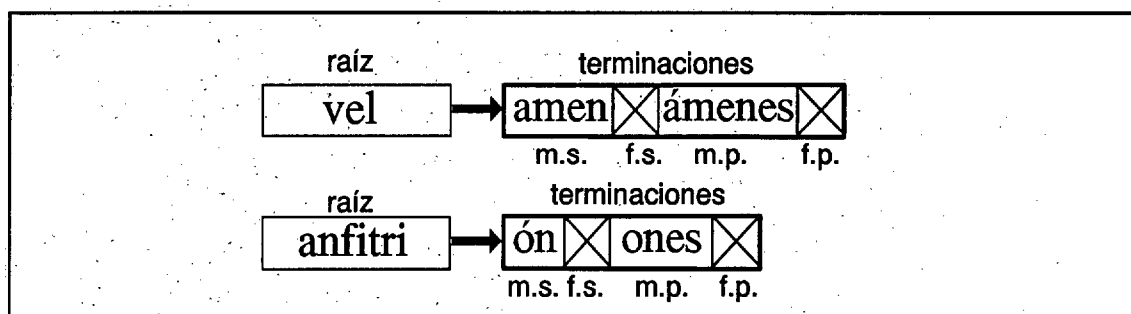


Figura 3: Raíces y terminaciones principales. Tildes

La lista de terminaciones principales se obtiene a partir de: la palabra, su categoría gramatical —determina qué tipo de flexiones puede sufrir, en particular si puede o no cambiar de género—, las reglas de cambio de género —si es de una o de dos terminaciones— y las reglas de cambio de número. Se cuenta con la aparición y desaparición de la tilde como consecuencia de una flexión; esta circunstancia se refleja mediante un cambio en la raíz y consecuentemente en las terminaciones, figura 3.

Las palabras que no admiten flexión tendrán como raíz toda la palabra y no tendrán terminaciones ni principales ni adicionales.

chafalmejas/-

2.2.2.- Sufijos apreciativos.

Los sustantivos y adjetivos tienen además listas de terminaciones para aumentativos, diminutivos y despectivos.

La derivación de aumentativos, diminutivos y despectivos se lleva a cabo mediante la concatenación de la raíz con las terminaciones de derivación correspondientes según las reglas de derivación. Para ello se clasifican las formas principales en función de: a) la lista de terminaciones principales que la caracteriza, b) el número de sílabas, la posición de la tónica y la aparición de diptongos —para la adición de interfijos en el caso de diminutivos— y c) la consonante que encabeza la última sílaba —*c*, *z*, *qu* y *g* tienen tratamiento especial por razones fonológicas.

Por ser tónicos todos estos sufijos se ha de eliminar la tilde de la raíz, si la hubiere.

árbol ⇒ *arbol-ito*

Se han tenido en consideración para la formación de aumentativos las terminaciones: *-azo*, *-ón*, *-ote* y *-acho*; para los diminutivos: *-ito*, *-illo*, *-ico*, *-ín* y *-uelo* con los interfijos *-c-*, *-ec-* o *-ecec-* cuando sea necesario; y para los despectivos: *-ejo* y *-ucho*. Cuando la forma es femenina o plural se realizan las modificaciones correspondientes.

2.2.3.- Grado superlativo.

Los adjetivos poseen una lista de terminaciones adicional para la formación del grado superlativo. Se tienen en cuenta las mismas circunstancias que para la formación de apreciativos —consonantes con tratamiento especial por razones fonológicas y desaparición de tilde por incorporación de sufijo tónico—,

fresc-o $\Rightarrow$ *fresqu-ísimo*

eriaz-o $\Rightarrow$ *eriac-ísimo*

gag-o $\Rightarrow$ *gagu-ísimo*

contigu-o $\Rightarrow$ *contigü-ísimo*

rabón $\Rightarrow$ *rabon-ísimo*

en caso de que el adjetivo acabe en diptongo *io* o *ia*, al unir con la terminación se reduce la doble *i*

ampli-o $\Rightarrow$ *ampl-ísimo*

además se considera la irregularidad genérica de los adjetivos terminados en *-ble* que intercalan una *i* —la terminación es *-bilísimo*—

aprecia-ble $\Rightarrow$ *aprecia-bilísimo*

y de irregularidades específicas para algunos adjetivos en la tabla que sigue se presentan algunas.

Superlativos irregulares			
Adjetivo	Superlativo	Adjetivo	Superlativo
acre	acérrimo	incierto	incertísimo, inciertísimo
amigo	amiciísimo, amiguísimo	infiel	infidelísimo
antiguo	antiquísimo	inico	iniquísimo
ardiente	ardentísimo	íntegro	integérrimo, integrísimo
áspero	aspérrimo, asperísimo	libre	libérrimo
benéfico	beneficentísimo	largo	longísimo, longuísimo
bueno	buenísimo, bonísimo	luciente	lucetísimo
caliente	calentísimo	magnificante	magnificentísimo
célebre	celebérrimo	magnífico	magnificentísimo
cierto	ciertísimo, certísimo	miser	misérrimo
cruel	cruelísimo, crudelísimo	negro	negrísimo, nigérrimo
diestro	destrísimo, diestrísimo	nuevo	novísimo, nuevísimo
enemigo	enemiguísimo, enemicísimo, inimicísimo	parco	parcísimo
ecuo	equísimo	pobre	pobrísim
ferviente	fervientísimo, ferventísimo	pulcro	pulquérrimo
fiel	fidelísimo	reciente	recentísimo, recientísimo
florecente	florentísimo	sabio	sapientísimo
fuerte	fortísimo, fuertísimo	sagrado	sacratísimo
frío	fríísimo, frigidísimo	salubre	salubérrimo
grueso	gruesísimo, grosísimo	simple	simplísimo, simplicísimo
		valiente	valentísimo

2.2.4.- Categorías gramaticales.

Para cada forma principal se dispone de las diferentes categorías gramaticales con las que puede ser utilizada; se obtienen por combinación de las que aparecen en la siguiente tabla.

Categorías gramaticales básicas para formas no verbales		
abreviatura	adverbio relativo de tiempo	locución prepositiva
adjetivo	artículo definido	nombre propio
adjetivo comparativo	artículo indeterminado	nombre propio femenino
adjetivo correlativo	conjunción	nombre propio masculino
adjetivo demostrativo	conjunción adversativa	palabra de otro idioma
adjetivo indefinido	conjunción causal	preposición
adjetivo indefinido plural	conjunción concesiva	pronombre
adjetivo interrogativo	conjunción condicional	pronombre correlativo
adjetivo plural	conjunción continuativa	pronombre demostrativo
adjetivo posesivo	conjunción copulativa	pronombre en plural
adjetivo pronominal	conjunción distributiva	pronombre indefinido
adjetivo relativo	conjunción ilativa	pronombre interrogativo
adjetivo superlativo	conjunción subordinante temporal	pronombre personal
adverbio	contracción de preposición y artículo	pronombre posesivo
adverbio afirmativo	elemento prefijal	pronombre relativo
adverbio comparativo	interjección	pronombre relativo posesivo
adverbio correlativo	locución	sustantivo
adverbio de cantidad	locución adjetiva	sustantivo género ambiguo
adverbio de lugar	locución adverbial	sustantivo género común
adverbio de modo	locución adverbial conjuntiva	sustantivo común en plural
adverbio de modo relativo comparativo	locución adverbial de modo	sustantivo femenino
adverbio de negación	locución adverbial de tiempo	sustantivo femenino plural
adverbio de orden	locución adverbial temporal	sustantivo masculino
adverbio de tiempo	locución conjuntiva	sustantivo masculino plural
adverbio demostrativo	locución conjuntiva causal	sustantivo masculino singular
adverbio interrogativo	locución conjuntiva condicional	sustantivo masculino y femenino
adverbio relativo	locución conjuntiva concesiva	sustantivo plural
adverbio relativo de cantidad	locución latina	

2.3.- TRATAMIENTO DE LOS VERBOS.

La morfología verbal tiene un tratamiento diferente de la nominal e incluye la conjugación —persona, número y tiempo. Las formas pronominalizadas se construyen directamente a partir de la forma conjugada correspondiente y los pronombres enclíticos que incorpore. El diminutivo del gerundio y la flexión del participio como adjetivo verbal se tratan con los mismos criterios que en la flexión nominal. La adjetivación, la sustantivación y la adverbialización se resuelven mediante relaciones con las formas no verbales.

La generación de las formas conjugadas se realizó partiendo de un módulo desarrollado por Santana, Hernández y Rodríguez que producía los tiempos verbales a partir de su infinitivo conociendo una serie de reglas que cumplen los distintos verbos. En un esfuerzo por afianzar la primera versión del conjugador: a) se retocaron algunas reglas para mejorar la conjugación, b) se añadieron nuevos verbos que antes no habían sido tenidos en consideración hasta alcanzar un total de 11.830 infinitivos, c) se llevó a cabo una revisión a fondo de los participios irregulares y d) la categoría gramatical de los verbos fue añadida como dato de salida del conjugador.

El primer paso es determinar si el verbo es regular o irregular ya que el tratamiento es diferente según el caso. Cuando el verbo es regular: a) se

extrae la raíz del verbo separándola de su terminación, b) a partir de la terminación se determina la conjugación a la que pertenece y c) en función de la conjugación y del tiempo deseado se obtienen las terminaciones que hay que añadir a la raíz para construir el tiempo en todas sus personas tanto en singular como en plural. Se presentan los siguientes problemas:

- Cambios ortográficos al unir la terminación a la raíz, se resuelven mediante las *reglas de unión* —comunes a todos los verbos regulares o irregulares.
- Verbos cuya única irregularidad consiste en la acentuación de algunas personas en determinados tiempos; esta irregularidad responde a reglas bien definidas.

Los verbos irregulares se apartan de la norma general aunque afortunadamente las variaciones que presentan poseen un alto grado de sistematización —no carente de excepciones que necesitan reglas específicas—; la solución propuesta se basa en dos pilares fundamentales:

- Establecer tres tipos de reglas para las irregularidades: reglas de *cambio de raíz*, de *excepción* y de *postexcepción*. Además existen las *reglas de unión* y una clasificación de terminaciones verbales que contempla las especificidades propias de los verbos irregulares.

- Determinar un grupo reducido de verbos, usados como modelos para los demás, que representa todas las posibilidades de combinación de los tres tipos de reglas anteriores.

2.3.1.- Reglas de cambio de raíz.

Las reglas de *cambio de raíz* son un conjunto de modificaciones que pueden experimentar las raíces de los verbos irregulares al conjugar los distintos tiempos. Básicamente se corresponden con las irregularidades vocálicas consonánticas y mixtas; como norma general, un verbo irregular utiliza hasta tres reglas de este tipo, una por cada grupo de tiempos irregulares. Indican un cambio en la secuencia de caracteres de la raíz. Ejemplos de reglas de este tipo son:

En la raíz del primer grupo de tiempos irregulares:

Reemplazar:

- ab por ep; *sab-er* $\Rightarrow$ *sep-a*, si hay una c delante ésta cambia por qu

cab-er $\Rightarrow$ *quep-a*

- c por g; *hac-er* $\Rightarrow$ *hag-o*

- e por i; *ceñ-ir* $\Rightarrow$ *ciñ-o*

- e por ie; *acert-ar* $\Rightarrow$ *aciert-o*, o por ye si es a comienzo de palabra

err-ar $\Rightarrow$ *yerr-o*

- ec por ig; *dec-ir* ⇒ *dig-o*
- i por ie; *adquir-ir* ⇒ *adquier-o*
- o por u; *dorm-ir* ⇒ *durm-amos*
- o por ue; *almorz-ar* ⇒ *almuerz-o*, o por **hue** si es a comienzo de palabra *ol-er* ⇒ *huel-o*
- u por ue; *jug-ar* ⇒ *jueg-o*

Añadir:

- e; *v-er* ⇒ *ve-o*
- g; *as-ir* ⇒ *asg-o*
- ig; *ca-er* ⇒ *caig-o*
- v; el verbo *ir* tiene raíz regular vacía, sin embargo para este grupo de tiempos irregulares la raíz es v-: *-ir* ⇒ *v-oy*

Eliminar:

- ab; *hab-er* ⇒ *h-e*

En la raíz del segundo grupo de tiempos irregulares:

Reemplazar:

- a por i; *hac-er* ⇒ *hic-e*
- a por u; *hab-er* ⇒ *hub-e*
- ab por up; *cab-er* ⇒ *cup-e*
- c por j; *produc-ir* ⇒ *produj-e*

- ec por ij; *dec-ir* $\Rightarrow$ *dij-e*
- en por uv; *ten-er* $\Rightarrow$ *tuv-e*
- er por is; *quer-er* $\Rightarrow$ *quis-e*
- on por us; *pon-er* $\Rightarrow$ *pus-e*

Añadir:

- j; *tra-er* $\Rightarrow$ *traj-e*
- uv; *and-ar* $\Rightarrow$ *anduv-e*
- v; la raíz del verbo *ir* para este grupo de tiempos irregulares es
fu-: -ir $\Rightarrow$ *fu-e*

En la raíz del tercer grupo de tiempos irregulares:

Reemplazar:

- dec por d; *dec-ir* $\Rightarrow$ *d-iré*

Añadir:

- d; *val-er* $\Rightarrow$ *vald-ría*

2.3.2.- Reglas de excepción.

En casi todos los verbos irregulares existen algunos tiempos y personas de un mismo grupo de tiempos irregulares que no siguen la regla de cambio de raíz del grupo; sin embargo las excepciones suelen reproducirse en los distintos tiempos del mismo grupo o ir acompañadas de

alguna otra excepción común a varios verbos. La existencia de estos patrones posibilita el establecimiento de determinados modelos de conjugación basados en las posibilidades de normalización que permiten instrumentar un prudente número de *reglas de excepción* —modelizan las excepciones. Podemos destacar entre otras las siguientes:

— Los tiempos
del primer
grupo de

Presente de indicativo				regular	regular	
Presente de subjuntivo				regular	regular	
Imperativo				regular	regular	

tiempos irregulares conjugan la primera y segunda persona del plural con la raíz regular sin alterar.

acert-ar ⇒ *acert-amos, acert-áis*

acert-ar ⇒ *acert-emos, acert-éis*

acert-ar ⇒ *acert-ad*

— Todo el pre-
sente de
indicativo

Presente de indicativo		regular	regular	regular	regular	regular
Imperativo		regular			regular	

excepto la primera persona del singular se conjuga con la raíz regular y también las segundas personas del singular y del plural del imperativo.

as-ir ⇒ *asg-o, as-es, as-e, as-imos, as-is, as-en*

as-ir ⇒ *as-e, as-id*

Además el

verbo *ver*

Pretérito imperfecto indicativo	añadir e	añadir e	añadir e	añadir e	añadir e	añadir e
---------------------------------	----------	----------	----------	----------	----------	----------

y los que se conjugan como él, forman el pretérito imperfecto de indicativo con la regla de cambio de raíz añadir e.

v-er ⇒ *ve-ía, ve-ías, ve-ía, ve-íamos, ve-íais, ve-ían*

Y el verbo

venir y los

que se

Presente de indicativo		e por ie	e por ie			e por ie
Gerundio	e por i					

conjugan como él para formar la segunda y tercera persona del singular y la tercera del plural en el presente de indicativo utilizan la regla de cambio de raíz reemplazar e por ie.

ven-ir ⇒ *vien-es, vien-e, vien-en*

y el gerundio se forma con la regla de cambio de raíz reemplazar e por i.

ven-ir ⇒ *vin-iendo*

— La primera y

segunda

persona

del plural

Presente de indicativo				regular	regular	
Imperativo					regular	
Pretérito indefinido	regular	regular		regular	regular	
Gerundio	e por i					

del presente de indicativo y la segunda del plural del imperativo, se

forman con la raíz regular.

re-ír ⇒ *re-ímos, re-ís*

re-ír ⇒ *re-íd*

Las dos primeras personas del singular y del plural del pretérito indefinido se forman con la raíz regular.

re-ír ⇒ *re-í, re-íste, re-ímos, re-ísteis*

Y el gerundio se forma aplicando la regla de cambio de raíz reemplazar e por i.

re-ír ⇒ *ri-endo*

— El presente
de indica-
tivo se

Presente de indicativo		e por i	e por i	regular	regular	e por i
Imperativo					regular	
Gerundio	e por i					

conjuga utilizando en la primera y segunda persona del plural la raíz regular; en la segunda y tercera del singular y en la tercera del plural se aplica a la raíz reemplazar e por i.

dec-ir ⇒ *dic-es, dic-e, dec-imos, dec-ís, dic-en*

La segunda persona del plural del imperativo se forma con la raíz regular.

dec-ir ⇒ *dec-id*

Y el gerundio se forma aplicando la regla de cambio de raíz reemplazar e por i.

dec-ir ⇒ *dic-iendo*

— La primera y
la segunda
persona
del plural
del pre-

Presente de indicativo				regular	regular	
Presente de subjuntivo				o por u	o por u	
Imperativo				o por u	regular	
Pretérito indefinido	regular	regular		regular	regular	
Gerundio	o por u					

sente de indicativo utilizan la raíz regular.

dorm-ir ⇒ *dorm-imos, dorm-ís*

Para formar la primera y segunda personas del plural del presente de subjuntivo se utiliza la regla de cambio de raíz reemplazar o por u.

dorm-ir ⇒ *durm-amos, durm-aís*

El imperativo usa la regla de cambio de raíz reemplazar o por u para la primera persona del plural y la raíz regular para la segunda.

dorm-ir ⇒ *durm-amos, dorm-id*

Las dos primeras personas, tanto del singular como del plural del pretérito indefinido se forman con la raíz regular.

dorm-ir ⇒ *dorm-í, dorm-iste, dorm-imos, dorm-isteis*

Y el gerundio se forma cambiando o por u en la raíz.

dorm-ir ⇒ *durm-iendo*

— La primera
persona
del singu-

Presente de indicativo	sin ab	regular	regular	regular	regular	regular
Imperativo		regular			regular	

lar del presente de indicativo se forma con la raíz regular sin los dos últimos caracteres y las demás personas del tiempo utilizan la raíz regular.

sab-er ⇒ *s-e, sab-es, sab-e, sab-emos, sab-éis, sab-en*

Las segundas personas del singular y del plural del imperativo aplican la raíz regular.

sab-er ⇒ *sab-e, sab-ed*

— La dos pri-
meras per-
sonas del
plural

Presente de indicativo				regular	regular	
Presente de subjuntivo				e por i	e por i	
Imperativo				e por i	regular	
Pretérito indefinido	regular	regular	e por i	regular	regular	e por i
Gerundio	e por i					

usan para el presente de indicativo la raíz regular, para el presente de subjuntivo la regla de cambio de raíz reemplazar e por i, y para el imperativo la misma regla en la primera y la raíz regular en la segunda.

sent-ir ⇒ *sent-imos, sent-ís*

sent-ir $\Rightarrow$ *sint-amos, sint-áis*

sent-ir $\Rightarrow$ *sint-amos, sent-id*

El pretérito indefinido utiliza la raíz regular, salvo las terceras personas en donde se usa la regla de cambio de raíz reemplazar e por i. Esta misma regla se usa para el gerundio.

sent-ir $\Rightarrow$ *sent-í, sent-iste, sint-ió,*

sent-imos, sent-isteis, sint-ieron

sent-ir $\Rightarrow$ *sint-iendo*

— En el presente de indicativo la

Presente de indicativo	irregular	e por ie	e por ie	regular	regular	e por ie
Imperativo		regular			regular	

primera persona del singular mantiene el cambio de raíz propio de su grupo de tiempos irregulares; la primera y la segunda persona del plural utilizan la raíz regular y el resto aplica la regla de cambio de raíz reemplazar e por ie.

ten-er $\Rightarrow$ *teng-o, tien-es, tien-e, ten-emos, ten-éis, tien-en*

Las segundas personas del imperativo utilizan la raíz regular.

ten-er $\Rightarrow$ *ten-, ten-ed*

- La segunda
persona
del plural

Presente de indicativo					regular	
Imperativo			añade ay	añade ay	regular	añade ay

del presente de indicativo utiliza la raíz regular.

hab-er ⇒ hab-éis

La segunda persona del plural del imperativo se forma con la raíz regular; la tercera del singular, la primera y la tercera del plural usan la regla de cambio de raíz añadir ay.

hab-er ⇒ hay-a, hay-amos, hab-ed, hay-an

- El pretérito
imperfecto

Pretérito imperfecto indicativo	o por u	o por u	o por u	o por u	o por u	o por u
Gerundio	o por u					

de indicativo y el gerundio utilizan la regla de cambio de raíz reemplazar o por u.

podr-ir ⇒ pudr-ía, pudr-ías, pudr-ía,

pudr-íamos, pudr-íais, pudr-ían

podr-ir ⇒ pudr-iendo

- La segunda persona del plural del imperativo no posee raíz (regla exclusiva para el verbo *ir*).

-ir ⇒ -id

2.3.3.- Reglas de unión.

Las *reglas de unión* son fundamentalmente reglas ortográficas y anomalías debidas al sistema fonológico español de aplicación general a los verbos, regulares e irregulares. Indican qué variaciones se deben realizar al concatenar una raíz verbal —modificada o no— con la terminación correspondiente. Por ejemplo:

Reemplazar:

- c por **qu** cuando la raíz acaba en c y la terminación comienza por e o é —verbos de la primera conjugación

marc-ar : marc + é $\Rightarrow$ marqué

- c por z cuando la raíz acaba en **erc** y la terminación comienza por a o por o —verbos de la segunda y tercera conjugación

torc-er : tuerc + a $\Rightarrow$ tuerza

- c por z cuando la raíz acaba en **hic** y la terminación comienza por o

rehac-er : rehic + o $\Rightarrow$ rehízo

- g por j cuando la raíz acaba en g y la terminación comienza por a o por o —verbos acabados en **ger** o **gir**

coleg-ir : colig + a $\Rightarrow$ colija

— **i** por **y** cuando la raíz es vacía y la terminación comienza por **ie**

-ir : - + iendo ⇒ yendo

— **u** por **ü** cuando la raíz acaba en **gu** y la terminación comienza por

e o **é** —verbos de la primera conjugación

averigu-ar : averigu + e ⇒ averigüe

— **u** o **ü** por **uy** cuando la raíz acaba en **u** o **ü** y la terminación

comienza por vocal abierta, salvo que la raíz acabe en **gu** o **fu**

—verbos de la tercera conjugación

argü-ir : argü + o ⇒ arguyo

— **üi** por **uy** cuando la raíz acaba **ü** y la terminación comienza por

i seguida de vocal

argü-ir : argü + iera ⇒ arguyera

— **z** por **c** cuando la raíz termina en **z** y la terminación empieza por

e o **é**

caz-ar : caz + é ⇒ cacé

Añadir:

— **u** cuando la raíz acaba en **g** y la terminación comienza por **e** o **i**

(**é** o **í**) —verbos de la primera conjugación

azog-ar : azog + e ⇒ azogue

- y cuando la raíz acaba en **o** y la terminación comienza con **e**
—verbos de la tercera conjugación

o-ír : o + es ⇒ oyes

- y cuando la raíz está compuesta por una única letra o es **est** y la terminación es **o**

-ir : v + o ⇒ voy

est-ar : est + o ⇒ estoy

- **z** cuando la raíz acaba en **ac**, **oc**, **uc** o **ec** y la terminación comienza por **a**, **á** u **o** —verbos de la segunda y tercera conjugación

complac-er : complac + a ⇒ complazca

conoc-er : conoc + o ⇒ conozco

traduc-ir : traduc+ áis ⇒ traduzcáis

agradec-er : agradec+ as ⇒ agradezcas

Eliminar:

- **e** cuando la raíz acaba en **b** (que no sea **eb** o **rb**) o en **ha**, la terminación tiene más de dos caracteres y empieza por **er**

hab-er : hab + ería ⇒ habría

- e cuando la raíz acaba en **d** y no es la del infinitivo o acaba en **od**, la terminación tiene más de dos caracteres y empieza por **er**

pod-er : pod + ería ⇒ podría

val-er : vald + eré ⇒ valdré

pero *mord-er : mord + eré ⇒ morderé*

- e cuando la raíz acaba en **r**, la terminación tiene más de dos caracteres y empieza por **er**

quer-er : quer + ería ⇒ querría

salvo que la raíz acabe en **rr**

socorr-er : socorr + ería ⇒ socorrería

- i cuando la raíz acaba en **d** y no es la del infinitivo, la terminación tiene más de dos caracteres y empieza por **ir**

sal-ir : sald + iré ⇒ saldré

pero *alud-ir : alud + iré ⇒ aludiré*

- i cuando la raíz acaba en **fu** y la terminación comienza por **i** o **í** seguida de vocal

-ir : fu + iera ⇒ fuera

- i cuando la raíz acaba en **j** y la terminación comienza por **ie** o **ié**

dec-ir : dij + iera ⇒ dijera

- **i** cuando la raíz acaba en **ñ** o **ll** y la terminación comienza por **i** seguida de otra vocal

ceñ-ir : ciñ + iendo ⇒ ciñendo

mull-ir : mull + iera ⇒ mullera

- **u** cuando la raíz acaba en **gu** y la terminación comienza por **a** o por **o** —verbos de la tercera conjugación

segu-ir : sigu + a ⇒ siga

distingu-ir : distingu + o ⇒ distingo

- **doble vocal**, si la raíz acaba en la misma vocal con la que empieza la terminación se simplifica salvo que sea la vocal final de la raíz del infinitivo.

v-er : ve + es ⇒ ves

pero *carte-ar : carte + es ⇒ cartees*

2.3.4.- Reglas de postexcepción.

Las *reglas de postexcepción* entran en juego con posterioridad al resto de las reglas y cubren los casos atípicos. Afectan tanto a la raíz como a la terminación, se aplican después de haber realizado la unión y dependen de la regla de excepción que se utilice. Se pueden citar:

- Los verbos cuya raíz irregular acaba en **ng** forman la segunda

persona del singular del imperativo con la raíz regular desprovista de terminación.

ven-ir : veng- ⇒ ven-

- Los verbos cuya raíz irregular acaba en *lg* forman la segunda persona del singular del imperativo eliminando la *g* y con la terminación correspondiente.

val-er : valg- ⇒ val-e

- Los verbos cuya raíz irregular acaba en *ag* forman la segunda persona del singular del imperativo cambiando *g* por *z* y sin añadir terminación.

hac-er : hag- ⇒ haz-

- La segunda persona del singular del imperativo del verbo *decir* es *di* y los verbos que siguen el modelo de *decir* usan la raíz regular con la regla reemplazar e por i.

predec-ir : predig- ⇒ predic-e

- Los verbos que en el presente de indicativo la primera persona del singular mantiene el cambio de raíz propio de su grupo de tiempos irregulares, la primera y la segunda persona del plural utilizan la raíz regular y el resto aplica a la raíz reemplazar e por ie; forman la segunda persona del singular del imperativo

con la raíz del infinitivo sin terminación.

ten-er : teng- ⇒ ten-

— El presente de indicativo del verbo *ser* es:

soy, eres, es, somos, sois, son

— El pretérito imperfecto de indicativo del verbo *ser* es:

era, eras, era, éramos, erais, eran

— El pretérito imperfecto de indicativo del verbo *ir* es:

iba, ibas, iba, íbamos, ibais, iban.

2.3.5.- Participios irregulares.

Hay verbos que presentan *participio irregular*, tanto por irregularidad en la raíz del participio como por no presentar las terminaciones regulares *-ado, -ido*

poner ⇒ puesto

e incluso algunos poseen dos participios uno regular y otro irregular.

imprimir ⇒ imprimido / impreso

Las posibilidades de sistematización parecen escasas; sin embargo son de aplicación algunas reglas como:

— En los verbos tales que el presente de indicativo salvo la primera persona del singular y las segundas personas del singular y del

plural del imperativo se conjugan con la raíz regular ocurre que:

+ si la raíz del infinitivo termina en *hac* o *fac*, el participio se forma con la terminación **echo** y la raíz pierde la **ac** final.

satisfac-er $\Rightarrow$ *satisf-echo*

+ si la raíz del infinitivo termina en *pon*, el participio se forma con la terminación **uesto** y la raíz pierde la **on** final.

superpon-er $\Rightarrow$ *superp-uesto*

+ si la raíz del infinitivo termina en *v*, el participio se forma con la terminación **isto**.

prev-er $\Rightarrow$ *prev-isto*

— En los verbos tales que el presente de indicativo se conjuga utilizando en la primera y segunda persona del plural la raíz regular y en la segunda y tercera del singular y en la tercera del plural se aplica la regla de cambio de raíz reemplazar e por **i**, forman el participio con la terminación **ho** y modifican la raíz según reemplazar e por **i**.

dec-ir $\Rightarrow$ *dic-ho*

2.3.6.- Reglas de acentuación.

Ciertos verbos presentan irregularidades en la acentuación ortográfica: la aparición de una tilde sobre *u* o *i* en determinados tiempos y personas. Siguen una norma: la ruptura de un diptongo (*enraízo*); debe extenderse el término diptongo a aquellos casos en que existe una *h* intercalada en el grupo vocálico (*cabrahígo*). Para las terminaciones desprovistas de tilde estos verbos cumplen las siguientes reglas de unión:

- Si la raíz de segunda persona del plural del imperativo termina en vocal fuerte aparece una tilde sobre la *i* de la terminación —verbos de la tercera conjugación.

re-ír : re + id ⇒ reíd

- Si el tiempo pertenece al primer grupo de tiempos irregulares, la penúltima vocal de la unión (raíz+terminación) es la vocal débil de un diptongo, no es el primer carácter de la raíz y no proviene de la regla de cambio de raíz añadir *ig*, se tilda.

actu-ar : actu + o ⇒ actúo

- Si el tiempo es el pretérito indefinido o se trata de una forma impersonal, la raíz termina en vocal abierta no acentuada y la terminación comienza con vocal cerrada no acentuada, se tilda ésta.

le-er : le + iste $\Rightarrow$ leíste

- Existen verbos que no siguen estas reglas.

cambi-ar : cambi + o $\Rightarrow$ cambio

2.3.7.- Terminaciones.

El estudio de las terminaciones que utilizan los verbos a la hora de formar sus tiempos revela que existen 33 conjuntos de terminaciones usados en la conjugación de los verbos regulares —un conjunto de seis terminaciones para cada tiempo en cada una de las tres conjugaciones.

PRIMERA CONJUGACIÓN						
Primer grupo de tiempos irregulares			Segundo grupo de tiempos irregulares			
Presente de Indicativo	Presente de Subjuntivo	Imperativo	Pretérito Indefinido	Pretérito Imperfecto de Subjuntivo		Futuro Imperfecto de Subjuntivo
o as a amos áis an	e es e emos éis en	* a e emos ad en	é aste ó amos asteis aron	ara aras ara áramos arais aran	ase ases ase ásemos aseis asen	are ares are áremos areis aren
Tercer grupo de tiempos irregulares			Pretérito Imperfecto de Indicativo		Formas Impersonales	
Futuro Imperfecto de Indicativo	Condicional Simple					
aré arás ará aremos aréis arán	aría arías aría aríamos aríais arían	aba abas aba ábamos abais aban	ar ando ado * * *			
" * " Significa que no existe la forma correspondiente						

SEGUNDA CONJUGACIÓN						
Primer grupo de tiempos irregulares			Segundo grupo de tiempos irregulares			
Presente de Indicativo	Presente de Subjuntivo	Imperativo	Pretérito Indefinido	Pretérito Imperfecto de Subjuntivo		Futuro Imperfecto de Subjuntivo
o es e emos éis en	a as a amos áis an	* e a amos ed an	í iste ió imos isteis ieron	iera ieras iera iéramos ierais ieran	iese ieses iese iésemos ieseis iesen	iere ieres iere iéremos iereis ieren
Tercer grupo de tiempos irregulares			Pretérito Imperfecto de Indicativo		Formas Impersonales	
Futuro Imperfecto de Indicativo	Condicional Simple					
eré erás erá eremos eréis erán	ería erías ería eríamos eríais erían	ía ías ía íamos íais ían	er iendo ido * * *			
" * " Significa que no existe la forma correspondiente						

TERCERA CONJUGACIÓN						
Primer grupo de tiempos irregulares			Segundo grupo de tiempos irregulares			
Presente de Indicativo	Presente de Subjuntivo	Imperativo	Pretérito Indefinido	Pretérito Imperfecto de Subjuntivo		Futuro Imperfecto de Subjuntivo
o es e imos ís en	a as a amos áis an	* e a amos id an	í iste ió imos isteis ieron	iera ieras iera iéramos ierais ieran	iese ieses iese iésemos ieseis iesen	iere ieres iere iéremos iereis ieren
Tercer grupo de tiempos irregulares			Pretérito Imperfecto de Indicativo		Formas Impersonales	
Futuro Imperfecto de Indicativo	Condicional Simple					
iré irás irá iremos iréis irán	iría irías iría iríamos iríais irían		ía ías ía íamos íais ían		ir iendo ido * * *	
" * " Significa que no existe la forma correspondiente						

Se han determinado otros conjuntos de terminaciones para los verbos irregulares, aunque muchos usan el mismo sistema de terminaciones que los regulares.

TERMINACIONES IRREGULARES																	
Presente de Indicativo							Presen. Subjun.	Imperativo				Pretérito Indefinido			Pretérito Imperfecto Subjuntivo		Futuro Imperf. Subjun.
oy	é	oy	é	é	e	oy	aya	*	*	*	*	e	i	i	era	ese	ere
as	es	ás	és	es	as	es	ayas	á	az	i	e	iste	iste	iste	eras	eses	eres
a	é	á	é	e	a	es	aya	é	a	a	aya	o	io	e	era	ese	ere
amos	emos	amos	emos	emos	emos	omos	ayamos	amos	amos	amos	ayamos	imos	imos	imos	éramos	ésemos	éremos
áis	eis	áis	éis	éis	éis	ois	ayáis	ad	ed	id	id	isteis	isteis	isteis	erais	eseis	ereis
an	en	án	en	en	an	on	ayan	én	an	an	ayan	ieron	ieron	ieron	eran	esen	eren

Las terminaciones se encuentran en un fichero agrupadas de seis en seis por cada tiempo.

2.3.8.- Modelos de conjugación utilizados.

Todos los verbos modelo se tienen en un fichero en el que se dispone de cada uno de ellos con el siguiente formato:

verbo

regla de cambio de raíz del primer grupo de tiempos irregulares

regla de cambio de raíz del segundo grupo de tiempos irregulares

regla de cambio de raíz del tercer grupo de tiempos irregulares

terminaciones

regla de excepción

La lista de verbos modelo que se ha usado es la siguiente:

Verbos modelo					
acertar	caber	dar	hacer	podrir	ser
adquirir	caer	decir	ir	poner	tener
aducir	cambiar	dormir	jugar	producir	traer
almorzar	ceñir	entender	mover	querer	valer
andar	concebir	errar	oír	reír	venir
asir	concluir	estar	oler	saber	ver
auxiliar	conocer	haber	poder	sentir	yacer

2.3.9.- Conjugación.

El tratamiento se resuelve a través de tres módulos independientes: el conjugador propiamente dicho y dos preprocesadores encargados de preparar las estructuras a partir de la información que acompaña a los verbos. La figura 4 muestra la relación entre los tres módulos.

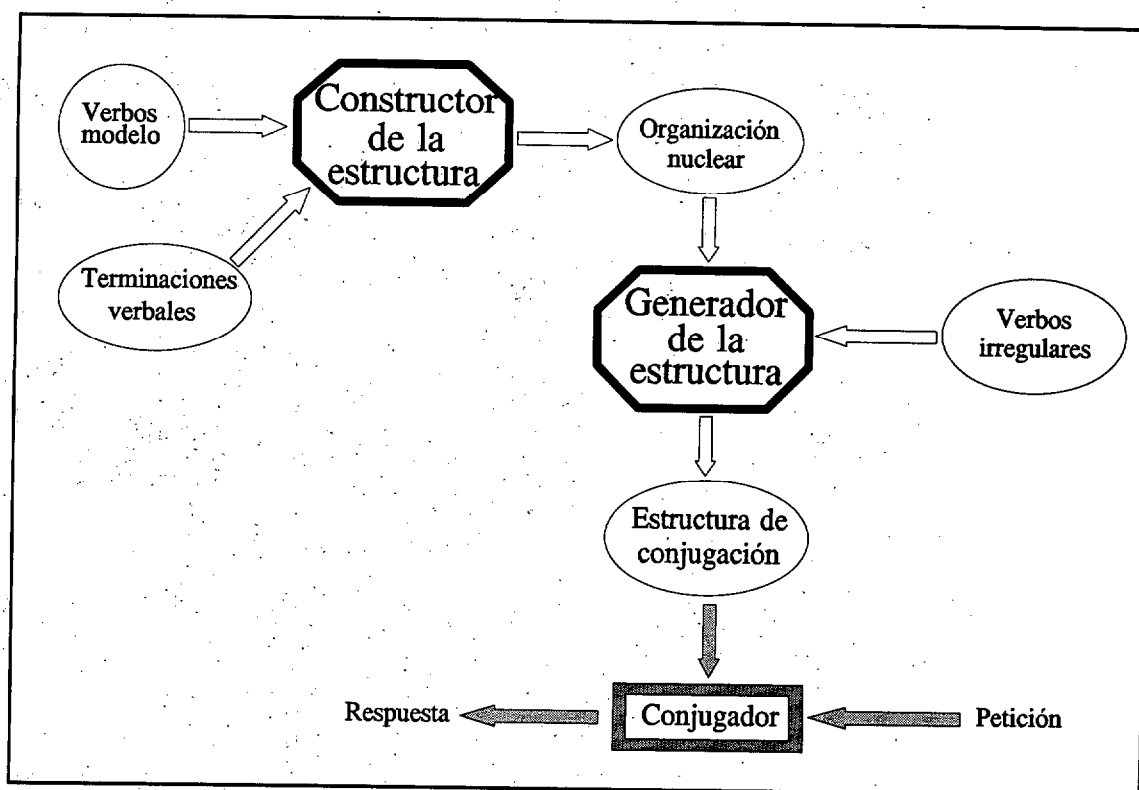


Figura 4: Desarrollo del Conjugador

El primero de los preprocesadores se encarga de construir una estructura nuclear formada por los verbos modelo y las terminaciones verbales.

El segundo preprocesador trata el fichero que contiene los verbos irregulares relacionados con sus modelos, de la forma:

verbo irregular:verbo modelo

y con la organización nuclear genera la estructura de conjugación lista para ser usada por el módulo conjugador. Nótese que es preciso alimentar además al conjugador con el infinitivo y el tiempo que se desea conjugar.

La estructura de conjugación permite determinar para cada verbo irregular: sus reglas de cambio de raíz, las terminaciones que usan los distintos tiempos de cada grupo de tiempos irregulares y las reglas de excepción y postexcepción que son de aplicación, figura 5. Los verbos

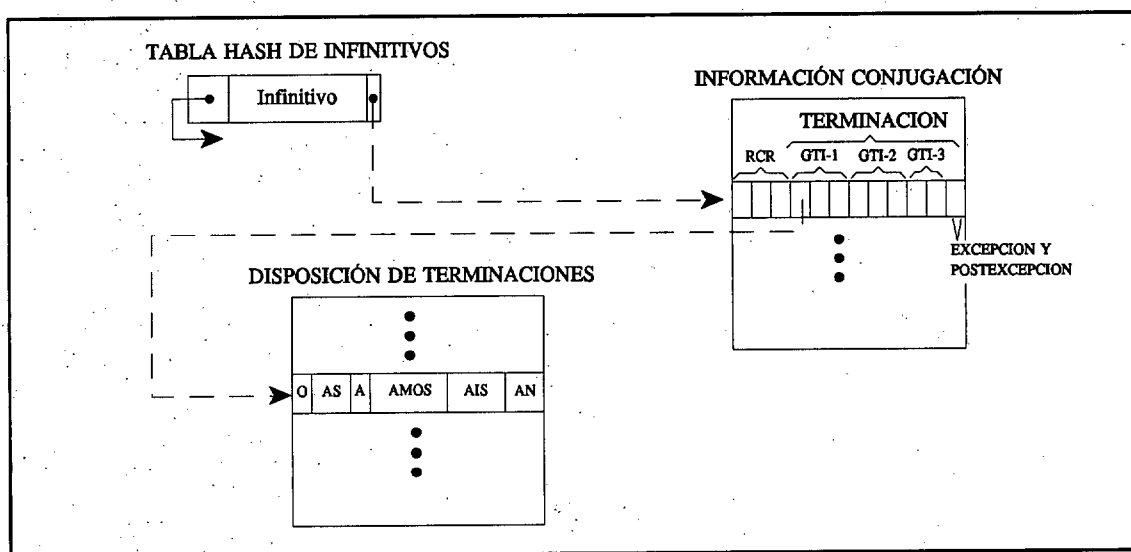


Figura 5: Estructura de Conjugación

regulares sólo se almacenan a efectos de fijar su existencia puesto que el tiempo y la conjugación determinan las terminaciones que han de usarse. El acceso se lleva a cabo mediante una tabla hash que direcciona a la información de conjugación; allí se apunta a las terminaciones que procedan.

Gracias a la estructura, el conjugador determina si se trata de un verbo regular o irregular. En caso de que sea irregular no hay más que seguir la referencia contenida para determinar las reglas que se aplican y el conjunto de terminaciones que se utiliza según el tiempo que se conjuga.

Entran entonces en juego la secuencia de acciones: a) se efectúa el cambio de raíz, b) se aplican las excepciones y postexcepciones, c) se realiza la unión con las terminaciones y d) se llevan a cabo las postexcepciones. La figura 6 muestra el proceso seguido en la conjugación.

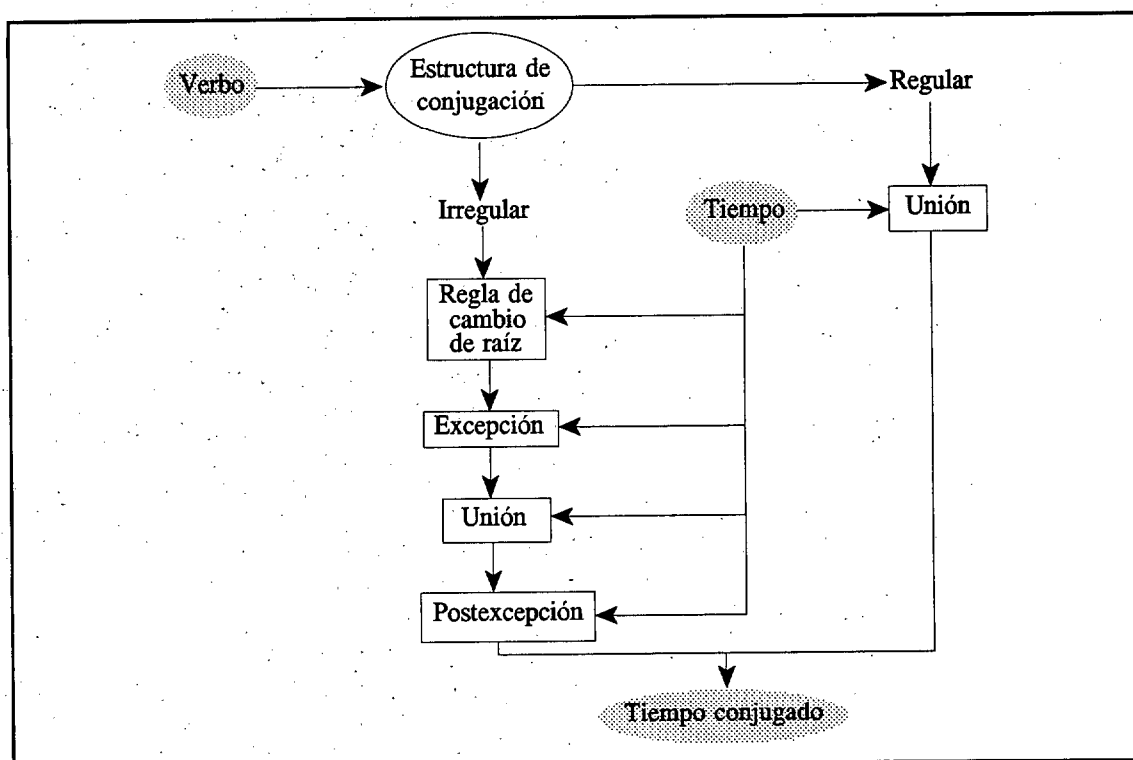


Figura 6: Proceso de conjugación

Si por el contrario es regular, no hay que aplicar cambio de raíces y, en consecuencia, tampoco excepciones o postexcepciones; las únicas acciones a tomar son las de realizar las uniones correctamente.

Para los verbos defectivos se mantiene una señalización que informa de qué personas y números no existen en cada uno de los tiempos de manera que no sean conjugados.

2.3.10.- Categorías gramaticales.

Para cada verbo se almacenan las diferentes categorías con las que puede ser usado. En esta tesis se consideran las siguientes: **absoluto, factitivo, frecuentativo, impersonal, intransitivo, pasivo, pronominal, recíproco, transitivo y unipersonal**. Así por ejemplo: *comer* es intransitivo, transitivo o pronominal.

2.4.- ESTABLECIMIENTO DE RELACIONES MORFOLÓGICAS ENTRE DIFERENTES FORMAS PRINCIPALES.

Hasta ahora se ha expuesto cómo obtener para cada forma principal sus formas flexionadas y derivadas. A continuación se expone cómo conseguir relaciones morfológicas entre diferentes formas principales que permiten obtener otras formas relacionadas morfológicamente con una dada; en particular posibilitan realizar derivaciones entre las diferentes categorías gramaticales: sustantivaciones, adjetivaciones, adverbializaciones o verbalizaciones.

Las relaciones que se establecen son:

- Para los verbos: sustantivos y adjetivos relacionados y verbos relacionados mediante cambio prefijal.
- Para los sustantivos: verbos relacionados.
- Para los adjetivos: verbos relacionados y adverbio de modo correspondiente.
- Para los adverbios de modo: el adjetivo del que provienen.

A fin de conseguir las citadas listas de formas relacionadas se tienen en cuenta las siguientes consideraciones:

- Todo adjetivo terminado en **-ble** posee un sustantivo femenino relacionado cuya terminación es **-bilidad** y viceversa.

— Con la terminación *-mente* se forman adverbios de modo a partir de la forma femenina de los adjetivos.

Existen terminaciones que, a partir de las raíces de los verbos, posibilitan su sustantivación y su adjetivación; básicamente, son las expuestas en la tabla siguiente. A veces deben tenerse en cuenta modificaciones en función de la conjugación a la que pertenece el verbo a cuya raíz se unen, posibles desinencias femeninas adicionales y aparición en ocasiones de una tilde.

Terminaciones básicas sustantivación y adjetivación					
-a	-ara	-dero	-ez	-isto	-rro
-aba	-arado	-dicio	-eza	-istrajo	-sco
-abreña	-arajo	-dizo	-ezco	-itín	-tad
-áceo	-ardo	-do	-ezno	-itis	-tano
-achal	-aredo	-dor	-ezuelo	-ito	-terio
-achar	-arín	-dura	-í	-itud	-tivo
-achín	-arino	-e	-ia	-ivo	-tor
-acho	-ario	-edal	-iano	-izal	-torio
-achuelo	-arrado	-edar	-icio	-izar	-triz
-aco	-arrajo	-edero	-ico	-izco	-ucho
-aico	-arro	-ejo	-icoso	-izo	-uco
-aino	-aruco	-el	-iego	-mento	-udo
-aje	-asco	-ellín	-iendo	-miento	-uelco
-ajo	-astre	-encho	-iense	-o	-ueldo
-ajón	-astro	-encia	-iente	-ol	-uelo
-ajoso	-atario	-eno	-lento	-olento	-uelto
-al	-ático	-ense	-ierto	-olo	-uesto
-alla	-ato	-ente	-iesto	-ón	-ugo
-ambre	-azal	-ento	-igüeño	-onazo	-ujado
-ambra	-azar	-eño	-ijo	-oncho	-ujo
-amen	-azco	-er	-il	-ongo	-ujón
-án	-azgo	-ería	-illo	-or	-ulento
-anchín	-azo	-ericio	-imbre	-orio	-ullo
-ancho	-bilidad	-erio	-imo	-oro	-umbre
-ancia	-ble	-erizo	-ín	-orral	-uno
-anco	-bundo	-ero	-ina	-orrar	-uría
-ando	-cia	-erro	-íneo	-orrial	-uro
-áneo	-cía	-és	-ino	-orriar	-urro
-ango	-cico	-esa	-ío	-orrio	-urrón
-ano	-cillo	-esco	-irente	-orritín	-usco
-ante	-cín	-esino	-irorio	-orro	-uzco
-antez	-ción	-estre	-irritín	-orrotín	-uzo
-anza	-cionario	-ete	-isco	-oso	-yente
-año	-cito	-etín	-isión	-otazo	-zuelo
-ar	-dad	-eto	-ismo	-ote	...

También hay terminaciones para obtener verbos a partir de nombres —verbalización— (*-achar, -ajar, -ar, -azar, -ear, -ecer, -edar, -eguear, -ejar, -ificar, -iguar, -iquear, -iscar, -itar, -izar, -orrrar, -orriar, -otear, -ucar, -ujar, -uscar, -uzar,...*).

A partir de estos conjuntos de terminaciones se realiza el siguiente procedimiento: a) de cada verbo se elimina su terminación —*-ar, -er* o *-ir*— a fin de obtener su raíz gramatical, b) si es posible se eliminan aquellas terminaciones de verbalización que pudieran poseer a fin de obtener una raíz primitiva, c) a todas las raíces obtenidas se unen las terminaciones de sustantivación y adjetivación teniendo en cuenta los cambios ortográficos que se pueden presentar en la unión (cambios de *c* por *qu*, de *z* por *c*, de *g* por *j*, de *g* por *gu*,...), d) se considera la incorporación o eliminación de algunos prefijos —*a, de, en*,...— que a veces se utilizan en los procesos de sustantivación, adjetivación o verbalización y e) se busca el resultado de la unión en la base de datos de formas no verbales para su toma en consideración si guarda conexión semántica con el infinitivo. De esta manera se produce para cada verbo un conjunto de sustantivos y de adjetivos que constituyen sus relaciones y además un conjunto de verbos para cada sustantivo y adjetivo. Adicionalmente, para cada adjetivo se forma su adverbio de modo correspondiente.

afortunar $\Rightarrow$ *afortunado, fortuna*

humedecer $\Rightarrow$ *humedad, humedal, humedecido, humedecimiento, húmedo*

humanizar $\Rightarrow$ *humanal, humanidad, humanista, humanitario, humanización, humanizado, humanizador, humano*

Por otra parte, se eliminan los prefijos que los verbos puedan poseer según las reglas de prefijación comentadas en la sección 1.10; si el resultado es un verbo de la base y guarda relación semántica con el originario se establece la relación prefijal entre ambos.

retransmitir $\Rightarrow$ *transmitir, emitir*

Todas las relaciones se almacenan en la estructura léxica según el esquema expuesto en la figura 1; se usan en la operaciones de derivación que implican cambio de categoría gramatical.

3.- SOLUCIÓN AUTOMATIZADA DEL PROBLEMA.

3.1.- DISEÑO E IMPLEMENTACIÓN DEL RECONOCEDOR.

El reconocedor morfológico actúa sobre una palabra de entrada por medio de un *segmentador* que la descompone en: a) los posibles pares raíz-terminación, b) los prefijos que pudieran poseer y c) los pronombres enclíticos que llevara en caso de ser una forma conjugada, figura 4. La raíz

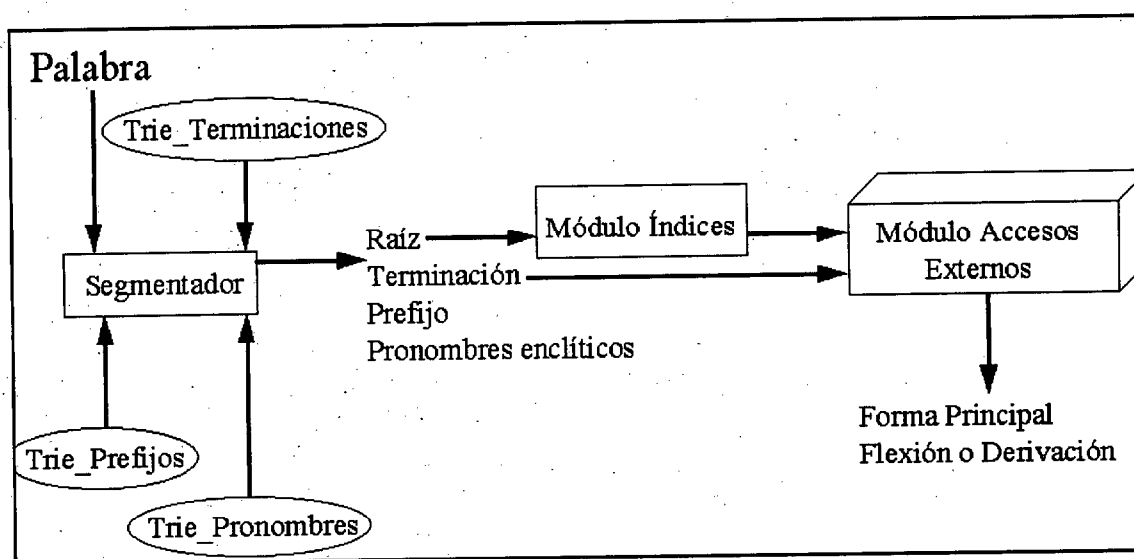


Figura 4: Esquema del Reconocedor

pasa al *módulo de índices* que determina su localización. El *módulo de accesos externos* comprueba: a) si la raíz admite la terminación, b) a qué flexión o derivación corresponde y c) cuál es la forma principal de la que proviene.

El *segmentador* se apoya en tres estructuras de árbol trie, una para las terminaciones, otra que soporta los prefijos y otra para los pronombres enclíticos. Los posibles pares raíz-terminación se obtienen tras confirmar la existencia de una terminación en el trie de terminaciones.

amare $\Rightarrow$ *amar+e* o *am+are*

Tal estructura utiliza todas las terminaciones necesarias para poder obtener la raíz de cualquier forma flexionada o derivada; se reúnen seguidamente las terminaciones según su aplicación a formas verbales o no verbales.

LISTA DE TERMINACIONES							
PARA VERBOS				PARA FORMAS NO VERBALES			
a	ases	hay	irías	a	ejo	cito	ecicos
á	aste	i	ís	e	eja	cita	ecicas
aba	asteis	í	iste	o	ejos	citos	ecín
abais	aya	ía	isteis	s	ejas	citas	ecina
ábamos	ayáis	íais	jera	as	ucho	cillo	ecines
aban	ayamos	íamos	jeráis	es	ucha	cilla	ecinas
abas	ayan	ían	jéramos	os	uchos	cillos	ezuelo
ad	ayas	ías	jeran	mente	uchas	cillas	ezuela
adamente	az	íd	jeras	amente	ito	cico	ezuelos
ado	e	íd	jere	emente	ita	cica	ezuelas
ais	é	idamente	jereis	idad	itos	cicos	ececito
áis	ed	idamente	jéremos	idades	itas	cicas	ececita
al	eis	ido	jerén	ísimo	illo	cín	ececitos
amente	éis	ído	jeres	ísima	illa	cina	ececitas
amos	emos	iendo	jerón	ísimos	illos	cines	ececillo
an	en	iera	jese	ísimas	illas	cinas	ececilla
án	én	ierais	jeseis	azo	ico	zuelo	ececillos
ando	er	iéramos	jésemos	aza	ica	zuela	ececillas
ar	era	ieran	jesen	azos	icos	zuelos	ececico
ara	erá	ieras	jeses	azas	icas	zuelas	ececica
ará	erais	iere	o	ón	ín	ecito	ececicos
arais	éramos	iereis	ó	ona	ina	ecita	ececicas
áramos	eran	iéremos	ois	ones	ines	ecitos	ececón
aran	erán	ieren	omos	onas	inas	ecitas	ececina
arán	eras	ierés	on	ote	uelo	ecillo	ececines
aras	erás	ieron	ón	ota	uela	ecilla	ececinas
arás	ere	iese	os	otes	uelos	ecillos	ecezuelo
are	eré	ieseis	oy	otas	uelas	ecillas	ecezuela
aré	ereis	iésemos	yendo	acho		ecico	ecezuelos
areis	eréis	iesen	yera	acha		ecica	ecezuelas
aréis	eremos	ieses	yerais	achos			
aremos	éremos	imos	yéramos	achas			
áremos	eren	ímos	yeran				
aren	eres	io	yeras				
ares	ería	ió	yere				
aría	eríais	ir	yereis				
aríais	eríamos	ír	yéremos				
aríamos	erían	irá	yeren				
arían	erías	irán	yeres				
arías	eron	irás	yeron				
aron	es	iré	yese				
as	és	iréis	yeseis				
ás	ese	iremos	yésemos				
ase	eseis	iría	yesen				
aseis	ésemos	iríais	yeses				
ásemos	esen	iríamos	yó				
asen	eses	irían					

Si la localización de la palabra de entrada resulta infructuosa, el reconocedor realiza el tratamiento de los prefijos: los aísla situándolos en el índice correspondiente y procede a buscar la palabra desprovista de ellos.

archimalo $\Rightarrow$ *a+rchimalo* o *arc+himalo* o *archi+malo*

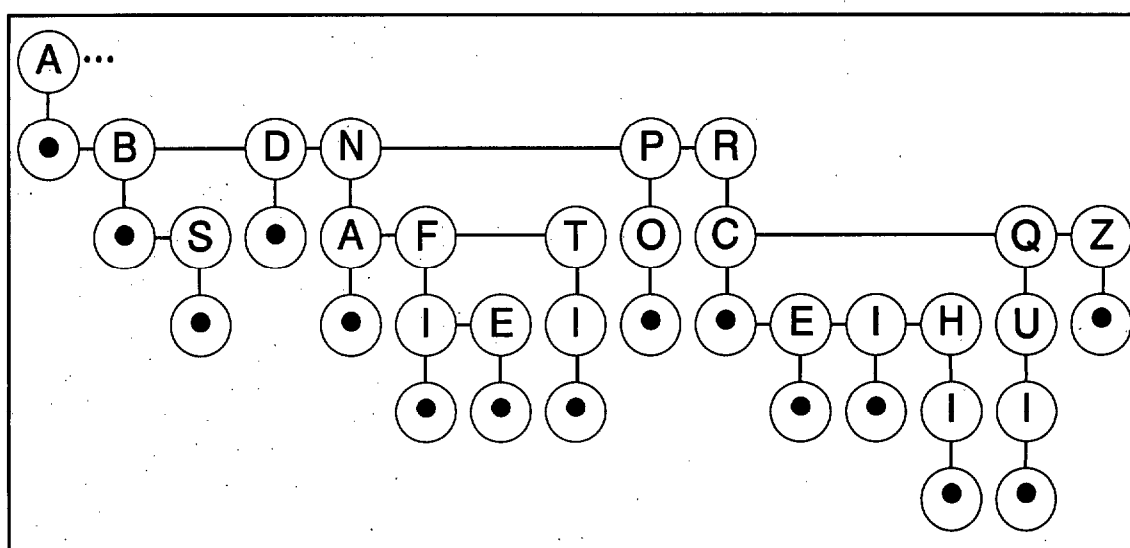


Figura 5: Trie de prefijos

La figura 5 presenta una vista del trie de prefijos que comienza por a. El "•" señala el final de la cadena de caracteres correspondiente a un prefijo.

El trie para los prefijos se construye tomando los caracteres desde el principio hacia el final de la palabra, mientras que los trie de pronombres y terminaciones se organizan desde el último carácter hacia el primero —el sentido lo impone la segmentación por evolucionar en tales direcciones. Se estructuran como prefijos los enumerados en la sección 1.10 dedicada a la prefijación y que se detallan seguidamente.

LISTA DE PREFIJOS				
a-	bis-	entre-	ob-	sota-
ab-	cata-	entro-	opisto-	soto-
abs-	circun-	epi-	para-	sub-
ad-	cis-	equi-	pos-	super-
ana-	citra-	es-	post-	supra-
anfi-	co-	ex-	pre-	sus-
ante-	con-	exo-	pro-	tele-
anti-	contra-	extra-	proto-	trans-
apo-	de-	hemi-	re-	tras-
arc-	des-	hiper-	requete-	ultra-
arce-	di-	hipo-	rete-	vi-
arci-	dia-	in-	retro-	vice-
archi-	dis-	inter-	semi-	viz-
arqui-	ecto-	intro-	sin-	yuxta-
arz-	en-	iso-	so-	za-
bi-	endo-	meta-	sobre-	zam-

Cuando trata con verbos busca además la existencia de pronombres enclíticos en el trie correspondiente y aplica las reglas de corte para obtener la forma correcta.

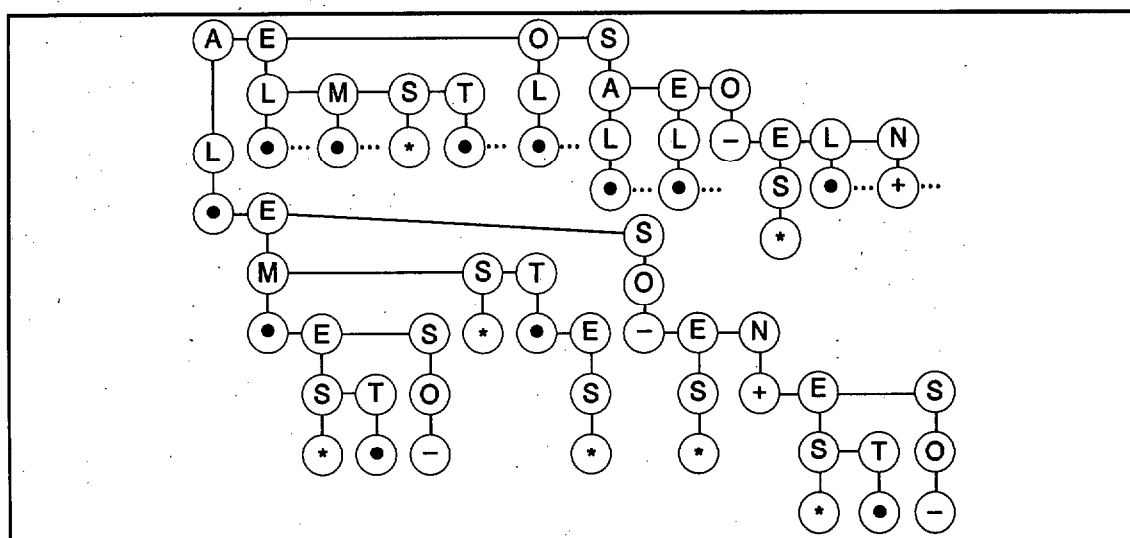


Figura 6: Trie de pronombres enclíticos

La figura 6 muestra una vista de la organización del trie de

pronombres donde se representan: *-la, -las, -le, -les, -lo, -los, -me, -mela, -nos, -nosla, -os, -osla, -osmela, -osnosla, -se, -sela, -semela, -senosla, -seos, -seosla, -setela, -te, -tela, -temela* y *-tenosla*; los símbolos "•, -, +, *" indican: a) la finalización de la cadena y b) la regla de unión-corte que se ha de aplicar a la forma conjugada según el pronombre o conjunto de pronombres del que se trate.

encontréla ⇒ *encontré+la* (•)

amaos ⇒ *amad+os* (-)

temámonos ⇒ *temamon+os* o *temamos+nos* (+)

partámosela ⇒ *partámose+la* o *partamos+sela* (*)

Las combinaciones de pronombres —individuales, pares o ternas— que se indizan son las estudiadas con anterioridad en el apartado 1.8.3 dedicado a los pronombres enclíticos y que se listan a continuación.

LISTA DE PRONOMBRES ENCLÍTICOS							
Un pronombre	Dos pronombres			Tres pronombres			
se	seme	mele	teme	semele	setele	temele	osmele
me	senos	mela	tenos	semela	setela	temela	osmela
te	sete	melo	tele	semelo	setelo	temelo	osmelo
le	seos	meles	tela	semeles	seteles	temeles	osmeles
la	sele	melas	telo	semelas	setelas	temelas	osmelas
lo	sela	melos	teles	semelos	setelos	temelos	osmelos
nos	selo	nosle	telas	senosle	seosme	tenosle	osnosle
os	seles	nosla	telos	senosla	seosnos	tenosla	osnosla
les	selas	noslo	osme	senoslo	seosle	tenoslo	osnoslo
las	selos	nosles	osnos	senosles	seosla	tenosles	osnosles
los		noslas	osle	senoslas	seoslo	tenoslas	osnoslas
		noslos	osla	senoslos	seosles	tenoslos	osnoslos
			oslo	seteme	seoslas		
			osles	setenos	seoslos		
			oslas				
			oslos				

3.2.- DISEÑO E IMPLEMENTACIÓN DEL GENERADOR.

La entrada al generador la constituye a) una forma principal, b) la flexión o derivación correspondiente, c) los prefijos y d) los pronombres, figura 7.

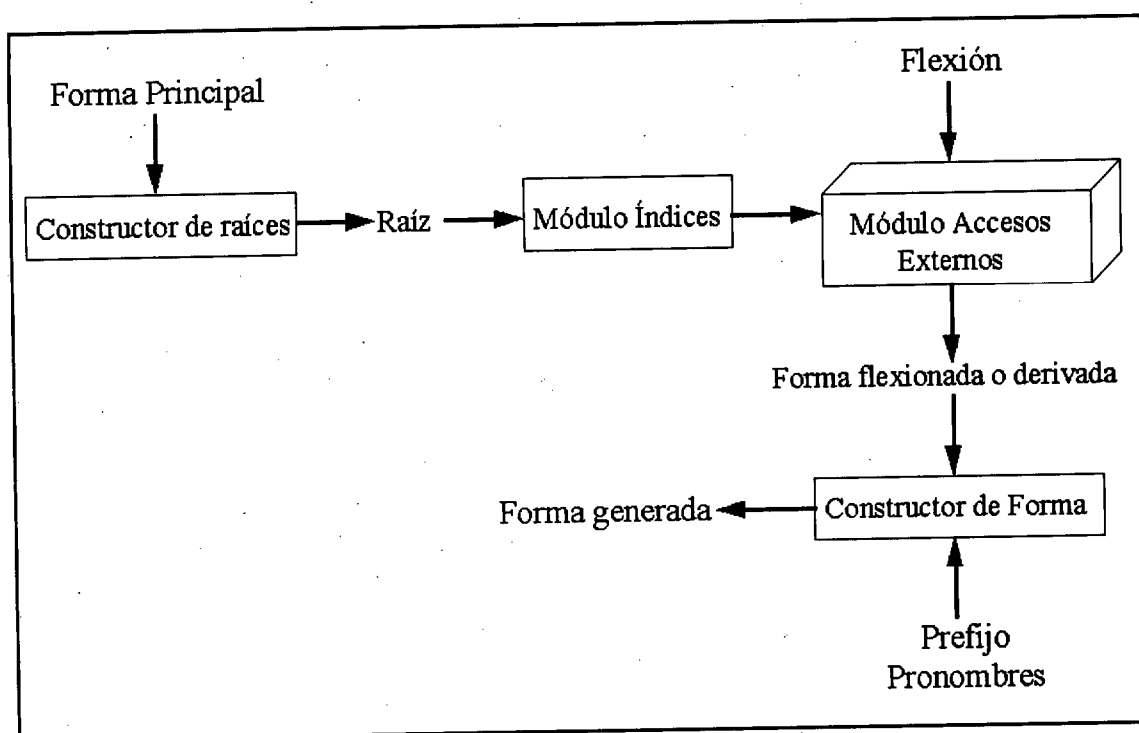


Figura 7: Esquema del Generador

De la forma principal obtiene su raíz mediante un *constructor de raíces* que elimina la terminación.

probar ⇒ *prob-*

ir ⇒ *ir-*

amigo ⇒ *amig-*

El *módulo de índices* localiza la raíz en la base adecuada. A través del *módulo de accesos externos* se obtiene la forma flexionada o derivada.

Para determinar la *forma generada*, el *constructor de forma* añade el prefijo o conjunto de prefijos aplicando las reglas de unión detalladas en la sección 1.10 que pueden afectar al último carácter del prefijo,

in- + lógico $\Rightarrow$ ilógico

al primero de la palabra,

anti- + robo $\Rightarrow$ antirrobo

obligar a una tilde

pre- + di $\Rightarrow$ predí

o sus combinaciones;

in- + racional $\Rightarrow$ irracional

si la incorporación de un prefijo produce ambigüedad en la tilde —prefijos *tele-* y *para-*— trata de resolver esta situación consultando el universo.

tele- + fono $\Rightarrow$ teléfono

En caso de verbos, el *constructor de forma* se encarga de añadir los pronombres enclíticos y trata las modificaciones del último carácter de la forma conjugada y la cuestión de la tilde según las reglas descritas en el apartado 1.8.4.

divertid + -os $\Rightarrow$ divertíos

olimos + -nos $\Rightarrow$ olímonos

cantamos + -sela $\Rightarrow$ cantámosela

3.3.- MÓDULO DE ÍNDICES.

El *módulo de índices* se utiliza tanto en el reconocimiento como en la generación, figura 8; es el encargado de obtener la dirección correspondiente en la base de raíces mediante una función de Hash de tipo división —*hash nuclear*.

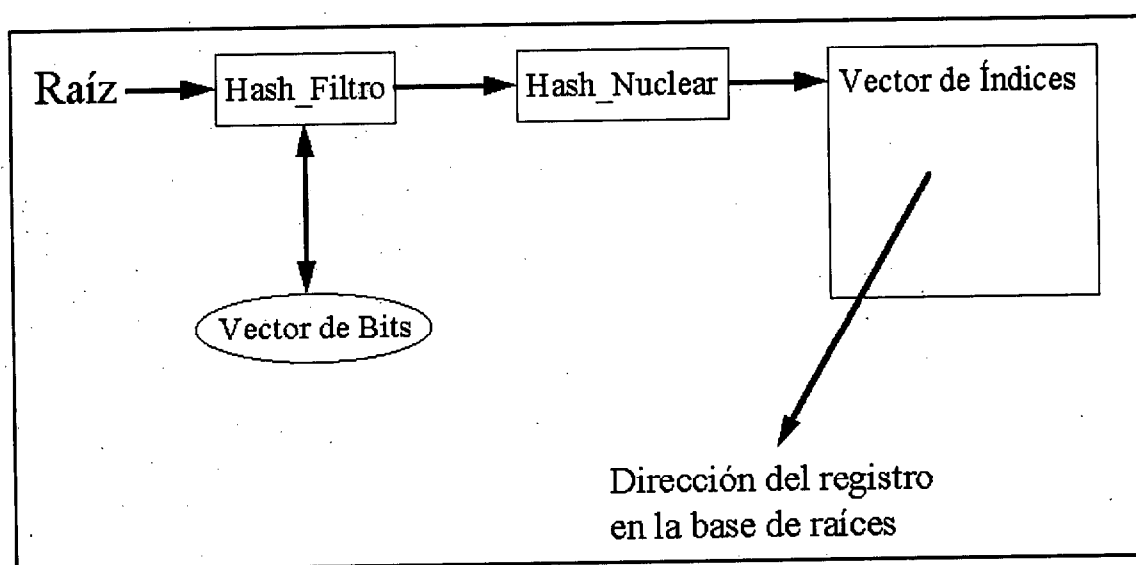


Figura 8: Módulo de índices

A fin de obtener una función que produzca pocas colisiones, se transforma la cadena de caracteres en un valor numérico mediante un tratamiento adecuado de la representación binaria de sus caracteres

```

long Hash(String &s)
{
    int i,l=s.Longitud();
    long h=0;
    for (i=0;i<l;i++) h += ((long) s[i]) << (7*(i%4));
    return labs((h)%HASH_LON);
}
  
```

Porque las colisiones disminuyen al crecer el tamaño de la tabla se sitúa en memoria principal un vector de bits gestionado por otra función de Hash —*hash filtro*— con módulo muy grande a fin de reducir accesos infructuosos a dispositivos externos —sus posiciones libres aseguran que tales raíces no existen y hacen innecesario el *hash nuclear*.

A partir de la dirección proporcionada por el *hash nuclear*, un *vector de índices* determina la dirección del registro en la base de raíces; este mecanismo trabaja con una base compacta aun cuando el módulo de la función de hash sea relativamente grande, porque aprovecha todos los registros sin tener que dejar ninguno intermedio vacío —la reestructuración de los registros proporciona una recuperación óptima de las colisiones.

3.4.- MÓDULO DE ACCESOS EXTERNOS.

Este módulo se usa tanto en el reconocimiento como en la generación de formas flexionadas o derivadas, figura 9. Gobierna las bases que guardan

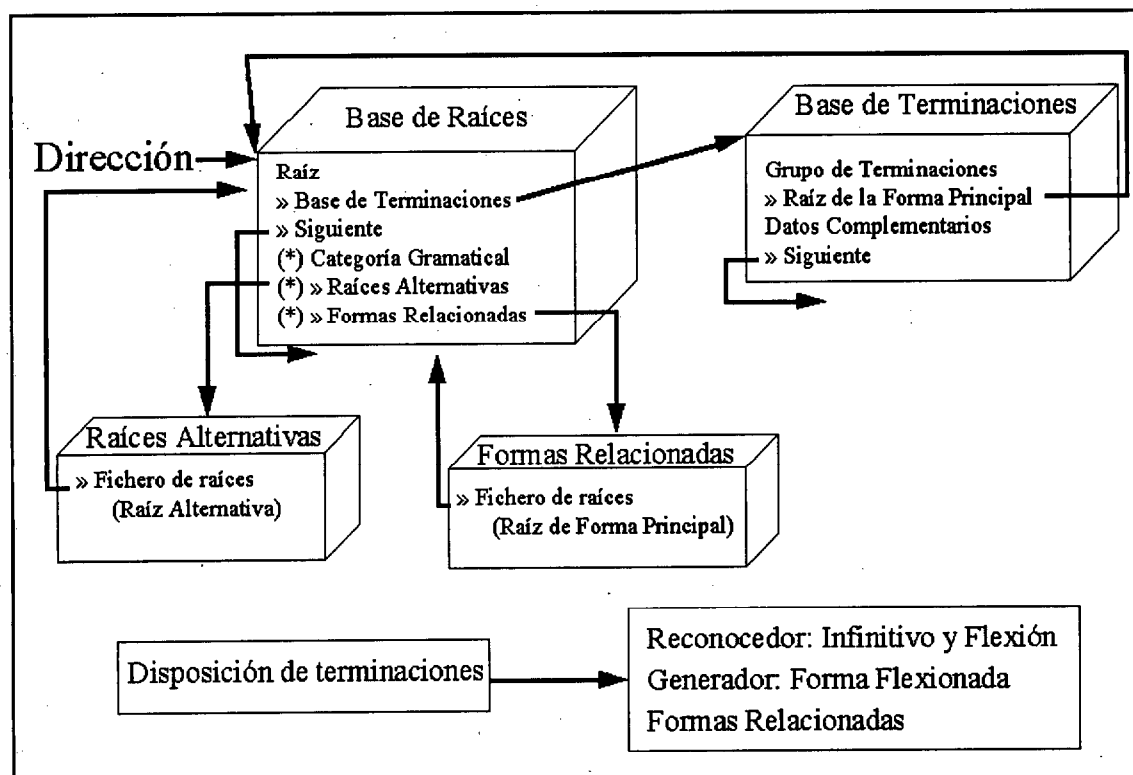


Figura 9: Módulo de accesos externos

la información referente a raíces, terminaciones y las flexiones o derivaciones que representan sus uniones; también permite obtener las relaciones de derivación entre las diferentes formas principales.

A partir de la dirección de una raíz accede a la base de raíces, de la que obtiene su posición en la base de terminaciones. Junto a la raíz de una forma principal aparecen su categoría gramatical, las raíces alternativas que pudiera poseer y las otras formas principales con las que guarda alguna

relación morfológica. El acceso a las raíces alternativas posibilita la generación de formas flexionadas o derivadas que no posean la misma raíz que la forma principal.

Superlativo(blanco) ⇒ blanc- + Terminacion_Superlativo ⇒

⇒ blanqu- + Terminacion_Superlativo ⇒ blanqu-ísimo

La relación morfológica con otras formas principales faculta los procesos de derivación mediante sufijos que otorgan cambio de categoría gramatical: sustantivación, adjetivación, adverbialización y verbalización; así como el acceso a formas relacionadas por derivación prefijal.

Sustantivación(abonar) ⇒ abonado, abonador, abonamiento,

abonaré, abono

Relación_Prefijal(acalorar) ⇒ calorar, desacalorar

La base de terminaciones contiene la información acerca de las terminaciones que admite una raíz y de la formación de su forma principal. El grupo de terminaciones permite averiguar qué terminaciones puede llevar una raíz y qué flexión representa para esa raíz cada terminación. Los datos complementarios aportan información para la reconstrucción de la forma principal a partir de su raíz y detalles del tipo de unión entre raíz y terminación.

La existencia de raíces alternativas y la formación de los grupos de

terminaciones que cada una de ellas admite obedece al conjunto de terminaciones que se considera en el proceso de construcción. Las posibilidades varían desde no considerar ninguna terminación y por tanto cada forma sería una raíz, hasta considerar un conjunto de terminaciones lo suficientemente amplio como para poder generar cualquier forma derivada a partir de una única raíz por simple concatenación. Dadas las características morfológicas del español como lengua flexiva, se opta por una solución que proporciona un número de terminaciones suficiente para que: a) en los casos regulares se puedan obtener todas las formas a partir de una raíz única porque las operaciones de flexión y derivación no afectan a la raíz gramatical de la palabra y b) cuando la raíz gramatical se ve afectada aparezcan raíces alternativas. Se consigue así un equilibrio entre abundancia de raíces alternativas y número de grupos de terminaciones.

3.5.- MEDIDAS DE REALIZACIÓN.

Con el propósito de probar la eficacia del modelo se realiza una aplicación desarrollada en C++ bajo Windows para el reconocimiento y la generación de formas verbales y no verbales —11.830 verbos, 19.767 sustantivos, 10.445 adjetivos, 374 adverbios de modo y 414 formas invariantes. Las formas verbales pueden verse afectadas por pronombres enclíticos y todas las formas admiten la prefijación. Los datos utilizados por el sistema ocupan un total de 6,64 Mb en disco. Sobre un procesador Pentium a 100 Mhz con 32 Mb de memoria RAM y un disco con velocidad de transferencia de 1.120 Kilobytes/segundo, se reconocen las 111.690 palabras de un texto literario a razón de 190 por segundo. La generación de 1.054.988 formas no verbales se realiza a una velocidad promedio de 320 palabras por segundo y para las 827.856 formas verbales a 480 palabras por segundo.

3.6.- INTERFAZ DEL GENERADOR MORFOLÓGICO.

Se ha diseñado una interfaz de comunicación para facilitar la generación morfológica de un vocablo —la ventana dispone de una *Entrada* que puede ser libremente editada.

The screenshot shows the 'MORFOLOGIA' software interface. It has a menu bar at the top. Below it is an 'Entrada:' text field. The interface is divided into two main sections: 'VERBAL' on the left and 'NOMINAL' on the right. The 'VERBAL' section includes a 'CONJUGACIÓN' area with options for 'Tiempo simple' and 'Tiempo compuesto', and 'Modo indicativo' and 'Modo subjuntivo'. It also has a 'DERIVACIÓN' section with options for 'Sustantivo', 'Adjetivo', and 'Adverbio'. The 'NOMINAL' section includes a 'DECLINACIÓN' area with options for 'Forma principal', 'Masculino singular', 'Masculino plural', 'Femenino singular', 'Femenino plural', and 'Grado superlativo'. It also has a 'DERIVACIÓN' section with options for 'Sufijos apreciativos' and 'Categoría gramatical'. At the bottom of the 'VERBAL' section is a 'PREFIJACIÓN' section with a 'sin prefijo' button and an 'Avanzado' checkbox. The bottom of the window features a logo for '2D GELSA', the text 'Grupo de Investigación en Estructuras de Datos', and a 'Cerrar' button.

A partir del reconocimiento de la palabra deduce su forma principal 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*.

Si el reconocimiento da lugar a más de un infinitivo o a más de una forma principal y ninguna coincide con la *Entrada*, abre una ventana y solicita la elección de una.

The screenshot shows a window titled "MORFOLOGÍA". It has two main tabs: "VERBAL" and "NOMINAL". The "Entrada" field contains the text "fui". Below the "VERBAL" tab, there are sections for "CONJUGACIÓN" and "DECLINACIÓN". The "CONJUGACIÓN" section lists "Infinitivo:", "Tiempo simple", and "Gerundio". The "DECLINACIÓN" section lists "principal:", "1.º singular", and "2.º singular". The "NOMINAL" tab is also visible. At the bottom, there is a message: "Elija entre las posibles formas principales o infinitivos" and a "Cerrar" button.

Para facilitar la interpretación del resultado del reconocimiento mientras se recorre cada opción aparece un mensaje que indica la flexión o derivación a que está sujeta la *Entrada* en ese caso.

The screenshot shows the same "MORFOLOGÍA" window. The "Entrada" field still contains "fui". The "CONJUGACIÓN" section is now active, and the "DECLINACIÓN" section shows the result: "FUI es 1.º per. sing. pret. indefinido del verbo.SER". The "Cerrar" button is still present at the bottom.

Si la entrada es reconocida como forma verbal conjugada —afectada o no por prefijos y pronombres enclíticos— habilita la parte de la morfología verbal, y si es reconocida como flexión o derivación de una

forma no verbal —con prefijos o no— habilita la sección correspondiente a la morfología nominal.

MORFOLOGIA

Entrada: **deber**

VERBAL

Infinitivo: **deber**

CONJUGACIÓN

☒ Tiempos simples ☒ Tiempos compuestos
☒ Gerundio ☒ Participio
 Modo indicativo Modo subjuntivo
☒ Presente ☒ Presente
☒ Pretérito indefinido ☒ Pretérito imperfecto
☒ Condicional simple ☒ Futuro imperfecto
☒ Pretérito imperfecto ☒ Modo imperativo
☒ Futuro imperfecto ☒ Presente

Pronombres enclíticos

☒ no ☒ me ☒ te ☒ os

DERIVACIÓN

☒ Sustantivo ☒ Adjetivo ☒ Adverbio

PREFIJACIÓN

sin prefijo

☐ Acumular ☐ Eliminar

NOMINAL

Forma principal: **deber**

DECLINACIÓN

☒ Masculino singular ☒ Masculino plural
☒ Femenino singular ☒ Femenino plural
☒ Grado superlativo

DERIVACIÓN

Sufijos apreciativos Categoría gramatical
☒ Aumentativos ☒ Sustantivo
☒ Diminutivos ☒ Adjetivo
☒ Despectivos ☒ Adverbio
☒ Verbo

RESULTADO

DEBER es infinitivo del verbo DEBER y es principal (masculino singular) del sustantivo DEBER

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. Por comodidad y sencillez de uso genera de una sola vez las seis personas y números de un tiempo verbal

—entraña una sobrecarga muy poco relevante. El sistema permite cambiar de tiempo y de pronombre de manera tan simple e independiente que hace muy amigable su manejo. En todo momento estarán inhabilitadas las opciones no permitidas.

The screenshot shows the 'MORFOLOGIA' software interface. At the top, there's a title bar 'MORFOLOGIA' and a text input field 'Entrada: deber'. Below this, the interface is divided into two main sections: 'VERBAL' on the left and 'NOMINAL' on the right.

VERBAL Section:

- Infinitivo:** deber
- CONJUGACIÓN:**
 - Tiempos simples:** Gerundio, Participio
 - Tiempos compuestos:** Presente, Pretérito imperfecto, Futuro imperfecto
 - Modo indicativo:** Presente, Pretérito indefinido, Condicional simple, Pretérito imperfecto, Futuro imperfecto
 - Modo subjuntivo:** Presente, Pretérito imperfecto, Modo imperativo, Presente
 - Pronombres enclíticos:** no, uno, dos, tres, and a dropdown menu.
- DERIVACIÓN:** Sustantivo, Adjetivo, Adverbio
- PREFIJACIÓN:** sin prefijo, Acumular, Eliminar

NOMINAL Section:

- Forma principal:** deber
- DECLINACIÓN:** Masculino singular, Masculino plural, Femenino singular, Femenino plural, Grado superlativo
- DERIVACIÓN:**
 - Sufijos apreciativos:** Superlativos, Diminutivos, Despectivos
 - Categoría gramatical:** Sustantivo, Adjetivo, Adverbio, Verbo
- RESULTADO:**
 - Presente de indicativo:**

1ª s.	debo
2ª s.	debes
3ª s.	debe
1ª p.	debemos
2ª p.	debéis
3ª p.	deben

At the bottom, there's a logo for 'GELSA', the text 'Grupo de Investigación en Estructuras de Datos', and a 'Cerrar' button.

En cuanto se elige un tiempo —incluyendo *Infinitivo*, *Gerundio* y *Participio*— se presenta su conjugación en la ventana de resultados y los pronombres quedan habilitados. A pesar de que el infinitivo del verbo siempre está presente, activar *Infinitivo* permite a) visualizarlo en la ventana de resultado, b) incorporarle pronombres enclíticos, c) añadirle prefijos y

d) formar su tiempo compuesto.

Pronombres enclíticos		VERBOS	
no	uno	dos	tres
DERIVACIÓN		*RESULTADO*	
Sustantivo	Adjetivo	Presente de indicativo	
PREFIJACIÓN			
	mela	1ª s.	debo
	melas	2ª s.	debes
	mele	3ª s.	debe
	meles		
	melo		
	melos		

La incorporación de pronombres enclíticos está organizada en base al número de pronombres a considerar. Al elegir el número —*uno*, *dos* o *tres*— abre una ventana desplegable que contiene las opciones válidas correspondientes en orden alfabético —su localización es sencilla y para evitar errores no acepta edición.

MORFOLOGIA

Entrada:

VERBAL

◆ Infinitivo: deber

CONJUGACIÓN

◆ Tiempos simples

◆ Gerundio

Modo indicativo

◆ Presente

◆ Pretérito indefinido

◆ Condicional simple

◆ Pretérito imperfecto

◆ Futuro imperfecto

◆ Tiempos compuestos

◆ Participio

Modo subjuntivo

◆ Presente

◆ Pretérito imperfecto

◆ Futuro imperfecto

Modo imperativo

◆ Presente

Pronombres enclíticos

◆ no ◆ uno ◆ dos ◆ tres

DERIVACIÓN

◆ Sustantivo ◆ Adjetivo ◆ Adverbio

PREFIJACIÓN

☐ Acumular

NOMINAL

◆ Forma principal: deber

DECLINACIÓN

◆ Masculino singular

◆ Femenino singular

◆ Masculino plural

◆ Femenino plural

◆ Grado superlativo

DERIVACIÓN

Sufijos apreciativos

◆ Augmentativos

◆ Diminutivos

◆ Despectivos

Categoría gramatical

◆ Sustantivo

◆ Adjetivo

◆ Adverbio

◆ Verbo

*** RESULTADO ***

Presente de indicativo

1ª s. débosela

2ª s. debesela

3ª s. debesela

1ª p. debémosela

2ª p. debeisela

3ª p. débensela



Grupo de Investigación en Estructuras de Datos

En cuanto se escoge un pronombre aparece añadido al tiempo conjugado sin ninguna otra acción; si se desea otro, basta con abrir la lista y elegirlo. La alternativa *no* desproveye de pronombres enclíticos a la conjugación. Para mayor comodidad de uso es posible modificar el tiempo verbal sin afectar al pronombre.

Cuando se opta por *Tiempos compuestos*: a) cambian los títulos de los diferentes tiempos por sus correspondientes para una mejor disposición de la información —la simetría entre tiempos simples y compuestos aconseja esta solución—, b) inhabilita el *Presente del Modo imperativo* ya que no posee tiempo compuesto y c) si había sido activado algún tiempo conjuga su compuesto —facilita la tarea de visualización de ambos. El proceso es totalmente reversible al pasar de *Tiempos compuestos* a *Tiempos simples*. En estos tránsitos se conservan los pronombres.

MORFOLOGIA

Entrada: **deber**

VERBAL

◆ Infinitivo: deber

CONJUGACIÓN

◆ Tiempos simples

◆ Gerundio compuesto

Modo indicativo

◆ Pretérito perfecto compuesto

◆ Pretérito anterior

◆ Condicional perfecto

◆ Pretérito pluscuamperfecto

◆ Futuro perfecto

◆ Tiempos compuestos

◆ Participio compuesto

Modo subjuntivo

◆ Pretérito perfecto

◆ Pretérito pluscuamperfecto

◆ Futuro perfecto

Modo imperativo

◆ Presente

Pronombres enclíticos

◆ no ◆ uno ◆ dos ◆ tres

DERIVACIÓN

◆ Sustantivo ◆ Adjetivo ◆ Adverbio

PREFIJACIÓN

sin prefijo

Acumular Eliminar

NOMINAL

◆ Forma principal: deber

DECLINACIÓN

◆ Masculino singular

◆ Femenino singular

◆ Masculino plural

◆ Femenino plural

◆ Grado superlativo

DERIVACIÓN

Sufijos apreciativos

◆ Aumentativos

◆ Diminutivos

◆ Despectivos

Categoría gramatical

◆ Sustantivo

◆ Adjetivo

◆ Adverbio

◆ Verbo

RESULTADO

Preterito perfecto compuesto de indicativo

1^a s. he debido

2^a s. has debido

3^a s. ha debido

1^a p. hemos debido

2^a p. habeis debido

3^a p. han debido


 Grupo de Investigación en Estructuras de Datos
 Cerrar

Las posibilidades de derivación —*Sustantivo, Adjetivo y Adverbio*— permiten obtener formas que pertenecen a otra categoría gramatical, pero que están morfológicamente relacionadas con el verbo —sustantivaciones, adjetivaciones o adverbializaciones. Al elegir una de estas opciones desactiva tiempo y pronombre ya que estas formas no verbales no presentan tales posibilidades de flexión. Mientras se recorren las formas presentadas en la ventana *RESULTADO*, el área de mensajes refleja el detalle de su categoría gramatical para mayor abundamiento en la forma generada.

The screenshot shows the 'MORFOLOGIA' window with the input 'deber'. The interface is divided into several sections:

- VERBAL**: Includes 'Infinitivo: deber' and a 'CONJUGACIÓN' section with options for 'Tiempos simples' and 'Tiempos compuestos'. Under 'Tiempos simples', there are checkboxes for 'Gerundio', 'Participio', 'Modo indicativo', and 'Modo subjuntivo'. The 'Modo indicativo' section includes 'Presente', 'Pretérito indefinido', 'Condicional simple', 'Pretérito imperfecto', and 'Futuro imperfecto'. The 'Modo subjuntivo' section includes 'Presente', 'Pretérito imperfecto', 'Futuro imperfecto', and 'Modo imperativo'. There is also a 'Pronombres enclíticos' section with options for 'no', 'me', 'te', 'os', and 'los'.
- NOMINAL**: Includes 'Forma principal: deber' and a 'DECLINACIÓN' section with options for 'Masculino singular', 'Masculino plural', 'Femenino singular', and 'Femenino plural'. There is also a 'Grado superlativo' option.
- DERIVACIÓN**: Includes 'Sufijos apreciativos' (with options for 'Superlativo', 'Diminutivo', and 'Despectivo') and 'Categoría gramatical' (with options for 'Sustantivo', 'Adjetivo', 'Adverbio', and 'Verbo').
- PREFIJACIÓN**: Includes a 'sin prefijo' option and a 'Acumular' button.
- RESULTADO**: Shows the 'Sustantivación verbal' results: 'deber' and 'debe'.

At the bottom, a message states: 'DEBER es verbo pronominal o impersonal o pasivo o sustantivo masculino'.

Como generalmente estas formas obtenidas por derivación admiten

flexión de tipo nominal se puede continuar operando sobre cualquiera de ellas mediante la acción de un *doble clic* que la sitúa automáticamente en *Entrada* —esta opción facilita y mejora las posibilidades de generación.

Si el verbo es defectivo, sólo aparecerán aquellas personas y números que admite en el tiempo correspondiente.

Infinitivo: llover		Forma principal:	
CONJUGACIÓN		DECLINACIÓN	
☒ Tiempos simples	☒ Tiempos compuestos	☒ Masculino singular	☒ Masculino plural
☒ Gerundio	☒ Participio	☒ Femenino singular	☒ Femenino plural
Modo indicativo	Modo subjuntivo	☒ Grado superlativo	
☒ Presente	☒ Presente	DERIVACIÓN	
☒ Pretérito indefinido	☒ Pretérito imperfecto	Sufijos apreciativos	Categoría gramatical
☒ Condicional simple	☒ Futuro imperfecto	☒ Ampliativos	☒ Sustantiva
☒ Pretérito imperfecto	Modo imperativo	☒ Diminutivos	☒ Adjetivo
☒ Futuro imperfecto	☒ Presente	☒ Despectivos	☒ Adverbio
Pronombres enclíticos		☒ Verbo	
☒ no ☒ uno ☒ dos ☒ tres		*RESULTADO*	
DERIVACIÓN		Condicional simple	
☒ Sustantivo	☒ Adjetivo	1ª s.	
☒ Adverbio		2ª s.	
PREFIJACIÓN		3ª s. llovería	
sin prefijo		1ª p.	
☐ Acumular ☐ Eliminar		2ª p.	
		3ª p.	

☐ Pretérito imperfecto ☐ Futuro imperfecto	Modo imperativo ☐ Presente	☐ Dialectales ☐ Despectivos	☐ Adjetivo ☐ Adverbio ☐ Verbo
Pronombres enclíticos ☐ no ☐ uno ☐ dos ☐ tres			
DERIVACIÓN			
☐ Sustantivo ☐ Adjetivo ☐ Adverbio			
PREFIJACIÓN			
sin prefijo			
☐ Acumular ELIMINAR			
* RESULTADO *			
Presente de imperativo			
1ª s.			
2ª s.			
3ª s.			
1ª p.			
2ª p.			
3ª p.			
Atención: LLOVER no admite Presente de imperativo			

Si el tiempo no permite ninguna forma, un mensaje —con señal acústica— indicará esa circunstancia; en lugar de inhabilitar a priori el tiempo que el verbo no posee, se opta por esta solución ya que tal comprobación tendría un coste significativo y sólo la aprovecharía una baja población de defectivos; además los tiempos inexistentes se solicitan muy raramente.

MORFOLOGÍA

Entrada: **bosques**

VERBAL

Infinitivo: ☐

CONJUGACIÓN

☐ Tiempos simples ☐ Tiempos compuestos
☐ Gerundio ☐ Participo
Modo indicativo **Modo subjuntivo**
☐ Presente ☐ Presente
☐ Pretérito indefinido ☐ Pretérito imperfecto
☐ Condicional simple ☐ Futuro imperfecto
☐ Pretérito imperfecto **Modo imperativo**
☐ Futuro imperfecto ☐ Presente
☐ Pronombres enclíticos
☐ no ☐ aya ☐ day ☐ tres

DERIVACIÓN

☐ Sustantivo ☐ Adjetivo ☐ Adverbio

PREFIJACIÓN

sin prefijo

☐ Acumular

NOMINAL

Forma principal: **bosque**

DECLINACIÓN

☐ Masculino singular ☐ Masculino plural
☐ Femenino singular ☐ Femenino plural
☐ Grado superlativo

DERIVACIÓN

Sufijos apreciativos **Categoría gramatical**
☐ Superlativos ☐ Sustantivo
☐ Diminutivos ☐ Adjetivo
☐ Despectivos ☐ Adverbio
☐ Verbo

RESULTADO

BOSQUES es masculino plural del sustantivo BOSQUE

Si reconoce una forma principal no verbal habilita la flexión nominal; con ella consigue la forma principal y los diferentes géneros y números que admita. Estarán inhabilitadas las declinaciones que no se puedan obtener; como existe una alta tasa de palabras que no admiten todas las declinaciones y el coste de evaluar las opciones posibles es bajo anticipa esta información. A pesar de que la *Forma principal* está continuamente presente, activar esta opción faculta la incorporación de prefijos a la misma.

Muestra las cuatro opciones de declinación completas: *Masculino singular, Femenino singular, Masculino plural y Femenino plural*, porque mantener el género independiente del número a) exige dos acciones en lugar de una y b) dificulta la visualización de las declinaciones inválidas al no encontrarse todas las opciones presentes.

Forma principal: bosque	
DECLINACIÓN	
☒ Masculino singular ☐ Femenino singular	☐ Masculino plural ☐ Femenino plural
☐ Grado superlativo	
DERIVACIÓN	
Sufijos apreciativos ☐ Aumentativos ☐ Diminutivos ☐ Despectivos	Categoría gramatical ☐ Sustantivo ☐ Adjetivo ☐ Adverbio ☐ Verbo
RESULTADO	
Masculino singular	
bosque	

Forma principal: bosque	
DECLINACIÓN	
☒ Masculino singular ☐ Femenino singular	☐ Masculino plural ☐ Femenino plural
☐ Grado superlativo	
DERIVACIÓN	
Sufijos apreciativos ☐ Aumentativos ☒ Diminutivos ☐ Despectivos	Categoría gramatical ☐ Sustantivo ☐ Adjetivo ☐ Adverbio ☐ Verbo
RESULTADO	
Diminutivo Masculino singular	
bosquecito bosquencillo bosquecico bosquecim bosquezuelo	

Una vez seleccionado un género y número permite elegir derivaciones mediante sufijos apreciativos —*Aumentativos*, *Diminutivos* y *Despectivos*—, y si la entrada es adjetivo su grado superlativo.

Forma principal: blanco	
DECLINACIÓN	
☐ Masculino singular ☐ Femenino singular	☐ Masculino plural ☐ Femenino plural
☐ Grado superlativo	
DERIVACIÓN	
Sufijos apreciativos ☐ Aumentativos ☐ Diminutivos ☐ Despectivos	Categoría gramatical ☐ Sustantivo ☐ Adjetivo ☐ Adverbio ☐ Verbo
* RESULTADO *	
Superlativo Femenino plural	
blanquísimas	

Forma principal: blanco	
DECLINACIÓN	
☐ Masculino singular ☐ Femenino singular	☐ Masculino plural ☐ Femenino plural
☐ Grado superlativo	
DERIVACIÓN	
Sufijos apreciativos ☐ Aumentativos ☐ Diminutivos ☐ Despectivos	Categoría gramatical ☐ Sustantivo ☐ Adjetivo ☐ Adverbio ☐ Verbo
* RESULTADO *	
Despectivo Femenino plural	
blanquejas blancuchas	

La *DERIVACIÓN* por cambio de *Categoría gramatical* de las entradas sustantivas da sus adjetivaciones, de los adjetivos sus sustantivaciones y adverbializaciones y de los adverbios sus adjetivaciones; en cualquier caso ofrece los verbos relacionados morfológicamente. La activación de cualquier derivación que conlleve cambio de categoría gramatical —*Sustantivo*, *Adjetivo*, *Adverbio* o *Verbo*— desactiva la declinación, ya que puede existir incompatibilidad con la nueva categoría.

MORFOLOGÍA

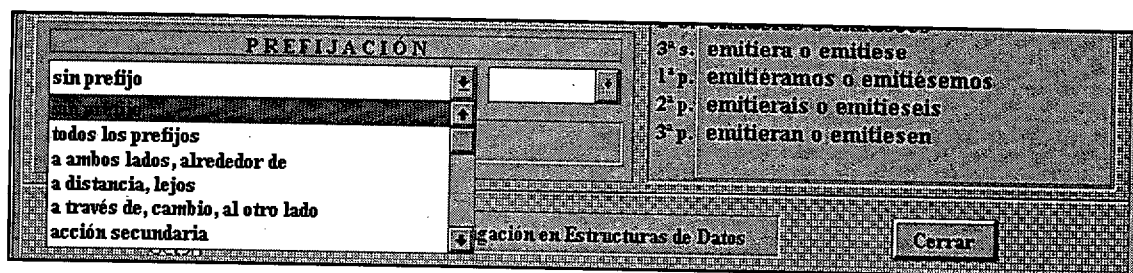
Entrada: fabricación

VERBAL	NOMINAL										
◆ Infinitivo: CONJUGACIÓN ◆ Tiempos simples ◆ Tiempos compuestos ◆ Gerundio ◆ Participo Modo indicativo Modo subjuntivo ◆ Presente ◆ Presente ◆ Pretérito indefinido ◆ Pretérito imperfecto ◆ Condicional simple ◆ Futuro imperfecto ◆ Pretérito imperfecto Modo imperativo ◆ Futuro imperfecto ◆ Presente Pronombres enclíticos ◆ me ◆ te ◆ os ◆ les DERIVACIÓN ◆ Sustantivo ◆ Adjetivo ◆ Adverbio PREFIJACIÓN sin prefijo ↓ ☐ Acumular Eliminar	◆ Forma principal: fabricación DECLINACIÓN ◆ Masculino singular ◆ Masculino plural ◆ Femenino singular ◆ Femenino plural ◆ Grado superlativo DERIVACIÓN <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 50%;">Sufijos apreciativos</th> <th style="width: 50%;">Categoría gramatical</th> </tr> <tr> <td>◆ Superlativos</td> <td>◆ Sustantivo</td> </tr> <tr> <td>◆ Diminutivos</td> <td>◆ Adjetivo</td> </tr> <tr> <td>◆ Despectivos</td> <td>◆ Adverbio</td> </tr> <tr> <td></td> <td>◆ Verbo</td> </tr> </table> * RESULTADO * Adjetivación fabricado fabricador fabricante	Sufijos apreciativos	Categoría gramatical	◆ Superlativos	◆ Sustantivo	◆ Diminutivos	◆ Adjetivo	◆ Despectivos	◆ Adverbio		◆ Verbo
Sufijos apreciativos	Categoría gramatical										
◆ Superlativos	◆ Sustantivo										
◆ Diminutivos	◆ Adjetivo										
◆ Despectivos	◆ Adverbio										
	◆ Verbo										

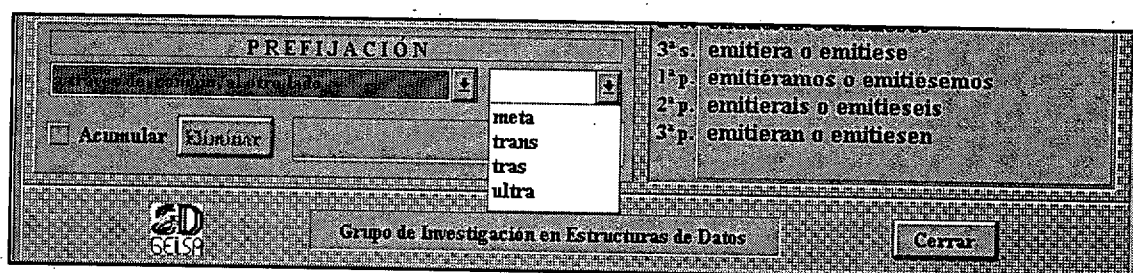
FABRICADOR es adjetivo usado también como sustantivo

Mediante un *clic* en la ventana *RESULTADO*, el área de mensajes muestra el detalle de la categoría gramatical para un mejor conocimiento de la forma obtenida; un *doble clic* realiza la transferencia a *Entrada* para continuar con la flexión o derivación.

Tanto en la morfología verbal como en la nominal se dispone de la facultad de incorporar prefijos que se seleccionan en las listas clasificadas según los distintos significados que poseen. La abundancia de significados



sugiere su ubicación en una ventana desplegable donde aparecen los nombres de los grupos ordenados alfabéticamente: al fijar uno abre la lista correspondiente.



Incluye la opción de *sin prefijo* al objeto de eliminarlos con comodidad y la de *todos los prefijos* para escoger alfabéticamente uno cualquiera con independencia de su significado —el área de mensajes informa de sus significados.

En cuanto se elige uno modifica la respuesta en *RESULTADO*.

MORFOLOGÍA

Entrada: emitir

VERBAL **NOMINAL**

Infinitivo: emitir Forma principal:

DERIVACIÓN

Sustantivo Adjetivo Adverbio

PREFIJACIÓN

todos los prefijos

☐ Acumular trans-

RESULTADO

Preterito imperfecto de subjuntivo

1ª s. transmitiera o transmitiese
 2ª s. transmitieras o transmitieses
 3ª s. transmitiera o transmitiese
 1ª p. transmitiéramos o transmitiésemos
 2ª p. transmitierais o transmitieseis
 3ª p. transmitieran o transmitiesen

TRANS es del tipo a través de, cambio, al otro lado.

A diferencia de lo que ocurre con los pronombres enclíticos, el desconocimiento de reglas que rijan la secuencialización de prefijos aconseja la opción *Acumular* que añade por la izquierda cualquier prefijo a los ya existentes; cuando está desactivada un nuevo prefijo sustituye al anterior.

no uno dos tres

DERIVACIÓN

Sustantivo Adjetivo Adverbio

PREFIJACIÓN

aumento, encarecimiento, refuerzo

☒ Acumular re-trans-

RESULTADO

Preterito imperfecto de subjuntivo

1ª s. retransmitiera o retransmitiese
 2ª s. retransmitieras o retransmitieses
 3ª s. retransmitiera o retransmitiese
 1ª p. retransmitiéramos o retransmitiésemos
 2ª p. retransmitierais o retransmitieseis
 3ª p. retransmitieran o retransmitiesen

La opción de *sin prefijo* elimina todos los prefijos de una vez; como resulta interesante poder descartarlos de uno en uno, el botón *Eliminar* lo hace de izquierda a derecha.

DERIVACIÓN	
☐ no	☐ uno ☐ dos ☐ tres
☐ Sustantivo	☐ Adjetivo ☐ Adverbio

PREFIJACIÓN	
compañía	co
☒ Acumular	Eliminar co-trans

RESULTADO	
Futuro imperfecto de indicativo	
1ª s.	cotransmitiré
2ª s.	cotransmitirás
3ª s.	cotransmitirá
1ª p.	cotransmitiremos
2ª p.	cotransmitireis
3ª p.	cotransmitirán

Por varios motivos la salida se muestra en una misma ventana tanto para la parte *VERBAL* como para la *NOMINAL*: a) dada la gran cantidad de información que contiene el diálogo ahorra espacio al disponer de una sola ventana *RESULTADO*, b) dirige mejor la observación de los resultados y c) resulta sencillo su adaptación a los diferentes tipos de salida. El encabezamiento de *RESULTADO* con la flexión o derivación presente ayuda a su interpretación. El número de líneas que muestra simultáneamente es de seis debido a que la conjugación mínima seleccionable la compone un tiempo verbal y el resto de las opciones rara vez supera ese tamaño. Cuando la respuesta es más amplia que el espacio de *RESULTADO*, dispone el desplazamiento horizontal o vertical necesario.

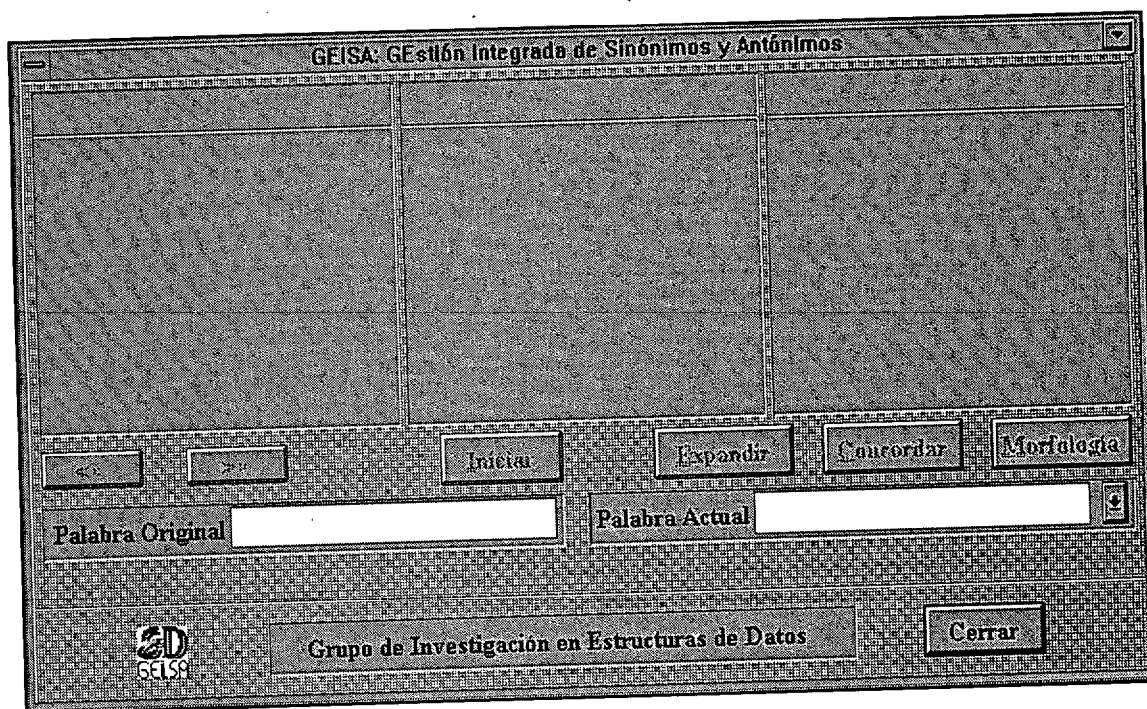
El área de mensajes a) usa la zona muerta en la parte inferior del diálogo, b) sólo está presente cuando es imprescindible, c) llama la atención

de manera discreta, d) no oculta ninguna información relevante y e) permite continuar sin ningún requerimiento extra.

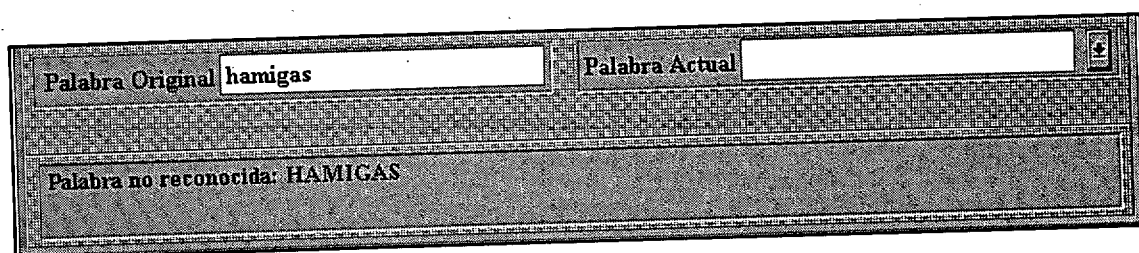
4.- DEL RECONOCIMIENTO Y LA GENERACIÓN EN UN DICCIONARIO DE SINÓNIMOS Y ANTÓNIMOS.

Se ha desarrollado una aplicación de estos conceptos al universo dado por un diccionario compuesto de 36.918 entradas, cada una de las cuales 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 19.767 sustantivos, 10.445 adjetivos, 6.317 verbos, 375 adverbios de modo y otras 413 formas invariantes. El objetivo que se persigue es obtener una versión informatizada del diccionario de sinónimos 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) 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 salida a cualquier inconsistencia en la concordancia mediante un abundamiento en la flexión y derivación de la respuesta.

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

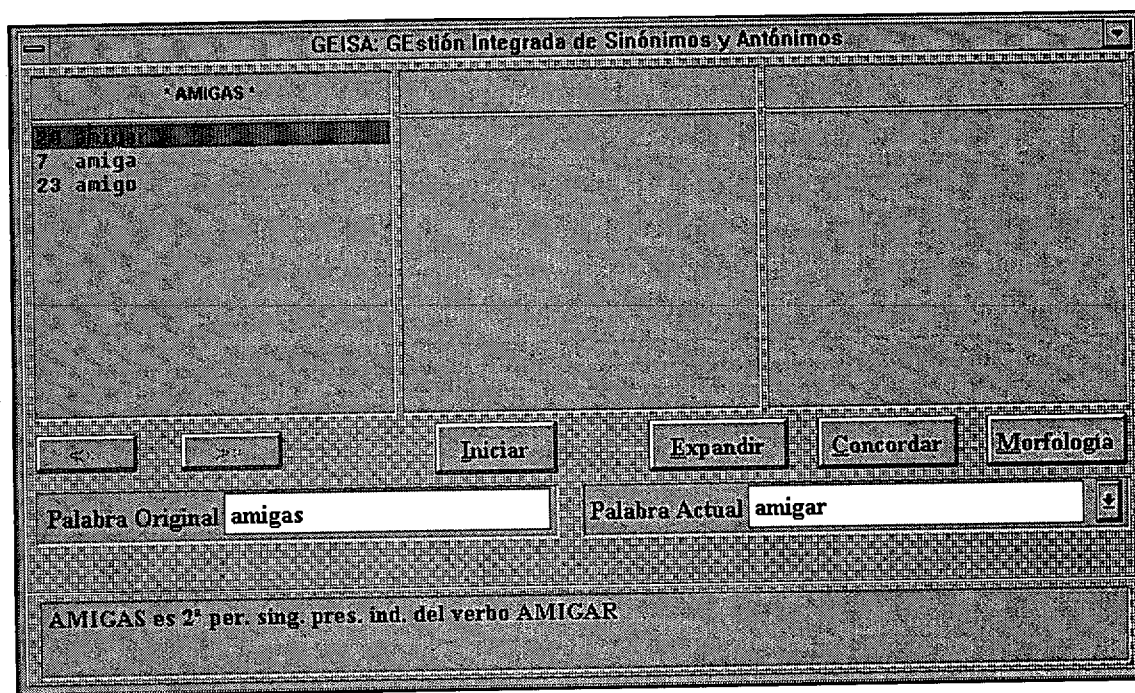


En esta ventana se encuentran inhabilitados todos los botones salvo *Cerrar* —se habilitan a medida que tenga sentido su uso. Permite escribir en *Palabra Original* para iniciar el proceso —activa automáticamente el botón *Iniciar*. Actuar sobre *Iniciar* implica realizar el análisis morfológico de *Palabra Original* —en caso de que no sea reconocida aparece un mensaje



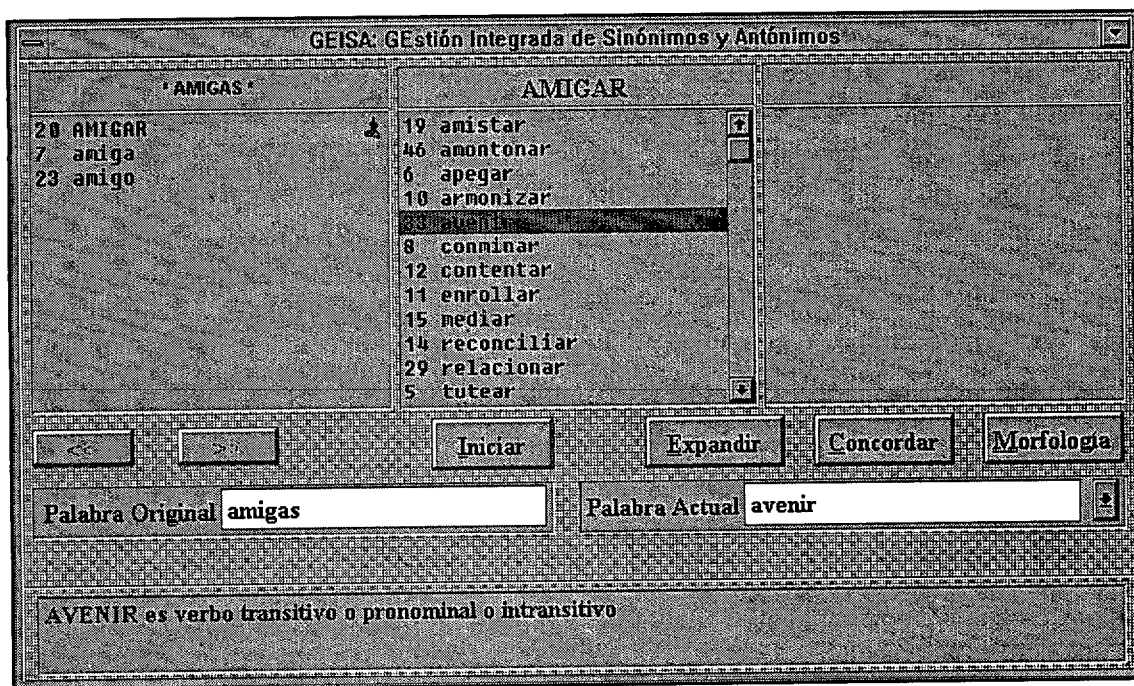
indicándolo—; admite así no sólo formas principales o infinitivos que sean

cabeceras en el diccionario, sino cualquiera de sus formas flexionadas o derivadas.



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 principal o un infinitivo 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

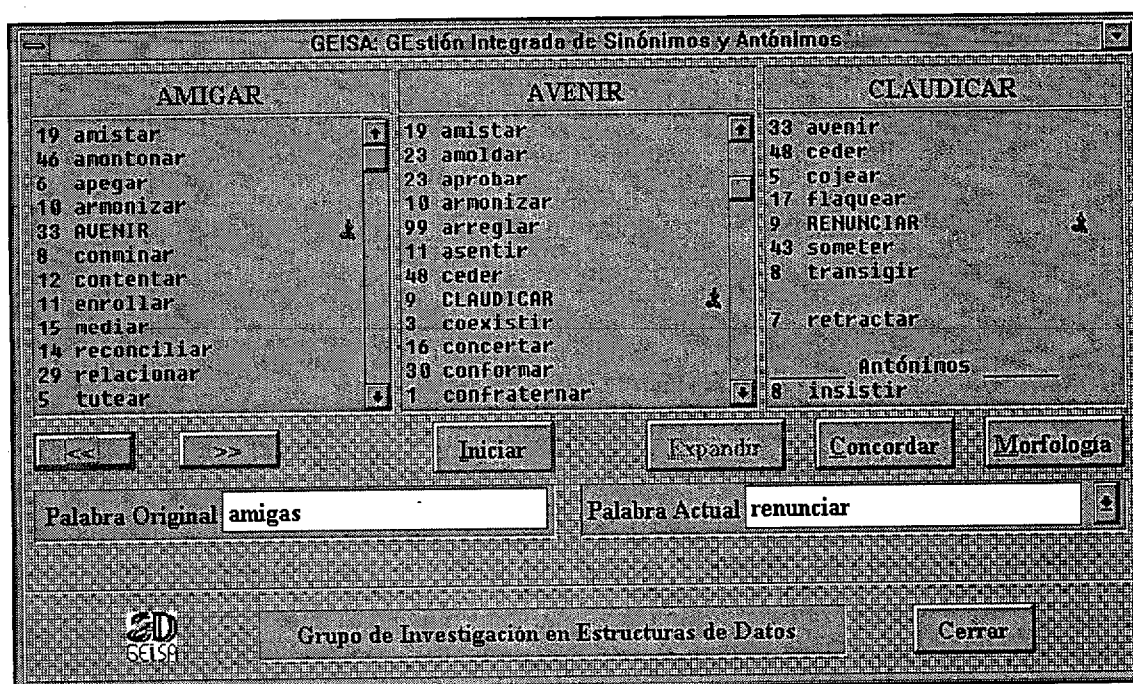
Morfología.



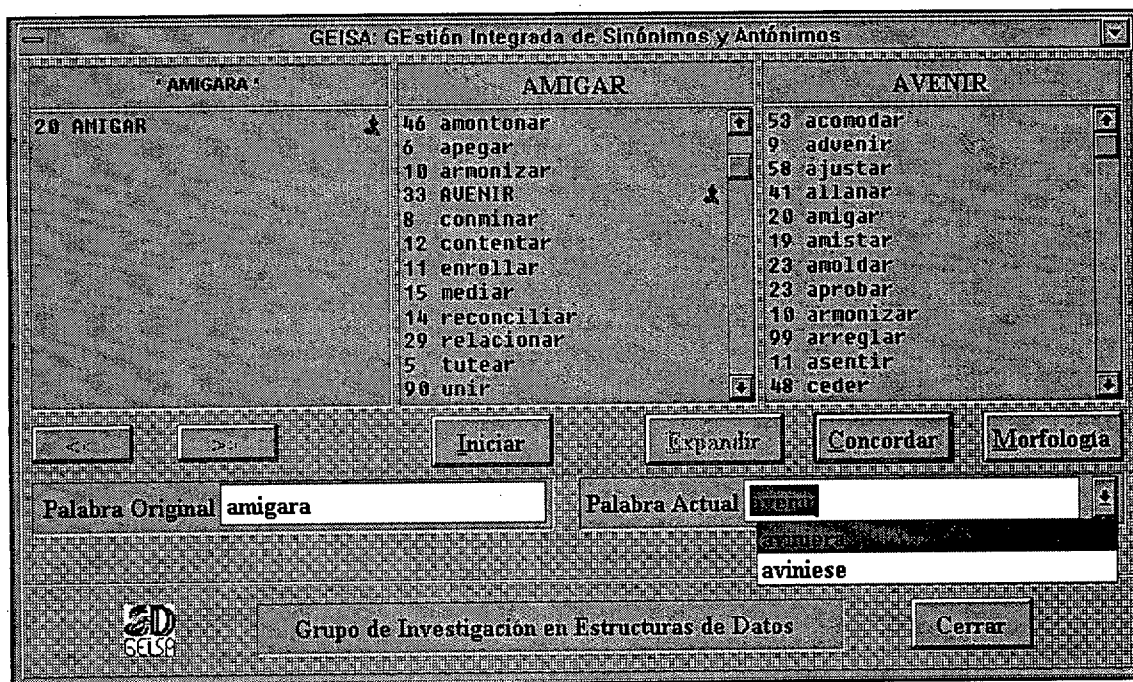
A la izquierda de cada palabra aparece el número de elementos que posee su lista de sinónimos y antónimos al objeto de anticipar información que ayude a decidir por donde continuar la investigación. Un *doble clic* o la *tecla de entrada* hace progresar la navegación por el diccionario; consiste en exponer en la siguiente ventana la lista de sinónimos y antónimos que corresponde. Para destacar la navegación señala la palabra seleccionada con una marca y la sitúa como cabecera de la ventana que contiene su lista; si tal información no cabe, dispone su desplazamiento vertical u horizontal.

Al resaltar una palabra en cualquier ventana que no sea la inicial aporta su categoría gramatical como información general —de particular interés para la concordancia. Dado que el proceso puede extenderse más allá

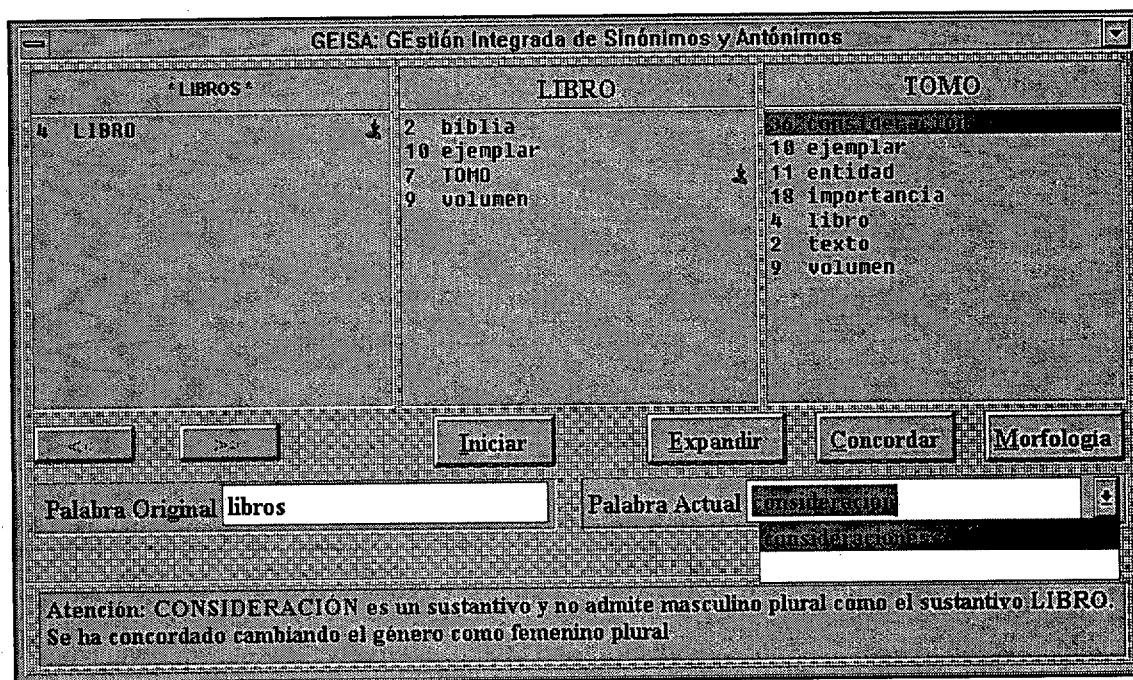
de las tres ventanas que están visibles simultáneamente habilita los botones << o >> para visualizar las listas que han quedado fuera. La marca de la tercera ventana apunta la cabecera que se oculta por la derecha y la cabecera de la primera coincide con la palabra marcada en la ventana escondida por la izquierda.



Actuar sobre 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.



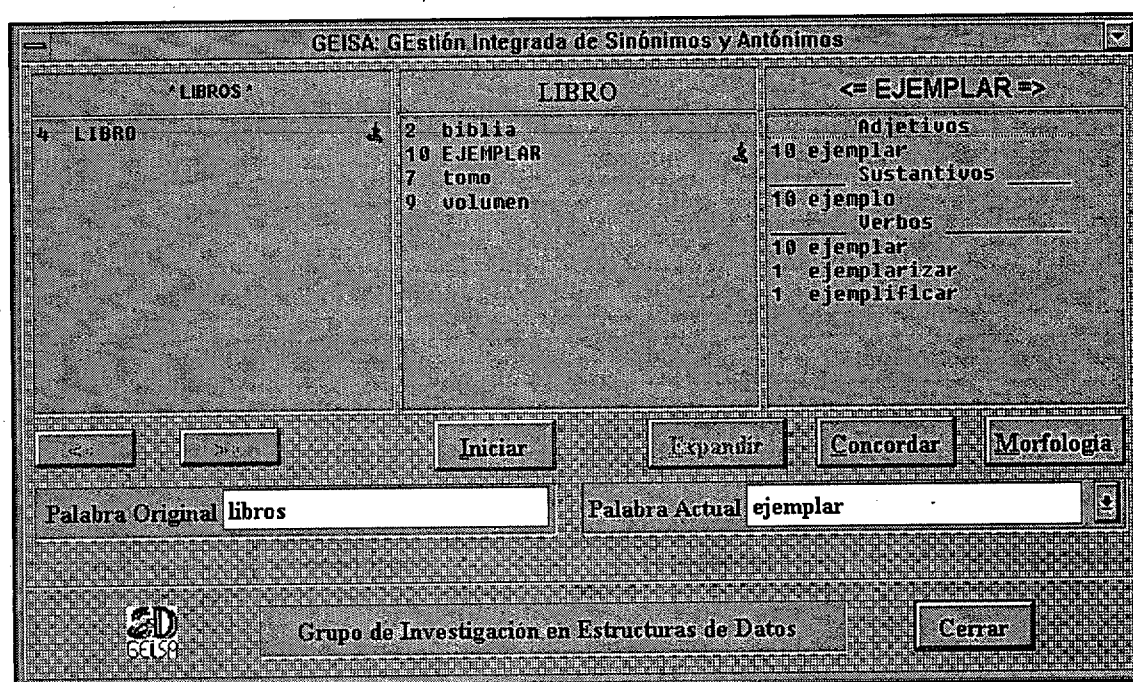
Si existe alguna dificultad o se produce una violación gramatical y no es factible conseguir la concordancia muestra dos tipos de mensaje:



a) *Atención*, indica como ha procedido —normalmente ajuste de la flexión

a la categoría gramatical o cambio de género— y b) *Fallo*, señala el error y aconseja pasar a la ventana de *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.

5.- ORGANIZACIÓN DEL DICCIONARIO DE SINÓNIMOS.

Como la normal utilidad de un diccionario de sinónimos no requiere una alta velocidad de acceso a sus listas, se optó por almacenarlo en memoria externa; el acceso a las cabeceras se lleva a cabo mediante una función de hash 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.

En la óptica de un diccionario estático puede considerarse otro tipo de soluciones tendentes a minimizar la redundancia sin perjudicar el número de accesos —utilizar referencias en lugar de ristas de caracteres para representar las palabras y combinar el uso de memoria principal con secundaria.

Las distintas alternativas deben evaluarse bajo los siguientes parámetros:

- Redundancia: Número de veces que una misma palabra se repite y espacio que se invierte en ello.
- Accesos: Número promedio de lecturas sobre dispositivos externos para obtener una lista de sinónimos.

— Requerimientos: Memoria principal y secundaria, la solución debe ser susceptible de realizarse sobre una gran variedad de equipos.

La solución actual usa dos ficheros externos uno de *Control* (384Kb) y el *Diccionario* (2,18Mb). Necesita hacer tantos accesos a *Control* como colisiones haya y un acceso a *Diccionario* por cada una. Una opción que disminuiría el almacenamiento secundario a costa de memoria principal y reduciría los accesos consiste en disponer *Control* en memoria principal; dado el bajo número de colisiones no se ha optado por esta solución que con un costo de almacenamiento no demasiado alto podría mejorar la velocidad ya que accedería a memoria secundaria una vez en lugar de dos por colisión.

Es posible disminuir el tamaño del *Diccionario* almacenando en lugar de las palabras de las listas de sinónimos su dirección en un vector dispuesto en memoria principal. Esta opción tendría sentido si se puede situar *Diccionario* en memoria principal; en este caso *Diccionario* ocuparía en torno a 1,5Mb y el vector de palabras menos de 0,5Mb. Esta opción tiene sentido si se puede localizar en memoria principal *Control*, *Diccionario* y vector de palabras (aproximadamente 2,3Mb) los accesos a memoria secundaria serían nulos y la velocidad de navegación sería mucho

mayor. No se ha elegido esta solución ya que el aumento de velocidad no justifica para esta aplicación el uso de memoria necesario.

6.- CONCLUSIONES.

1) Esta tesis aporta una solución al problema del reconocimiento y la generación de la morfología del español basándose en un universo de formas flexionadas y derivadas obtenidas a partir de un conjunto de formas principales y de un preprocesador morfológico articulado sobre reglas y modelos.

2) Una de las principales aportaciones consiste en el enfoque orientado a datos más que a reglas —prefijos y pronombres enclíticos— por su mejor realización. El núcleo consiste en una estructura para las raíces y las terminaciones capaz de soportar tal volumen de información gracias a la considerable reducción de espacio que se consigue al tener en cuenta las características morfológicas del español.

3) Sobre tal organización se lleva a cabo tanto el reconocimiento como la generación. El análisis presenta la forma principal de la que proviene, su relación de flexión o derivación y la categoría gramatical a la que pertenece. La producción permite obtener cualquier forma flexionada o derivada a partir de una forma principal.

4) Además de considerar el género y el número en sustantivos y adjetivos y la conjugación en los verbos, en esta tesis se considera el grado superlativo de los adjetivos, la incorporación de pronombres enclíticos en

las formas verbales, la derivación mediante sufijos apreciativos y el uso de prefijos.

5) A pesar de no disponer de una base firme para la sustantivación, la adjetivación, la adverbialización y la verbalización, su solución por vía de las relaciones morfológicas estructurales pone de manifiesto que es posible resolver un problema práctico sin la totalidad de sus fundamentos; por tanto, supone una contribución a la consecución de la teoría. La ampliación de la derivación al caso de cambio en la categoría gramatical se sitúa en la frontera entre morfología y sintaxis y ensancha posibilidades imaginativas.

6) 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 alta interacción hombre-máquina muy amigable fundada en ventanas de diálogo, botones y demás herramientas que caracterizan a una interfaz gráfica.

7) La solución alcanzada proporciona buenos resultados con un universo apreciable; no obstante, se trabaja ya en la adecuación de los datos para la ampliación del universo que no supondrá esfuerzo de programación ni pérdida de realización significativa.

8) La gestión de un diccionario de sinónimos y antónimos constituye una práctica apropiada de la utilidad conseguida; aporta una herramienta que funciona bien con pocos requisitos informáticos. También se ha llevado a efecto mediante una interfaz gráfica y amigable.

9) Entre otras posibilidades, la solución propuesta evolucionará hacia:

- a) la implantación sobre otros soportes como Unix, X-Windows o entorno de red, b) la consideración de sufijos apreciativos adicionales, c) el estudio de la heteronimia y de otras relaciones complejas entre palabras y d) la aplicación a otras lenguas flexivas.

10) Los universos de palabras logrados y la experiencia alcanzada proporcionarán tempranos beneficios en otras aplicaciones entre las que cabe destacar los correctores ortográficos y las búsquedas complejas en bases textuales.

BIBLIOGRAFÍA.

- AÏT-MOKHTAR, S.; RODRIGO MATEOS, J.L.: "Segmentación y análisis morfológico de textos en español utilizando el sistema SMORPH". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*. Nº 17: 29/41. Septiembre, 1995.
- ALARCOS LLORACH, E.: *Gramática de la Lengua Española*. Edit. Espasa-Calpe. Madrid, 1994.
- ALSINA, R.: *Todos los Verbos Castellanos Conjugados*. 17ª edic. Edit. Teide. Barcelona, 1990.
- ALVAR EZQUERRA, M.: *La formación de palabras en español*. Edit. Arco/Libros. Madrid, 1993.
- ALVAR EZQUERRA, M.: *Lexicología y lexicografía: Guía bibliográfica*. Edit. Almar. Salamanca, 1983.
- ALVAR EZQUERRA, M.: "Obtención automática de índices de rimas y de sufijos". *Revista de Dialectología y Tradiciones Populares*. Tomo XXXII: 35/41. Madrid, 1976.
- ALVAR EZQUERRA, M.; MEDINA GUERRA, A.M.: *Manual de Ortografía de la Lengua Española*. Edit. Biblograf. Barcelona, 1995.
- ARONOFF, M.: *Word Formation in Generative Grammar*. Edit. M.I.T. Press. Cambridge, 1976.
- BEAR, J.: "A Morphological Recogniser with Syntactic and Phonological Rules". *11th International Conference on Computational Linguistics (COLING-86)*. Bonn, 1986.
- BELLO, A.: *Gramática de la Lengua Castellana*. Edit. EDAF. Madrid, 1984.
- BOSQUE, I.; PÉREZ FERNÁNDEZ, M.: *Diccionario Inverso de la Lengua Española*. Edit. Gredos. Madrid, 1987.
- CALDER, J.: "Paradigmatic Morphology". *Proceedings of 4th Conference of European Chapter of ACL*. 58/65. Manchester, 1989.
- CHOMSKY, N.: *Aspectos de la Teoría de la Sintaxis*. Edit. Aguilar. Madrid, 1970.
- CHOMSKY, N.: *Estructuras Sintácticas*. Edit. Siglo XXI. Madrid, 1974.

- Diccionario Actual de la Lengua Española*. VOX. Edit. Biblograf. Barcelona, 1990.
- GAZDAR, G.: "Review article: Finite State Morphology". *Linguistics*. Nº 23: 597/607. Berlín, 1985.
- GILI GAYA, S.: *Curso superior de sintaxis española*. VOX. Edit. Biblograf. Barcelona, 1985.
- GÓMEZ TORREGO, L.: *El buen uso de las palabras*. Edit. Arco/Libros. Madrid, 1992.
- GÓMEZ TORREGO, L.: *Manual de Español Correcto*. Edit. Arco/Libros. Madrid, 1991.
- GONZÁLEZ COLLAR, A. L.; GOÑI MENOYO, J. M.; GONZÁLEZ CRISTÓBAL, J. C.: "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, Noviembre 1995.
- Gran Diccionario de Sinónimos y Antónimos*. 4ª edic. Edit. Espasa-Calpe. Madrid, 1991.
- GROSS, M.: *Modelos matemáticos en Lingüística*. Edit. Gredos. Madrid, 1976.
- GUERRERO RAMOS, G.: *Neologismos en el español actual*. Edit. Arco/Libros. Madrid, 1995.
- HALLE, M.: "Prolégomena to a Theory of Word Formation". *Linguistic Inquiry*. Nº 4: 3/16. 1973.
- HALLEBECK, J.: "Una gramática de afijo extendida para el español". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*. Nº 7: 147/157. 1989.
- KAY, M.: "Nonconcatenative Finite-State Morphology". *Proceedings of the Third Conference of the European Chapter of the ACL*. 2/10. Copenhagen, 1973.
- LYONS, J.: *Nuevos horizontes de la Lingüística*. Edit. Alianza. Madrid, 1975.
- LYONS, J.: *Introducción a la Lingüística Teórica*. Edit. Teide. Barcelona, 1979.
- MARCOS MARÍN, F.: *Informática y Humanidades*. Edit. Gredos. Madrid, 1995.
- MARTÍ, M.A.: "Un sistema de análisis morfológico por ordenador". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*. Nº 4: 104/110. 1986.
- MARTINET, A.: *Elementos de lingüística general*. Edit. Gredos. Madrid, 1978.

- MATTHEWS, P.H.: *Morfología: Introducción a la Teoría de la Estructura de la Palabra*. Edit. Paraninfo. Madrid, 1980.
- MEYA, M.: "Análisis morfológico como ayuda a la recuperación de información". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*. Nº 4: 91/103. 1986.
- MORENO SANDOVAL, A.: *Un Modelo Computacional basado en la Unificación para el Análisis y la Generación de la Morfología del Español*. Tesis Doctoral. Universidad Autónoma de Madrid. 1992.
- QUILIS MORALES, A.: *Fonética acústica de la Lengua Española*. Edit. Gredos. Madrid, 1981
- QUILIS MORALES, A.; FERNÁNDEZ, J.A.: *Curso de fonética y fonología españolas para estudiantes angloamericanos*. Edit. Consejo Superior de Investigaciones Científicas. Madrid, 1992
- REAL ACADEMIA ESPAÑOLA: *Esbozo de una Nueva Gramática de la Lengua Española*. Edit. Espasa-Calpe. Madrid, 1989.
- RODRÍGUEZ, G.; HERNÁNDEZ, Z.; SANTANA, O.: "Agrupaciones de Tiempos Verbales en un Texto". *Anales de las II Jornadas de Ingeniería de Sistemas Informáticos y de Computación*, 132/137. Quito (Ecuador). Abril, 1993.
- RODRÍGUEZ MAGRO, C.; SOPEÑA, L.; VALLADARES, C.; VILLAR, C.: "Clasificación morfológica del léxico castellano para un analizador en ordenador". *Actas del VII Congreso Nacional de Lingüística Aplicada*, 491/503. Sevilla, 1990.
- SANTANA, O.; HERNÁNDEZ, Z. J.; RODRÍGUEZ, G.: "Conjugaciones Verbales". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*. Nº 13: 443/450. Febrero, 1993.
- SANTANA, O.; RODRÍGUEZ, J.C.; GONZÁLEZ, J.D.: "FRECTEXT: Una Aplicación de Ayuda a la Elaboración de Documentos". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)*. Nº 13: 451/462. Febrero, 1993.
- SAUSSURE, F.: *Curso de lingüística general*. Edit. Alianza. Madrid, 1987.
- SCALISE, S.: *Morfología Generativa*. Edit. Alianza. Madrid, 1987.
- SECO, M.: *Diccionario de dudas y dificultades de la lengua española*. 9ª edic. Edit. Espasa-Calpe. Madrid, 1991.