Reconocedor y generador automático de formas nominales

Autores: Santana, O.; Pérez, J.; Carreras, F.; Duque, J.D.; Hernández, Z.; Rodríguez, G.
Departamento de Informática y Sistemas
Universidad de las Palmas de Gran Canaria
http://protos.dis.ulpgc.es

RESUMEN

El propósito del presente trabajo consiste en presentar una aplicación informática útil para lematizar las diferentes formas nominales al identificar su forma canónica, categoría gramatical y la flexión o derivación que la produce, y obtiene las formas nominales correspondientes a partir de una forma canónica y de la flexión o derivación solicitada; tanto el reconocimiento como la generación operan sobre una misma estructura de datos —recorrerla en sentidos contrarios implica que la herramientas funciones en una u otra modalidad. Considera: género y número en los sustantivos, adjetivos, pronombres y artículos; heteronimia por cambio de sexo en los sustantivos; grado superlativo en los adjetivos y adverbios; adverbialización y adverbialización del superlativo en los adjetivos; derivación apreciativa en los sustantivos, adjetivos y adverbios; formas canónicas múltiples en todas las categorías gramaticales; formas invariantes tales como preposiciones, conjunciones, exclamaciones, palabras de otros idiomas y locuciones o frases. El sistema incluye composición por prefijación.

0.– INTRODUCCIÓN.

A pesar de la popularización de los ordenadores no se ha resuelto la comunicación entre las personas y las máquinas en lenguaje natural. La ingeniería lingüística constituye un campo de investigación cada vez más estratégico; con el reciente desarrollo de las industrias de la lengua, los usuarios de la tecnología lingüística disponen de recursos que tienden hacia la sociedad de la información multilingüe. El tratamiento automatizado de la morfología del español es la piedra angular sobre la que construir cualquier procesador de lenguaje natural que habrá de considerar ulteriormente la sintaxis y la semántica. La disponibilidad de un procesador morfológico facilita además adecuada solución a una amplia gama de problemas actuales, tales como consultas en bases de datos documentales, corrección ortográfica, lematización, etiquetado, análisis y generación de textos escritos, tratamiento de corpus, etcétera.

El presente trabajo se enmarca dentro de un conjunto de herramientas orientadas a servir de ayuda en la elaboración de documentos escritos —dedicadas a analizar el texto que se produce y a ofrecer facilidades que enriquezcan la expresión— tales como frecuencia de aparición de palabras, empleo de formas verbales y no verbales, corrección ortográfica, búsquedas en texto libre, detección de vicios, depuración de estilos, análisis sintáctico, diccionarios de sinónimos o ideológicos, etc.

Se ha desarrollado un instrumento con la capacidad de:

1. Lematizar una forma no verbal que obtiene junto a la forma canónica, su categoría gramatical y la correspondiente relación de flexión o derivación.
2. Generar una determinada forma flexionada o derivada a partir de la forma canónica.

Considera las siguientes flexiones y derivaciones:

- Género y número en los sustantivos, adjetivos, pronombres y artículos.
- Heteronimia por cambio de sexo en los sustantivos.
- Grado superlativo en los adjetivos y adverbios.
- Adverbialización y adverbialización del superlativo en los adjetivos
- Derivación apreciativa en los sustantivos, adjetivos y adverbios.
- Formas canónicas múltiples en todas las categorías gramaticales.
- Formas invariantes tales como preposiciones, conjunciones, exclamaciones, palabras de otros idiomas y locuciones o frases.

Trata los prefijos que sean necesarios en el análisis y los incorpora en la generación de la forma solicitada.

1.– MORFOLOGÍA NO VERBAL.

Para el estudio de la morfología no verbal 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, pronombres, artículos, adverbios y otras formas que carecen de flexión —preposiciones, conjunciones, exclamaciones, palabras de otros idiomas y locuciones o frases.

Los sustantivos poseen las desinencias flexivas de género y número y, mediante los correspondientes sufijos apreciativos, pueden derivar en sus aumentativos, diminutivos o peyorativos. En general, y con respecto al género, se clasifican en masculinos o femeninos; algunos cambian de género, con o sin repercusiones semánticas, y varían su morfología bien en la desinencia, bien en toda su forma —heteronimia—, o conservan la 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 peyorativos; en general pueden ser adverbializados añadiendo a la forma femenina —de su forma original o de la superlativizada— la terminación *-mente*. Hay adjetivos de dos terminaciones —una para cada género—, y de una terminación —no cambian su morfología con el género.

En los pronombres y artículos se consideran las desinencias flexivas de género (masculino, femenino y neutro) y de número (singular y plural).

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

Las preposiciones, conjunciones, exclamaciones, palabras de otros idiomas y locuciones o frases se tratan como invariantes, no admiten flexión ni derivación.

Aunque sean cuestiones bien sabidas, no está de más recordar cuáles son las reglas morfológicas que rigen estos procesos, pues de esta manera se entenderá mejor la formalización a la que se quiere llegar.

1.1.– LA FORMACIÓN DEL FEMENINO EN SUSTANTIVOS Y ADJETIVOS.

Son de uso genérico las siguientes reglas básicas:

1. Para los sustantivos:
   - 1.1. Los terminados en ***o*** cambian la ***o*** por ***a***: ***niño / niña***
   - 1.2. Los terminados en consonante añaden una ***a***: ***horticultor / horticultora***
   - 1.3. Los terminados en ***e*** permanecen invariables: ***el conferenciante / la conferenciante***, aunque en ocasiones cambian la ***e*** por ***a***: ***franchute / franchuta*** …
   - 1.4. Los terminados en ***a*** permanecen invariables: ***el malabarista / la malabarista***, salvo excepciones: ***poeta / poetisa*** …
2. Para los adjetivos:
   - 2.1. Los gentilicios que acaban en consonante añaden una ***a***: ***andaluz / andaluza***
   - 2.2. Los que acaban en ***o***, ***ote*** o ***ete*** cambian la vocal final por ***a***: ***grandote / grandota***
   - 2.3. Los que acaban en ***an***, ***on*** u ***or*** añaden una ***a***: ***bribón / bribona***, salvo los comparativos latinos: ***exterior***, ***inferior***, ***menor***, ***peor***,…
   - 2.4. El resto permanecen invariantes: ***amable***, ***azul***, ***decente***, ***verde***,…, salvo excepciones: ***el cliente / la clienta*** …

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

3. Excepciones a la formación del femenino:
   - 3.1. Muchos sustantivos son sólo masculinos: ***libro***, ***objeto***,…
   - 3.2. Muchos sustantivos son sólo femeninos: ***casa***, ***legaña***,…
   - 3.3. Hay sustantivos femeninos que acaban en ***o***: ***la dinamo***, ***la bonoloto***,…
   - 3.4. Hay sustantivos masculinos que acaban en ***a***: ***el califa***, ***el fotograma***,…
   - 3.5. Hay sustantivos de género común y adjetivos de una terminación, cuya morfología no cambia al usarlos en masculino o en femenino: ***el cónyuge / la cónyuge***,…
   - 3.6. Hay sustantivos de género ambiguo —aunque el uso va reduciéndolos o dándoles una nueva distribución, se crean otros casos por analogía, ignorancia, etc.—, normalmente pueden ser usados en ambos géneros: ***el mar / la mar,***…
   - 3.7. En casos especiales pueden aparecer otras terminaciones como ***esa***, ***isa***, ***ina***, ***iza*** o ***triz***: ***sacerdote / sacerdotisa,***…

3.8. Algunos sustantivos tienen dos formas para el femenino: ***actor / actora / actriz,***…

3.9. Algunos sustantivos poseen heteronimia por cambio de sexo, cambian la forma de la palabra y no sólo su desinencia: ***toro / vaca,***…

3.10. Se ha de resaltar aquí el problema que surge de la incorporación de la mujer a trabajos, oficios, etc., tradicionalmente de hombres —unas veces la Academia admite el femenino y otras no, unas veces los admite la sociedad y otras no—; en un esfuerzo por lograr la máxima generalidad en este trabajo se recogen tanto los aceptados por la Academia como los admitidos por el uso (ya incorporados en otros diccionarios).

3.11. También se han tenido en cuenta, aún siendo más raros, los masculinos regresivos: ***comadrona / comadrón*** …

4. Conviene tener en cuenta que existen voces que cambian su significado con el género: ***el frente / la frente,***…, o según la forma del femenino: ***la directora / la directriz***,…; y que el género *no es igual* al sexo: ***la foca macho / el cocodrilo hembra***.

1.2.– LA FORMACIÓN DEL PLURAL EN SUSTANTIVOS Y ADJETIVOS.

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

1. Para las palabras terminadas en consonante:

1.1. Distinta de ***z***, ***c***, ***x***, ***n***, ***s*** añaden ***es***: ***árbol / árboles,*** con excepciones que añaden ***s***: ***réquiem / réquiems***,…; y, en este caso, si terminan en ***y*** se hace la corrección ortográfica: ***guirigay / guirigáis***,…

1.2. Las palabras terminadas en ***z***, precisan corrección ortográfica, la sustituyen por ***ces***: ***matriz / matrices***

1.3. Las palabras terminadas en ***c***, precisan corrección ortográfica, la sustituyen por ***ques***: ***ruc / ruques***, con excepciones que añaden ***s***: ***coñac / coñacs***,…

1.4. Las palabras terminadas en ***x*** permanecen invariables: ***el clímax / los clímax***, con excepciones que añaden ***es*** y precisan modificación ortográfica: ***fénix / fénices***,…

1.5. Las palabras terminadas en ***n*** o ***s*** que sean agudas añaden ***es***: ***gas / gases***, si llevan tilde la pierden: ***camión / camiones***, salvo que sea para formar hiato: ***mohín / mohínes*** o diacrítica: ***quién / quiénes***.

1.3. Las palabras terminadas en ***n*** que no sean agudas añaden ***es***: ***oxímoron / oxímorones***, y las llanas ganan una tilde al pasar a esdrújulas: ***velamen / velámenes***.

1.4. Las palabras terminadas en ***s*** que no sean agudas permanecen invariables: ***el guardagujas / los guardagujas***.

2. Para las palabras terminadas en vocal:

2.1. No acentuada añaden ***s***: ***copa / copas***, excepto los monosílabos que añaden ***es***: ***la i / las íes***; salvo las notas musicales que añaden ***s***.

2.2. ***á*** u ***ó***, añaden ***es*** y pierden la tilde: ***abacá / abacaes***, con excepciones que añaden ***s***: ***dominó / dominós***,…

2.3. ***é***, añaden ***s***: ***café / cafés***

2.4. ***í*** o ***ú***, añaden ***es***: ***baladí / baladíes***, con excepciones que añaden ***s***: ***canesú / canesús***,…

Además, hay un número importante de excepciones que pueden concretarse en:

3. Excepciones a la formación del plural:

3.1. Palabras que sólo se usan en singular: ***cariz***, ***cenit***, ***salud***,…

3.2. Palabras que sólo se usan en plural: ***albricias***, ***ambages***, ***anales***,…, aunque los hablantes ponen algunas de ella bajo la forma singular: ***braga, gafa***,…

3.3. Palabras invariantes: ***el quórum / los quórum,***…

3.4. La desinencia del plural no aparece al final de la palabra: ***medianoche / mediasnoches***,…

3.5. Cambian la sílaba tónica al formar el plural: ***carácter / caracteres***,…

3.6. Otras irregularidades: ***desiderátum / desiderata***, ***vermut / vermús***,…

3.7. Palabras que tienen más de una forma para el plural: ***accésit / accésit / accésits / accesis***,…

3.8. Palabras que cambian el género al formar el plural: ***la orina / los orines***,…

3.9. Palabras de género ambiguo que forman el plural con un solo género: ***el mar***, ***la mar / los mares***,…

4. Conviene tener en cuenta que hay palabras que pueden cambiar su significado con el número: ***celo / celos,***…

1.3.– EL GÉNERO Y EL NÚMERO EN LOS PRONOMBRES.

Los pronombres personales, demostrativos, indefinidos, relativos, interrogativos y posesivos se tratan como en la mayoría de las gramáticas.

1.4.– EL GÉNERO Y EL NÚMERO EN LOS ARTÍCULOS.

El artículo tiene las formas sabidas ***el***, ***la***, ***los***, ***las*** y el neutro singular ***lo***.

1.5.– EL GÉNERO Y EL NÚMERO EN LOS NUMERALES: CARDINALES, ORDINALES, FRACCIONARIOS Y PROPORCIONALES.

Los numerales son adjetivos, nombres o pronombres, y, en ocasiones, adverbios según la función gramatical que desempeñen. Se clasifican en cardinales, ordinales, partitivos o fraccionales y múltiplos o proporcionales.

Los cardinales no tienen variación de número: ***uno*** es singular y el resto son plurales; salvo ***millón/millones, billón/billones***, etcétera. Son invariantes con el género salvo ***uno/una***, ***veintiuno/veintiuna*** y los acabados en ***-ientos*** que para el femenino usan ***-ientas***. Los ordinales y los fraccionarios o partitivos admiten variación de género y

número con las desinencias ***o/a/os/as***. En los proporcionales o múltiplos la variación de género sólo la tienen los terminados en ***-o*** que forman el femenino con ***-a***, todos forman el plural añadiendo ***-s***.

1.6.– EL GRADO SUPERLATIVO.

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

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

amable / amabilísimo / amabilísima / amabilísimos / amabilísimas

— Se precisan las correcciones ortográficas pertinentes:

— Si acaba en ***ca***, ***co*** o ***cu*** se convierte la ***c*** en ***qu***: ***ri-c-o / ri-qu-ísimo***

— Si acaba en ***z***, ***za***, ***zo*** o ***zu*** se convierte la ***z*** en ***c***: ***efica-z / efica-c-ísimo***

— Si acaba en ***ga***, ***go*** o ***gu*** se convierte la ***g*** en ***gu***: ***va-g-a / va-gu-ísima***

— Existen adjetivos con irregularidades en la formación del superlativo:

inicuo / iniquísimo, ***sabio / sapientísimo***,…

— Otros que además de la forma regular admiten una o varias formas irregulares

enemigo / enemiguísimo / enemicísimo / inimicísimo,…

— No todos los adjetivos admiten el grado superlativo en una formación regular:

— Por su significado: ***absoluto***, ***omnipotente***, ***infinito***,…

— Por tener una marca superlativizadora: ***buenísimo***, ***mínimo***, ***óptimo***,…

— Por ser gentilicios: ***asturiano***, ***canario***,…

— Por otros motivos: ***exiguo***, ***político***, ***público***,…

— Aunque el grado superlativo es una característica propia de los adjetivos, existen adverbios que lo admiten: ***cerca*** / ***cerquísima***, ***lejos*** / ***lejísimos***, ***tarde*** / ***tardísimo***,…

1.7.– LA ADVERBIALIZACIÓN.

Los adjetivos permiten la formación de adverbios de modo mediante la concatenación de su forma femenina con la terminación ***-mente***: ***irónico*** / ***irónicamente***; si son de una terminación, se añade directamente: ***afable*** / ***afablemente***. Sin embargo, hay adjetivos que no admiten esta formación: ***mucho***, ***ninguno***,…

También es posible adverbializar el superlativo de los adjetivos, añadiendo la terminación ***-mente*** a su forma femenina: ***claro / clarísimo / clarísimamente***.

1.8.– SUFIJOS: CONSIDERACIONES GENERALES Y REGLAS DE DERIVACIÓN.

Una manera habitual 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. Por muy conocido que sea, a continuación se hará una detallada recapitulación de la

derivación, con el objeto de ejemplificar bien todo lo que el sistema de lematización y flexión que se ha desarrollado tiene en cuenta.

— Si la primitiva termina en vocal pierde las letras finales ***a***, ***e***, ***o***: ***verd-e / verd-or***, en algunas ocasiones pierde todo el diptongo final: ***palac-io / palac-ete*** y puede alterar la forma del lexema: ***rab-ia / ráb-ico***; cuando termina en ***u***, ***i*** normalmente el sufijo se añade sin sufrir modificaciones: ***cursi / cursi-lería*** y es frecuente la reducción del diptongo resultante: ***tribu / tribual / tribal*** o la vocal repetida: ***dandi / dandi-ismo / dandismo***.

— Si termina en consonante, normalmente se añade el sufijo sin más: ***verbal / verbal-ismo***; las terminadas en ***-dad*** pierden el ***-ad*** final: ***humed-ad / humed-ecer***, sin embargo, existen excepciones: ***virus / viro-logía***,…

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

— Si acaba en ***c*** con sonido */z/*, ante ***a***, ***o*** o ***u*** se convierte en ***z***: ***dul-c-e / dul-z-ura***

— Si acaba en ***c*** con sonido */k/*, ante ***e*** o ***i*** se convierte en ***qu***: ***taba-c-o / taba-qu-ería***

— Si acaba en ***qu*** ante ***a***, ***o*** o ***u*** se convierte en ***c***: ***miriña-qu-e / miriña-c-ote***.

— Si acaba en ***g*** con sonido */x/*, ante ***a***, ***o*** o ***u*** se convierte en ***j***: ***esfin-g-e / esfin-j-ucha***

— Si acaba en ***g*** con sonido */g/*, ante ***e*** o ***i*** se añade una ***u***: ***tra-g-o / tra-gu-ear*** y si el sufijo comienza por ***u*** seguida de vocal débil se debe añadir la diéresis para conservar el diptongo: ***lla-g-a / lla-g-üela***.

— Si acaba en ***gu***

a) tras eliminar ***a*** u ***o***, ante ***e*** o ***i*** se pone diéresis para mantener el sonido */gu/*: ***ambi-gu-o / ambi-gü-edad***.

b) tras eliminar ***e***, ante ***a***, ***o*** o ***u*** se elimina la ***u*** para mantener el sonido /g/: ***meren-gu-e / meren-g-ada*** y si el sufijo comienza por ***u*** seguida de vocal débil se debe reducir la doble vocal y añadir la diéresis para conservar el diptongo: ***pira-gu-a / pira-g-üela***.

Una clase importante de sufijos la constituyen los apreciativos: se caracterizan por imprimir un matiz semántico subjetivo sin alterar normalmente la categoría gramatical y son tónicos. Según el tipo de eufemismo que producen se clasifican en: aumentativos, diminutivos o peyorativos. Además de los sufijos apreciativos discrecionales en cada caso, se tiene en cuenta una importante cantidad de sufijos adicionales extraídos de las fuentes estudiadas; se recogen formas derivadas con más de un matiz semántico —***abogadillo***, diminutivo despectivo de ***abogado***—; en este trabajo no se han tratado los meliorativos ya que en ninguna palabra de los diccionarios consultados figura tal característica. No suelen llevar apreciativos las sustantivos abstractos —la mayoría de los terminados en ***ad*** lo son, ***libertad***, ***igualdad***, ***fraternidad***,… en cambio ***ciudad*** no lo es.

1.8.1.– La formación de aumentativos.

Morfena normalmente subfijo, que añade al significado de la base léxica a la que se une la noción de magnitud o agrandamiento. Puede aportar, a la vez, otros valores,

especialmente el de desprecio. Forman aumentativos los sustantivos, los 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.

— Las palabras que terminan en vocal tónica, añaden el interfijo ***-z-*** para la formación de sus aumentativos (mantienen la vocal, aunque pierden la tilde): ***papá / papa-z-ote***.
— Además de éstos existe un importante número de sufijos que se utilizan con menor frecuencia para la formación de aumentativos, algunos son combinaciones de otros: ***hues-o / hues-arrón***, ***pícar-o / picar-onazo***, ***grand-e / grand-ullón***,…
— Existen algunas voces que reducen un diptongo en el lexema: ***fuert-e / fort-achón***.
— En ocasiones, sustantivos femeninos forman aumentativos en masculino: ***la vel-a / el vel-ón***,… y puede darse además la reducción de diptongo en el lexema: ***cazuel-a / cazol-ón***,…
— A veces la formación de aumentativos es muy irregular: ***nariz / narigón***,…
— Aunque no es frecuente, en algunos adverbios se forman aumentativos: ***antañ-o / antañ-azo***, ***lej-os / lej-otes***,…

1.8.2.– LA FORMACIÓN DE DIMINUTIVOS.

Morfena normalmente subfijo, que añade al significado de la base léxica a la que se une la noción de pequeñez en cantidad o tamaño. Puede aportar, a la vez, valores apreciativos, especilamente de afecto, aunque también irónicos y de desprecio. Los sustantivos, los adjetivos 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 utilizan ***-ico*** (aunque más bien es regional: Aragón, Navarra, Murcia y algunas zonas de Andalucía y Sudamérica), ***-ín*** (frecuente en Asturias) y ***-uelo*** (de aplicación más restringida y en ocasiones con carácter peyorativo) —admiten femeninos y plurales: ***-ica***/***-icos***/***-icas***, ***-ina***/***-ines***/***-inas***, ***-uela***/***-uelos***/***-uelas***. 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 de dos o más sílabas
  — agudas terminadas en ***n*** o en ***r***: ***camión / camion-c-ito***
  — acabadas en vocal tónica: ***mamá / mama-c-ita***
  — y las llanas acabadas en ***n***: ***dictamen / dictamen-c-illo***.
— Añaden el interfijo ***-ec-***
  — los monosílabos acabados en consonante: ***son / son-ec-illo***
  — los bisílabos terminados en ***e***: ***cort-e / cort-ec-ito***
  — 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***: ***savi-a / savi-ec-illa***, aunque en algunos casos no se añade el interfijo: ***len-gu-a / len-gü-eta***

— y las voces con dos o más sílabas terminadas en ***io***: ***cenobi-o / cenobi-ec-ito*** y en ocasiones hay reducción de diptongo: ***pie-z-a / pe-c-ez-uela***.

— Añaden el interfijo ***-ecec-*** los monosílabos acabados en vocal: ***té / t-ecec-ito***

— Además de los mencionados, existe un importante número de sufijos que se utilizan con menor frecuencia para la formación de diminutivos, algunos son combinaciones de otros: ***alegr-e / alegr-ete***, ***cuerp-o / corp-iño***, ***caf-é / caf-etín***, ***bob-o / bob-irrinchín***, ***chic-o / chiqu-itín / chic-orrotín / chiqu-irritín***,…

— Hay algunas voces que pierden el diptongo del lexema al formar diminutivos: ***cuern-o / corn-ecito***, ***viej-o / vej-ecillo***,…

— Algunos sustantivos femeninos forman diminutivos en masculino: ***la faj-a / el faj-ín***,…

— A veces la formación de diminutivos es muy irregular: ***azúcar / azuquítar***,…

— Aunque no es frecuente, con algunos adverbios se forman diminutivos: ***apen-as / apen-itas***,… y en ocasionas de manera irregular: ***ahor-a / ahor-ita / hor-itica***.

1.8.3.– LA FORMACIÓN DE PEYORATIVOS.

Morfena normalmente subfijo, que añade al significado de la base léxica a la que se une el valor de desprecio. Se unen a sustantivos, adjetivos y muy raramente a adverbios.

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

— Las palabras que terminan en vocal tónica, añaden el interfijo ***-c-*** para la formación de sus peyorativos (mantienen la vocal, aunque pierden la tilde): ***puré / pure-c-ejo***.

— Además de éstos existe un importante número de sufijos que se utilizan con menor frecuencia: ***bich-o / bich-arraco***, ***cald-o / cald-ibache***, ***cur-a / cur-ángano***,…

— Existen algunas voces que pierden un diptongo intermedio: ***cuerv-o / corv-ucho***,…

— En ocasiones, sustantivos femeninos forman peyorativos en masculino: ***la alde-a / el alde-orro***, ***la cam-a / el cam-astro***,…

— A veces la formación de peyorativos es muy irregular: ***mezcla / mezcolanza***, ***francés / franchute***, llegando a la heternonimia: ***animal / alimaña*** e incluso a la composición ***cojo / cojitranco***,…

— Son infrecuentes los adverbios que forman peyorativos: ***arrib-a / arrib-ota***,…

1.9.– FORMAS INVARIANTES.

El resto de categorías gramaticales se consideran invariantes —no poseen flexión ni derivación—; entre ellas se cuentan las preposiciones, las conjunciones, las exclamaciones, las locuciones o frases y las palabras de otros idiomas.

1.10.– FORMAS CANÓNICAS MÚLTIPLES.

Algunas palabras vacilan entre varias posibles grafías —***cardiaco***/***cardíaco***, ***gambuj***/***gambujo***/***gambux***,…—; para reflejar este heteromorfismo se establece el concepto de forma canónica múltiple que relaciona estas formas entre sí.

1.11.– LA PREFIJACIÓN.

La prefijación es una operación de derivación o de composición —según la teoría— que normalmente matiza, corrige o modifica el significado de la palabra, con independencia de la flexión, sin cambiar su categoría gramatical. No son aplicables en las formas consideradas invariantes, en los pronombres ni en los artículos.

Los prefijos considerados se pueden agrupar según su significación:

1. A ambos lados o alrededor de: ***anfi-***, ***circun-***.
   - — El prefijo ***anfi-*** se convierte en ***anfí-*** ante palabras bisilábicas llanas: ***anfi- + podo / anfípodo***.
   - — El prefijo ***circun-*** toma la forma ***circum-*** ante ***b***, ***p*** o ***m***: ***circun- + molar / circummolar*** y en ocasiones ante ***n***: ***circun- + nutación / circumnutación***. En ocasiones elimina la ***e*** ante palabras que comienzan por esta vocal: ***circun- + escrito / circunscrito***.
2. A distancia o lejos: ***tele-***.
   - — El prefijo ***tele-*** puede convertirse en ***telé-*** ante palabras bisilábicas llanas: ***tele- + fono / teléfono*** y reduce la ***e*** ante palabras que comienzan por esta vocal: ***tele- + espectador / telespectador***.
3. A través de, cambio o al otro lado: ***tras-***, ***trans-***, ***ultra-***, ***meta-***.
   - — Los prefijos ***tras-*** y ***trans-*** reducen la ***s*** ante palabras que comienzan por esta letra: ***trans- + sexual / transexual*** y pueden hacer desaparecer la inicial en palabras que comienzan por ***e***: ***trans- + emisión / transmisión***; la reducción de la ***s*** puede ocurrir tras la desaparición de la ***e***: ***tras- + esquilador / trasquilador***. Pueden convertirse en ***trás-*** y ***tráns-***: ***tras- + fuga / trásfuga***.
4. Acción secundaria o atenuación del significado: ***so-***, ***sub-***, ***za-***, ***zam-***, ***sus-***, ***entre-***.
   - — El prefijo ***sus-*** a veces hace perder la ***e*** a palabras que comienzan por esta vocal: ***sus- + estrato / sustrato***.
5. Arriba, en alto o sobre: ***ana-***, ***epi-***, ***supra-***.
6. Alejamiento, separación y privación: ***ab-***, ***abs-***, ***dis-***, ***dia-***, ***di-***.
7. Aumento, encarecimiento o refuerzo del significado: ***re-***, ***rete-***, ***requete-***, ***archi-***, ***super-***, ***sobre-***, ***hiper-***, ***ultra-***, ***extra-***.
   - — El prefijo ***sobre-*** puede admitir además la forma con reducción de ***e*** ante palabras que comienzan por esta letra: ***sobre- + esdrújulo / sobreesdrújulo / sobresdrújulo***.
8. Bajo o debajo: ***so-***, ***sub-***, ***za-***, ***zam-***, ***sus-***, ***sota-***, ***soto-***.
9. Compañía: ***co-***, ***con-***.
   - — El prefijo ***con-*** cambia la ***n*** por ***m*** ante ***b*** o ***p***: ***con- + paternidad / compaternidad***.
10. Conforme a: ***ana-***.
    - — El prefijo ***ana-*** no se usa ante palabras que comienzan por ***a*** y se apocopa ante vocal: ***ana- + ion / anión***.
11. Contra, contrariedad, contrario, oposición, opuesto o rechazo: ***contra-***, ***anti-***, ***di-***, ***dis-***, ***para-***, ***re-***.
    - — El prefijo ***anti-*** puede convertirse en ***antí-***: ***anti- + tesis / antítesis***.

12.Delante, anterioridad en tiempo y espacio o prioritario: ***pro-***, ***pre-***, ***ante-***, ***proto-***.

13.Dentro: ***intro-***, ***endo-***, ***en-***.

— El prefijo ***endo-*** en ocasiones hace perder la ***e*** a palabras que comienzan por esta vocal: ***endo- + esfera / endosfera***.

— El prefijo ***en-*** cambia la ***n*** por ***m*** ante ***b*** o ***p***: ***en- + bolso / embolso***.

14.Después de en el sentido de detrás: ***pos-***, ***post-***, ***meta-***.

15.Doble o dos: ***bi-***, ***bis-***, ***anfi-***, ***di-***.

16.En vez de o por substitución: ***pro-***.

17.Entre o en medio: ***inter-***, ***dia-***.

18.Extensión o dilatación: ***di-***.

19.Exterior, junto o próximo: ***para-***, ***epi-***, ***yuxta-***, ***ad-***.

— El prefijo ***para-*** puede convertirse en ***pará-***: ***para- + metro / parámetro***.

20.Fuera, más allá o externo: ***extra-***, ***ex-***, ***es-***, ***des-***, ***ecto-***, ***meta-***, ***supra-***, ***ultra-***, ***exo-***.

— El prefijo ***exo-*** reduce la ***o*** ante palabras que comienzan por esta vocal: ***exo- + oftálmico / exoftálmico***, si aparece tildada, puede mantenerse o no la tilde: ***exo- + ósmosis / exósmosis*** / ***exosmosis*** y ante bisílabas llanas en ocasiones se convierte en ***exó-***: ***exo- + gamo / exógamo*** pero en otras no: ***exo- + dermis / exodermis***. A veces hace perder la ***e*** a las palabras que comienzan por esta vocal: ***exo- + espora / exospora***.

21.Hacia atrás, de nuevo, tiempo anterior o inversión de la acción: ***retro-***, ***des-***, ***re-***, ***ana-***.

22.Igual: ***equi-***, ***iso-***.

23.Inferioridad: ***hipo-***.

— A veces se convierte en ***hipó-***: ***hipo- + tesis / hipótesis*** y en ocasiones se pierde la ***e*** en palabras que comienzan por esta vocal: ***hipo- + estasis / hipostasis***.

24.Lejos de o separado de: ***apo-***.

25.Medio, casi o mitad: ***semi-***, ***hemi-***.

— El prefijo ***hemi-*** a veces hace perder la ***e*** a las palabras que comienzan por esta vocal: ***hemi- + esfera / hemisferio***.

26.Origen o procedencia: ***di-***, ***ab-***.

27.Parte de acá: ***cis-***, ***citra-***.

28.Parte de atrás: ***opisto-***.

29.Preeminencia, primacía, superioridad o prioridad: ***archi-***, ***proto-***.

30.Privación, negación o ausencia: ***re-***, ***a-***, ***des-***, ***de-***, ***dis-***, ***ex-***, ***ana-***, ***in-***.

— El prefijo ***a-*** se convierte en ***an-*** ante vocal: ***a- + estesia / anestesia***.

— El prefijo ***in-*** se convierte en ***im-*** ante ***b*** o ***p***: ***in- + borrable / imborrable***, en ***i-*** ante ***l*** o ***r***: ***in- + legítimo / ilegítimo***, pero en el caso de **r** ésta se dobla para mantener el sonido: ***in- + realidad / irrealidad***.

31.Progreso, continuidad de acción o hacia adelante: ***pro-***, ***para-***.

32.Que suple, hace las veces de, ocupa el segundo lugar en categoría o subalterno: ***vice-***, ***viz-***, ***vi-***, ***sota-***, ***soto-***, ***sub-***.

33.Situación o calidad intermedia: ***entre-***, ***entro-***.

34.Unión: ***co-***, ***con-***, ***sin-***.

— Los prefijos ***con-*** y ***sin-*** cambian la ***n*** por ***m*** ante ***b*** o ***p***: ***con- + patriota / compatriota***

— El prefijo ***sin-*** en ocasiones se convierte en ***sín-***: ***sin- + tesis / síntesis***

Además, el proceso de unión entre prefijos y formas debe tener en cuenta las siguientes reglas:

1. Cuando se añade un prefijo a un monosílabo sin tilde que acaba en vocal, ***n*** o ***s*** debe tildarse: ***requete- + bien / requetebién***.
2. Cuando se añade un prefijo terminado en vocal a una forma que comienza por ***r*** debe duplicarse la ***r*** para mantener el sonido fuerte: ***contra- + reloj / contrarreloj***.
3. Cuando se añade un prefijo terminado en vocal fuerte a una forma que comienza por vocal débil tónica no tildada, ésta debe tildarse porque se forma un hiato: ***entre- + ido / entreído***, aunque vaya precedida de ***h***: ***re- + hilo / rehílo***.

2.– PRODUCCIÓN DE LAS FORMAS NO VERBALES FLEXIONADAS Y DERIVADAS.

En este trabajo, se parte de una base inicial con un total de 109 194 formas canónicas. Se han incluido todas las entradas no verbales del *Diccionario de la Lengua Española* de la Real Academia Española (70 056), del *Diccionario General de la Lengua Española* Vox (83 709), del *Diccionario de Uso del Español* de María Moliner (66 099), del *Gran Diccionario de la Lengua Española* de Larousse Planeta (58 605), del *Diccionario de voces de uso actual* dirigido por Manuel Alvar Ezquerra (4 644), del *Gran Diccionario de Sinónimos y Antónimos* de Espasa-Calpe (31 011) y del *Diccionario Ideológico* de Julio Casares (56 533).

Cada registro de la base inicial contiene: a) la forma canónica representante del registro, b) su categoría gramatical, c) la parte invariante, d) las terminaciones que permiten los cambios de género y de número —producto de la aplicación de las reglas descritas en los apartados que van del 1.1 al 1.5 considerando todas las excepciones e irregularidades—, e) la información sobre irregularidades y excepciones en la formación de apreciativos y en la del superlativo, f) las excepciones a la adverbialización y g) las formas canónicas relacionadas por multiplicidad o heteronimia.

Se construye el léxico mediante un proceso de generación que opera sobre los registros de la base inicial y produce las formas flexionadas y derivadas asociadas a cada forma canónica. Tal expansión se lleva a cabo gracias a la información contenida en el registro y a la aplicación de las reglas estudiadas en los apartados 1.8 y 1.9 —cada elemento del léxico conserva su relación de flexión o derivación con la forma canónica de la que procede.

El léxico obtenido se compone de más de tres millones y cuarto de formas (3 328 283). No se ha considerado la exagerada ampliación que produciría la prefijación (se cuenta con más de ochenta prefijos y con que una misma forma puede admitir varios); se deja a la discrecionalidad del usuario en la aplicación final ya que las reglas del apartado 1.11 se ejecutan con eficacia y no se justifica el desproporcionado aumento del volumen de información.

3.– ESTRUCTURACIÓN DE LOS DATOS.

La solución aportada se orienta a datos más que a reglas, con el fin de obtener unos mejores resultados. Dado el considerable volumen de datos, se ha diseñado una estructura para su almacenamiento en memoria secundaria que consigue un adecuado equilibrio entre ocupación y velocidad de recuperación.

Debido al carácter flexivo de la lengua española, se opta por un conjunto de terminaciones que permita la generación de todas las formas flexionadas y derivadas por simple concatenación a partir de la raíz de la forma canónica —las irregularidades y los cambios ortográficos se manifiestan con la aparición de un cambio en la raíz. Se usa un criterio de corte que genera un número mínimo de raíces, aunque no siempre dé lugar a la raíz lingüística; tal conjunto de raíces tiene una cardinalidad bastante menor que el de formas —180 775 frente a 3 328 283—; por tanto, resulta más favorable afrontar la solución haciendo la partición de las formas en raíces y terminaciones. Se organiza la estructura por raíces y cada registro contiene la información de las terminaciones de flexión o derivación que esa raíz acepta y la referencia al registro donde se encuentra la raíz de su forma canónica —útil para el reconocimiento en caso de raíces múltiples. Los registros correspondientes a formas canónicas incluyen además la categoría gramatical y la referencia a los registros de sus raíces alternativas —útil para la generación.

A partir de una raíz, se accede a la base de raíces de la que se obtiene su posición en la base de terminaciones. Junto a la raíz de una forma canónica, aparecen su categoría gramatical —útil para el proceso de lematización— y las raíces alternativas —útil para la generación de formas flexionadas o derivadas que no posean la misma raíz que la forma canónica. La base de terminaciones contiene la información acerca de las terminaciones que admite una raíz y de la formación de su forma canónica. El grupo de terminaciones permite averiguar qué terminaciones puede llevar una raíz y qué flexión representa para esa raíz cada una de ellas. La existencia de raíces alternativas y de los grupos de terminaciones que admiten obedece al conjunto de terminaciones que se considera en el proceso de construcción.

4.– LEMATIZACIÓN.

El proceso de identificación actúa sobre una palabra de entrada por medio de un *segmentador* que la descompone en: a) los posibles pares raíz-terminación y b) los prefijos que pudieran poseer, figura 1. La raíz pasa al *módulo de índices* que determina su localización. El *módulo de accesos externos*: a) comprueba si la raíz admite la terminación, b) determina a qué flexión o derivación corresponde, c) deduce la forma canónica de la que proviene y d) proporciona su categoría gramatical.

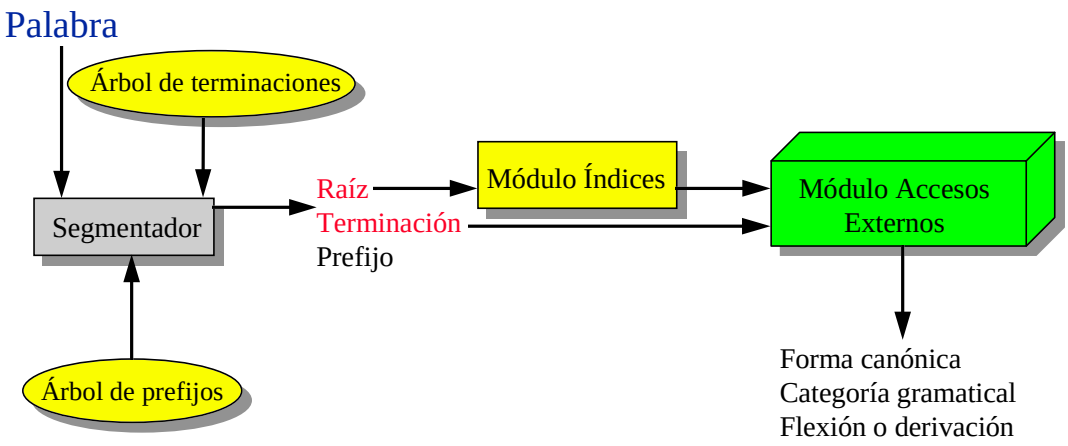

Figura 1: Esquema del lematizador

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

c***rédulamente***: ***crédulament-e***, ***crédula-mente***, ***crédul-amente***

En el *módulo de accesos* externos se rechazan los pares no localizados con el *módulo de índices*, pero en los casos exitosos se pasa a lematizar.

crédulament-e: forma canónica del advervio de modo ***crédulamente***

crédul-amente: adverbialización del adjetivo ***crédulo***

Para tratar los prefijos se aíslan situando sus cadenas de caracteres en el índice correspondiente y se procede a segmentar en pares raíz-terminación la palabra descargada de prefijos: ***archimaleado / a+rchimaleado*** o ***arc+himaleado*** o ***archi+maleado***.

5.– GENERACIÓN.

Al disponer de una estructura de datos que permite conocer para cada palabra cuál es su forma canónica, qué raíces tiene, qué terminaciones admite cada raíz y qué flexión o derivación presenta cada una de ellas, es posible generar con poco esfuerzo una forma a partir de la canónica y de la flexión o derivación propuesta. Basta con disponer de la capacidad de acceder a las distintas raíces que aparecen al flexionar o derivar una determinada forma canónica y conocer qué terminaciones corresponden a la flexión o derivación dada.

La entrada al generador está constituida por a) una forma canónica, b) la flexión o derivación correspondiente y c) los prefijos, figura 2.

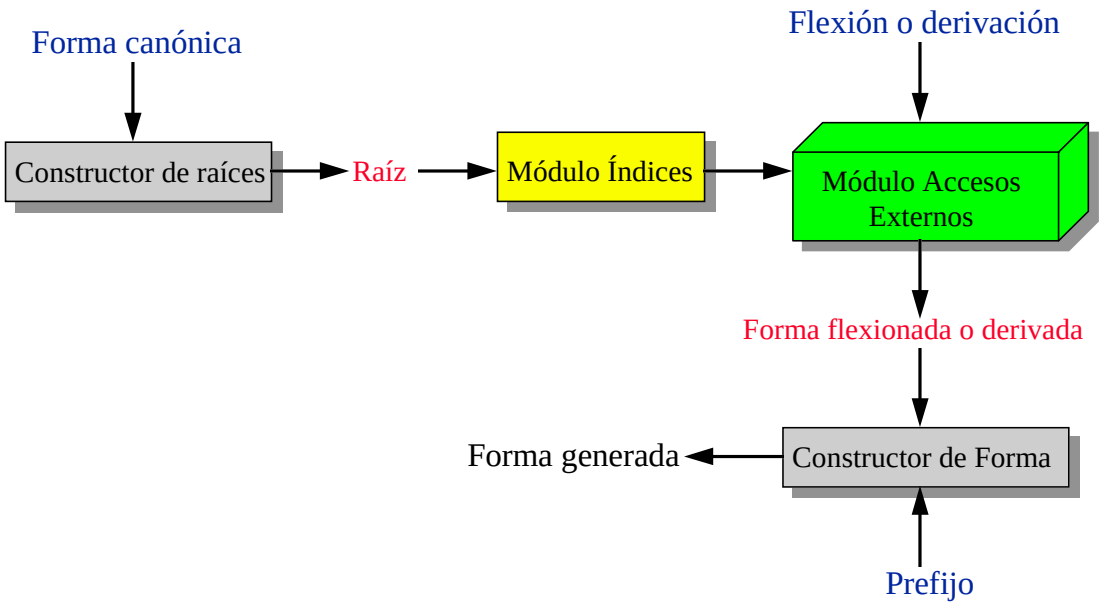

Figura 2: Esquema del generador

El *constructor de raíces* descarta la terminación y obtiene la raíz de la forma canónica, el *módulo de índices* localiza esa raíz en la base de raíces y gracias al *módulo de accesos externos* se llega a 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 1.11 dedicada a la prefijación.

6.– RESULTADOS EXPERIMENTALES.

La aplicación se ha desarrollado en C++. Los requerimientos para su funcionamiento son los que necesita Windows y el espacio mínimo en memoria externa para datos y código no supera los 20 Mb. Las estructuras de datos utilizadas por el programa ocupan un total de 19,1 Mb en disco; contienen compactadamente los 53,4 Mb que ocupan las 3 328 283 formas flexionadas y derivadas, además de la flexión o derivación que representan, la reseña sintáctica sobre categorías gramaticales y toda la información referente a prefijación.

El universo de formas tanto reconocibles como generables se compone de: 3 328 283 formas. Gran parte de este universo puede ser multiplicado por un factor cercano a 80, tantas veces como prefijos se combinen, por ejemplo combinando con un prefijo el universo supera los 200 millones de palabras —este corpus ocuparía cerca de tres gigabytes.

Sobre un procesador Pentium II a 300 Mhz con 128 Mb de memoria RAM, se generan las formas flexionadas o derivadas de las formas canónicas a una velocidad de 1010 formas por segundo y se identifican a razón de 480 formas por segundo. Si se incorporan prefijos a las formas se generan 670 formas por segundo. Todos los casos, tanto de generación como de reconocimiento, se han presentado de forma aleatoria para

evitar los efectos, favorables o desfavorables, de seguir un orden. Sobre un texto literario constituido por 111 690 palabras, el reconocimiento —sin considerar prefijos— se efectúa a razón de 590 formas por segundo; si no se consideran las formas verbales —constituyen el 10% del texto— la velocidad de lematización es de 570 formas por segundo, ya que se detectan las formas no reconocibles a razón de 800 palabras por segundo; lematizar con tratamiento de prefijos supone 450 formas por segundo y se reconoce un 1,3% más del texto; el 0,7% del texto lo constituyen 527 palabras que no son reconocidas por tratarse de nombres propios, toponimia, jerga o numeración.

8.– CONCLUSIONES.

Se ha logrado una herramienta que permite identificar a partir de una forma no verbal la forma o formas canónicas de las que proviene, su categoría gramatical y la flexión o derivación; incluye: el género y el número en los sustantivos, adjetivos, pronombres y artículos; la derivación apreciativa —aumentativo, diminutivo y peyorativo— en los sustantivos, adjetivos y adverbios; el grado superlativo en los adjetivos y adverbios; la adverbialización y la adverbialización del superlativo en los adjetivos; la heteronimia por cambio de sexo; las formas canónicas múltiples; los prefijos que pueda poseer. Se tienen en cuenta formas invariantes tales como preposiciones, conjunciones, exclamaciones, locuciones o frases y palabras de otros idiomas. El sistema es capaz de generar formas derivadas o flexionadas a partir de una forma canónica, una flexión o derivación y los prefijos que se deseen incorporar.

Tanto el lematizador como el generador operan sobre una única estructura de datos de manera bidireccional —recorrerla en sentidos contrarios implica pasar de la operación lematizadora a la flexionadora—; se aporta una solución equilibrada entre grado de operatividad, tiempo de respuesta y cantidad de almacenamiento.

Ya se está trabajando en el enriquecimiento del léxico con palabras cultas con elementos compositivos muy variados como las que aparecen en el *Diccionario etimológico de helenismos españoles* de Eseberri Hualde o en el *Diccionario de raíces griegas léxico castellano científico y médico* de Quintana Cabanas.

AGRADECIMIENTOS.

Queremos agradecer al profesor Dr. Manuel Alvar Ezquerra de la Universidad Complutense de Madrid y a la profesora Dra. María Auxiliadora Castillo Carballo de la Universidad de Sevilla su colaboración en cuantas consultas le hemos formulado a lo largo del desarrollo del presente trabajo.

REFERENCIAS:

[Als90] Alsina, R.: *Todos los Verbos Castellanos Conjugados*. 17ª Edición. Teide. Barcelona, 1990.

[Alv93] Alvar Ezquerra, M.: “La formación de palabras en español”. Arco/Libros. Madrid, 1993.

[Alv94] Alvar Ezquerra, M.: *Diccionario de voces de uso actual*. Arco/Libros. Madrid, 1994.

[Cas90] Casares, J.: *Diccionario Ideológico de la Lengua Española*. 2ª Edición. Ed. Gustavo Gili, s.a. Barcelona, 1990.

[DGLE97] *Diccionario General de la Lengua Española VOX en CD-ROM*. Biblograf, s.a. Barcelona, 1997.

[DLE95] *Diccionario de la Lengua Española*. Edición electrónica. Versión 21.1.0. Real Academia Española y Espasa-Calpe. Madrid, 1995.

[Ese79] Eseberri Hualde, C.: *Diccionario etimológico de helenismos españoles*. Ediciones Aldecoa. Burgos, 1979.

[Fer87] Fernández Ramírez, S.: *Gramática Española*. Arco/Libros, S.A. Madrid, 1987.

[DUE96] *Diccionario de Uso del Español* de María Moliner. Edición en CD-ROM. Gredos. Madrid, 1996.

[Gar97] García Platero, J. M.: "Sufijación apreciativa y prefijación intensiva en español actual". *Lingüística Española Actual*, XIX/1, 1997, págs. 51-61.

[GDL96] *Gran Diccionario de la Lengua Española*. Larousse Planeta, s.a. Barcelona, 1996.

[GDS91] *Gran Diccionario de Sinónimos y Antónimos*. 4ª edic. Espasa-Calpe. Madrid, 1991.

[Gil85] Gili Gaya, S.: *Curso superior de sintaxis española* Vox. Biblograf, s.a. Barcelona, 1985.

[Góm91] Gómez Torrego, L.: *Manual de Español Correcto*. Arco/Libros, s.a. Madrid, 1991.

[Góm92] Gómez Torrego, L.: *El buen uso de las palabras*. Arco/Libros, s.a. Madrid, 1992.

[Lyo75] Lyons, J.: *Nuevos horizontes de la lingüística*. Alianza Editorial. Madrid, 1975.

[Mar78] Martinet, A.: *Elementos de lingüística general*. Gredos. Madrid, 1978.

[Per96] Pérez Aguiar, J. R.: "Reconocimiento y generación integrada de la morfología del español: Una aplicación a la gestión de un diccionario de sinónimos y antónimos". Tesis Doctoral bajo la dirección del Dr. O. Santana Suárez. Universidad de Las Palmas de Gran Canaria, 1996.

[Qui97] Quintana Cabanas, J. M.: *Diccionario de raíces griegas léxico castellano científico y médico*. 2ª edic. Editorial Dykinson. Madrid, 1997.

[RAE89] Real Academia Española: *Esbozo de una nueva gramática de la lengua española*. 1ª edic. Espasa-Calpe. Madrid, 1989.

[RHS93] Rodríguez, A.; Hernández, Z.; Santana, O.: "Agrupaciones de Tiempos Verbales en un Texto". *Anales de las II Jornadas de Ingeniería de Sistemas Informáticos y de Computación*, Quito (Ecuador). Abril, 1993, págs. 132-137.

[Sec91a] Seco, M.: *Diccionario de dudas y dificultades de la lengua española*. 9ª Edición. Espasa-Calpe. Madrid, 1991.

[Sec91b] Seco, M.: *Gramática esencial del español: Introducción al estudio de la lengua*. 2ª edición, revisada y aumentada. Espasa-Calpe. Madrid, 1991.

[SHR93] Santana, O.; Hernández, Z. J.; Rodríguez, G.: "Conjugaciones Verbales". *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural* (SEPLN). Nº 13. Febrero, 1993, págs. 443-450.

[SHRPCB94] Santana, O.; Hernández, Z.; Rodríguez, G.; Pérez, J.; Carreras, F.; Bogliani, S.: "Reconocedor automático de formas verbales que trata conjugación y pronombres enclíticos". *Lingüística Española Actual*, XVI-1, 1994, págs. 125-133.

[SPCSRH97] Santana, O.; Pérez, J.; Carreras, F.; Santos, S.; Rodríguez, G. Hernández, Z.: "GEISA: Un diccionario de sinónimos en formato electrónico". *Revista de Lexicografía*. La Coruña 1997.

[SPHCR97] Santana, O.; Pérez, J.; Hernández, Z.; Carreras, F.; Rodríguez, G.: "FLAVER: Flexionador y lematizador automático de formas verbales". *Lingüística Española Actual*, XIX-2, 1997, págs. 229/282.

[SRG93] 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, Febrero 1993, págs. 451-462.

[Ver95] *VerbiCard: Todos los verbos castellanos conjugados*. Castellnou Editorial. Barcelona, 1995.