

Morfología del Español: Reconocimiento y generación automáticos. Desarrollos del Grupo de Estructuras de Datos y Lingüística Computacional de la Universidad de Las Palmas de Gran Canaria (GEDLC)

<http://www.gedlc.ulpgc.es/>

Santana, O.; Pérez, J.; Hernández, Z.; Carreras, F.; Rodríguez, G.; Losada, L.; Duque, J.

Resumen.

Se expone la línea de investigación del GEDLC orientada a la elaboración de utilidades de ayuda para el tratamiento de documentos. A partir del procesador morfológico desarrollado, se han resuelto la generación automática de respuestas en análisis morfológico, la gestión integrada de un diccionario de sinónimos y la búsqueda morfológica de palabras en un texto; también se logró una herramienta para el manejo de diccionarios ideológicos. Se encuentran en curso una estación de trabajo lexicológica orientada a Internet, la ampliación de las relaciones morfosemánticas del procesador morfológico, un desambiguador morfológico, un analizador sintáctico, una herramienta de ayuda al análisis de estilo; también se trabaja en torno a las colocaciones, la composición, los regímenes preposicionales y la clasificación semántica de los verbos.

0.- Introducción.

La ingeniería lingüística constituye un campo de investigación cada vez más estratégico —no se ha resuelto la comunicación entre las personas y las máquinas en lenguaje natural. La elaboración de documentos es un proceso creativo que exige del autor una notable capacidad de abstracción y estructuración para asegurar una calidad aceptable del resultado; tal tarea puede ser eficazmente asistida mediante diversas herramientas. 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.

Los trabajos del GEDLC se orientan hacia el desarrollo de un conjunto de herramientas que facilitan la adecuada solución a una amplia gama de problemas actuales en la elaboración y explotación de documentos escritos —analizan el texto y ofrecen ayudas para enriquecer la expresión— tales como frecuencia de aparición de palabras, empleo de formas verbales y no verbales, corrección ortográfica, consultas en bases de datos documentales, búsquedas en texto libre, detección de vicios, depuración de estilos, análisis sintáctico, diccionarios de sinónimos o ideológicos, lematización, etiquetado, análisis y generación de textos escritos, tratamiento de corpus, etcétera.

El tratamiento automatizado de la morfología del español constituye la primera piedra sobre la cual construir cualquier procesador de lenguaje natural que contemple la sintaxis y la semántica. Tanto por la complejidad de las reglas de formación de las palabras, como por la abundancia de excepciones, el tratamiento por ordenador de la morfología del español constituye un campo de investigación con serias dificultades.

TRABAJOS DESARROLLADOS

1.- Procesador morfológico.

Lematiza cualquier palabra del español al identificar su forma canónica, categoría gramatical y la flexión o derivación que la produce, y obtiene las formas

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 herramienta funcione en una u otra modalidad.

En los verbos, trata la conjugación simple y compuesta, los pronombres enclíticos, la flexión del participio como adjetivo verbal (género, número, grado superlativo y adverbialización) y el diminutivo del gerundio. Con las formas no verbales, 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; variantes gráficas en todas las categorías gramaticales; formas invariantes tales como preposiciones, conjunciones, exclamaciones, palabras de otros idiomas y locuciones o frases. Tanto en la lematización como en la generación se contempla la pre fijación.

The screenshot shows the MORFOLOGÍA software interface. At the top, the window title is 'MORFOLOGÍA'. Below it, there are two tabs: 'VERBAL' (selected) and 'NOMINAL'. The input field contains 'Entrada: comer'. The interface is divided into several sections:

- VERBAL**
 - Infinitivo: comer
 - CONJUGACIÓN**
 - Tiempos simples / Tiempos compuestos
 - Gerundio compuesto / Participio compuesto
 - Diminutivo del gerundio
 - Modo subjuntivo
 - Modo indicativo
 - Pretérito perfecto compuesto / Pretérito perfecto
 - Pretérito anterior / Pretérito pluscuamperfecto
 - Condicional perfecto / Futuro perfecto
 - Pretérito pluscuamperfecto / Modo imperativo
 - Futuro perfecto / Presente
 - Pronombres enclíticos: no, uno, dos, tres, **selas** (dropdown)
 - DERIVACIÓN**
 - Sustantivo / Adjetivo / Adverbio
 - PREFIJACIÓN**
 - delante, anterioridad en tiempo y espacio (dropdown) / pre (dropdown)
 - Acumular / Eliminar / pre-
- NOMINAL**
 - Forma canónica: comer
 - FLEXIÓN**
 - Masculino singular / Masculino plural
 - Femenino singular / Femenino plural
 - Neutro singular / Neutro plural
 - Grado superlativo / Adverbializa superlat.
 - DERIVACIÓN**
 - Sufijos apreciativos: Aumentativo, Diminutivo, Peyorativo
 - Categoría gramatical: Sustantivo, Adjetivo, Adverbio, Verbo
 - * RESULTADO ***
 - Gerundio compuesto: **habiéndoselas precomido**

At the bottom, there is a logo for 'SD MORFO' and the text 'Grupo de Investigación en Estructuras de Datos. <http://protos.dis.ulpgc.es> Cerrar'. A footer box contains the text: 'COMER es infinitivo de verbo transitivo o pronominal o intransitivo COMER y es forma canónica (masculino singular) de sustantivo masculino usado también en locución adverbial COMER'.

Sin sumar la ampliación inherente a los prefijos y a los pronombres enclíticos, el universo está compuesto por alrededor de cuatro millones y medio de formas flexionadas y derivadas obtenidas a partir de las más de ciento veinticuatro mil formas canónicas con que opera el sistema que abarcan todas las entradas del Diccionario de la Lengua Española de la Real Academia [DLE95], del Diccionario General de la Lengua Española Vox [DGLE97], del Diccionario de Uso del Español de María Moliner [DUE96], del Gran Diccionario de la Lengua Española de Larousse Planeta [GDL96], del Diccionario de Uso del Español Actual Clave SM [DCI97], del Diccionario de voces de uso actual dirigido por Manuel Alvar Ezquerro [Alv94], del Gran Diccionario de

Sinónimos y Antónimos de Espasa-Calpe [GDS91] y del Diccionario Ideológico de la Lengua Española de Julio Casares [Cas90].

Existe una versión de este procesador morfológico disponible en Internet para su uso gratuito y discrecional en <http://www.gedlc.ulpgc.es/morfolog/morfolog.htm> que permite a los estudiosos hacer consultas y observar su funcionamiento. También se dispone de un lematizador para textos cortos que responde automáticamente a vuelta de correo electrónico.

Para el universo de las cuarenta mil formas canónicas iniciales mantiene las relaciones morfosemánticas inherentes a la derivación con cambio de categoría gramatical que supone la sustantivación y adjetivación deverbales y la verbalización de sustantivos y adjetivos.

2.- Generación automática de respuestas en análisis morfológico.

Se modela el espacio de aplicación a partir de un conjunto de resultados simples de análisis morfológico en lenguaje natural. Se elimina toda redundancia en los datos de entrada para extraer su información esencial y obtener la estructura lógica de la información. La posibilidad de decir lo mismo de diferentes formas obliga a seleccionar el léxico, a observar el contexto y a reforzar el control mediante criterios de selección para asegurar que el texto obtenido sea correcto y resulte formal, objetivo y escueto. La principal aportación radica en pasar de la estructura lógica a una frase con la que representar la información de forma cohesionada, coherente, comprensible y con cierto estilo.

Así, por ejemplo, partiendo de:

- '1ª persona singular presente de indicativo'
- '2ª persona singular presente de indicativo'
- '3ª persona plural presente de indicativo'
- '2ª persona singular futuro imperfecto de indicativo'
- '2ª persona singular imperativo'

se obtiene el mensaje:

"1ª y 2ª persona singular y 3ª plural presente de indicativo y 2ª persona singular futuro imperfecto de indicativo e imperativo".

Y partiendo de:

- 'forma canónica'
- 'forma del masculino plural'
- 'forma del masculino singular'
- 'forma del femenino plural'
- 'forma del femenino singular'
- 'forma del diminutivo femenino singular'
- 'forma del diminutivo masculino singular'
- 'forma del aumentativo femenino singular'
- 'forma del aumentativo masculino plural'

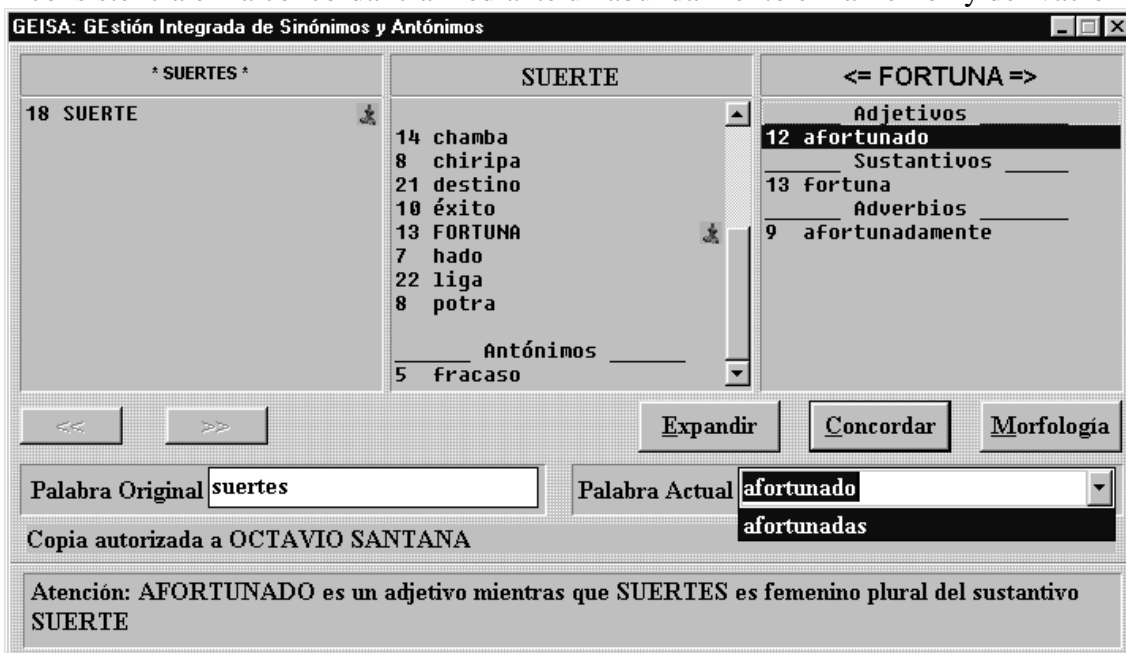
se obtiene el mensaje:

"forma canónica (masculino o femenino, singular o plural) o diminutivo (femenino o masculino, singular) o aumentativo (femenino singular o masculino plural)".

3.- Geisa: Gestión Integrada de Sinónimos y Antónimos.

Soporta un interfaz de uso fácil y claro que: a) admite como entradas las cabeceras del diccionario y cualquiera de sus formas flexionadas o derivadas —gracias a la función lematizadora del procesador morfológico—, b) permite la navegación

clásica en estos diccionarios —de cabecera a lista y de un elemento de la lista a cabecera—, c) ofrece formas de exploración alternativas por medio de entradas relacionadas morfosemánticamente que permite opciones de exploración mejores que las que ofrece el diccionario impreso —las relaciones sufijales eliminan cabeceras molestas alfabéticamente próximas y las prefijales proponen alternativas interesantes que aparecen en páginas distantes—, d) facilita la recuperación de la historia de una navegación apoyada en hasta tres entradas simultáneamente visibles, e) considera la salida con los accidentes gramaticales que afectan a la entrada tanto cuando se trata de formas con la misma categoría gramatical —conjugación con pronombres enclíticos, género, número, apreciativos, superlativos— como cuando es distinta —cambios de género o de número, de adjetivo a verbo a través del participio y otras derivaciones con cambio de categoría gramatical— y f) facultad para dar respuesta a cualquier inconsistencia en la concordancia mediante un abundamiento en la flexión y derivación.



4.- Búsqueda morfológica de palabras en un texto.

En todo editor de texto, desde el sencillo *edit* del MSDOS o el *bloc de notas* de Windows 95/98 hasta el más completo *Microsoft Word* o *Corel WordPerfect*, es posible encontrar una herramienta para búsquedas de ristas de caracteres dentro de un texto, pero todas estas herramientas se caracterizan por la realización de búsquedas de cadenas de caracteres. Aquí se presenta una perspectiva novedosa, en el campo de las búsquedas en editores de texto y otras aplicaciones, que permite la búsqueda de las apariciones de las palabras de una manera mucho más flexible, tanto en su forma canónica como afectada de las flexiones y derivaciones que se deseen —de entre las múltiples opciones que proporciona el procesador morfológico. En las búsquedas que impliquen varias formas debe especificarse la distancia máxima entre sus ocurrencias.

Búsquedas Avanzadas

Entrada Palabra Forma canónica

Todas sus flexiones y derivaciones

Todas las formas verbales
 Desactivar opciones

Formas no personal
 Infinitivo Gerundio Participio

Número y persona
 1ª persona 2ª persona 3ª persona Singular Plural

Modo indicativo
 Presente Pretérito indefinido Condicional simple
 Pretérito imperfecto Futuro imperfecto

Modo Subjuntivo
 Presente Pretérito imperfecto Futuro imperfecto

Modo imperativo
 Presente

Otras flexiones
 Flexión de participio como adjetivo verbal Diminutivo de Gerundio

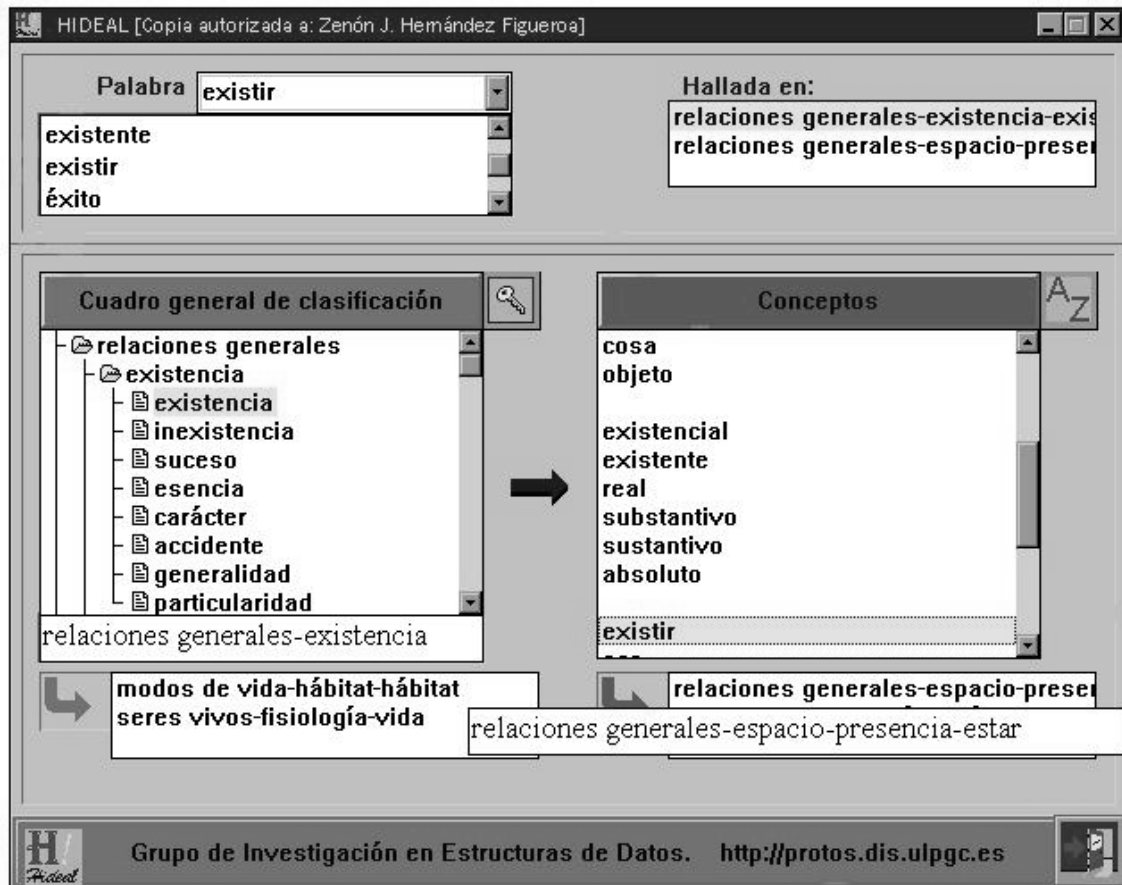
Pronombres enclíticos
 Ninguno Uno Dos Tres

Derivación
 Verbo Sustantivo Adjetivo Adverbio

Prefijación
 Sin prefijar Acumular
 Primero Segundo Tercero

5.- Herramienta para el manejo de diccionarios ideológicos.

El esquema de funcionamiento puede considerarse dividido en tres módulos independientes: a) el *módulo de relaciones* está concebido como un almacén de términos y relaciones directas entre términos, b) el *módulo de reglas* permite establecer relaciones lógicas entre términos en virtud de la evaluación de un conjunto de reglas de inferencia instituidas con anterioridad —en la derivación de las relaciones indirectas a partir de las establecidas, se aplica conocimiento añadido para que el módulo de relaciones pueda ser minimal sin pérdida de potencialidad— y c) el *módulo de interfaz* establece la comunicación con el sistema. Los tres módulos se conciben con tan alto nivel de independencia que hace posible distintas combinaciones de uso: desde un administrador general de relaciones a múltiples herramientas de propósito más específico —manejadores independientes de diccionarios ideológicos, de sinónimos, tesauros, etc.



TRABAJOS EN CURSO

6.- Estación de trabajo lexicológica orientada a Internet.

Consiste en la explotación de las posibilidades de realización de estaciones de trabajo lexicológicas orientadas a Internet, entendidas como un conjunto integrado de aplicaciones informáticas especializadas en tareas de análisis de textos contenidos en documentos electrónicos que se encuentren disponibles en la metared. Confluyen, de una parte, Internet, como ente generador y suministrador de recursos lingüísticos, de otra, la investigación lingüística como cliente y beneficiaria del nuevo nicho de oportunidades de estudio abierto por la red y, de una tercera, las técnicas informáticas de gestión y presentación de información compleja, como vehículo de intermediación y catalización entre las dos anteriores. El usuario interactúa con los corpora para buscar formas específicas de palabras, formas de palabras que casen con una cadena específica, coocurrencias de formas de palabras y/o cadenas en un espacio determinado del texto, etc. En general, el lingüista hará muchas otras cosas: localizar concordancias, realizar estudios estadísticos sobre el uso de las palabras, de sus componentes, de segmentos de texto formados por varias palabras casuales, frases hechas, perífrasis, etc.

Estado de realización del trabajo: 85%.

Archivos Buscar Anotaciones Opciones Acerca de...

Anterior Siguiente Parar Recargar

Dirección Analizar

Página Texto HTML Lematización Header Info Anotaciones

SD
Grupo de Estructuras de Datos y Lingüística Computacional

Lo más nuevo

Investigación
Artículos
Aplicaciones
Proyectos
Tesis leídas

El **Grupo de Estructuras de Datos y Lingüística Computacional** del Departamento de Informática y Sistemas de la Universidad de Las Palmas de Gran Canaria ha estado trabajando desde 1986 en el análisis de estructuras de datos aplicadas a la recuperación asociativa de información.
A partir de 1990 ha ampliado sus áreas de interés al procesamiento

Aplicaciones en línea

Lematizador

Categorías Canónicas Ordenaciones Resultados Distribución

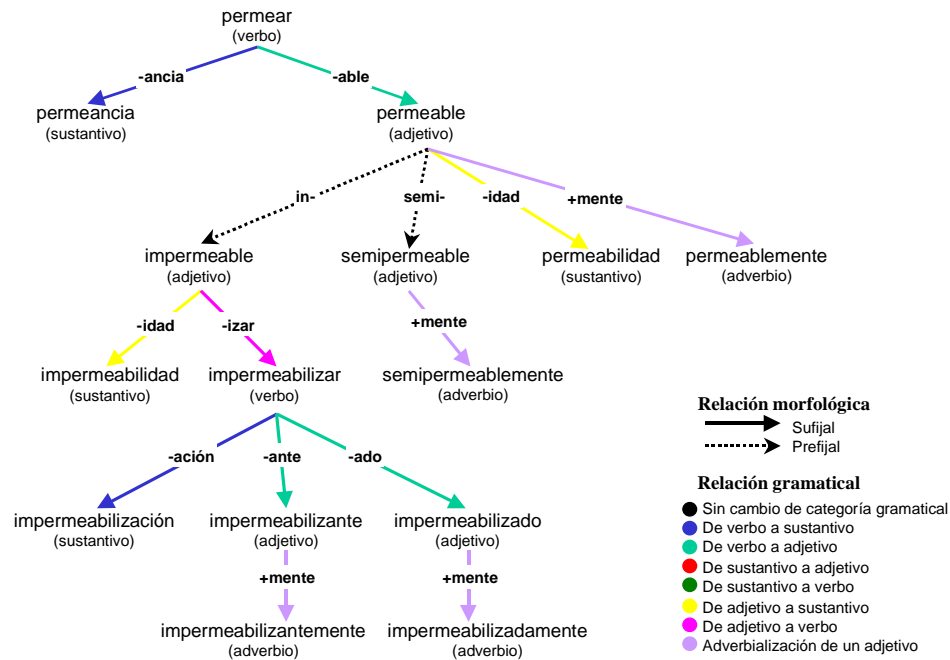
No reconocido	Verbos	Sustantivos	Adjetivos	Otras formas	No alfabéticas
Carlos	>adecuar	Francisco	adecuadamente	a	04
Díaz	adecuadamente	García	administrador	adecuadamente	15
Domínguez	>ampliar	generación	Amigos	al	19
EUROMAP	ampliado	Gestión	ampliado	con	1986
Figueroa	>aplicar	Grupo	antónimos	de	1987
FLAPE	aplicadas	Habana	aplicadas	del	1990
18	48	195	82	31	11

7.- Relaciones morfosemánticas.

A partir de los procesos de formación de las palabras en español, se establece un conjunto de relaciones morfosemánticas entre formas canónicas de utilidad para aplicaciones de procesamiento automático del lenguaje natural. Se consideran dos fenómenos complementarios básicos: alteraciones sufijales y alteraciones prefijales. La mayoría de los procesos derivativos del español se incluyen entre las alteraciones sufijales y conllevan, aunque no siempre, cambio en la categoría gramatical de la forma canónica. La prefijación generalmente no produce cambios de categoría gramatical y es independiente de la flexión de la palabra. La inclusión de la prefijación, junto a el resto de los procesos formativos (parasíntesis, regresión, suplección, abreviación,...), completa las relaciones morfosemánticas entre los elementos del léxico. La composición da lugar a formas de relación morfosemántica de distinta naturaleza y no es tenida en cuenta en este trabajo.

Estado de realización del trabajo: 80%.

Estructura arbórea de relaciones morfosemánticas



8.- Desambiguación morfológica.

Uno de los problemas que surge a la hora de realizar tratamientos informatizados de textos es la aparición de ambigüedades de tipo morfológico. Se dan cuando aparecen palabras que tienen más de una interpretación; se dividen en dos tipos: cuando la palabra puede proceder de más de una forma canónica con iguales o distintas categorías gramaticales. En las ambigüedades del segundo tipo, para determinar la categoría gramatical correcta se atiende al contexto en las estructuras sintácticas mínimas —sintagma nominal, adverbial, etcétera. Se consideran todas las ambigüedades del español que implican diferentes categorías gramaticales, se generan todas las combinaciones posibles en grupos de tres palabras y se aplican un conjunto de reglas de desambiguación en función de las subestructuras locales implicadas. Durante el proceso de análisis sintáctico se realiza inicialmente el proceso de desambiguación morfológica para evitar la construcción de gran cantidad de árboles erróneos.

Ejemplo: En el resultado del análisis morfológico de "esa" se obtiene que puede ser verbo, pronombre o adjetivo demostrativo, si va seguido de "carta" que es un sustantivo, se deduce que sólo puede ser adjetivo demostrativo.

Estado de realización del trabajo: 75%.

9.- Análisis sintáctico.

Se basa en un sistema de reglas simples que representan las estructuras sintácticas básicas. Se ha construido una primera aplicación para el aprendizaje de las reglas a partir de corpora etiquetados con el análisis sintáctico correspondiente; durante el proceso de análisis se parte de las reglas obtenidas y se construyen los árboles de representación de los posibles análisis. Se impone el problema de la desambiguación en

el que deben eliminarse las estructuras incorrectas mediante la adquisición de nuevas reglas, es aquí donde se introduce información como la concordancia, información semántica, reconocimiento de verbos copulativos, etcétera. Las reglas de eliminación de ambigüedades se aplican durante la construcción de los árboles para evitar en lo posible su propagación a niveles superiores.

Fase de aprendizaje:

"La lluvia trajo alivio al campesino"

(1)modificador_directo(2)núcleo_nominal(3)núcleo_verbal(4)núcleo_nominal(5)preposición subordinante + actualizador(6)núcleo_nominal(7)

(1)sintagma_nominal_homogéneo(3)(4)sintagma_nominal_homogéneo(5)sintagma_preposicional(7)

(1)sujeto_simple_explicito(3)(4)objeto_directo(5)objeto_indirecto(7)

(3)predicado_verbal_simple(7)

(1)sentencia(7)

Fase de análisis:

<i>Sentencia</i>					
		Predicado verbal simple			
Sujeto simple explícito				Objeto directo	Complemento indirecto
Sintagma nominal homogéneo		Núcleo verbal	Sintagma nominal homogéneo	Sintagma preposicional	
Actualizador	Núcleo nominal	Forma verbal simple	Núcleo nominal	Preposición subordinante + actualizador	Núcleo nominal
Artículo	Sustantivo	Verbo transitivo pronominal	Sustantivo	Contracción	Sustantivo adjetivo
La	lluvia	trajo	alivio	al	campesino

Estado de realización del trabajo: 65%.

10.- Herramienta de ayuda al análisis de estilo.

En el análisis de textos surgen diferentes cuestiones sobre sus aspectos de forma. Entre las diversas maneras que puede adoptar el estudio de la forma de un texto, se resalta el punto de vista morfológico y el estadístico. Se podrá decidir entre un conjunto de análisis predefinidos o establecer perfiles de análisis propios. El primer caso fijará una serie de opciones entre las que elegir: 1) problemas morfosintácticos (concordancias, anacolutos, pronombres y determinantes —situación, incorrecciones—, localización de perífrasis verbales, detección de errores en el uso de las formas verbales, abuso de la voz pasiva, usos impropios o anómalos de elementos relacionales o adverbiales,...), 2) problemas léxicos (extranjerismos, préstamos, palabras comodín,...) y 3) problemas estilísticos (cacofonías, repeticiones, redundancias, muletillas, tópicos, refranes, frases hechas, abusos de locuciones, de adverbios, de pronombres, de gerundios,...). El segundo caso mostrará una serie de alternativas combinables con las que realizar la oportuna clasificación del texto; cada opción responderá a una de estas cuatro ideas genéricas: 1) rasgos de flexión —verbales y no verbales—, 2) categorías gramaticales —verbales y no verbales— 3) familias morfosemánticas —flexión y derivación— y 4) familias semánticas —ideológicas y de sinonimia y antonimia.

Estado de realización del trabajo: 40%.

11.- Colocaciones.

Los trabajos encaminados al estudio de las colocaciones léxicas se han programado para una ejecución en tres fases consecutivas: 1) fase de recolección de colocaciones léxicas mediante su búsqueda en fuentes documentales —tanto manualmente como usando herramientas de localización específicas— y su almacenamiento en una base de datos, 2) fase de análisis de las relaciones semánticas entre las unidades léxicas componentes que se fijaron en la fase de recolección para establecer hipótesis acerca de la formación de colocaciones en función de los rasgos semánticos de los componentes y 3) fase de proyección de los resultados obtenidos en la fase 2 que dará lugar, entre otros posibles logros, al desarrollo de herramientas con capacidad para hacer predicciones y realizar sugerencias en entornos tales como el análisis textual, la ayuda a la elaboración de documentos y de diccionarios modernos del español en los que se de cuenta de este tipo de estructuras —de capital importancia para la enseñanza del español como lengua extranjera.

Estado de realización del trabajo: 30%.

12.- Composición.

Para recoger las reglas de formación que permitan la identificación automatizada de palabras compuestas con unión gráfica, se extraen de diferentes fuentes los distintos compuestos y se estudia cómo se unen las palabras que los forman según sus categorías gramaticales: sustantivo + adjetivo, adjetivo + sustantivo, sustantivo + sustantivo, verbo + verbo, sustantivo + i + adjetivo, adjetivo + adjetivo, adverbio + adjetivo, etcétera. El reconocimiento automático de los compuestos se realiza mediante la aplicación de estas reglas a sus posibles divisiones; se necesita añadir un conjunto de excepciones que eliminen las palabras que no sean compuestas —en *novelar*, “no” es un adverbio y “velar” es un adjetivo.

Estado de realización del trabajo: 30%.

13.- Regímenes preposicionales.

Hasta consolidar una colección de 4754 verbos junto a las preposiciones que los acompañan y ejemplos propios de su uso, se ha construido una base de datos, a partir de una recopilación de material de diversas fuentes, que se ha ido comparando con otras existentes: "Diccionario de dudas y dificultades de la lengua española" de Seco [Sec91], "Diccionario Práctico. Incorrecciones: Dudas y Norma Gramatical" de Corripio [Cor95], "Diccionario de Uso del Español" de María Moliner [DUE96], "Gran Diccionario de la Lengua Española" de Larousse [GDL96], "Vox Diccionario para la Enseñanza de la Lengua Española" [DELE95], "Diccionario de construcciones sintácticas del español. Preposiciones" de Nández [Náñ01], "Pequeño Diccionario de Construcciones Preposicionales" de Slager [Sla97] y "Archivo Gramatical de la Lengua Española. Partículas" de Fernández [Fer95]. Además de que todos los interesados tengan acceso a través de Internet al conocimiento del régimen preposicional, será útil en la optimización del análisis sintáctico automático.

Estado de realización del trabajo: 55%.

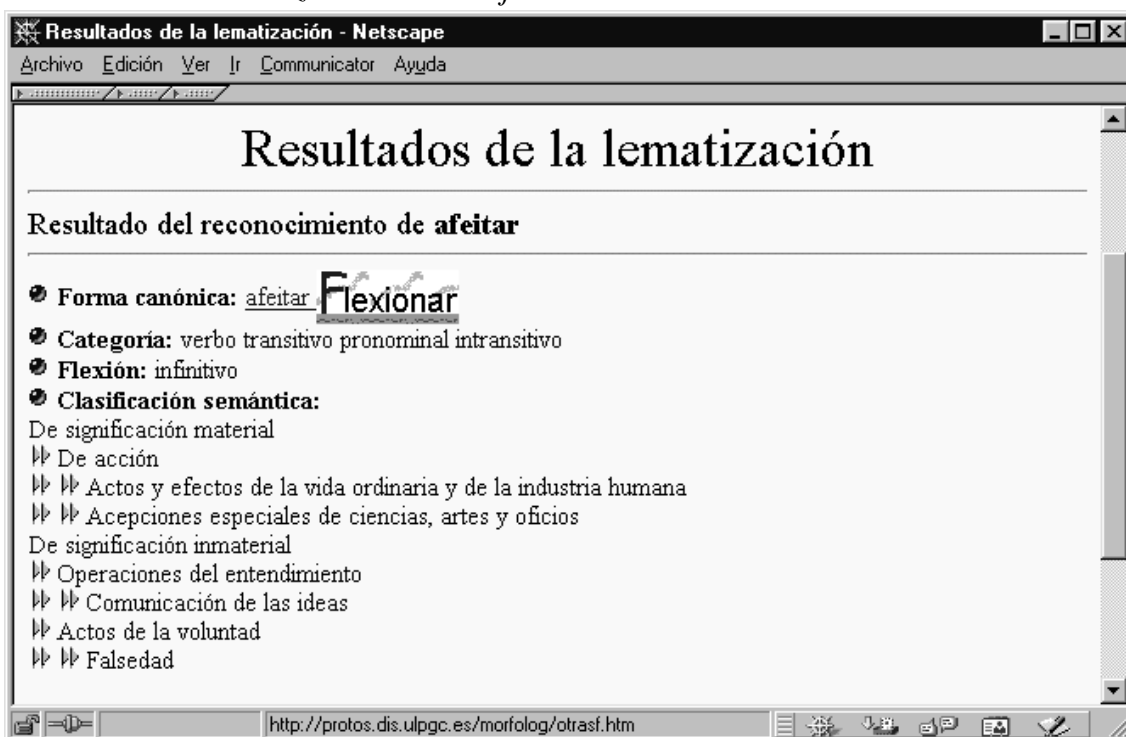
14.- Clasificación semántica de los verbos.

Disponer de una clasificación ideológica de verbos, en la que intervengan entre otras fuentes las obras: "Inventario de la Lengua Castellana. Índice ideológico de la Academia. Verbos" de Ruiz León [Rui79], "Diccionario Ideológico de la Lengua

Española" de Casares [Cas90] y "Diccionario Ideológico de la Lengua Española" de Vox [DILE95], proporcionará una base —con más de diez mil verbos— fundamental para alcanzar ulteriores aplicaciones que resuelvan problemas de índole tanto sintáctica como semántica (representación del conocimiento) en el tratamiento automático del lenguaje natural. En el procesador morfológico disponible en <http://www.gedlc.ulpgc.es/> figura una clasificación semántica de más de ocho mil verbos.

Del estudio de la relación entre la semántica y los regímenes preposicionales de los verbos saldrá a la luz un ahondamiento en la clasificación ideológica y quizá nuevas propuestas preposicionales.

Estado de realización del trabajo: 60%.



Referencias:

- [Als90] Alsina, R.: *Todos los Verbos Castellanos Conjugados*. 17ª Edición. Teide. Barcelona, 1990.
- [Alv93] Alvar Ezquerro, M.: *La formación de palabras en español*. Arco/Libros. Madrid, 1993.
- [Alv94] Alvar Ezquerro, M.: *Diccionario de voces de uso actual*. Arco/Libros. Madrid, 1994.
- [Alv99] Alvar Ezquerro, M.; Castillo Carballo, A.; García Platero, J. M.; Jiménez Cuenca, M. A.; Medina Guerra, A. M.: *Manual de redacción y estilo*. Istmo s.a.. Madrid, 1999.
- [Cas90] Casares, J.: *Diccionario Ideológico de la Lengua Española*. 2ª Edición. Ed. Gustavo Gili, s.a. Barcelona, 1990.
- [Cor95] Corripio Pérez, F.: *Diccionario Práctico. Incorrecciones: Dudas y Norma Gramatical*. Ed. Larousse Planeta. Barcelona, 1995.
- [DALE90] *Diccionario Actual de la Lengua Española Vox*. Bibliograf, s.a. Barcelona, 1990.
- [DCI97] *Diccionario de Uso del Español Actual. Clave*. Edición en CD-ROM. SM. Madrid, 1997.
- [DELE95] *Diccionario para la Enseñanza de la Lengua Española Vox*. Bibliograf, s.a., Universidad de Alcalá de Henares. Barcelona, 1995.
- [DGLE97] *Diccionario General de la Lengua Española Vox*. Edición en CD-ROM. Bibliograf, s.a. Barcelona, 1997.
- [DILE95] *Diccionario Ideológico de la Lengua Española Vox*. Bibliograf, s.a. Barcelona, 1995.

- [DLE95] *Diccionario de la Lengua Española*. Edición electrónica. Versión 21.1.0. Real Academia Española y Espasa-Calpe. Madrid, 1995.
- [DUE96] *Diccionario de Uso del Español* de María Moliner. Edición en CD-ROM. Gredos. Madrid, 1996.
- [Fer95] Fernández Ramírez, Salvador: *Archivo Gramatical de la Lengua Española. I. Partículas*. Edición en CD-ROM. Instituto Cervantes. Madrid, 1995.
- [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.
- [Gut78] Gutierrez Araus, María Luz: *Estructuras sintácticas del español actual*. Sociedad General Española de Librería, S.A.. Madrid, 1978.
- [Lyo75] Lyons, J.: *Nuevos horizontes de la lingüística*. Alianza Editorial. Madrid, 1975.
- [Mar78] Martinet, A.: *Elementos de lingüística general*. Gredos. Madrid, 1978.
- [Náñ01] Náñez Fernández, E.: *Diccionario de construcciones sintácticas del español. Preposiciones*. Universidad Autónoma de Madrid Ediciones. Madrid, 2001.
- [Oni90] Onieva Morales, Juan Luis: *Cómo dominar el análisis gramatical superior*. 7ª edición. Editorial Playor, S.A.. Madrid, 1990.
- [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.
- [RAE89] Real Academia Española: *Esbozo de una nueva gramática de la lengua española*. 1ª edic. Espasa-Calpe. Madrid, 1989.
- [Que97] Quesada Moreno, José F.: *El algoritmo SCP de análisis sintáctico mediante propagación de restricciones. Aplicación a la traducción automática en tiempo real*. Tesis Doctoral. Facultad de Matemáticas. Universidad de Sevilla. Sevilla, 1997.
- [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.
- [Rui79] Ruiz León, José: *Inventario de la Lengua Castellana. Índice ideológico de la Academia. Verbos*. Imprenta de Fortanet. Madrid, 1879.
- [Sec91] Seco, M.: *Diccionario de dudas y dificultades de la lengua española*. 9ª Edición. 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*. 16-1, Ed. Arco/Libros, S.L., 1994. Págs. 125-133.
- [Sla97] Slager, E.: *Pequeño Diccionario de Construcciones Preposicionales*. Visor Libros, S.L. Madrid, 1997.
- [SPCDHR99] Santana, O.; Pérez, J.; Carreras, F.; Duque, J.; Hernández, Z.; Rodríguez, G.: “FLANOM: Flexionador y lematizador automático de formas nominales”. *Lingüística Española Actual*. 21-1, Ed. Arco/Libros, S.L., 1999.
- [SPHCR97] Santana, O.; Pérez, J.; Hernández, Z.; Carreras, F.; Rodríguez, G.: “FLAVER: Flexionador y lematizador automático de formas verbales”. *Lingüística Española Actual*, 19-2, Ed. Arco/Libros, S.L., 1997. Págs. 229/282.
- [SPL00] Santana, O.; Pérez, J.; Losada, L.: “Generación automática de respuestas en análisis morfológico”. *Estudios de lingüística*. Universidad de Alicante, 14. Departamento de Filología Española, Lingüística General y Teoría de la Literatura. 2000. Págs. 245/257.

- [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). 13. Febrero, 1993. Págs. 451-462.
- [Ver95] *VerbiCard: Todos los verbos castellanos conjugados*. Castellnou Editorial. Barcelona, 1995.
- [ZUK95] Zaenen, A.; Uszkoreit, H.; Karlsson, F.; Karttunen, L.; Sanfilippo, A.; Pulman, S. G.; Pereira, F.; Briscoe, T.: *Language Analysis and Understanding. Survey of the Art in Human Language Technology*. Capítulo 3. <http://www.ces.ogi.edu/CSLU/HLT/survey>. 1995.