

## **Diseño e implementación de una estrategia de asignación de funciones sintácticas para un analizador del castellano**

Antonio Sánchez Valderrábanos (WordPerfect Development, Bélgica)

J. Gabriel Amores Carredano (Departamento de Lengua Inglesa, Universidad de Sevilla)

### **1. Planteamiento del problema**

La asignación automática de valores y funciones sintácticas a los constituyentes oracionales plantea serios problemas en lenguas como el castellano en las que, al contar con mayor información morfológica, el orden de aparición de los constituyentes que realizan esas funciones puede considerarse semilibre.

Generalmente, los analizadores automáticos se han probado sobre lenguas como el inglés, que imponen un orden más rígido a la estructura oracional. Básicamente, cualquier analizador para estas lenguas opera en dos fases simultáneas: se identifica un constituyente y se le asigna una función teniendo en cuenta lo analizado hasta ese momento y la información aportada por el diccionario.

Esta estrategia no es directamente aplicable al castellano, ya que podemos no tener información suficiente para tomar una decisión correcta en el momento en que hemos identificado un constituyente. Resulta, pues, preferible esperar a contar con toda la información contextual para tomar una decisión acertada.

El problema alcanza mayores proporciones si tenemos en cuenta que cada verbo puede tener varias acepciones que habrían de ser consideradas una a una en combinación con los constituyentes que se van analizando.

Además, si utilizamos un formalismo declarativo para la expresión de las reglas gramaticales en un lenguaje de programación como Prolog, habría que incluir en la gramática tantas posibilidades de orden de constituyentes como sean requeridas en castellano, con sus correspondientes algoritmos de asignación de funciones.

Presentamos aquí una estrategia implementada en un analizador para el castellano (ROMEIO) cuyo objetivo es ofrecer una solución a estos problemas.

Antes de implementar la estrategia de análisis, hemos definido un modelo de complementación verbal del castellano, que da cuenta de la fundamentación lingüística de la estrategia de análisis. Este modelo describe el comportamiento formal, sintáctico y morfológico, de cada función y de los constituyentes que la desempeñan. Es decir, describe el contexto o contextos en los que aparecen las distintas funciones. A partir de este modelo hemos clasificado un conjunto amplio de verbos, alrededor de dos mil, según las funciones que pueden acompañarlo. El marco de subcategorización de cada

verbo contiene además información sobre las preferencias semánticas de cada función. Esta parte del trabajo no será objeto de este artículo.

## 2. Estrategia de análisis

El sistema consta de una gramática libre de contexto (DCG en Prolog) que realiza el análisis de constituyentes. A cada constituyente se le asigna una estructura funcional de pares atributo-valor obtenida mediante un algoritmo de unificación similar al usado en implementaciones de gramáticas léxico-funcionales (LFG). La diferencia estriba en que las etiquetas que se asignan a los constituyentes son temporales y no definitivas, como ocurre con las ecuaciones funcionales en una gramática LFG.

Una vez que se ha llegado a la regla gramatical más alta, aquella que expande la oración, un algoritmo de decisión toma la estructura funcional con etiquetas temporales y asigna funciones a los constituyentes teniendo en cuenta el contexto oracional completo. Este algoritmo también está escrito en Prolog.

El problema se resuelve por tanto en dos fases consecutivas no simultáneas:

- a) Determinación de la estructura de constituyentes
- b) Asignación de funciones y valores

## 3. Análisis de constituyentes

Como hemos dicho, el análisis de constituyentes lo realiza una gramática libre de contexto para el castellano formalizada en una DCG en Prolog.

A cada nodo, terminal o no terminal, se le asocia una estructura funcional obtenida por unificación de pares atributo-valor. Los nodos que expanden proyecciones máximas de sintagmas no asignan ninguna función sintáctica, como se haría en una gramática LFG, sino que asignan etiquetas temporales que son manipuladas en un estadio posterior del análisis.

Ejemplo de regla gramatical

```
o(o(SN,SV),Ffinal) --> sn(SN,Fsn),          % analiza un sintagma nominal
                      sv(SV,Fsv,Num),      % analiza un sintagma verbal
                      {unify([sn_prv1:Fsn],Fsv,Ftemporal),
                       % unifica en estructura
                       % funcional temporal
                       asigna_funcion(Ftemporal,Ffinal)}.
                      % asigna funciones finales
```

Ejemplo de árbol de constituyentes y resultado de unificación parcial para la oración *se aleja el peligro de contagio*.

```
o(vg(pro(se),
  v(aleja)),
  np(det(el),
    n1(n(peligro))),
  pp(pre(de),
    np(n1(n(contagio)))))
```

```
verbo:alejar
num:sing
per:3
pro_prv1:p_at:se
  semf:[anim]
  agr:per:3
  corref:si
sn_psv1:pos:posv
  sust:peligro
  semf:[no_anim]
  agr:num:sing
  gen:masc
  per:3
  det:el
sp_psv1:pos:posv
  sust:contagio
  semf:[no_anim]
  agr:num:sing
  gen:masc
  per:3
  prep:de
```

Las etiquetas parciales contienen en sí información referente a categoría (sintagma nominal, preposicional), su posición con respecto al verbo (preverbal o postverbal), y número de categoría de este tipo en el orden lineal de la frase (primero, segundo, etc.), de acuerdo con el siguiente esquema (CATEGORIA\_POSICION\_NUMERO). Esta información será usada posteriormente por el algoritmo de asignación de funciones. La regla superior, por ejemplo, nos dice que lo analizado es un sintagma nominal (sn\_POSICION\_NUMERO) en posición preverbal (sn\_prv\_NUMERO), y es el primero (sn\_prv1).

#### 4. El algoritmo de asignación de funciones.

Una vez que hemos obtenida una representación temporal de la oración,

procedemos a asignar funciones a esos constituyentes. El algoritmo de asignación de funciones no está basado en ninguna teoría lingüística concreta. El tipo de representación que obtendremos será equivalente al obtenido con una gramática libre de contexto aumentada.

Desde el punto de vista teórico, este algoritmo tiene dos partes bien diferenciadas que, en la práctica, pueden superponerse en determinadas ocasiones. En primer lugar el algoritmo determina la acepción verbal presente en la oración que se analiza. En segundo lugar el algoritmo asigna funciones y valores. La primera parte es específica de cada tipo de verbo y la segunda es común a todos los verbos. Un verbo puede tener hasta cinco acepciones diferentes, en dos casos exclusivamente.

```
asigna_funcion(Ftemporal,Ffinal) :-  
    member(verbo:Verbo,Ftemporal),           % identifica el verbo  
    busca_tipo_verbo(Verbo,Acepciones),      % busca las acepciones  
    asigna(Modelo,Acepciones,Verbo,Ftemporal,Final).  
                                             % asigna funciones
```

Determinar la acepción antes de asignar funciones nos permite comenzar la asignación de funciones sólo cuando sabemos el tipo de acepción que aparece en la oración. Además, nos permite hacerlo conociendo los rasgos de subcategorización de cada acepción. Es importante hacerlo así porque los rasgos de subcategorización pueden diferir de una acepción a otra. Determinar la acepción antes de asignar funciones supone un importante ahorro de tiempo de ejecución.

La consulta al diccionario para determinar las acepciones existentes se realiza una sola vez. De esta manera conseguimos tener disponible a la vez la información referente a todas las acepciones sin tener que consultar el diccionario de nuevo.

Para ambas tareas, determinación de la acepción y asignación de funciones, utilizamos el siguiente conjunto de características para clasificar los constituyentes que es necesario tratar en cada caso.

1. Constituyentes obligatorios de cada acepción
  - 1.1. exclusivos de cada acepción
    - 1.1.1. no ambiguos dentro de cada acepción
    - 1.1.2. ambiguos dentro de cada acepción
  - 1.2. comunes a las acepciones posibles
    - 1.2.1. no ambiguos dentro de cada acepción
    - 1.2.2. ambiguos dentro de cada acepción
2. constituyentes opcionales de cada acepción

- 1.1. exclusivos de cada acepción
  - 1.1.1. no ambiguos dentro de cada acepción
  - 1.1.2. ambiguos dentro de cada acepción
- 1.2. comunes a las acepciones posibles
  - 1.2.1. no ambiguos dentro de cada acepción
  - 1.2.2. ambiguos dentro de cada acepción

Generalmente, en la determinación de la acepción nos serviremos de las características del apartado (1) y en la asignación de funciones y valores utilizamos las del apartado (2).

La diferencia entre utilizar esta jerarquía de características para la determinación de la acepción y utilizarla para asignación de funciones está en que la apliquemos, en el primer caso, entre los constituyentes que pueden acompañar a cada una de las acepciones de un verbo; y, en el segundo caso, entre los constituyentes que pueden desempeñar dos funciones. Dado que son características relativas podemos aplicarlas entre dos entidades cualesquiera.

#### 4.1. Determinación de la acepción verbal

Para determinar la acepción que analizamos, en primer lugar consideramos qué funciones o valores pueden acompañar a cada una de las acepciones posibles y qué constituyentes pueden realizar estas funciones o valores. Una vez obtenido el conjunto de constituyentes, los clasificamos según las características citadas.

El objeto de esta clasificación es delimitar el conjunto de constituyentes a partir de los que es posible determinar qué acepción aparece en la oración analizada, ya que no todos los constituyentes son útiles para este propósito.

En lo que se refiere a la determinación de la acepción, en primer lugar tratamos los constituyentes obligatorios porque nos permiten determinar la acepción de forma inmediata, sabemos que estos constituyentes están presentes siempre en la oración y por tanto nos ahorramos intentos de encontrar constituyentes que pueden o no aparecer. Si no existen constituyentes obligatorios, pasamos a tratar los opcionales.

Dentro de los obligatorios, tratamos en primer los que son exclusivos de una de las acepciones posibles. Si no existen, trataremos los comunes, es decir, los que pueden aparecer en ambas acepciones y son por tanto ambiguos. Dentro de cada uno de estos tipos trataremos primero los no ambiguos dentro de cada acepción y después los ambiguos. Dentro de los ambiguos trataremos primero los desambiguables por medios formales exclusivamente.

De esta forma, trataremos primero los constituyentes que, dentro de cada acepción, son no ambiguos. A continuación trataremos aquéllos que son desambiguables teniendo en cuenta la estructura de constituyentes que hemos obtenido. El motivo es que determinados constituyentes ambiguos exigen estar acompañados de ciertos constituyentes o funciones para poder desempeñar una de las posibles funciones o valores, mientras que otros no. Por ejemplo, "se" reflexivo para transitivos pronominales debe estar acompañado de refuerzo reflexivo.

Como se puede apreciar, existirá un algoritmo de determinación de la acepción para cada tipo de verbo, según las acepciones que posea. Si el número de tipos de verbo existente fuera excesivamente elevado, la estrategia de análisis que proponemos podría resultar inviable. Sin embargo, tras la fase de clasificación verbal realizada, sabemos que las combinaciones posibles de acepciones diferentes no superan las quince.

#### 4.2. La estrategia de asignación de funciones y valores

Esta estrategia está basada en dos hechos fundamentales:

4.2.1. Un orden de preferencia para las funciones: el tipo de verbo que esté analizando el algoritmo determinará cuál es el orden en que el algoritmo debe tratar de asignar las funciones y valores.

Así por ejemplo, para los verbos transitivos el algoritmo tratará de asignar en primer lugar el objeto directo y después el sujeto.

4.2.2. Un orden de preferencia para los constituyentes: el tipo de verbo y la función o valor que esté tratando determinará cuál es el orden en que el algoritmo debe tratar los constituyentes.

Así por ejemplo, para determinar el objeto directo de un verbo transitivo el algoritmo tratará en primer lugar los constituyentes no ambiguos (lo, la, los, las).

Como se puede apreciar por el tipo de estrategia utilizada, el algoritmo de asignación de funciones y valores está fuertemente determinado por la información verbal.

El orden de preferencia para las funciones está determinado fundamentalmente por el grado de cohesión que existe entre verbo y función o valor.

El orden de preferencia para asignar cada función a los constituyentes que la pueden desempeñar se define mediante la jerarquía de características descrita antes. Como hemos dicho, la diferencia radica en que en este caso aplicamos las características descritas no entre acepciones sino entre funciones. Es decir, definimos las características

de los constituyentes entre la función que asignamos y el resto de funciones que se deben asignar después de ésta y que pueden estar desempeñadas por los mismos constituyentes. Por ejemplo, si tratamos de asignar el objeto directo con un verbo transitivo definiremos las características de los constituyentes en relación con el sujeto y con el complemento preposicional.

De esta forma, igual que antes, conseguimos tratar los constituyentes de menos a más ambiguos. Para determinar el grado de ambigüedad de un constituyente tenemos en consideración las funciones que éste puede desempeñar dentro de la acepción que estamos analizando. Por ejemplo, el sintagma nominal con un verbo transitivo puede ser sujeto y objeto directo, dentro de los constituyentes regidos. Por ello intentaremos antes con los pronombres átonos y con el sintagma preposicional introducido por "a".

Puesto que conocemos las posibles acepciones del verbo que tratamos, podemos considerar no ambiguos, dentro del tipo de verbo que tratamos, ciertos constituyentes que sí lo son con otros verbos. Por ejemplo, un sintagma preposicional introducido por "a" y con un sustantivo animado podría ser objeto indirecto, entre otras cosas, pero con un verbo transitivo no puede serlo.

## 5. Un ejemplo: análisis de los verbos transitivos y pronominales

Para mostrar el funcionamiento de la estrategia de análisis descrita vamos a tomar como ejemplo los verbos transitivos y pronominales por dos motivos: en primer lugar, estos verbos representan un porcentaje alto de los verbos del castellano, alrededor del 25 por ciento de los clasificados; y, en segundo lugar, pueden aparecer con todos los valores de "se" que consideramos, por lo que su análisis resulta interesante.

La versión que tenemos implementada actualmente considera los constituyentes cuyos núcleos son sustantivos y pronombres: sintagma nominal, sintagma preposicional, pronombres átonos (lo, la, los, las, le, les, me, te, nos, os y se) y pronombres tónicos (yo, tú, él, ella, ello, nosotros/as, vosotros/as, ellos/as, mí, ti, sí). No están incorporadas aún las oraciones subordinadas ni coordinadas. Esta versión considera todas las funciones y valores que pueden desempeñar los constituyentes citados: sujeto, complemento directo complemento indirecto, complemento preposicional regido por el verbo, complemento predicativo y atributo; valor pronominal, valor reflexivo y recíproco, valor impersonal, valor pasivo y dativo ético.

### 5.1. Determinación de la acepción los verbos transitivos y pronominales

La primera tarea que debemos realizar es determinar el conjunto de constituyentes a partir de los que podemos determinar la acepción. Para ello, aplicamos a este caso las características definidas antes.

## 1. Constituyentes obligatorios.

acepción pronominal: me, te, nos, os, se

acepción transitiva: no existen

### 1.1. Constituyentes exclusivos.

Dentro de los obligatorios, no existen constituyentes exclusivos de ninguna de las acepciones, ya que los pronombres átonos "me, te, nos, os, se" pueden aparecer también con la acepción transitiva.

#### 1.1.1. Constituyentes no ambiguos.

Dentro de la acepción pronominal, "me, te, nos, os, se" son no ambiguos puesto que sólo pueden tener valor pronominal. Dentro de la acepción transitiva, "me, te, nos, os" pueden desempeñar la función de objeto directo y, en determinados contextos, este objeto directo puede tener valor reflexivo. En lo que respecta a "se" puede además tener valor impersonal o pasivo. Por ello, trataremos "se" por un lado y "me, te, nos, os" por otro.

Por tanto, para determinar la acepción en el caso de los verbos pronominales y transitivos la primera pregunta que es necesario hacer es si existe un pronombre "me, te, nos, os, se".

Si no existe ninguno de estos constituyentes, la acepción es transitiva, puesto que estos constituyentes son obligatorios para la acepción pronominal. En este caso pasamos asignar funciones a la acepción transitiva según el orden determinado.

Si existe uno de estos constituyentes, la acepción puede ser pronominal o transitiva. Puesto que los constituyentes a partir de los que es posible determinar la acepción son comunes a ambas, es necesario definir un procedimiento de desambiguación. Este procedimiento está basado en modelo de complementación verbal que hemos definido previamente.

La información formal que necesitamos para determinar la acepción está contenida en la estructura de constituyentes que hemos obtenido en la fase anterior del análisis. Por ello resulta necesario completar la fase de análisis de constituyentes antes de comenzar la de asignación de funciones, cuyo primer paso es la determinación de la acepción.

#### 5.1.1. Determinación de la acepción en caso de "se"

Los valores que puede tener "se" son:



acepción pronominal: "se" pronominal  
acepción transitiva: "se" impersonal  
acepción transitiva: "se" reflexivo (objeto directo)

En todos los casos, "se" concuerda en persona y número con el verbo; por ello no podemos desambiguar a partir de esta información.

En este contexto no es posible desambiguar a partir de información formal exclusivamente. Es necesario tener información funcional, es decir, necesitamos saber cuáles son las funciones que acompañan a la secuencia "se + verbo pronominal y transitivo". Las funciones que nos permiten desambiguar son el objeto directo, el sujeto y el refuerzo reflexivo.

Para ejemplificar la estrategia de análisis tomaremos el verbo *alejar* que en su acepción pronominal puede tener sujeto animado o no y que en su acepción transitiva puede tener tanto un sujeto como un objeto directo animados o no. Evidentemente, éste es el caso más complicado que se nos puede presentar, puesto que cualquier restricción semántica referente a cualquiera de las funciones tratadas limita las posibilidades de análisis. Señalaremos los casos en que determinado valor de "se" no es posible debido a las restricciones semánticas de las funciones regidas por el verbo.

En primer lugar comprobamos si existe un sujeto. La detección del sujeto la lleva a cabo la regla de asignación de esta función, que se aplica a todos los constituyentes existentes en la oración. El funcionamiento de esta regla se describe más adelante.

En caso de que exista un sujeto, como en

*alguien se aleja (de algo / alguien)*  
*algo se aleja (de algo / alguien)*

el valor impersonal no es posible. Es necesario desambiguar entre el valor pronominal, acepción pronominal, y el reflexivo, acepción transitiva. El valor reflexivo no es posible para los verbos transitivos cuyo sujeto es no animado, como "averiar".

Para desambiguar entre estos valores utilizamos el modelo de complementación verbal definido. Según este modelo, el valor reflexivo necesita de un refuerzo reflexivo, un sintagma preposicional introducido por "a" que tiene por núcleo un pronombre tónico ("a sí") correferencial con el pronombre átono "se", que puede estar incrementado por "mismo". Si no existe este refuerzo, el valor de "se" no puede ser reflexivo y, por tanto, es pronominal.

*alguien se aleja a sí mismo (de algo / alguien)*  
*alguien se aleja (de algo / alguien)*

Para motivar esta forma de distinguir entre ambas acepciones es necesario

describir con detalle el comportamiento de los verbos transitivos y pronominales cuando su objeto directo tiene valor reflexivo. Al contrario de lo que sucede con las acepciones que son exclusivamente transitivas (*me lavo*), con estos verbos la duplicación pronominal es obligatoria para expresar la reflexividad. La frase *me alejo* no puede tener valor reflexivo si no incluye el refuerzo, *me alejo a mí mismo*.

En segundo lugar comprobamos si existe un objeto directo, desempeñado por cualquier constituyente que no sea "se", como sucede en el caso de que el valor de "se" sea reflexivo. De la misma forma que sucedía antes con el sujeto, la regla que asigna la función de objeto directo se aplica a todos los constituyentes existentes. El funcionamiento de esta regla se describe más adelante.

Si no existe un sujeto y existe un objeto directo,

*se aleja a alguien (de algo / alguien)*

*se aleja algo (de algo / alguien)*

el valor de "se" es impersonal, ya que la acepción pronominal no puede estar acompañada de objeto directo y la acepción transitiva, en la que "se" tiene función de objeto directo con valor reflexivo, no puede llevar otro objeto directo. El valor impersonal no es posible para los verbos transitivos cuyo sujeto es no animado, como "agostar".

Tras este análisis, existe un tipo de oraciones que puede tener más de una interpretación: son aquéllas en las que aparece un sintagma nominal que puede ser interpretado como sujeto, "se" con valor pronominal, o como objeto directo, "se" con valor impersonal. Son oraciones del tipo:

*algo se alejó*

*se alejó algo*

Esta situación sólo se puede producir para aquellos verbos cuya acepción pronominal puede tener un sujeto no animado y cuya acepción transitiva pueda tener sujeto animado y objeto directo no animado. Para desambiguar estas oraciones es necesario, tanto para un lingüista como para un analizador automático, tener conocimiento del contexto extraoracional.

En caso de que no exista ninguna de las funciones o valores que nos permiten desambiguar, es decir, en caso de que tengamos la secuencia "se + verbo"  
*se alejó*

son posibles dos valores: pronominal e impersonal. Sin embargo, para admitir la posibilidad del valor impersonal es necesario sobrentender un objeto directo omitido, ya que este valor aparece con la acepción transitiva. Para interpretar el valor de "se" como pronominal es necesario sobrentender un sujeto, función que se omite con frecuencia en

castellano. Por ello, en este caso el valor que asignamos a "se" por defecto es el pronominal.

### 5.1.2. Determinación de la acepción en caso de "me, te, nos, os"

En este caso, la situación es menos compleja que con "se". Los valores posibles de estos constituyentes son:

acepción pronominal: "me" pronominal

acepción transitiva: "me" objeto directo

acepción transitiva: "me" reflexivo (objeto directo)

En este caso sí es posible determinar la acepción a partir exclusivamente de información formal. No es necesaria la información funcional como lo es en el caso de "se". Para ello es necesario comprobar si existe concordancia en número y persona entre pronombre y verbo. Si no existe concordancia, la acepción es transitiva y la función del pronombre átono es objeto directo. En este caso consideramos la posibilidad de que aparezca un pronombre tónico redundante:

*alguien me aleja (a mí)*

*algo me aleja (a mí)*

En caso de que exista concordancia, el pronombre puede tener valor pronominal o reflexivo. Para que tenga valor reflexivo o recíproco es necesario que exista un refuerzo reflexivo, como en el caso de "se". Si no existe un refuerzo reflexivo, la acepción es pronominal y el valor de "me" es pronominal. Si existe un refuerzo reflexivo, la acepción es transitiva y la función de "me" es objeto directo con valor reflexivo.

*yo me alejo*

*yo me alejo a mí mismo*

De esta forma, con dos preguntas, concordancia y refuerzo reflexivo, podemos decidir, con certeza, cuál es la acepción presente en la oración que analizamos.

Una vez determinada la acepción, el algoritmo pasa a la fase de asignación de funciones. Como hemos visto, en muchos casos la determinación de la acepción nos permite asignar algunas de las funciones. Naturalmente, en este caso el algoritmo controla cuál es la siguiente función que se debe buscar.

### 5.2. Asignación de funciones

Como dijimos, la asignación de funciones y valores a los constituyentes se lleva a

cabo según un orden predefinido, tanto en lo que respecta a las funciones y valores como a los constituyentes que las realizan. Este orden resulta especialmente relevante para las funciones o valores que se pueden realizar mediante los mismos constituyentes, como el sujeto y el objeto directo. El orden de preferencia de funciones y valores para verbos transitivos es:

1. objeto directo
2. valor reflexivo o redundancia pronominal
3. sujeto
4. complemento preposicional regido

Como dijimos, para determinar el orden en que cada regla intenta asignar una función o valor a los distintos constituyentes que los pueden realizar, utilizamos la misma jerarquía de características que para determinar la acepción. La diferencia es que en este caso la aplicamos entre funciones y valores y no entre acepciones. Este orden nos permite tratar los problemas de ambigüedad más complicados sólo cuando estamos seguros de que existen. Ejemplificaremos esta estrategia mediante la regla de asignación del objeto directo.

#### 5.2.1. La regla de asignación del objeto directo

En primer lugar, es necesario describir el modo en que funcionan ésta y otras reglas. Cuando esta regla intenta asignar la función de objeto directo a un determinado tipo de constituyente prueba con todos los constituyentes de ese tipo existentes en la oración, puesto que contamos con la estructura de constituyentes completa. Sin embargo, el orden en que lo intenta no es lineal, no va desde el primero al último. Se establece un orden de preferencia respecto a la posición en que el algoritmo debe buscar cada tipo de constituyente. Este orden no es el mismo para todas las funciones ni para todos los constituyentes.

Así por ejemplo, la regla que trata de asignar la función de objeto directo a un sintagma nominal, cuya posición típica es la posverbal, intentará asignarle esta función en primer lugar al primer sintagma nominal que aparece en posición posverbal, luego al segundo y así sucesivamente. Cuando agote los sintagmas nominales que aparecen en posición posverbal procederá de igual manera con los que aparecen en posición preverbal. Sin embargo, no será éste el orden que se aplique para asignar la función objeto directo a los pronombres átonos "lo, la, los, las", puesto que éstos no se disponen en la oración de la misma manera que el sintagma nominal.

Pasamos a ver el orden en que las reglas tratan los distintos constituyentes que pueden desempeñar la función de objeto directo. Puesto que los constituyentes "me, te, nos, os, se" han sido ya tratados en la fase de determinación de la acepción, la regla pasa

a tratar los constituyentes opcionales y exclusivos. Dentro de éstos, la regla trata de asignar la función de objeto directo, en primer lugar, a los constituyentes no ambiguos. En este caso los únicos constituyentes no ambiguos que pueden desempeñar esta función son: "lo, la, los, las" y para casos de léismo "le, les". En este caso consideramos la posibilidad de que aparezca un pronombre tónico redundante:

*alguien lo aleja (a él)*  
*algo lo aleja (a él)*

En caso de que no exista ninguno de los constituyentes tratados hasta ahora, pasamos a tratar los constituyentes ambiguos, los que pueden desempeñar otras funciones además de la de objeto directo: sintagma preposicional y sintagma nominal.

Por la estrategia que seguimos no son posibles confusiones con el sintagma nominal, que puede ser, por ejemplo, sujeto de la acepción pronominal u objeto directo de la transitiva.

El siguiente constituyente al que la regla trata de asignar la función de objeto directo es el sintagma preposicional. Obviamente esta regla sólo se aplicará para los verbos transitivos cuyo objeto directo puede ser animado; quedan excluidos los que tienen objeto directo exclusivamente animado.

Esta regla comprueba que el núcleo del sintagma preposicional no es un pronombre tónico porque el objeto directo no puede realizarse exclusivamente como pronombre tónico; éste sólo puede aparecer como duplicación de un objeto directo en forma de pronombre átono.

El sintagma preposicional, además de objeto directo, puede desempeñar las funciones de complemento preposicional y complemento circunstancial. Asignamos antes el objeto directo que el complemento preposicional y que el complemento circunstancial porque el primero posee un mayor grado de cohesión con el verbo que los otros dos. De esta manera, si en las reglas anteriores hemos detectado un objeto directo no es necesario plantearse por ejemplo la ambigüedad entre objeto directo y complemento preposicional en forma de sintagma preposicional con sustantivo como núcleo.

De la misma manera, intentamos antes el objeto directo en forma de sintagma preposicional que en forma de sintagma nominal porque éste puede ser también sujeto. Es decir, antes de intentar asignar la función de objeto directo a un sintagma nominal agotamos las posibilidades de que tenga otra forma puesto que el sintagma nominal podría ser tanto objeto directo como sujeto o complemento circunstancial.

El constituyente más ambiguo, el sintagma nominal, es el último que intenta la regla de asignación del objeto directo. Si en las reglas anteriores hemos detectado un objeto directo no es necesario plantearse por ejemplo la ambigüedad entre objeto directo

y sujeto o complemento circunstancial en forma de sintagma nominal.

En el caso del sintagma nominal distinguimos dos tipos: el sintagma nominal cuyo núcleo es no animado y el sintagma nominal cuyo núcleo es animado. El sintagma nominal cuyo núcleo es no animado sólo puede aparecer con los verbos transitivos que pueden tener objeto directo no animado, pero no puede aparecer con los verbos transitivos que tienen objeto directo exclusivamente animado. El sintagma nominal cuyo núcleo es animado sólo puede aparecer con los verbos transitivos que pueden tener objeto directo animado o no; no puede aparecer ni con los verbos transitivos que tienen objeto directo exclusivamente no animado ni con los que lo tienen exclusivamente animado.

En caso de que el verbo transitivo pueda tener objeto directo animado o no, único caso en que pueden aparecer ambos tipos de sintagma nominal, la regla intentará primero el no animado y luego el animado. El motivo es que los verbos transitivos tienen sujeto animado en la mayoría de los casos y éste puede estar desempeñado por un sintagma nominal cuyo núcleo es animado. Sin embargo, la regla intenta antes el objeto directo que el sujeto porque el primero posee un mayor grado de cohesión con el verbo que el segundo.

## 6. Estado actual del sistema

Como hemos dicho, en esta versión consideramos exclusivamente constituyentes formados por sustantivos y pronombres: sintagma nominal, sintagma preposicional, pronombres átonos (lo, la, los, las, le, les, me, te, nos, os y se) y pronombres tónicos (yo, tú, él, ella, (ello), nosotros/as, vosotros/as, ellos/as, mí, ti, sí). No están consideradas las oraciones subordinadas ni coordinadas. Hemos decidido reducir el tipo de constituyentes tratados porque el objetivo que nos planteamos es demostrar que una estrategia de este tipo es eficaz en la resolución de ambigüedades.

Por un lado, el tratamiento de las oraciones subordinadas implica problemas de otro orden que no son actualmente objeto de nuestro estudio, fundamentalmente el de la determinación de la estructura de constituyentes. Por otro lado, en el ámbito definido existen problemas de ambigüedad suficientemente relevantes como para que nos sirva para determinar la validez de la estrategia que definimos.

Respecto a las funciones sintácticas, consideramos todas aquéllas que pueden desempeñar estos constituyentes y que son dependientes del verbo; no consideramos por el momento las funciones que dependen de otros constituyentes. Las funciones son: sujeto, complemento directo, complemento indirecto, complemento preposicional regido, complemento predicativo, atributo y complemento circunstancial. Los valores son: valor pronominal, valor reflexivo y recíproco, valor impersonal, valor pasivo y valor medio.

Actualmente el sistema analiza todos los tipos de verbos que resultan del

conjunto de dos mil verbos clasificados. Es decir, en lo que se refiere a tipología verbal, podemos decir que el sistema cubre prácticamente todos los posibles tipos de verbos existentes en español, aunque es posible encontrar un tipo de verbo que no haya aparecido en los clasificados. No obstante, es poco probable que esto suceda puesto que los verbos poco utilizados, en cualquier lengua, suelen tener un número reducido de acepciones; son los verbos de uso frecuente los que pueden tener más diversidad de acepciones.

## 7. Conclusiones y discusión

Lo primero que puede saltar a la vista al pensar sobre la estrategia utilizada es una aparente contradicción. De una parte se utiliza un lenguaje de programación declarativo como Prolog en el que se define una gramática libre de contexto y, por otra, el algoritmo de resolución es procedimental.

La contradicción no es tal si se tiene en cuenta lo siguiente. Las reglas de la gramática quedan reducidas al mínimo, ya que éstas no tienen que tomar ninguna decisión acerca de la estructura interna de los constituyentes. De otra parte, hemos diseñado este prototipo con la idea de investigar una posible interacción entre Prolog y C. La siguiente versión del prototipo pretende utilizar Prolog para la definición de la gramática y dejar que C resuelva el algoritmo de asignación de funciones.

El segundo problema que se puede plantear es que el algoritmo de asignación de funciones es poco universal. Esto es, está diseñado sólo para el castellano. De hecho, pretendemos demostrar que la asignación de funciones en lenguas con orden semilibre se puede efectuar de forma más rápida y con un índice de éxito más alto con una estrategia de retraso frente a una estrategia basada en reglas declarativas.

Desde un punto de vista psicolingüístico, pensamos que, si se verifican los resultados que hemos obtenido hasta ahora, podemos ofrecer una serie de principios válidos para la asignación de funciones en castellano que quizá se puedan extender a otras lenguas romances. Por ejemplo, al contrario que todos los analizadores conocidos, que operan de izquierda a derecha, nuestro algoritmo busca asignar primero el objeto directo, y no el sujeto.

El tercer aspecto susceptible de discusión es que, pese a la tendencia actual en formalismos lingüísticos, que tienden a una especificación localista de las restricciones y fenómenos lingüísticos, parece que esto no es posible en el caso de la ambigüedad estructural que ofrece el castellano. O, si se quiere, se puede redefinir el problema diciendo que el contexto en el que hay que resolver esa ambigüedad es la oración entera. Este contexto está determinado en primer lugar por el tipo de verbo y los constituyentes presentes en la oración.

