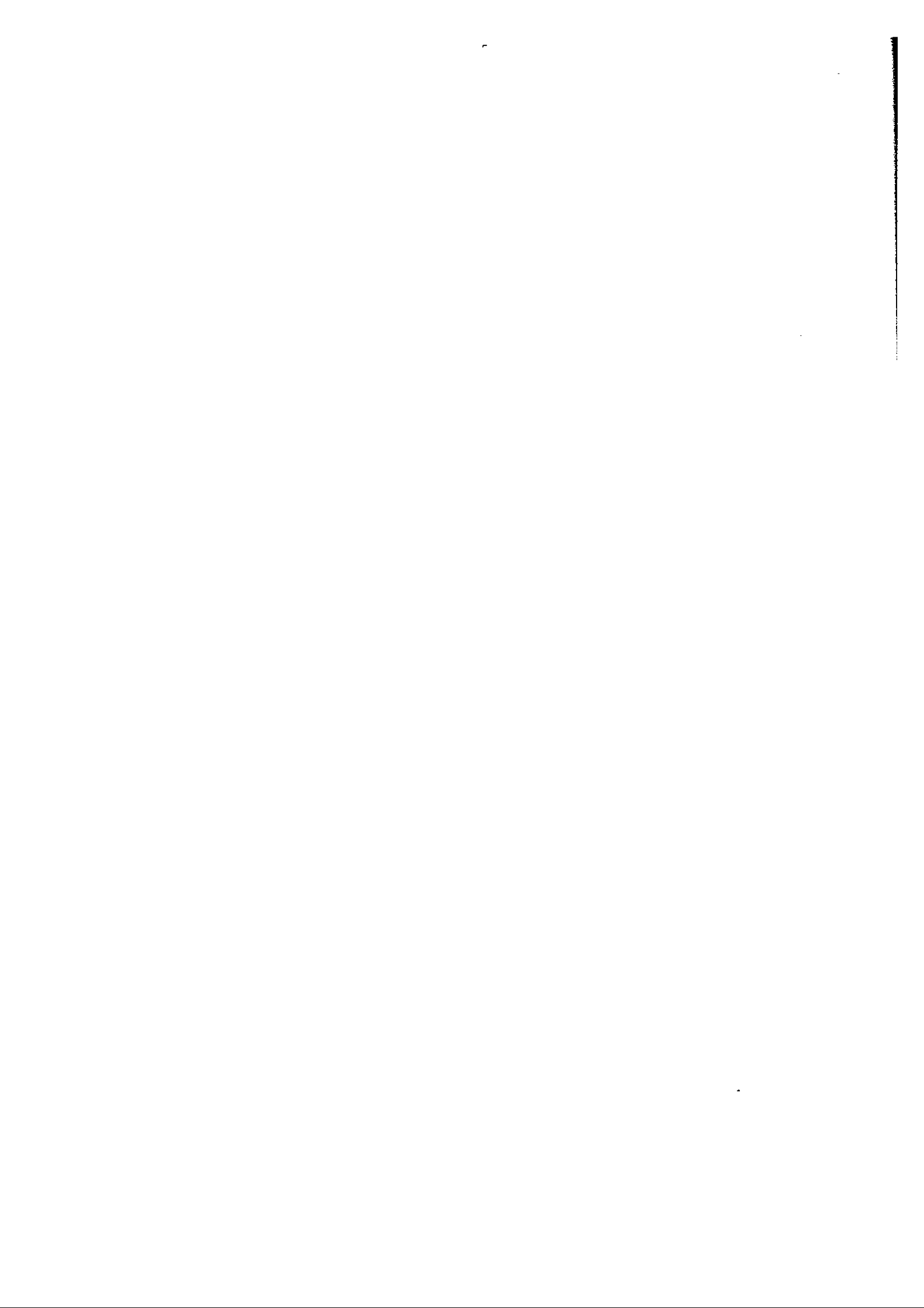


5. Traducción automática



La naturaleza sintáctica del nombre

Ana Aguilar-Amat

FBG/UB

ana@gilcub.es

Resumen:

El nombre es la categoría léxica menos voluble semánticamente hablando. Por ese motivo parece importante fijar una clasificación que controle sus relaciones sintagmáticas antes de clasificar los complementos del nombre y/o los adjetivos, clases semánticamente mucho más conflictivas.

La subcategorización sistemática del léxico en este sentido no sólo permite determinar mucho más fácilmente la pluralidad conceptual de una entrada, sino también elaborar reglas que procedan al análisis y la desambiguación del nombre y de los adjetivos o sintagmas preposicionales que le acompañen.

En cuanto a la selección léxica, y en el presupuesto de que un diccionario de coocurrencias facilitaría en extremo el control de las contextualizaciones, la correcta clasificación de las categorías implicadas en las colocaciones proporciona economía e inteligencia al diccionario.

1. Introducción

La clasificación del nombre que se ofrece en este trabajo ha sido concebida para poder ejercer un control sobre el valor temático de sus complementos.

En el proceso de análisis esta clasificación permite identificar a los complementos nominales de los nombres predicativos como argumentos (la entrada del fontanero), y como modificadores (una entrada de 25 pesetas/de piedra) o posesivos (una entrada de la casa) en el caso de nombres no predicativos.

Ayuda por lo tanto a construir una estructura abstracta correcta del sintagma (ya sea un árbol o una red semántica).

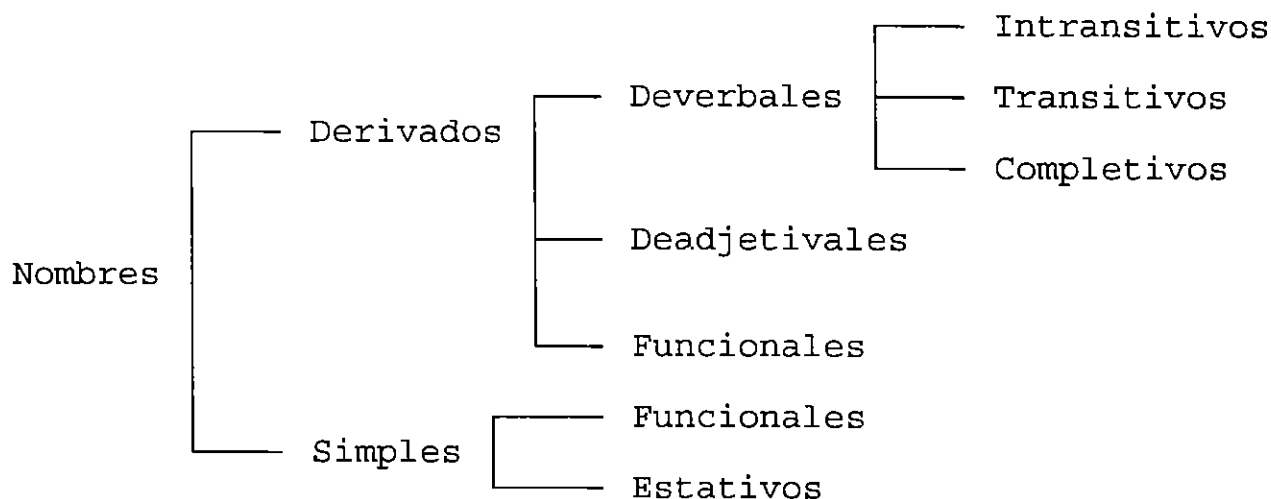
En cuanto a los adjetivos, su propia tipología, en estrecha relación con la clasificación del nombre, permite eliminar la ambigüedad provocada por polisemia del nombre y/o del adjetivo.

Ahora bien, en el proceso de generación el problema no es el de la desambiguación de estructuras sino el de la selección léxica. Por ejemplo, el hecho de saber que el nombre que nos ocupa es un nombre predicativo como pueda serlo "respiración" (la respiración de Pepe -> Pepe respira) y que el adjetivo que le acompaña es, pertinentemente, del tipo A3, no ayuda a efectuar la selección de "entrecortada" de entre los adjetivos *discontinuo*, *entrecortado*, *desacompañado*, o, en caso de que se generara en inglés, no ayuda a escoger entre los vocablos "laboured" que se aplica a *respiración* y "faltering", "hesitant" o "confused" que se aplican a *habla*.

A este efecto, la clasificación que se ofrece es parte de un estudio de las relaciones del nombre con sus complementos en aras de fijar ciertas "etiquetas" que caractericen el comportamiento sintagmático de una palabra en un diccionario de colocaciones.

2. Clasificación sintáctica del nombre.

La clasificación que concebimos después de atender a las consideraciones de otros autores y buscando su máxima aplicabilidad en el corpus con el que trabajamos responde al siguiente cuadro:



Esta clasificación resulta en siete grupos de nombres, los cuales llevan una marca subcategorizando cada una de las entradas. Estas marcas son siglas que utilizaremos a partir de ahora y son las siguientes:

NDVI	Nombres derivados deverbiales intransitivos
NDVT	Nombres derivados deverbiales transitivos
NDVC	Nombres derivados deverbiales completivos
NDD	Nombres derivados deadjetivales
NDF	Nombres derivados funcionales
NSF	Nombres simples funcionales
NSE	Nombres simples estativos

2.1. Nombres derivados deverbiales intransitivos (NDVI).

Esta denominación agrupa aquellos nombres que proceden de verbos que poseen un sólo argumento, independientemente de que éste sea agente o experimentante, pero no de aquellos verbos cuyo único argumento tiene rol de objeto (verbos inacusativos), casos que se contemplan como NDVT.

Son por ejemplo: *salida, entrada, llegada, caída, llanto, sonrisa, muerte, paso*, etc.

La característica principal de estos nombres es que pueden y suelen llevar un complemento nominal con 'de' que cumple una función de sujeto: la salida del tren, la entrada del conferenciante, la llegada de Pedro, la caída del abuelo, el llanto del bebé, etc.. Este complemento sólo se omite en casos de elisiones anafóricas o debido a usos específicos o pragmáticamente traslaticios como 'dar la salida (de los corredores)', 'limitar la entrada (del público) (al local)', etc.

Existen nombres que proceden de verbos terciopersonales como los ambientales (llover, nevar -> lluvia, nevada) que carecen por completo de argumentos (exceptuando la lectura figurativa de *lluvia* en *lluvia de piedras, lluvia de confetti*, etc. que será tratada como NDVT). Estos casos, irán marcados con un rasgo adicional indicativo de su excepcionalidad.

Otros nombres procedentes de verbos terciopersonales como son los existenciales no tienen este problema (hay un libro -> la presencia del libro, tiene una mancha -> la existencia de la mancha).

Hay que notar que muchos verbos tienen una doble lectura: una transitiva y otra inergativa cuando se los utiliza con 'se', como ocurre, por ejemplo, con *unirse* y *unir*. En estos casos se prevén dos entradas, con lo cual habrá también dos entradas para los nombres argumentales que de ellos se derivan: la unión de los directivos (NDVI) y la unión de las piezas por el zapatero (NDVT).

2.2. Nombres derivados deverbales transitivos (NDVT)

Esta denominación afecta a nombres derivados de verbos que poseen dos o más argumentos, uno con rol de agente, otro con rol de objeto y a veces un objeto indirecto con rol benefactivo o un complemento de lugar (verbos de movimiento).

Son, por ejemplo: *ejecución, amonestación, absorción, venta, compra, deglución, traslado, transporte, recorrido, manipulación, contemplación, decisión*, etc.

Estos argumentos se representan en forma de sintagmas preposicionales. Normalmente el objeto se expresa en primer lugar y siempre con la preposición 'de', mientras que el sujeto ocupa el segundo lugar (o se omite, dado que la nominalización es una topicalización de la acción frente al agente) con la preposición 'por': la ejecución del reo por el verdugo, la absorción del agua por la esponja, la venta de la producción por la empresa, etc.

Ocurre también, sin embargo, que el objeto puede omitirse, con lo que el primer complemento será el agentivo: *la omisión de Pedro, la declaración de la testigo*, etc. Naturalmente esto crea casos de ambigüedad temática que, sólo a veces, pueden ser resueltos por restricciones semánticas del verbo sobre su sujeto y/o objeto.

Por otro lado se da el caso de los objetos directos de rasgo humano o animal que llevan la preposición 'a' y son a la vez benefactivos de la acción verbal. En estos casos la nominalización puede provocar dos estructuras distintas:

la amonestación de los niños por el profesor
la amonestación del profesor a los niños

Esto significa que habrá que tener en cuenta la presencia y la preposición del segundo complemento para poder determinar si el primero es el sujeto o el objeto, sobre todo porque en este caso ambos poseen el mismo rasgo semántico.

Esto se une a la dificultad de que haya una preposición 'de' regida, caso en que suele dársele a ese complemento la preferencia de proximidad al término:

la acusación *de* prevaricación del fiscal (al ministro)
la acusación *de* prevaricación del ministro por el fiscal.

Cuando la preposición no es 'de' la ambigüedad desaparece, pero entonces el orden no está marcado:

la disertación *sobre* economía del ministro
la disertación del ministro sobre economía

la predisposición a soñar de los alumnos

la predisposición de los alumnos a soñar

la relación con los alumnos del maestro

la relación del maestro con los alumnos

Los nombres procedentes de verbos de movimiento tienen la posibilidad de tener un complemento de origen y otro de destino: *el transporte de las naranjas desde Valencia hasta la frontera, el traslado de los muebles de una casa a la otra, el trayecto de la flecha desde el arco hasta el objetivo*, etc.

2.3. Nombres deverbales derivados completivos (NDVC)

Existe un grupo de nombres que se originan como una nominalización a partir de una acción, pero que no se corresponden con la predicación propiamente dicha sino con alguno de los complementos o adjuntos del verbo; por lo que no son argumentales. Los sintagmas preposicionales o complementos nominales que acompañan a estos nombres no son argumentos sino modificadores o posesivos.

En general los nombres predicativos (NDVI y NDVT) suponen proceso, y son marcados en nuestro sistema con los rasgos 'evento' o 'estado', aunque existen nombres de procesos que no son deverbales (ruido, neumonía). Los deverbales completivos suelen ser nombres concretos o abstractos.

Los nombres derivados de los adjuntos del verbo y no del verbo en sí con sus complementos no son predicativos propiamente dichos, y no poseen argumentos. De esta manera, un nombre derivado deverbal como pueda serlo "entrada", que deriva del verbo "entrar" tiene varias entradas en el diccionario (léase "items") diferenciadas semánticamente que coinciden con la lectura predicativa y con las lecturas no predicativas derivadas de los adjuntos del verbo.

El patrón verbal de una de las entradas del verbo 'entrar' es el siguiente:

SN(AGT)/entrar/SPrep(en,PLACE)

alguien/entra/en algún lugar

Pepe/entró/en el local

Pero el entorno verbal completo incluye los siguientes adjuntos:

SN(AGT)/entrar/SPrep(en,PLACE)/Sprep(por,PLACE)/SPrep(por,IMPL)

alguien/entra/en algún lugar/por algún lugar/por un precio

Pepe/entró/en el local/por la puerta principal/por 25 pesetas

Las entradas derivadas de este verbo serán las siguientes:

entrada1_ DEVERB, PRED, EVENTO, INTRANS

acción (la entrada de Pepe en el local)
entrada2_ DEVERB, NO_PRED, CONCRETO
puerta principal (la entrada era de madera)
entrada3_ DEVERB, NO_PRED, CONCRETO
'ticket' (la entrada valía 25 pesetas)

Nótese que el verbo entrar tiene una lectura causativa:

SN(AGT) entrar SN(AGT) SPrep (en, PLACE)
alguien entra algo en algún lugar
(hace que algo entre)
Pepe entró las maletas en el auto
(hizo entrar las maletas)

Esto provocaría la presencia de un nombre derivado deverbal predicativo con dos argumentos y no uno solo, es decir, que habría dos lecturas predicativas de *entrada*, una SDVI y otra SDVT, la entrada cuatro:

entrada_4 DEVERB, PRED, EVENTO, TRANS
la entrada de las maletas en el auto (por Pepe)

En ocasiones se ha calificado de resultativos a los nombres deverbales que proceden del objeto directo del verbo: la producción (lo producido) ha sido vendida, lleva la cesta con la compra (lo comprado), etc. Naturalmente sólo poseeran un homófono resultativo los nombres derivados deverbales que sean transitivos.

2.4. Nombres derivados deadjetivales (NDD)

Los nombres derivados deadjetivales son nombres procedentes de adjetivos, y coinciden en parte con lo que los estudiosos han denominado "nombres de medida". Son nombres deadjetivales: *anchura, estrechez, altura, fortaleza, debilidad, intensidad...* y, también: *juventud, vejez, fealdad, belleza, sabiduría, estupidez, etc.*

A pesar de que los adjetivos de los cuales se derivan estos nombres son adjetivos bipolares (pares de antónimos), los correspondientes nominales a menudo están decantados hacia el componente de exceso, tal como ya indicó Vendler (1968). Es de uso infrecuente 'bajura' frente a 'altura', por ejemplo, y puede uno referirse a 'la anchura de algo estrecho', pero no a 'la estrechez de algo ancho'. Dicho de otra manera el nombre correspondiente al exceso es el genérico de la cualidad. Esto no sucede así, sin embargo, cuando lo medido no es una cualidad física sino abstracta o moral (*la bondad de alguien malo, *la belleza de alguien feo), por lo que cabe, o bien hacer dos grupos distintos, o bien crear una nueva entrada para el genérico de la dualidad en el caso de las medidas físicas.

El hecho de proceder de adjetivos relativos (subjativos) y por lo tanto graduables, es decir, que suelen admitir el superlativo, ya sea sintáctico (muy pequeño), morfológico (pequeñísimo) o léxico (ínfimo), hace que estos nombres admitan adjetivos de gradación o cuantificadores.

Estos nombres son denominados también de "cualidad" (Romero Gualda, 1988), quien afirma que los adjetivos que les acompañan suelen aportar un matiz superlativo-valorativo: *extraordinaria belleza, gran altura, deformante gordura*, lo que hace que a menudo se antepongan al nombre.

La probabilidad de que sean adjetivados por algún adjetivo de otro tipo es prácticamente nula, a excepción de los denominales: *belleza paisajística (del paisaje), fortaleza espiritual (del espíritu), anchura corporal (del cuerpo)*, que son en realidad transformaciones de estructuras del tipo 'el paisaje es bello, el espíritu es fuerte, el cuerpo es ancho, etc.

2.5. Nombres derivados funcionales (NDF)

Por nombres derivados funcionales nos referimos a los nombres que derivan del sujeto de ciertos verbos, como *corredor, nadador, bailarina, moribundo*, etc., y también a nombres que son agentivos pero que derivan del objeto de una acción implícita: *panadero (que hace pan), hormigonera (que hace hormigón), carnicero (que vende carne)* etc.

Los primeros pueden relacionarse morfológicamente con los verbos de los que son sujeto agente:

corredor > que corre
bailarina > que baila
jugador > que juega
investigador > que investiga
amante > que ama

pero se producen excepciones debido a cambios semánticos históricos:

orador > que habla
*que ora

Esto significa que el criterio morfológico no aporta el suficiente rigor, y que, de ser deseada la plasmación de las relaciones, estas deberán ser explicitadas de otra forma, con la enunciación del verbo original en forma de rasgo, por ejemplo (nom_agt_verb='hablar').

Los nombres derivados funcionales del segundo grupo señalado tienen más problemas en ese sentido. Obsérvese que estos nombres son susceptibles de ser sujeto no de uno sino de una pluralidad de verbos:

científico > que hace ciencia
que estudia, observa, investiga, escribe, publica, etc.

panadero > que hace pan
que mezcla, amasa, cuece, vende, etc.

Pero esto, claro está, está más cerca de una información enciclopédica que de una

propriadamente lingüista con un objetivo de control determinado, como puede ser el de relacionar los nombres con sus modificadores.

Pero esta particularidad de este segundo grupo hace que se asemeje o esté a medio camino entre los nombres funcionales derivados y los nombres funcionales simples.

2.6. Nombres simples funcionales (NSF)

Llamamos simples a los funcionales que no son derivados, es decir, que no proceden de un verbo ni de un nombre-objeto, pero que son susceptibles de realizar acciones y son agente de verbos implícitos:

sacerdote > que oficia
madre > que cuida, ama, educa, etc.
caballo > que corre, lleva, tira, etc.

Naturalmente que los nombres con rasgo semántico animado (humano o animal) pueden ser sujeto de verbos como 'comer', 'beber', etc., y junto a los vegetales también de verbos como 'reproducirse' o 'morir', pero este no es el caso. Cuando especificamos que estos nombres son funcionales nos referimos al hecho de que pueden ser sujeto de acciones que afectan a segundos o a terceros, y en general, que pueden ser modificados por adjetivos adverbializables, es decir, adjetivos que califican o valoran sus acciones y no a ellos mismos.

2.7. Nombres simples estativos (NSE)

Los nombres simples estativos son nombres que no son funcionales y tampoco son derivados. Constituyen lo que filosóficamente podría denominarse nombres de los entes o también nombres originales o primitivos. Naturalmente que esto es muy relativo puesto que muchos nombres derivados también son nombres de entes, pero ellos pueden relacionarse, al menos morfológicamente, con palabras de otras categorías y permitir cambios de estructuras. Por eso van subcategorizados de otra manera.

La calificación de 'estativo' ha sido escogida por oposición a 'funcional'.

Ejemplos de nombres simples estativos son: árbol, mofeta, lápiz, libro, tecla, nota, mueble, luz, etc.

Obsérvese que 'mueble', como 'inmueble', es un nombre que procede del latín *móbilis*, derivado de *movere* (mover), por lo que podría objetarse que se catalogara como simple y no como NDVT; pero lo cierto es que a efectos de su comportamiento sintáctico este nombre hace ya tiempo que perdió el vínculo con el verbo que lo originó. Además, este es un nombre que resulta de la nominalización de un adjetivo (mueble -> que puede moverse), por lo que en todo caso debiere ser un NDD. Hay que distinguir, sin embargo, entre los NDD morfológicos o derivados, que tienen un comportamiento sintáctico similar, y las nominalizaciones de adjetivos por 'economía' o elisión (dulce -> un dulce de nata, loco -> un loco peligroso), casos que provocan una duplicación en entradas homófonas, solución costosa que sólo podría ser simplificada en una concepción de la gramática que otorgara la información categorial según el comportamiento sintáctico en cada caso.

Por otro lado, cabe señalar que los nombres derivados deverbales completivos (NDVC) tienen un comportamiento similar a los nombres simples, y, desde luego, también

son nombres de entes.

Hechas estas aclaraciones, cabe decir que los nombres simples estativos se caracterizan primordialmente por que no van acompañados por adjetivos adverbializables, o, en todo caso, no pueden transformarse en una estructura de N V Adv:

flor aromática -> *flor que es aromáticamente

Estos nombres se combinan, según Vendler, con adjetivos del tipo A1. Es preciso evitar, sin embargo, la mezcla de criterios que caracteriza la clasificación de este autor. Afirma que este tipo de nombre se caracteriza por su neutralidad (nueva palabra para la extensibilidad o co-hiponimia), es decir, que la adjetivación puede transferirse a otro nombre:

flor carnívora -> planta carnívora

Restringir este fenómeno a los nombres simples no parece lo más acertado. Aunque la hiponimia pudiese servir para facilitar la tarea de establecer un cuadro de relaciones que permitiera, al igual que hacen las relaciones sintagmáticas, reducir el número de colocaciones por extrapolación de las relaciones de coocurrencia, se trata de un fenómeno general que no depende de la tipología nominal tanto como del carácter atributivo (intensional) o no atributivo (extensional) del adjetivo, distinción que consideramos básica:

flor oval -> *planta oval

Ahora bien existen más objeciones a una teoría general de combinaciones de N y A. Como ocurre con el ejemplo de 'bailarina grácil/flor grácil', la capacidad metafórica de los hablantes puede hacer que se produzcan combinaciones de nombres simples con adjetivos propios de nombres funcionales y viceversa, por lo que se hace patente que, cuando éstas se generalizan, deben estar caracterizadas en cada caso.

Referencias

ROMERO GUALDA, V., 1988, El nombre:sustantivo y adjetivo, Arco Libros, M., 1989

VENDLER, Z., Adjectives and nominalizations, 1968, The Hague, Mouton & Co. N.V. Publishers, 1968.

Módulo de generación en JULIETTA

Gabriel Amores Carredano
Departamento de Lengua Inglesa
Universidad de Sevilla
e-mail: gamores@sevax4.us.es

0. Introducción

En anteriores presentaciones (Amores Carredano 1993) he descrito el prototipo de traducción automática (TA) JULIETTA, un sistema basado en la gramática léxico funcional (LFG) implementado en Prolog.

En esas ocasiones nos ocupamos de describir la arquitectura general del sistema, su cobertura lingüística y el algoritmo de transferencia.

En este artículo nos ceñiremos al módulo de generación.

1. Fases de la generación

Dentro de la arquitectura general del sistema, el módulo de generación toma una representación de la oración en lengua destino y genera la correspondiente cadena de palabras.

En el caso de JULIETTA, la representación que se obtiene tras la transferencia es una estructura funcional (f-structure en LFG) para el español. De esta estructura funcional, siguiendo el camino inverso del análisis, se debe generar primero la estructura de constituyentes que le corresponde (c-structure en LFG) y después la cadena de palabras.

La figura 1 ilustra esta derivación.

Source c-structure

S --> NP VP
↑Subj=↓ ↑=↓

PRED	:	X
TENSE	:	Y
SUBJ	:	[Z]

Target c-structure

S' --> NP' VP'
↑Subj=↓ ↑=↓

Transfer

PRED	:	X
TENSE	:	Y
SUBJ	:	[Z]

Source f-structure

Target f-structure

ANALYSIS -----> TRANSFER -----> GENERATION

figura 1

2. Generación desde estructuras funcionales

El concepto de generación en el Procesamiento del Lenguaje Natural varía de una aplicación a otra. En sistemas de TA que siguen el enfoque de transferencia se asume que la selección léxica ya ha tenido lugar durante la transferencia y que la labor del generador consiste en la generación morfológica y ordenación de los constituyentes siguiendo una gramática de la lengua destino.

El desarrollo de algoritmos de generación para gramáticas basadas en unificación es un campo relativamente reciente, ya que casi todas las implementaciones se habían ceñido a problemas de análisis y/o transferencia.

Por lo que se refiere a LFG, Momma y Dörre (1987), Netter y Wedekind (1986), Wedekind (1986), y Gates y otros (1989) proponen algoritmos de generación desde estructuras funcionales, mientras que Calder y otros (1989) y Shieber y otros (1990) ofrecen ofrecen un algoritmo general de generación aplicable a gramáticas de unificación. Asimismo, Ruiz Antón (1993) hace una propuesta de generación desde estructuras funcionales aplicando ideas de la gramática funcional.

Nuestro algoritmo sigue la propuesta de Momma y Dörre (1987) y Gates y otros (1989).

Podemos definir el problema de la generación como el *descubrimiento de un árbol de análisis* para una estructura funcional dada como insumo (**input**). Este proceso encierra varios problemas.

Primero, la generación es, por definición, no determinística, ya que se puede generar un número infinito de árboles posibles para una estructura funcional dada.

Considérese, por ejemplo, la siguiente estructura funcional parcial, de la cual se quiere generar la cadena *la presión sanguínea sistólica*. Si nuestro generador opera con prioridad de profundidad (**depth-first**), no hay manera de asegurar que obtendremos el resultado deseado ya que se pueden generar varios árboles de los que resultarían, al menos, las cadenas de palabras que se muestran a la derecha de la representación.

pred:presión

spec:el
mods:pred:sistólico
temp:no
tail:pred:sanguíneo
agr:gen:fem
num:sing

presión, la presión, presión sanguínea,
la presión sanguínea, presión sistólica,
la presión sistólica, presión
sanguínea sistólica, la presión
sanguínea sistólica

figura 2

Este hecho descarta en principio un algoritmo con prioridad de profundidad, como ha sido demostrado también por Momma y Dörre (1987) y Shieber y otros (1990).

El segundo problema surge del tipo de representación que codifica una estructura funcional en LFG. La estructura funcional carece de información categorial y de información acerca del orden en que aparecen los constituyentes. Se supone que este tipo de información, al ser más dependiente de la lengua, no debe estar presente en la estructura funcional y se descarta durante el análisis. Esto ya fue apuntado por Kaplan y otros (1989) como argumento para defender un enfoque de **co-descripción** para sistemas de TA basados en LFG.

Otras gramáticas basadas en unificación, como la **Functional Unification Grammar (FUG)**, (Kay 1985), obvian este problema incluyendo información categorial y de precedencia lineal en la estructura funcional¹.

Así, por ejemplo, la estructura funcional parcial anterior tendría el aspecto siguiente en una representación basada en FUG.

pred:presión
pattern:<spec,pred,mods>
cat:noun
spec:el
mods:pred:sistólico
pattern:<tail,pred>
cat:adj
temp:no
tail:pred:sanguíneo
cat:adj
agr:gen:fem
num:sing

figura 3

¹ En FUG, el atributo **pattern** se usa para indicar precedencia lineal. Es interesante hacer notar que FUG fue la única gramática de unificación con una aplicación directa a la TA. Tal vez la inclusión de esta información en la estructura funcional se viera necesaria como un requisito impuesto por la TA.

Wedekind (1986) y Netter y Wedekind (1986), también en un entorno de TA, proponen una representación de un solo nivel (frente a los dos niveles tradicionales en LFG) para solventar el problema.

Las estructuras de su versión consisten en estructuras funcionales aumentadas con información adicional acerca de los símbolos derivados y su orden lineal. A continuación proponen un concepto de **derivación** que puede usarse tanto en modo análisis como en modo generación, con idea de que las gramáticas sean reversibles. Si se usan en generación, la estructura funcional insumo lidera el proceso y las reglas libres de contexto se satisfacen si las ecuaciones anotadas propias de LFG están realizadas con material no vacío en la estructura parcial que se está generando.

Sin embargo, no mencionan nada acerca del no determinismo del proceso de generación, ya que sólo comprueban la existencia de valores no vacíos (ya sean de tipo atómico o complejo) en los rasgos relevantes.

Momma y Dörre (1986) solucionan el problema imponiendo un algoritmo con prioridad de extensión (**breadth-first**) a la satisfacción del proceso de generación. Así, en vez de proceder simplemente en modo descendente (**top-down**), el algoritmo ha de comprobar que **toda la información** contenida en la estructura funcional se ha consumido en el proceso. De esta forma se asegura que se intenta primero la regla de generación más larga.

3. Nuestra propuesta

Si aplicamos estas ideas a nuestro ejemplo de estructura funcional parcial, el generador observará la presencia de un especificador (atributo **spec**), un nombre (atributo **pred** en este caso) y un modificador complejo (atributo **mods**), y tratará de aplicar la regla que tiene en cuenta a los tres, y sólo a los tres, en dos llamadas recursivas.

En JULIETTA, cada regla de generación toma como primer argumento una estructura funcional y devuelve como salida (**output**) la porción de árbol que esta regla genera. Por ejemplo, las reglas que se necesitan para generar el ejemplo anterior serían las siguientes.

```
generate(EstructuraF,sn(Art,Stbar)) :-  
  ( member(spec:_EstructuraF)  
  ; member(quant:_EstructuraF)  
  ; member(det:_EstructuraF)  
  ; member(dem:_EstructuraF) ),  
  generate_rs(EstructuraF,Art,EstructuraF1),  
  generate_stbar(EstructuraF1,Stbar).
```

```

generate_stbar(F,stbar(st(St),Sadj)) :-
    getargs(EstructuraF,List),
    perm(List,[agr,X]),
    member(X,[mod,mods,vmod]),
    member(X:Mod,EstructuraF),
    \+ member(stayput:yes,Mod),
    member(agr:A,EstructuraF),
    gener_sadj([agr:A|Mod],Sadj),
    delete(X:Mod,EstructuraF,EstructuraF1),
    reduce(EstructuraF1,st,St).

```

figura 4

La primera regla comprueba la presencia de cualquier especificador, y aplica recursivamente la regla de generación de sintagmas determinantes (**generate_rs/2**). Este predicado devuelve la estructura funcional original sin el especificador, y **generate_stbar/2** se aplica para obtener un sintagma nominal de un nivel jerárquico inferior.

La estrategia con prioridad de extensión (**breadth-first**) opera de la siguiente manera: el predicado **getargs/2** devuelve una lista con todos los valores no atómicos en la estructura funcional. En el ejemplo anterior instanciaría **List** a la lista [mods,agr]. **perm(ute)/2** comprueba que estamos tratando con los valores pertinentes, ya que los rasgos en la estructura funcional no vienen en un orden pre-establecido. Se puede asimismo observar que esta regla se utiliza para generar sintagmas nominales modificados por otro tipo de material, concretamente **mod**, **mods** o **vmod**.

De hecho existe cierto isomorfismo entre la estructura de constituyentes que se generará y el tipo de rasgo con el que estamos tratando. Así, **pobj**, **pmod**, **padj** y **padj1** generan todos sintagmas preposicionales, y así sucesivamente.

Finalmente, si todos los predicados tuvieron éxito, se hace una llamada al generador morfológico para sintetizar la forma apropiada del sustantivo, verbo, adjetivo, etc.

Nótese que los predicados más costosos computacionalmente hablando (la recursión sobre reglas de generación y el generador morfológico) sólo se invocan si se cumplen las restricciones impuestas a la regla de generación.

4. El generador morfológico

El generador morfológico está controlado por el predicado **reduce/3**. Éste toma como insumo una lista de pares atributo-valor y genera el lexema correspondiente teniendo en cuenta la categoría deseada. Un ejemplo de llamada sería el siguiente

?- reduce([pred:niño,agr:[gen:masc,num:plur]],st,L).

L = niños

yes

La versión actual del generador morfológico sintetiza todas las formas regulares de los verbos, sustantivos y adjetivos. Las irregularidades han de ser declaradas como excepciones a la base de datos en Prolog. No obstante, no es necesario contar con un diccionario en lengua destino. Toda la información léxica necesaria para la generación se encuentra a la salida de la transferencia.

Por último, una vez que la regla superior **top_generate/3** ha generado el nodo más alto de la gramática, este término se pasa a una gramática DCG del español, que convertirá directamente el árbol de análisis en una lista de palabras. Esto es posible gracias al carácter reversible de los predicados en Prolog.

Referencias

- Amores Carredano, J.G. (1993) "JULIETTA: Un prototipo de TA basado en LFG" Procesamiento del Lenguaje Natural 13, pp. 305-317
- Calder, J., M. Reape y H. Zeevat (1989) "An Algorithm for Generation in Unification Categorical Grammar" 4th Conference of the European Chapter of the ACL. Manchester, pp. 233-240.
- Gates, D., K. Takeda, T. Mitamura, L. Levin y M. Kee. (1989) "Analysis and Generation Grammars." Machine Translation 4 (1) Special Issue on Knowledge-Based Machine Translation. pp. 53-66.
- Kay, M. (1985) "Parsing in Functional Unification Grammar." En Dowty, D.R., L. Karttunen and A. Zwicky eds. Natural Language Parsing. Cambridge University Press, Cambridge, pp. 251-278.
- Momma, S. y J. Dörre (1987) "Generation from f-structures." En Klein and van Benthem eds. Categories, Polymorphism and Unification. Centre for Cognitive Science, University of Edinburgh, Edinburgh, 1987, pp. 148-167.
- Netter, K. y J. Wedekind (1986) "An LFG-based Approach to Machine Translation." Proceedings of IAI-MT86, Dudweiler, Germany, pp. 199-209.
- Ruiz Antón, J.C. (1993) "Un tratamiento funcional de la síntesis (en traducción automática)" Procesamiento del Lenguaje Natural 13, pp. 349-359.
- Shieber, S., F.C.N. Pereira, G. van Noord y R.C. Moore (1990) "Semantic-Head-Driven Generation" Computational Linguistics 16 (1), pp. 30-42.
- Wedekind, J. (1986) "A Concept of Derivation for LFG." En Coling 86. Bonn, pp. 487-489.

