



Procesamiento del Lenguaje Natural

ISSN: 1135-5948

secretaria.sepln@ujaen.es

Sociedad Española para el  
Procesamiento del Lenguaje Natural  
España

Tordera Yllescas, Juan Carlos

Propuesta de traducción sintáctico-semántica: el tratamiento anáforico a través de la LFG  
y la SDRT

Procesamiento del Lenguaje Natural, núm. 48, marzo, 2012, pp. 13-20

Sociedad Española para el Procesamiento del Lenguaje Natural  
Jaén, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=515751748001>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

# Propuesta de traducción sintáctico-semántica: el tratamiento anáforico a través de la LFG y la SDRT

## *Proposal for a syntactic-semantic translation: anaphoric approach by LFG and SDRT*

Juan Carlos Tordera Yllescas

Universitat de València

Blasco Ibañez, 13

juan.tordera@uv.es

**Resumen:** La Gramática Léxico-Funcional (LFG) y la Teoría Representacional del Discurso Segmentado (SDRT) han demostrado ser teorías formales considerablemente explicativas para el tratamiento de las lenguas naturales. El objetivo del presente trabajo es proponer una posible traducción que permita obtener fórmulas lógicas de la SDRT a partir de la estructura sintáctica proporcionada por la LFG.

**Palabras clave:** LFG, SDRT, traducción sintáctico-semántica, anáfora de sentido

**Abstract:** The Lexical-Functional Grammar (LFG) and the Segmented Discourse Representation Theory (SDRT) have proved to be two considerably explanatory formal theories for the treatment of natural languages. The aim of this paper is to propose a possible translation allowing to get SDRT logical formulas from syntactic structures provided by the LFG.

**Keywords:** LFG, SDRT, syntax-semantic translation, laziness anaphora

## 1 Introducción

Las gramáticas de unificación y las lógicas dinámicas desempeñan un papel importante en el desarrollo de la Lingüística computacional, tal como se describe en Allen (1995), Moreno Sandoval (1998) o Martí Antonín y Castellón Masalles (2000). Entre estas, la Gramática Léxico-Funcional (LFG) de Bresnan (2001) y la Teoría Representacional del Discurso Segmentado (SDRT) de Asher y Lascarides (2003) se han revelado teorías muy explicativas que dan cuenta de aspectos lingüísticos harto complejos como la pasiva, los elementos “desplazados”, las construcciones de infinitivo o la anáfora (Bresnan, 2001; Tordera Yllescas, 2008). Este trabajo persigue proponer algunas reglas de traducción que permitan obtener fórmulas lógicas propias de la SDRT a partir de la estructura sintáctica de la LFG para un fragmento del español (aunque es viable la futura inclusión de nuevas reglas) y esbozar algunas de las aplicaciones que podría conllevar para el tratamiento de algunos fenómenos lingüísticos (v.gr.: la anáfora).

## 2 Hacia una traducción de la LFG a la SDRT

Siguiendo las propuestas de Halvorsen (1983) o de Dalrymple (2001: 217-254), quienes proponen una descripción declarativa de reglas de correspondencia (RC) syntax-semántica para obtener fórmulas lógicas a partir de la descripción sintáctica de la LFG, pretendemos presentar una descripción declarativa de RC que nos permita obtener unas fórmulas lógicas infradeterminadas, (ULF), de la SDRT a partir de la estructura-f de la LFG.

En primer lugar señalaremos que, en este trabajo, no se discriminará categorialmente entre pronombres y determinantes, como es corriente ya en las gramáticas generativas.

Nuestras RC parten de los siguientes supuestos: es necesario que todo nombre y pronombre tenga asignado en el lexicón de la LFG una variable de individuo (i.e.,:  $x_1, x_2, \dots, x_n$ ) así como una etiqueta de la ULF (i.e.,:  $l_1, l_2, \dots, l_n$ ). Asimismo, los cuantificadores han de presentar una etiqueta de la ULF  $e$ , igualmente, los verbos han de presentar en el lexicón una

variable de evento (v.gr.:  $e_1, e_2 \dots e_n$ ) y una etiqueta de la ULF ( $l_1, l_2 \dots l_n$ ).

- (1) a. Ella: Pron. ( $\uparrow$  REFERENTE) =  $x_i$   
 ( $\uparrow$  ETIQUETA) =  $l_i$   
 b. Él: Pron. ( $\uparrow$  REFERENTE) =  $x_i$   
 ( $\uparrow$  ETIQUETA) =  $l_j$   
 c. Pocos: Cuant ( $\uparrow$  ETIQUETA<sub>q</sub>) =  $l_i$   
 d. niño: N ( $\uparrow$  REFERENTE) =  $x_i$   
 ( $\uparrow$  ETIQUETA) =  $l_k$   
 e. corre: V ( $\uparrow$  REFERENTE) =  $e_i$   
 ( $\uparrow$  ETIQUETA<sub>or</sub>) =  $l_1$

De este modo, una oración como «El niño corre» proyectará una estructura-f con los siguientes nuevos atributos:

$$(2) \quad f: \left[ \begin{array}{l} \text{SUJETO} \left[ \begin{array}{l} \text{PREDICADO "niño"} \\ \text{REFERENTE } x_1 \\ \text{ETIQUETA } l_1 \end{array} \right] \\ \text{PREDICADO "correr } \langle f \text{ SUJ} \rangle" \\ \text{REFERENTE } e_1 \\ \text{ETIQUETA}_{or} l_2 \end{array} \right]$$

Como se puede observar, tanto los pronombres como los sustantivos presentan atributos sobre el referente y la etiqueta (i.e., REFERENTE =  $x_1$ , o ETIQUETA =  $l_x$ ). Aunque sea redundante, esto tiene que ser de este modo, ya que no siempre los pronombres exigen sintagmas nominales (v.gr.: «Este corre», «Él corre»...) y no siempre los sintagmas nominales se presentan como argumentos del SPRON (v.gr.: «Pocos niños corren», «Ladran perros»...). En cualquier caso, la condición de unicidad de la LFG asegura que, para una misma estructura-f, el valor para REFERENTE y para ETIQUETA haya de ser único; por lo tanto, los valores que aporten el pronombre y el nombre no pueden ser diferentes.

A nuestra estructura-f todavía le falta un atributo-valor para que sea operativa para la obtención de una ULF. Dado que cada oración es un enunciado (aunque no todo enunciado sea una oración), la estructura-f máxima, esto es, aquella que no esté subordinada a ninguna otra estructura-f, ha de presentar la etiqueta  $\pi_i$ , que representa la etiqueta que nos permite conocer a qué SDRS (representación semántica del

enunciado) pertenece cada ULF. Por ello, se propone la siguiente regla:

- (3) Sea  $\langle \text{ETIQUETA}_{m\acute{a}x} \pi_i \rangle$  un  $\langle \text{atributo valor} \rangle$ . Añádase  $\langle \text{ETIQUETA}_{m\acute{a}x} \pi_i \rangle$  para toda estructura-f máxima  $f$  (i.e.,  $\langle \text{ETIQUETA}_{m\acute{a}x} \pi_i \rangle \in f$ ), definida como aquella estructura-f que no está incluida en otra estructura-f.

Así pues, por la regla de (3), nuestra estructura-f de (2) quedaría de la siguiente manera:

$$(4) \quad f: \left[ \begin{array}{l} \text{SUJETO} \left[ \begin{array}{l} \text{PREDICADO "niño"} \\ \text{REFERENTE } x_1 \\ \text{ETIQUETA } l_1 \end{array} \right] \\ \text{PREDICADO "correr } \langle f \text{ SUJ} \rangle" \\ \text{REFERENTE } e_1 \\ \text{ETIQUETA}_{or} l_2 \\ \text{ETIQUETA}_{m\acute{a}x} \pi_1 \end{array} \right]$$

Con estas adiciones de información, se pueden obtener ULFs a partir de la estructura-f. En primer lugar, es necesario que las palabras del léxico se enriquezcan con información conveniente para la traducción al lenguaje de la lógica infradeterminada. Para ello, se proponen las siguientes reglas de traducción:

### (5) Reglas de correspondencia para una lógica infradeterminada desde la LFG

- a. Sea  $\alpha$  un N. La traducción de  $\alpha$  será igual a:  
 $R_{[\uparrow \text{ PRED}]} (l_i, [\uparrow \text{ ETIQUETA}], [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \wedge R_{[\uparrow \text{ REFERENTE}]} (l_i, [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \Rightarrow \uparrow_\sigma$   
 b. Sea  $\beta$  un V cuyo predicado es “ $\beta$ ”  $\langle (\uparrow \text{ FANO}) \rangle$ . La traducción de  $\beta$  será igual a:  
 $R_{[\uparrow \text{ PRED}]} (l_i, l_m, [\uparrow \text{ ETIQUETA}_{or}], [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \wedge R_{[\uparrow \text{ REFERENTE}]} (l_i, [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \wedge R_{[\uparrow \text{ FANO REFERENTE}]} (l_m, [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \Rightarrow \uparrow_\sigma$   
 c. Sea  $\gamma$  un V cuyo predicado es “ $\gamma$ ”  $\langle (\uparrow \text{ FANO}_1), (\uparrow \text{ FANO}_2) \rangle$ . La traducción de  $\gamma$  será igual a:  
 $R_{[\uparrow \text{ PRED}]} (l_i, l_m, l_n, [\uparrow \text{ ETIQUETA}_{or}], [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \wedge R_{[\uparrow \text{ REFERENTE}]} (l_i, [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \wedge R_{[\uparrow \text{ FANO REFERENTE}]} (l_m, [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \wedge R_{[\uparrow \text{ FANO}_2 \text{ REFERENTE}]} (l_n, [(\uparrow x) \text{ ETIQUETA}_{m\acute{a}x}]) \Rightarrow \uparrow_\sigma$

d. Sea  $\delta$  un V cuyo predicado es “ $\delta$ ’  $\langle (\uparrow \text{FANO}_1), (\uparrow \text{FANO}_2), (\uparrow \text{FANO}_3) \rangle$ ”. La traducción de  $\delta$  será igual a:

$$R_{[\uparrow \text{PRED}]} (I_1, I_m, I_n, I_o, [(\uparrow \text{ETIQUETA}_{\text{or}}), [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]]) \wedge R_{[\uparrow \text{REFERENTE}]} (I_1, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge R_{[\uparrow \text{FANO REFERENTE}]} (I_m, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge R_{[\uparrow \text{FANO2 REFERENTE}]} (I_n, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge R_{[\uparrow \text{FANO3 REFERENTE}]} (I_o, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_{\sigma}$$

e. Sea  $\varepsilon$  un CUANT. La traducción de  $\varepsilon$  será igual a:

$$R_{\langle \uparrow \text{CUANT} \rangle} (I_m, [\uparrow \text{ETIQUETA}], I_n, [\uparrow \text{ETIQUETA}_q], [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge R_{\langle \uparrow \text{REFERENTE} \rangle} (I_m, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge \text{tiene-ámbito} ([\uparrow \text{ETIQUETA}_q], [(\uparrow \text{x})\text{ETIQUETA}_{\text{or}}], [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_{\sigma}$$

f. Sea  $\zeta$  un PRON[PREDI: PRO, DEF: +] cuyo PREDICADO en la estructura-f sea igual a PRO. La traducción de  $\zeta$  será igual a:

$$\exists Y [R_{\langle \uparrow \text{PRO} \rangle} (I_{x(i)}, I_y, I_o, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge Y (I_y, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}]) \wedge R_{[\uparrow \text{REFERENTE}]} (I_{x(i)}, [(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}])] \Rightarrow \uparrow_{\sigma}$$

De las RC presentadas, merecen realizarse algunos comentarios. En primer lugar, se ha de señalar que el símbolo FANO que aparece en estas reglas es una variable para las funciones nominales, esto es, SUJ, OBJ, OBJ $_{\theta}$  y OBL $_{\theta}$ . En segundo lugar, la expresión  $(\uparrow \text{x})\text{ETIQUETA}_{\text{máx}}$  designa una ecuación de incertidumbre funcional externa, según se describe en Bresnan (2001: 64-65) o en Tordera Yllescas (2008: 170-174). La expresión  $(\uparrow \text{x})$  designa una estructura-f si se encuentra un valor para  $x$ :  $x$  puede equivaler a un símbolo (SUJ), a una cadena de símbolos (COMP SUJ) o a nada. En tercer lugar, estas traducciones introducen una novedad frente a la formulación clásica de Bresnan (2001). En estas traducciones, se presentan expresiones tales como  $[\uparrow \text{ETIQUETA}]$ . Este tipo de elementos designan siempre un símbolo simple o una forma semántica (nunca una estructura-f). Por ejemplo, la expresión  $[\uparrow \text{PREDICADO}]$  de acuerdo con la estructura-f de (4) designará la forma semántica “correr  $\langle (f \text{ SUJ}) \rangle$ ”, mientras que  $[\uparrow \text{SUJ REFERENTE}]$  designará a  $x_i$ . Formalmente, este tipo de expresión ha de formularse del siguiente modo:

- (6) Dada una estructura-f  $f$  y una cadena de símbolos no vacía (al menos ha de existir un símbolo),  $\alpha$ , la expresión  $[f \alpha]$  designará el valor  $v$  que proyecta la estructura-f al tomar a  $\alpha$  como argumento.

Un último comentario va referido a la RC de (5f). Nosotros creemos que el pronombre (determinante) siempre introduce un referente en la estructura-f. Dicho referente será idéntico al que introduzca el sustantivo, si es que el pronombre exige algún sintagma nominal. En caso contrario, la estructura-f no estaría bien formada, dado que habría dos valores para el atributo REFERENTE. Y es el hecho de que sea definido o no (y no todo pronombre es definido) el que le confiere anaforicidad al pronombre. Por ello, se ha de distinguir en la traducción el carácter definido del pronombre. En concreto, creemos que todo pronombre (definido o indefinido) se traduciría simplemente como  $x$ , mientras que el carácter definido le vendrá dado por la expresión  $x = y$ , en el que  $y$  es el referente que se ha de buscar anafóricamente. Este tipo de distinción nos permitirá traducir una oración como «Un hombre ha venido. El hombre ha pedido una cerveza» en la lógica discursiva propia de la SDRT como (7a) y no como (7b) (semánticamente, igualmente verdadera pero no así textualmente):

- (7) a. hombre’ (x)  $\wedge$  ha-venido’ (x)  $\wedge$  hombre’ (y)  $\wedge$   $x = y$   $\wedge$  ha-pedido-cerveza’ (y)  
 b. hombre’ (x)  $\wedge$  ha-venido’ (x)  $\wedge$  hombre’ (x)  $\wedge$  ha-pedido-cerveza’ (x)

Al ser el artículo definido un elemento fórico, este ha de remitir a un elemento contextual. Y de esto da cuenta en (7a), la expresión “hombre (y)  $\wedge$   $x = y$ ”. En cambio, en la traducción de (7b), no se registra el carácter fórico del artículo de ninguna manera.

Las traducciones léxicas de (5), que han de constar en el lexicón, se insertan en la estructura-c; y en la estructura-f cada variable funcional ha de ser sustituida por una variable concreta de acuerdo con la estructura-f que proyecte la estructura-c, de tal modo que al final no quede ninguna función, sino únicamente símbolos (simples) y formas semánticas en las traducciones. Una vez hecho este proceso, las traducciones léxicas se pueden proyectar a la estructura- $\sigma$  (estructura semántica), como indica la expresión “ $\Rightarrow \uparrow_{\sigma}$ ”.

Con el fin de ilustrar cómo funcionaría este proceso, se ejemplificará el proceso de traducción y proyección a la estructura- $\sigma$  a

partir de tres oraciones: «Pocos niños corren», «El niño corre» y «Todos los niños lo vieron».

La ejemplificación comenzará por las dos primeras oraciones. En los ejemplos de (8abc) y (9abc), se recogen las fórmulas infradeterminadas en las que se han sustituido las metavariables (propias de la LFG) de las traducciones léxicas ( $\uparrow$ ), por las variables funcionales ( $f_1, f_2 \dots f_n$ ), de acuerdo con sus estructuras-c (vid. [10] y [11]).

$$(8) \text{ a. } R_{f_3 \text{ CUANT}} \langle I_m, [f_3 \text{ ETIQUETA}], I_n, [f_3 \text{ ETIQUETA}_q], [(f_3 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{f_3 \text{ REFERENTE}} \langle I_m, [(f_3 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge \text{tiene-}\acute{\text{a}}\text{mbito} \langle [f_3 \text{ ETIQUETA}_q], [(f_3 x) \text{ ETIQUETA}_{or}], [(f_3 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \Rightarrow \uparrow_\sigma$$

$$\text{b. } R_{[f_4 \text{ PRED}]} \langle I_i, [f_4 \text{ ETIQUETA}], [(f_4 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_4 \text{ REFERENTE}]} \langle I_i, [(f_4 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \Rightarrow \uparrow_\sigma$$

$$\text{c. } R_{[f_7 \text{ PRED}]} \langle I_1, I_m, [f_7 \text{ ETIQUETA}_{or}], [(f_7 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_7 \text{ REFERENTE}]} \langle I_1, [(f_7 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_7 \text{ FANO REFERENTE}]} \langle I_m, [(f_7 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \Rightarrow \uparrow_\sigma$$

$$(9) \text{ a. } \exists Y [R_- \langle [f_3 \text{ ETIQUETA}], I_y, I_o, [(f_3 x) \text{ E-TIQUETA}_{m\acute{a}x}] \rangle \wedge Y \langle I_y, [(f_3 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_3 \text{ REFERENTE}]} \langle [f_3 \text{ ETIQUETA}], [(f_3 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle] \Rightarrow \uparrow_\sigma$$

$$\text{b. } R_{[f_5 \text{ PRED}]} \langle I_i, [f_5 \text{ ETIQUETA}], [(f_5 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_5 \text{ REFERENTE}]} \langle I_i, [(f_5 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \Rightarrow \uparrow_\sigma$$

$$\text{c. } R_{[f_8 \text{ PRED}]} \langle I_1, I_m, [f_8 \text{ ETIQUETA}_{or}], [(f_8 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_8 \text{ REFERENTE}]} \langle I_1, [(f_8 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \wedge R_{[f_8 \text{ FANO REFERENTE}]} \langle I_m, [(f_8 x) \text{ ETIQUETA}_{m\acute{a}x}] \rangle \Rightarrow \uparrow_\sigma$$

Y a partir de las estructuras-f de (10) y (11), que se derivarían de las estructuras-c de [10] y [11] respectivamente (vid. la página siguiente), se pueden realizar los cambios pertinentes, con los que se obtienen las fórmulas infradeterminadas bien formadas de {10abc} y {11abc}, aptas para ser proyectadas a la estructura semántica, que corresponde a la lógica infradeterminada de la SDRT.

(10)

$$f_1, f_5, f_6, f_7: \left( \begin{array}{l} \text{CUANT poco} \\ \text{ETIQUETA}_q = I_1 \\ \text{PRED niño} \\ \text{ETIQUETA} = I_2 \\ \text{REFER.} = x \end{array} \right) \left( \begin{array}{l} \text{ETIQUETA}_{m\acute{a}x} = \pi_1 \\ \text{PRED} = \text{correr}' \langle (f_7 \text{ SUJ}) \rangle \\ \text{ETIQUETA}_{or} = I_3 \\ \text{REFERENTE} = e \end{array} \right)$$

(11)

$$f_1, f_6, f_7, f_8: \left( \begin{array}{l} \text{DEFIN +} \\ \text{PRED niño} \\ \text{ETIQUETA} = I_2 \\ \text{REFERENTE} = x \end{array} \right) \left( \begin{array}{l} \text{ETIQUETA}_{m\acute{a}x} = \pi_1 \\ \text{PRED} = \text{correr}' \langle (f_7 \text{ SUJ}) \rangle \\ \text{ETIQUETA}_{or} = I_3 \\ \text{REFERENTE} = e \end{array} \right)$$

$$\{10\} \text{ a. } R_{\text{poco}} \langle I_m, I_2, I_n, I_1, \pi_1 \rangle \wedge R_x \langle I_m, \pi_1 \rangle \wedge \text{tiene-}\acute{\text{a}}\text{mbito} \langle I_1, I_3, \pi_1 \rangle \Rightarrow \uparrow_\sigma$$

$$\text{b. } R_{\text{niño}} \langle I_i, I_2, \pi_1 \rangle \wedge R_x \langle I_i, \pi_1 \rangle \Rightarrow \uparrow_\sigma$$

$$\text{c. } R_{\text{correr}' \langle (f_7 \text{ SUJ}) \rangle} \langle I_1, I_m, I_3, \pi_1 \rangle \wedge R_e \langle I_1, \pi_1 \rangle \wedge R_x \langle I_m, \pi_1 \rangle \Rightarrow \uparrow_\sigma$$

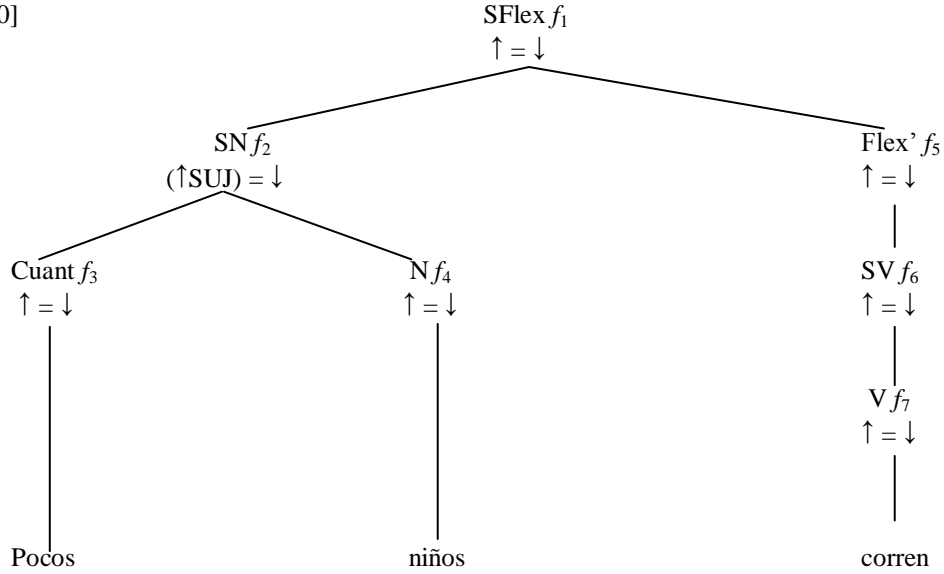
$$\{11\} \text{ a'. } \exists Y [R_- \langle I_x, I_y, I_o, \pi_1 \rangle \wedge Y \langle I_y, \pi_1 \rangle \wedge R_x \langle I_x, \pi_1 \rangle] \Rightarrow \uparrow_\sigma$$

$$\text{b'. } R_{\text{niño}} \langle I_i, I_2, \pi_1 \rangle \wedge R_x \langle I_i, \pi_1 \rangle \Rightarrow \uparrow_\sigma$$

$$\text{c'. } R_{\text{correr}' \langle (f_7 \text{ SUJ}) \rangle} \langle I_1, I_m, I_3, \pi_1 \rangle \wedge R_e \langle I_1, \pi_1 \rangle \wedge R_x \langle I_m, \pi_1 \rangle \Rightarrow \uparrow_\sigma$$

Igualmente, siguiendo estos pasos, se puede obtener la fórmula infradeterminada, ULF, de una oración algo más compleja desde el punto de vista anafórico. Así, «Todos los niños lo vieron» se procesaría en los siguientes pasos. En primer lugar, se sustituyen las metavariables funcionales ( $\uparrow$ ), cambio ya realizado en (12), por las variables funcionales ( $f_1, f_2 \dots f_n$ ) de acuerdo con la estructura-c en la que se insertan las distintas unidades léxicas (vid. [13] infra).

[10]

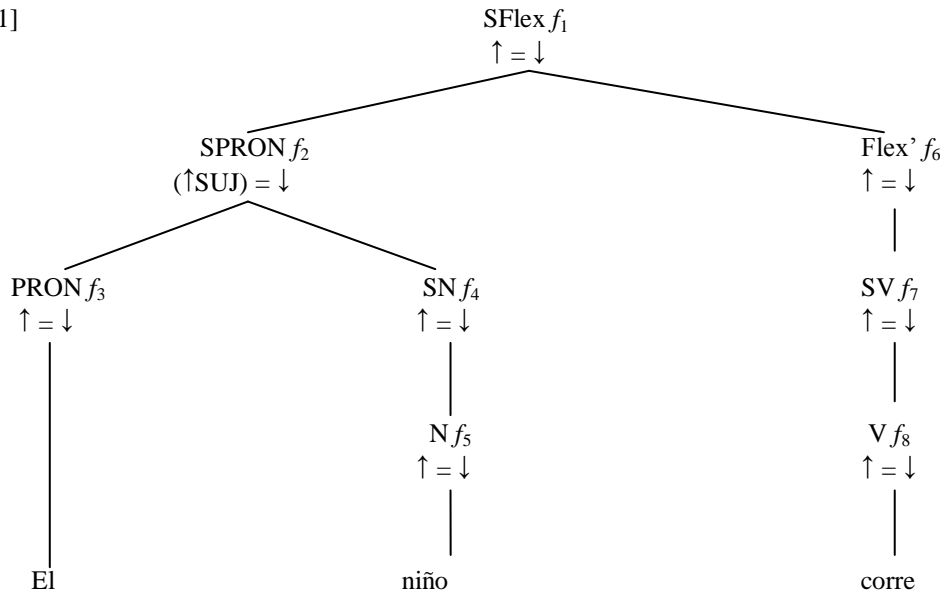


$(f_3 \text{ CUANT}) = \text{poco}$   
 $(f_3 \text{ ETIQUETA}_q) = l_1$   
 $R_{\langle f_3 \text{ CUANT} \rangle} (l_m, [f_3 \text{ ETIQUETA}], l_n, [f_3 \text{ ETIQUETA}_q], [(f_3 x) \text{ ETIQUETA}_{\text{máx}}] \wedge R_{\langle f_3 \text{ REFERENTE} \rangle} (l_m, [(f_3 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge \text{tiene-ámbito} (f_3 \text{ ETIQUETA}_q, [(f_3 x) \text{ ETIQUETA}_{\text{or}}], [(f_3 x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$

$(f_4 \text{ PRED}) = \text{niño}'$   
 $(f_4 \text{ ETIQUETA}) = l_2$   
 $(f_4 \text{ REFERENTE}) = x$   
 $R_{[f_4 \text{ PRED}]} (l_1, [f_4 \text{ ETIQUETA}], [(f_4 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_4 \text{ REFERENTE}]} (l_1, [(f_4 x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$

$(f_7 \text{ PRED}) = \text{correr}' \langle (f_7 \text{ SUJ}) \rangle$   
 $(f_7 \text{ ETIQUETA}_e) = l_3$   
 $(f_7 \text{ REFERENTE}) = e_1$   
 $R_{[f_7 \text{ PRED}]} (l_1, l_m, [f_7 \text{ ETIQUETA}_{\text{or}}], [(f_7 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_7 \text{ REFERENTE}]} (l_1, [(f_7 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_7 \text{ FANO REFERENTE}]} (l_m, [(f_7 x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$

[11]



$(f_3 \text{ DEF}) = +$   
 $(f_3 \text{ REFERENTE}) = x$   
 $(f_3 \text{ ETIQUETA}) = l_2$   
 $\exists Y [R_{=} (l_x, l_y, l_o, [(f_3 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge Y (l_y, [(f_3 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_3 \text{ REFERENTE}]} (l_x, [(f_3 x) \text{ ETIQUETA}_{\text{máx}}])] \Rightarrow \uparrow_\sigma$

$(f_5 \text{ PRED}) = \text{niño}'$   
 $(f_5 \text{ ETIQUETA}) = l_2$   
 $(f_5 \text{ REFERENTE}) = x$   
 $R_{[f_5 \text{ PRED}]} (l_1, [f_5 \text{ ETIQUETA}], [(f_5 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_5 \text{ REFERENTE}]} (l_1, [(f_5 x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$

$(f_8 \text{ PRED}) = \text{correr}' \langle (f_8 \text{ SUJ}) \rangle$   
 $(f_8 \text{ ETIQUETA}_e) = l_3$   
 $(f_8 \text{ REFERENTE}) = e$   
 $R_{[f_8 \text{ PRED}]} (l_1, l_m, [f_8 \text{ ETIQUETA}_{\text{or}}], [(f_8 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_8 \text{ REFERENTE}]} (l_1, [(f_8 x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{[f_8 \text{ FANO REFERENTE}]} (l_m, [(f_8 x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$

- (12) a.  $R_{\langle f_3 \text{ CUANT} \rangle} (l_m, [f_3 \text{ ETIQUETA}], l_n, [f_3 \text{ ETIQUETA}_q], [(f_3x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{\langle f_3 \text{ REFERENTE} \rangle} (l_m, [(f_3x) \text{ ETIQUETA}_{\text{máx}}]) \wedge \text{tiene-ámbito} ([f_3 \text{ ETIQUETA}_q], [(f_3x) \text{ ETIQUETA}_{\text{or}}], [(f_3x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$   
 b.  $\exists Y [R_{=} (l_{x1}, l_y, l_o, [(f_5x) \text{ ETIQUETA}_{\text{máx}}]) \wedge Y (l_y, [(f_5x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{\langle f_5 \text{ REFERENTE} \rangle} (l_{x1}, [(f_5x) \text{ ETIQUETA}_{\text{máx}}])] \Rightarrow \uparrow_\sigma$   
 c.  $R_{\langle f_7 \text{ PRED} \rangle} (l_i, [f_7 \text{ ETIQUETA}], [(f_7x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{\langle f_7 \text{ REFERENTE} \rangle} (l_i, [(f_7x) \text{ ETIQUETA}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$   
 d.  $\exists Y [R_{=} (l_{x2}, l_y, l_o, [(f_{10}x) \text{ ETIQUETA}_{\text{máx}}]) \wedge Y (l_y, [(f_{10}x) \text{ ETIQUETA}_{\text{máx}}]) \wedge R_{\langle f_{10} \text{ REFERENTE} \rangle} (l_{x2}, [(f_{10}x) \text{ ETIQUETA}_{\text{máx}}])] \Rightarrow \uparrow_\sigma$   
 e.  $R_{\langle f_{11} \text{ PRED} \rangle} (l_i, l_m, l_n, [\uparrow \text{ ETIQ}_{\text{or}}], [(f_{11}x) \text{ ETIQ}_{\text{máx}}]) \wedge R_{\langle f_{11} \text{ REF} \rangle} (l_i, [(f_{11}x) \text{ ETIQ}_{\text{máx}}]) \wedge R_{\langle f_{11} \text{ FANO REF} \rangle} (l_m, [(f_{11}x) \text{ ETIQ}_{\text{máx}}]) \wedge R_{\langle f_{11} \text{ FA-NO2 REF} \rangle} (l_n, [(f_{11}x) \text{ ETIQ}_{\text{máx}}]) \Rightarrow \uparrow_\sigma$

Y a partir de la estructura-f de (13), que se derivaría de la estructura-c de [13] respectivamente (vid. la página siguiente), se pueden realizar los cambios pertinentes, con los que se obtienen las fórmulas infradeterminadas bien formadas de {13abcde}, aptas para ser proyectadas a la estructura semántica, que corresponde a la lógica infradeterminada de la SDRT.

- (13)
- |                         |                                      |   |
|-------------------------|--------------------------------------|---|
| $f_1, f_8, f_9, f_{11}$ | SUJ $f_2, f_3, f_4, f_5, f_6, f_7$ : | $\left( \begin{array}{l} \text{DEF +} \\ \text{CUANT } \forall \\ \text{ETIQUETA}_q \ l_1 \\ \text{PRED } \text{'niño'} \\ \text{REFERENTE } x_1 \\ \text{ETIQUETA } l_2 \end{array} \right)$ |
|                         | ETIQUETA <sub>máx</sub> $\pi_1$      | PRED= ver'<(f <sub>11</sub> SUJ), (f <sub>11</sub> OBJ)>  |
|                         | ETIQUETA <sub>or</sub> $l_4$         | REFERENTE $e$   |
|                         | OBJ $f_{10}$ :                       | $\left( \begin{array}{l} \text{DEF +} \\ \text{PRED PRON} \\ \text{REFERENTE } x_2 \\ \text{ETIQUETA } l_3 \end{array} \right)$   |

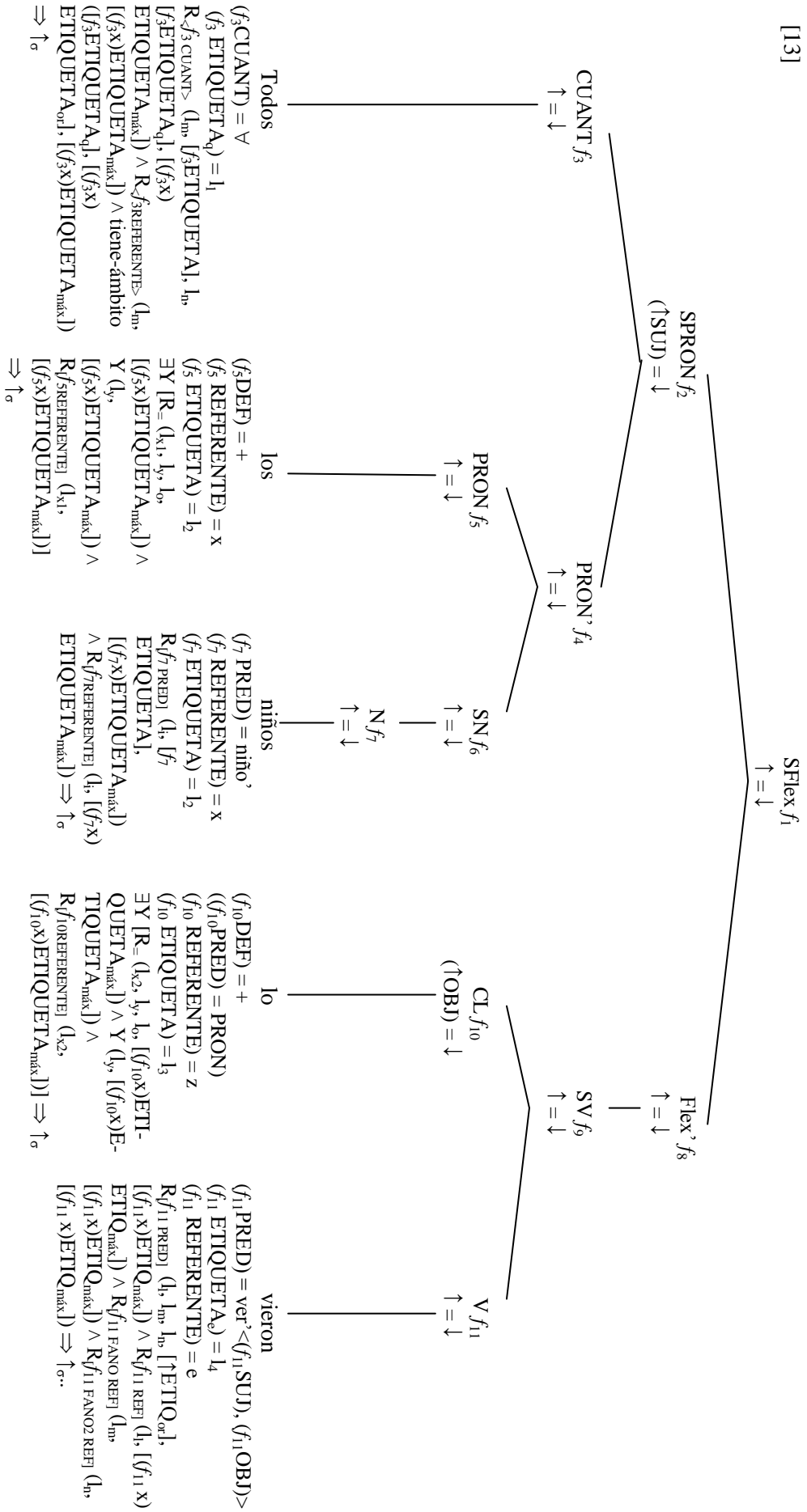
- {13} a.  $R_{\forall} (l_m, l_2, l_n, l_1, \pi_1) \wedge R_{x1} (l_m, \pi_1) \wedge \text{tiene-ámbito} (l_1, l_4, \pi_1) \Rightarrow \uparrow_\sigma$   
 b.  $\exists Y [R_{=} (l_{x1}, l_y, l_o, \pi_1) \wedge Y (l_y, \pi_1) \wedge R_{x1} (l_{x1}, \pi_1)] \Rightarrow \uparrow_\sigma$   
 c.  $R_{\text{niño}} (l_i, l_2, \pi_1) \wedge R_{x1} (l_i, \pi_1) \Rightarrow \uparrow_\sigma$   
 d'.  $\exists Y [R_{=} (l_{x2}, l_y, l_o, \pi_1) \wedge Y (l_y, \pi_1) \wedge R_{x2} (l_{x2}, \pi_1)] \Rightarrow \uparrow_\sigma$   
 e.  $R_{\text{ver}'\langle (f_{11}\text{SUJ}), (f_{11}\text{OBJ}) \rangle} (l_i, l_m, l_n, l_4, \pi_1) \wedge R_e (l_i, \pi_1) \wedge R_{x1} (l_m, \pi_1) \wedge R_{x2} (l_n, \pi_1) \Rightarrow \uparrow_\sigma$

### 3 Nuevos retos: la anáfora de sentido

Contar con unas RC que permitan obtener fórmulas lógicas de la SDRT a partir de las descripciones sintácticas de la LFG conlleva considerables ventajas. La LFG se ha revelado una gramática monoestratal explicativa para fenómenos tales como las construcciones de infinitivo, los elementos “desplazados”, las oraciones pasivas, la concordancia... Asimismo, la SDRT (teoría semántica dinámica al estilo de la *Discourse Representation Theory* de Kamp y Reyle (1993), aunque aquella propone una estructura discursiva no plana: jerarquía y orden de los enunciados textuales), ha aportado diferentes perspectivas de investigación valiosas para los nombres plurales, la presuposición, las anáforas asociativas (v.gr.: «Juan estuvo arreglando la habitación. Las paredes estaban plagadas de grietas»), la interpretación de los cuantificadores..., como se describe en Asher (1993) o Asher y Lascarides (2003 y 2007). Especialmente, ha resultado una teoría sumamente iluminativa en el tratamiento de las anáforas. Como la LFG impone restricciones sintácticas en la selección del antecedente de un pronombre (v.gr.: un elemento femenino plural no puede ser el antecedente de un pronombre masculino singular) y la SDRT ha prestado una especial atención a la determinación de los posibles antecedentes según factores semántico-pragmáticos (v.gr: determinados enunciados no pueden contener ningún antecedente posible si mantienen unas relaciones retóricas dadas con los enunciados en los que se sitúa el pronombre), la complementación de estas dos teorías puede ser harto fructífera.

Así por ejemplo, implementar una teoría sintáctica como la LFG con una teoría semántica como la SDRT puede aportar luz a la denominada anáfora de sentido, esto es, aquella anáfora que solo recupera el sentido de su antecedente pero no su referente.

[13]





(14) Juan le da **su sueldo** a su mujer y Pedro se *lo* da a la suya.

En casos como estos, el pronombre *lo* recupera el sentido de su antecedente (i.e., «sueldo»), pero no su referente, ya que el pronombre no hace referencia al sueldo de Juan sino al sueldo de Pedro.

A nuestro juicio, creemos que un primer paso para tratar este tipo de anáforas es incorporar *las relaciones retóricas* de la SDRT en la asignación de referente de la anáfora de sentido. En concreto, creemos que las anáforas de sentido se producen en aquellas relaciones retóricas que Asher y Lascarides (2003) denominan relaciones retóricas estructurales, a saber *Parelelismo* y *Contraste*. Estas relaciones que se establecen entre dos enunciados se caracterizan por no suponer una simple relación de coordinación entre los dos elementos y por que ambos enunciados entablan (prácticamente) un isomorfismo estructural. Esto es lo que ocurre con ejemplos como el de (14) en el que ambos enunciados parecen coimplicarse discursivamente y en el que el segundo enunciado repite el verbo «dar» y aparecen los mismos complementos. Creemos que este tipo de contexto sintáctico-discursivo es el que posibilita la aparición de anáforas de sentido y, a partir de teorías semánticas como la SDRT, se puede formalizar este tipo de fenómenos, tal como se recoge en Tordera Yllescas (2010).

### Conclusiones

Los logros de nuestro trabajo son humildes (pues las reglas propuestas dan cuenta solo de un fragmento del español), pero sus implicaciones futuras no son nada desdeñables.

En primer lugar, se ha de señalar que es relativamente fácil proponer nuevas RC, como las aquí presentadas, que puedan dar cuenta de estructuras sintácticas más complejas (oraciones completivas, oraciones de infinitivo...). Nuestras reglas presentan la ventaja de que no añaden ningún concepto teórico adicional respecto a la LFG o a la SDRT, por lo que dichas teorías no se ven distorsionadas.

En segundo lugar, este trabajo soluciona un problema no planteado en el trabajo de Asher y Lascarides (2003). *Grosso modo*, la SDRT es un conjunto de diferentes lógicas (Logic of Information Packaging, Glue Logic (GL)...), cada una encargada de un objetivo

(v.gr.: la GL asigna un antecedente para un pronombre, aunque este no esté establecido gramaticalmente). El hecho es que la SDRT asume que el punto de partida es la Lógica infradeterminada, lógica que recoge solo la información sintácticamente codificada. Sin embargo, los autores no abordan cómo se llega a este tipo de lógica desde una gramática. Así pues, con nuestras RC se resuelve el problema.

Finalmente, desde dos teorías formales, se trata de ofrecer los primeros pasos para resolver la anáfora de sentido: tener en cuenta el tipo de relaciones retóricas de la SDRT.

### Bibliografía

- Allen, J. 1995. *Natural Language Understanding*. Benjamin/Cummings, Redwood.
- Asher, N. 1993. *Reference to abstract objects in discourse*. Dordrecht: Kluwer
- Asher, N. y A. Lascarides. 2003. *Logics of conversation*. Cambridge University Press, Cambridge.
- Asher, N. y A. Lascarides. 2007. “Segmented Discourse Representation Theory: Dynamic Semantics with Discourse Representation”. En H. Bunt y R. Muskens, *Computing Meaning: Volumen 3*, págs 87-124. Kluwer Academic Publishers, Dordrecht/Londres/Boston.
- Bresnan, J. 2001. *Lexical-Functional Syntax*. Basil Blackwell, Oxford.
- Dalrymple, M. 2001. *Syntax and Semantics. Lexical Functional Grammar*. Academic Press, Londres y San Diego.
- Halvorsen, P.-K. 1983. Semantics por Lexical-Functional Grammar. *Linguistic Inquiry*. 14(4): 567-615.
- Kamp, H. y U. Reyle 1993. *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.
- Martí Antonín, M. A. e I. Castellón Masalles. 2000: *Lingüística computacional*. Edicions Universitat de Barcelona, Barcelona.
- Moreno Sandoval, A. 1998. *Lingüística computacional*. Síntesis, Madrid.
- Tordera Yllescas, J. C. 2008: *Introducción a la Gramática Léxico-Funcional*. Publicacions Universitat de València, Valencia.
- Tordera Yllescas, J. C. 2010: *Lingüística computacional y anáfora*. Publicacions Universitat de València, Valencia. Disponible en CD-ROM.