

ARTICULOS



La semántica del adjetivo y su representación mediante estructuras de rasgos.

Salvador Climent
climent@goliat.upc.es
Departament de Llenguatges i Sistemes Informàtics.
Facultat d'Informàtica.
Universitat Politècnica de Catalunya.

0. Abstract.

A general framework for representing the semantics of the adjective in a unification-based Lexical Knowledge Base has been set.

Main contributions are the design itself of a framework expressive enough to deal with many aspects of adjectival meaning, namely surface order, selectional restrictions, lexical semantics and formal semantics. Within this framework formal definitions are provided for (a) an adjectival lexical structure closely related to noun's Qualia Structure (Pustejovsky'91); and (b) logical types of adjectives attending to the kind of predication they are involved in.

0. Compendio.

Por el trabajo que se describe se ha formulado un marco general de representación semántica de los adjetivos en una base de conocimiento léxico basada en estructuras de rasgos en la que operan la herencia y la unificación.

Como principales aportaciones cabe destacar el propio diseño de un marco representativo suficientemente expresivo como para abarcar múltiples aspectos del significado del adjetivo, tales como orden superficial, restricciones selectivas, semántica léxica y semántica formal. En este marco, se han definido (a) una estructura léxica del adjetivo íntimamente relacionada con la Estructura de Qualia (Pustejovsky'91) de los nombres; y (b) tipos lógicos de adjetivo en base al tipo de predicación que establecen con el nombre.

1. Introducción.

El trabajo aquí descrito ha tenido como objetivo diseñar un marco global para la representación de adjetivos en la base de conocimiento léxico (BCL) desarrollada en el marco del proyecto *Acquilex-II*¹, para lo cual se ha procedido al estudio de las propiedades semánticas del adjetivo y al establecimiento de mecanismos para su representación formal.

La BCL representa el conocimiento en base a estructuras de rasgos tipadas -en el sentido de Carpenter'92- aumentadas con herencia por defecto y reglas léxicas (generadoras de estructuras léxicas derivadas). Su lenguaje de representación léxica está basado en la unificación, estando permitida la identidad de estructuras -representada mediante coindización de valores-.

El signo léxico consta de cuatro rasgos mayores: ortografía, sintaxis, semántica formal (SF) y semántica léxica (SL). La categoría sintáctica está descrita en términos de gramáticas categoriales (categorías básicas de tipo atómico o categorías complejas representadas mediante dos categorías relacionadas por un operador de dirección). La semántica formal se describe mediante rasgos correspondientes a predicado y argumentos, más un índice ontológico del argumento. La semántica léxica refleja la Teoría de Qualia (Pustejovsky'91) por la que se enriquece el signo léxico nominal mediante conocimiento general relativo a la entidad denotada (p.e. forma, color, etc.).

El signo adjetivo (*fig. 1*)² está representado categorialmente como N-dirección-N: la categoría que debe combinar con un signo nombre dando como resultado una categoría asimismo nominal. En SF la fórmula lógica que se le

¹ BRA 7315, programa ESPRIT de la CE. Participan las universidades de Amsterdam, Cambridge y Politécnica de Catalunya, los grupos editoriales Bibliograf, Van Dale Lexicografie y Cambridge University Press, y el Istituto di Linguistica Computazionale de Pisa. Su objetivo principal es el desarrollo de herramientas para la creación de una base de conocimiento léxico multilingue para PLN.

² En ésta y las subsiguientes figuras que aparecen en el texto debe tenerse en cuenta que las estructuras que se presentan no son las utilizadas en la BCL de *Acquilex-II* sino una simplificación de las mismas y que la nomenclatura utilizada para atributos y valores es una adaptación de la original, todo ello en aras de su inteligibilidad; en especial, el decisivo factor de la compartición de valores entre rasgos, se expresa mediante la identidad de variables.

atribuye por defecto es $[\text{adj}(X) \wedge \text{nombre}(X)]^3$; posteriores representaciones se han establecido atendiendo a las características semánticas de diversos tipos de adjetivos. La SL del adjetivo se ha desarrollado atendiendo a los efectos de éste en la SL del nombre con qué combinan.

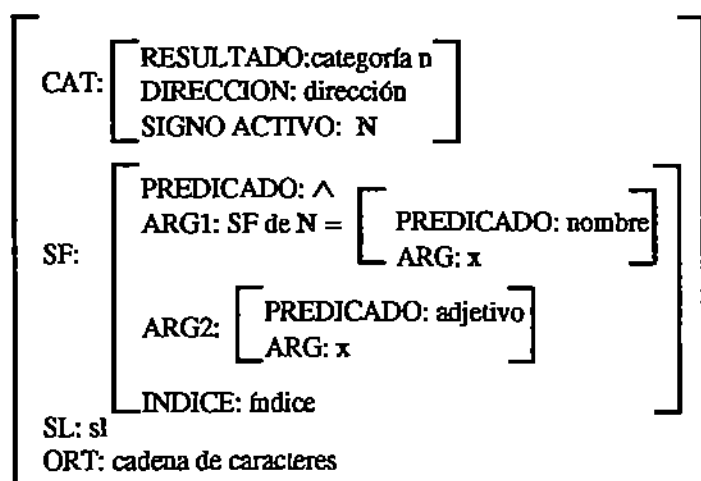


fig. 1

2. Aspectos semánticos considerados. Su representación.

2.1. Realización sintáctica.

El tipo semántico de las piezas léxicas suele tener su reflejo en su comportamiento sintáctico superficial; así, en muchos casos, diferentes realizaciones semánticas de un mismo adjetivo -correspondientes a diferentes sentidos de su entrada de diccionario⁴- tienen su correlato en su anteposición o posposición al nombre que modifican (p.e. "pobre hombre" vs. "hombre pobre"). Dicho fenómeno se representa atribuyendo al correspondiente sentido del adjetivo los valores apropiados de CAT:DIRECCION (vid. fig. 1) (p.e. para "pobre_X" ["=infeliz"], valor *forward*; y para "pobre_Y" ["=sin dinero"], *backward*).

2.2. Restricciones de selección.

³ P.e. $[\text{rojo}(X) \wedge \text{coche}(X)]$, refiriendo a las entidades X que son coche y son rojas.

⁴ En la BCL cada estructura léxica corresponde a un sentido de la palabra en cuestión. La distinción de sentidos es la del diccionario VOX'87, cuya versión magnética es la utilizada en Acquilex-II como fuente para adquisición semiautomática de conocimiento léxico del español.

Una de las características semánticas más notables del adjetivo es que, como predicador, selecciona el tipo semántico del nombre con el que combina. Esta propiedad se representa declarando para cada adjetivo el tipo del nombre en términos de la ontología de entidades existente en la BCL. Ello se realiza al declarar el índice de la SF del nombre (y su correlato en la SL, también estructurada jerárquicamente en tipos). Así, p.e., "pobre_1" [= "sin dinero"] es atribuible a nombres denotando seres humanos, mientras "pobre_2" [= "falto de algo para ser completo"] carece de dicha restricción; en el primer caso el índice denotativo del tipo de nombres seleccionados tendría el valor *humano*, mientras en el segundo, el valor más general *entidad* (vid. fig. 2).

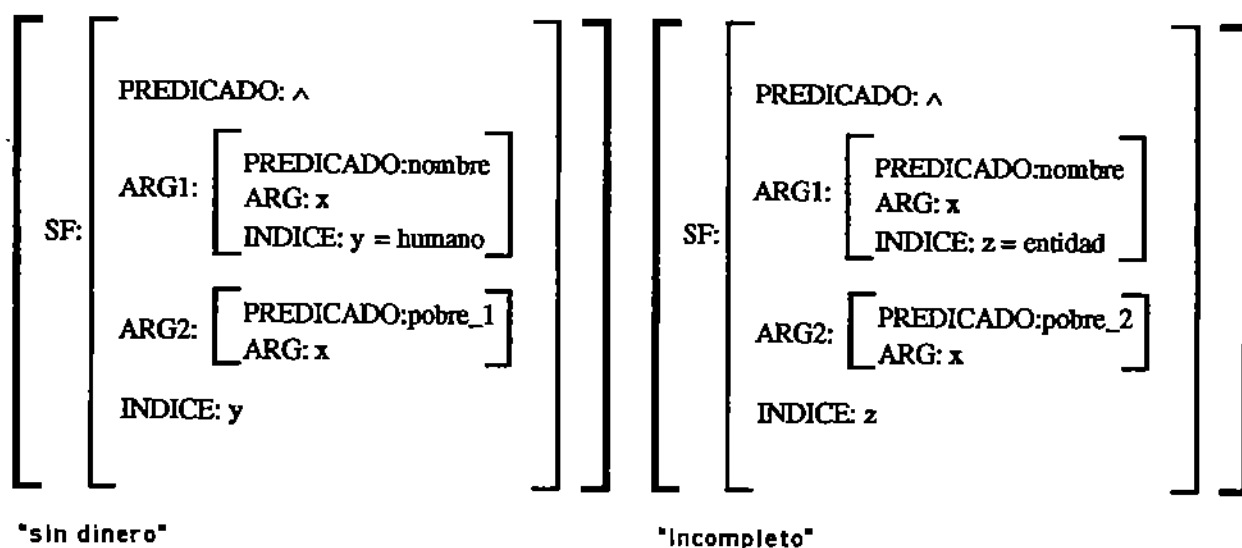


fig. 2

La expresividad del formalismo permite asimismo representar colocaciones restrictivas, p.e., el adjetivo "claudia" sólo atribuible a "ciruela", en cuyo caso el valor SF:ARG1:PREDICADO no se dejaría subespecificado sino que se instanciaría por "ciruela_1".

Es preciso tener en cuenta que los adjetivos que forman parte de lexía no son representables por estos medios ya que en ellas el significado no se

construye de forma composicional: una lexía constituirá una entrada unificada en el lexicón.

2.3. Semántica Léxica: Teoría de Qualia y función valorativa.

Pustejovsky'91 desarrolla la llamada Teoría de Qualia por la que atribuye a las entradas léxicas nominales una estructura que refleja múltiples aspectos del conocimiento general (no gramatical) atribuible a sus referentes, información potencialmente compartida por los hablantes y que es susceptible de jugar algún papel en determinados fenómenos lingüísticos.

El rasgo correspondiente a la SL en los signos léxicos nominales en la BCL es la especificación de su Estructura de Qualia (EQ), la cual depende de su clasificación ontológica. Ejemplos de atributos de la EQ son: área de origen, estado físico, color, transparencia, olor, sabor, textura, tamaño, similitud, constituyentes, volumen, peso, forma, edad o función típica.

Partiendo de la aceptación del marco de la EQ, notamos que un subconjunto de los adjetivos ejerce la evidente función de atribuir valores a rasgos de la EQ⁵ de los nombres con los que combinan (p.e. la entrada léxica de "coche" contiene el rasgo "color" de su EQ con valor indeterminado; al formarse el grupo "coche rojo", dicho rasgo se instanciará con el valor "rojo"). En consecuencia, postulamos una estructura SL del adjetivo en la que un primer rasgo informa de qué atributo de la EQ nominal es el seleccionado y con qué valor se instancia el mismo (vid. *fig 3*).

⁵ Por seguir la lista anterior de ejemplos, piénsese en "español" (área de origen), "gelatinoso" (textura), "verde" (color), "translúcido" (transparencia), "nauseabundo" (olor), etc.

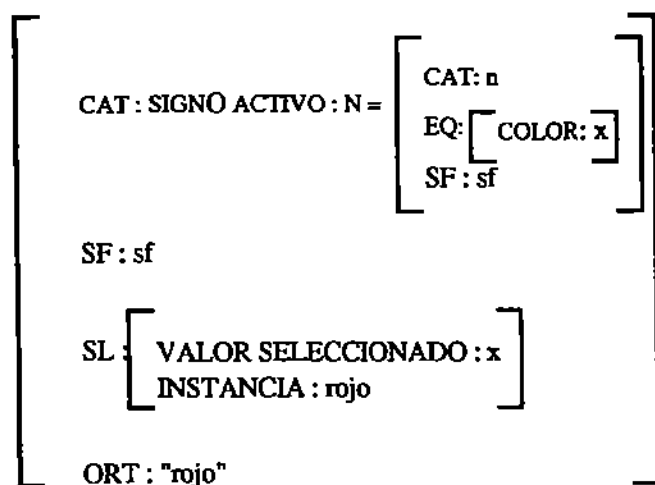


fig. 3

Por otra parte, a partir de Mel'çuk'84, advertimos que ciertos adjetivos cumplen una función de valoración del referente nominal (piénsese en ejemplos típicos como "bueno" o "pésimo"). Dicho autor ha desarrollado un riquísimo dispositivo teórico para representar ésta y otras muchas funciones del conocimiento léxico: las "funciones léxicas", que expresan conceptos tales como intensificación, valoraciones positiva o negativa, contradicción, etc. Decidimos incorporar información de ese tipo a la representación semántico-léxica del adjetivo, dotando a la SL del adjetivo de un rasgo indicador del tipo de valoración que éste aporta al nombre, representándose dicha valoración en términos de funciones léxicas.

Considerando simultáneamente la función valorativa y la Teoría de Qualia, puede notarse que ciertos adjetivos, muy cercanos en su significado, difieren entre sí debido a una diferente connotación valorativa. Pongamos por ejemplo "anciano", "viejo" y "decrépito": situándonos en el marco de la Estructura de Qualia, resultan semánticamente cercanos por causa de que los tres instancian un mismo valor de la EQ del nombre, en este caso "edad"; y difieren precisamente a causa de que efectúan una valoración distinta de la atribución de dicho rasgo (digamos, positiva para "anciano", neutra o algo peyorativa para "viejo" y claramente peyorativa para "decrépito").

En nuestra representación, conseguimos capturar dicho tipo de distinciones permitiendo que el rasgo valorativo que aporta el adjetivo al significado del grupo nominal pueda atribuirse, no tan sólo al referente del nombre como un todo, sino también en su caso únicamente a un rasgo de su EQ (vid. *fig 4*).

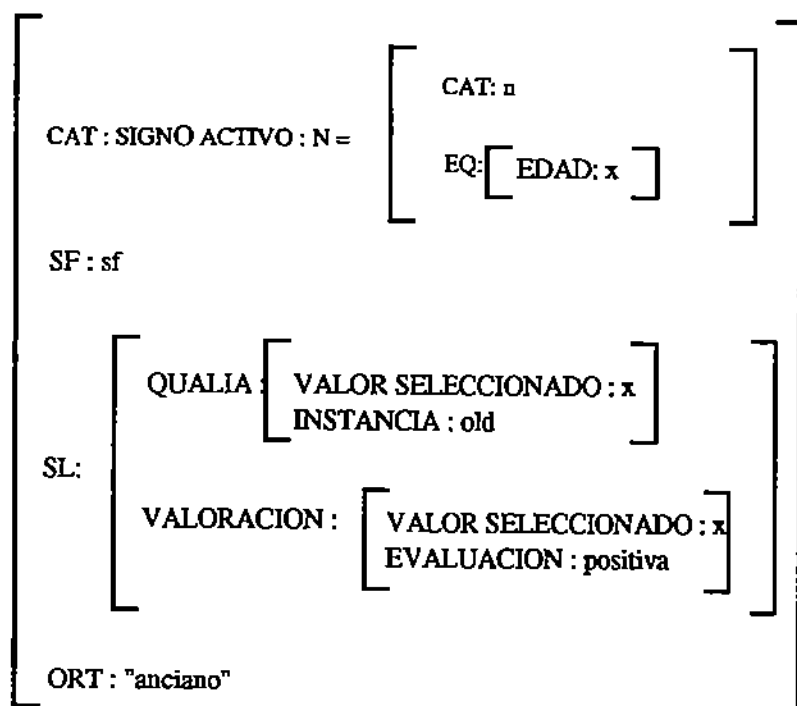


fig. 4

2.4. Semántica léxico-sintagmática.

Las descripciones propuestas en 2.3 atañen de forma directa a la estructura interna de la pieza léxica en cuestión. Aquí contemplaremos aspectos del significado que, aunque por supuesto siguen siendo explicativas del significado léxico adjetival, inciden además en las relaciones semánticas que se establecen entre los componentes del grupo nombre + adjetivo.

2.4.1. Derivación morfológica y descripción relacional.

Tradicionalmente se califica como relacionales a los adjetivos derivados de palabras de otras categorías en tanto que establecen algún tipo de relación entre el nombre con qué combinan y otro dominio externo a éste; p.e., en "crítica musical" o "alfombra voladora" el adjetivo aporta el significado de que

los referentes de "crítica" y "alfombra" se hallan relacionados de algún modo, respectivamente, con el dominio de la música y con la acción de volar.

Parece pues adecuado establecer una representación de este tipo de adjetivos que sustituya la relación lógica intersectiva básica [$\text{nombre}(X) \wedge \text{adj}(X)$] por otro tipo de representación formal que incluya el concepto de relación: [$\text{nombre}(X) \wedge \text{entidad_relacionada}(Y) \wedge \text{relación}(X,Y)$], siendo la entidad relacionada el nombre o verbo del que deriva el adjetivo en cuestión⁶ y la relación una entidad conceptual a determinar, que será de tipo diferente dependiendo de si el adjetivo es deverbal o denominal.

2.4.1.1. Adjetivos deverbales.

El tipo de relación semántica que establece un conjunto notable de adjetivos deverbales, típicamente los generados por sufijación de "-dor" o "-do", es de naturaleza temática, en el sentido de que se está indicando el que el nombre modificado ejerce un rol temático en el evento denotado por el verbo del cual deriva el adjetivo. Así en "poste indicador", entenderemos que "poste" ejerce un papel de tipo agentivo en la acción de "indicar".

Por consiguiente la relación formal utilizada para la composición de la forma lógica de estos adjetivos pertenecerá al conjunto de las relaciones temáticas verbales. Así, el rasgo SF en "indicador_1" representará la forma lógica [$\text{nombre}(X) \wedge \text{indicar_1}(E) \wedge \text{p-agt}(X,E)$]⁷ (vid. fig 5):

⁶ Para ser más precisos, el indicador del sentido apropiado de la entrada léxica de dicha palabra, ya que la BCL que se elabora consiste en un conjunto de sentidos y de relaciones entre éstos. P.e., en "alfombra voladora" la entidad relacionada será "volar_1" (moverse por el aire etc.) y no "volar_14" (hacer estallar una cosa).

⁷ En la estrategia adoptada en Acquilex-II para la representación de los roles temáticos, éstos pueden ser bien p-agt (proto-agente) o bien p-pat (proto-paciente), definidos respectivamente como los argumentos más y menos agentivo de un verbo. Ulteriores subtipificaciones proveen una especificación más detallada de su contenido semántico.

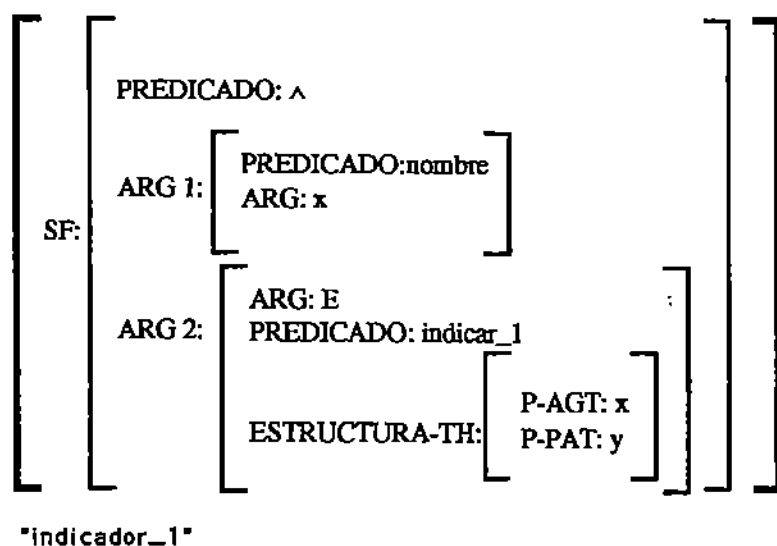


fig. 5

2.4.1.2. Adjetivos Denominales.

Las relaciones semánticas establecidas por este tipo de adjetivos son complejas y variadas, habiéndose propuesto diversas clasificaciones, entre las que cabe destacar la de Aarts & Calbert'79. Ejemplos típicos son las relaciones de causación ("benéfico", parafraseable por "que causa el bien") o experimentación ("envidioso", "que experimenta envidia").

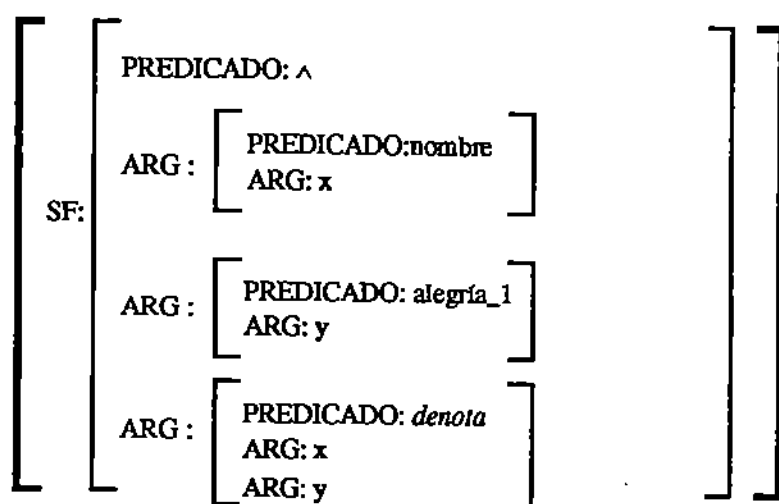
La aproximación que aquí planteamos es pre-teórica, no postulándose ningún conjunto estructurado de relaciones ni reclamando para las mismas estatus alguno de indescomponibilidad o realidad psicológica. La pauta seguida para su elección ha sido -toda vez que uno de los propósitos básicos de Aquilex-II es la recuperación de la información de diccionarios- la detección de patrones de definición existentes en la fuente utilizada (VOX'87) y el establecimiento a partir de éstos de las relaciones semánticas.

De ese modo, los sentidos cuya definición contiene los patrones "que manifiesta" y "que denota" dan lugar a la relación "denota"; aquéllos definidos mediante "que ocasiona", "que causa" y "que produce", la relación "causa"; o el muy abundante "relativo a", la relación "relativo_a".

Inciendiando en el proceso de extracción de la información de la definición del diccionario, hemos extendido el criterio de selección del nombre

denotativo del dominio con el que se establece la relación semántica más allá del que da origen por derivación al adjetivo, pasando a aceptarse aquél que proporciona la definición lexicográfica, el cual, aunque en la mayoría de casos suele ser dicha primitiva (p.e. "notarial_1: relativo al notario"), en algunos es otro nombre relacionado semántica pero no morfológicamente (p.e. "radiante_3: que denota alegría").

De este modo, el rasgo SF en "radiante_3" (p.e. en "sonrisa radiante") representará la forma lógica [nombre(X) $\wedge$ alegría_1(Y) $\wedge$ denota(X,Y)] (vid. fig 6.):



"radiante_3"

fig. 6

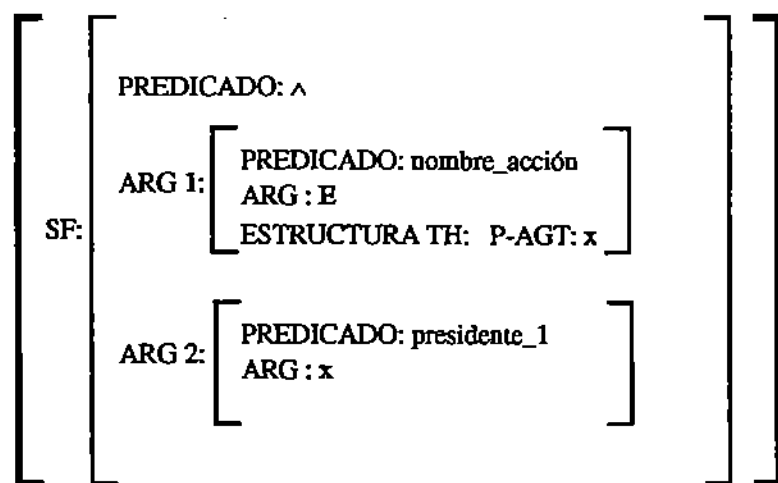
2.4.2. El adjetivo como depositario de papel temático: los adjetivos temáticos.

Bosque'93 revela la relación que se establece entre cierto tipo de adjetivos relacionales y el nombre, cuando éste es un nombre de acción. Estos ostentan, al igual que los verbos, estructura argumental. Así, el nombre "viaje" implica la existencia de un agente o experimentador de la acción de viajar.

En grupos nominales de ese tipo, como p.e. "viaje presidencial", el adjetivo resulta ser el depositario de uno de los roles temáticos atribuidos. En el caso propuesto, "presidencial" como adjetivo relacional hace referencia al

sustantivo "presidente", y es precisamente "presidente" el agente o experimentador de la acción denotada por "viaje".

En consecuencia, dicho tipo de adjetivos comportarán una forma lógica [nombre_de_acción (E) $\wedge$ entidad_relacionada (X) $\wedge$ rol_temático (X,E)]; p.e., para "presidencial" (en su acepción temática): [nombre_de_acción (E) $\wedge$ presidente_1 (X) $\wedge$ p-agt (X,E)] (vid. *fig 7*):



"presidencial" (temático)

fig. 7

3. Tipología.

En el actual estado de la BCL se han declarado cuatro tipos mayores de adjetivos, habiéndose realizado dicha clasificación en atención a la diferente estructura formal de los mismos. Los representaciones de los diferentes fenómenos abordados, y relatados en 2, tienen cabida dentro de un mismo tipo de marco de estructuras de rasgos, con excepción de los tres tipos de relaciones descritas en 2.4, que tienen su reflejo en tres estructuraciones distintas del rasgo SF del signo léxico adjetivo. Estas, junto con el tipo básico de forma lógica intersectiva (reservada para adjetivos conceptualmente primitivos, entendiéndose éstos como los no relacionales) dan lugar a la tipología, que representamos en el siguiente esquema (*fig 8*):

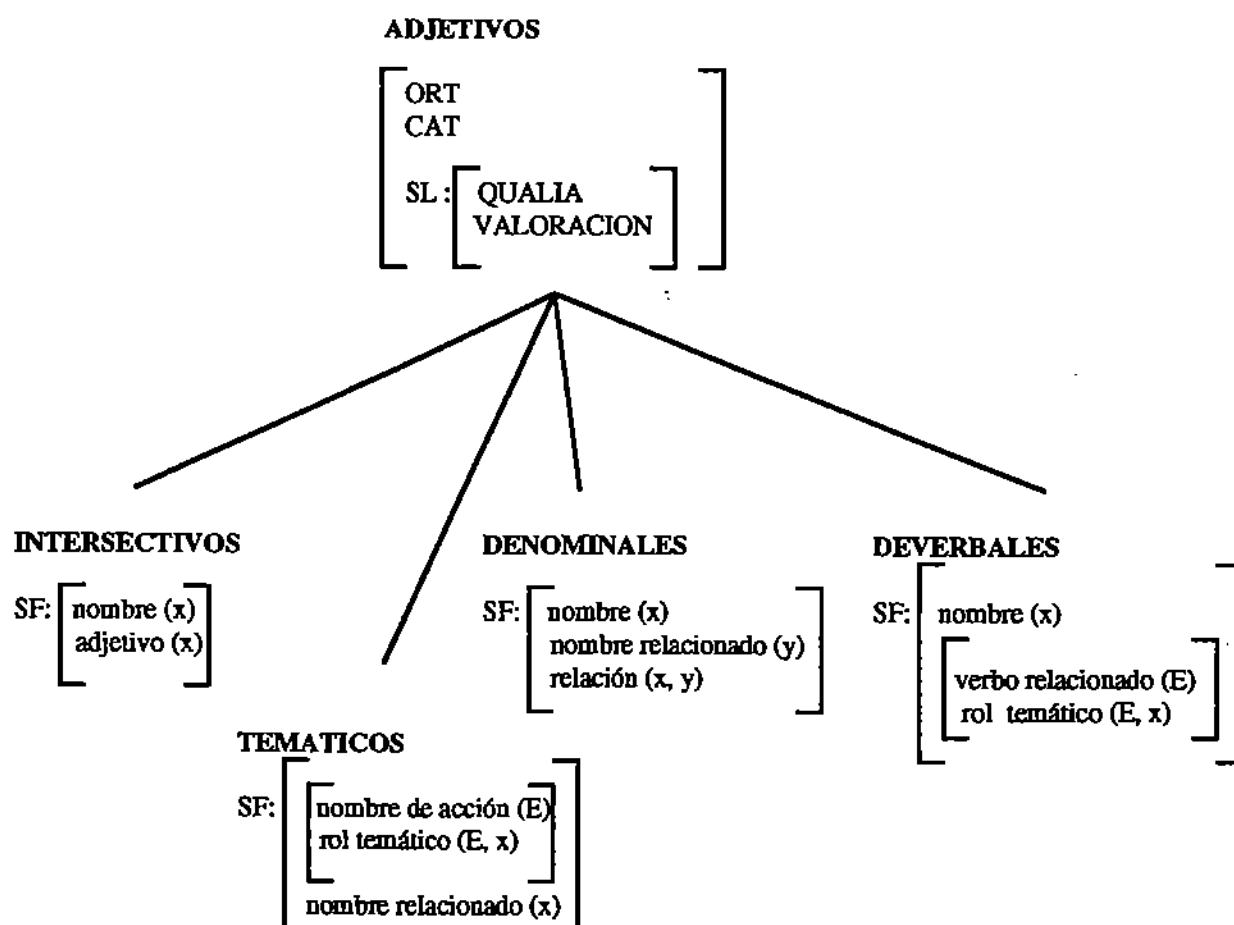


fig 8.

Los rasgos ORT, CAT y SL mantienen una estructura común a los cuatro tipos, difiriendo tan sólo en el tipo de valores que adoptarán. SL es un rasgo que típicamente servirá para la representación del significado de los adjetivos intersectivos; sin embargo, no les es privativo⁸, por lo que se atribuye -como ORT y CAT- al tipo común superior, siendo su estructura heredada por los subtipos lógicos, imbricándose en éstos con la SF para configurar la representación semántica. De no precisar de SL para la descripción de su significado, los subtipos simplemente dejarán dichos valores subespecificados.

4. Conclusiones.

⁸ P.e. el relacional anteriormente relatado, "benéfico", representado por relación de causación con "bien_3", aporta asimismo una valoración de tipo positivo, codificable en el rasgo valorativo de SL.

Creemos que el marco de representación formulado maneja de forma eficaz un rico abanico de componentes del significado del adjetivo: restricciones selectivas, orden superficial, distinción entre sentidos, potenciales calificativo y valorativo, estatus relacional-derivativo y tipología predicativa (semántica formal). Aunque evidentemente el inventario semántico no es exhaustivo, la confrontación del marco propuesto con los datos en el trabajo actualmente en proceso de carga de léxico en la BCL nos hace suponer que éste está dotado de capacidad expresiva suficiente como para afrontar la representación de nuevos aspectos del significado no específicamente abordados en el presente artículo.

Referencias.

- Aarts & Calbert. *Metaphor and Non-Metaphor. The Semantics of Adjective-Noun combinations*. Max Niemeyer Verlag, Tübingen. 1979.
- Bosque I. *Sobre las diferencias entre los adjetivos relacionales y los calificativos*. Revista Argentina de Lingüística vol. 1993.
- Carpenter B. *The Logic of Typed Feature Structures*. Cambridge University Press. 1992.
- Castellón I. *Lexicografía Computacional: Adquisición Automática del Conocimiento Léxico*. Tesis Doctoral. Report LSI-93-28-R Universitat Politècnica de Catalunya, Barcelona. 1993.
- Climent S. & Soler C. *Lexical Mismatches: A Semantic Representation of Adjectives in the LKB*. Acquilex-II Working Paper nº 28. 1994.
- Copestake A. *The Representation of Lexical Semantic Information*. Tesis Doctoral. Cognitive Science Research Papers 280. University of Sussex. 1992.
- Mel'çuk I. *Dictionnaire explicatif et combinatoire du français contemporain*. Les Presses de l'Université de Montréal. 1984.

Pustejovsky J. *The generative lexicon*. Computational Linguistics vol. 17 n. 4. 1991.

Soler C. & Martí MA. *Dealing with Lexical Mismatches*. Aquilex-II Working Paper nº 2.4. 1993.

VOX, *Diccionario General Ilustrado de la Lengua Española*. Bibliograf S.A. Barcelona, 1987.