

GENERACIÓN DE ALTERNANCIAS DE SUBCATEGORIZACIÓN MEDIANTE REGLAS LÉXICAS

Mariona Taulé Delor (Universitat de Barcelona)

marionat@lsi.upc.es

Irene Castellón Masalles (Universitat Politècnica de Catalunya)

cirene@lsi.upc.es

1.- Introducción

En las teorías lingüísticas actuales hay una tendencia generalizada a la lexicalización de la información lingüística. Entendemos lexicalización en el sentido de que la mayoría del contenido lingüístico reside en el léxico, convirtiéndolo así en un componente básico a partir del cual se deducirá la mayoría de información lingüística. El léxico no se concibe ya como una mera lista de entradas, sino como un componente complejo y estructurado.

Este nuevo enfoque permite la reducción de reglas sintagmáticas y, consecuentemente, la simplificación de la gramática, pero, por otro lado, el desplazamiento de la información al léxico puede comportar también la aparición de información redundante en las entradas léxicas. La cantidad y complejidad de la información puede crear, también, problemas de control de ésta. En consecuencia, se hace necesario no sólo decidir qué tipo de información deben contener las entradas léxicas sino también cómo estructurar toda esta información con el fin de evitar redundancia y capturar las similitudes entre clases de palabras que se comportan de igual forma.

Las Bases de Conocimiento Léxico (BCL) permiten abordar este problema de forma eficaz mediante la codificación estructurada del conocimiento. Las BCLs incorporan mecanismos de herencia y reglas léxicas que evitan tanto la redundancia como la inconsistencia de la información.

En el presente trabajo¹ se demuestra la viabilidad de las BCLs y, en concreto, de las reglas léxicas que soportan, como medio de representación y de generación de información lingüística a gran escala.

El objetivo principal es el tratamiento de las alternancias de subcategorización ("diatheses") del verbo mediante reglas léxicas. Los diferentes esquemas de subcategorización de cada verbo se deducirán a partir de la misma entrada verbal, la más básica o general, mediante la aplicación de reglas léxicas que reflejan los diversos cambios morfológicos, sintácticos y semánticos que se producen.

2.- La Base de Conocimiento Léxico

En la introducción se señalaba que el desplazamiento de la información lingüística a nivel léxico puede significar la presencia de redundancia de información en las entradas léxicas. Una propuesta interesante para resolver este problema es la planteada en el formalismo gramatical de la "Head-Driven Phrase Structure Grammar" (HPSG) (Pollard & Sag 1987 1992) donde los autores resuelven la redundancia de información con la postulación de tipos léxicos ("Lexical types") y reglas léxicas ("Lexical rules").

Los tipos léxicos son estructuras de rasgos que permiten organizar y clasificar jerárquicamente la información según las propiedades morfológicas, sintácticas y semánticas de las palabras. (Véase figura 1)

Las reglas léxicas, en cambio, se encargan de capturar las generalizaciones lingüísticas entre palabras pertenecientes a la misma clase léxica. (Véase figura 4 y 8)

Siguiendo el planteamiento teórico expuesto en la HPSG, la Base de

¹ El trabajo se ha realizado en el marco del proyecto ACQUILEX-II (Esprit 7315) en la Universitat Politècnica de Catalunya.

Gracias, también, a la ayuda del proyecto PB91-0854 de la DGICYT de la Universitat de Barcelona.

Conocimiento Léxico² diseñada en el marco del proyecto Acquilex representa los objetos mediante estructuras de rasgos tipificadas ordenadas jerárquicamente que permiten la herencia de la información. La operación básica de la BCL es la unificación. Por otro lado, la BCL soporta reglas léxicas que posibilitan derivar automáticamente entradas léxicas nuevas a partir de entradas ya existentes. Desde el punto de vista teórico, el desarrollo de las reglas léxicas en la BCL se basa en la concepción de "léxicos dinámicos" de J. Pustejovsky³ que considera la representación de un léxico básico a partir del cual se genera el resto de las entradas léxicas. En la BCL, la ampliación del léxico se realiza mediante la incorporación de nuevas entradas o mediante la aplicación de las reglas léxicas.

A continuación, se expone de forma resumida el tipo de información que contienen las entradas léxicas en la BCL y cómo se estructura esta información.

2.1.- Representación léxica del verbo

La cantidad y complejidad de información que contienen las entradas léxicas que se representan en la BCL implica una organización de ésta, es decir, es necesaria una clasificación léxica en función del tipo de información que se considere, ya sea de tipo morfológico, sintáctico o semántico.

En la BCL, de acuerdo a una aproximación basada en el signo (C. Pollard & I. Sag 1987), se distingue entre signos léxicos (palabras) y signos sintagmáticos, pero en el presente trabajo se hace referencia a la representación de los signos léxicos, en concreto de los signos verbales^{4,5}.

El tipo de información que incluyen los signos es básicamente de tipo

² Copestake, A. (1992), Sanfilippo, A. (1991), Ageno et al. (1992).

³ Pustejovsky, J. 1991.

⁴ Sanfilippo, A. (1991)

⁵ Los trabajos de Castellón, I. (1992), Copestake et al. (1992), etc. presentan un tratamiento completo de los signos nominales.

morfo-sintáctico y semántico. La información de tipo ortográfico se reduce a una secuencia ortográfica de la palabra (o sintagma). La información morfo-sintáctica indica el tipo de categoría léxica, las propiedades morfológicas inherentes y, en el caso de las entradas verbales, el número de argumentos que subcategorizan.

En cuanto a la información de tipo semántico se especifica la naturaleza de los argumentos subcategorizados mediante roles temáticos, las restricciones de selección así como el tipo de verbo según su comportamiento aspectual. La estructura que presentan los signos léxicos es la siguiente:

lex-sign (sign)	verb-sign (lex-sign complex-sign)
<orth> = orth	<orth> = orth
<cat> = cat	<cat> = complex-cat
<sem> = sem	<sem> = verb-sem
<rqs> = rqs	<rqs> = vrqs.
<sense-id> = sense-id ⁶ .	

Figura 1: Signo léxico y Signo verbal.

Los criterios que han determinado la clasificación de los verbos en la BCL han sido el número y tipo de argumentos que subcategoriza el verbo. En el rasgo <cat> de las entradas verbales se especifica la categoría ("Part of Speech") y el comportamiento sintáctico del verbo. Según el número de argumentos, los verbos se clasifican, fundamentalmente, en dos clases: los verbos con un argumento ('strict-intrans-cat'), y los verbos con dos o más argumentos ('x-sign-x-cat'). Según el tipo de argumento que subcategorizan, los tipos anteriores se subclasifican en diversas clases 'strict-trans-cat', 'ditrans-cat', 'obl-intr-cat', 'obl-trans-cat', 'comp-cat', etc., es decir, según si el argumento subcategorizado es una completiva, un complemento preposicional, una frase de gerundio, un adjetivo, etc. (En la figura 2 se

⁶ El rasgo 'sense-id' informa sobre el diccionario fuente, la lengua, el número de acepción, etc. de la entrada.

expone una parte de la jerarquía de tipos que representa la clasificación verbal según la categoría).

En el rasgo <sem> se incluye la información relativa a la naturaleza de los argumentos subcategorizados por el verbo, mediante la asignación de proto-roles temáticos ('p-agent', 'p-patient', 'p-ag-cause', etc.) basada en la idea originaria de D. Dowty (1988 1989) y elaborada por A. Sanfilippo (1991), que considera los proto-roles como "clusters" o conjuntos de propiedades que permiten definir un agente prototipo y un paciente prototipo en función de las propiedades características predominantes.

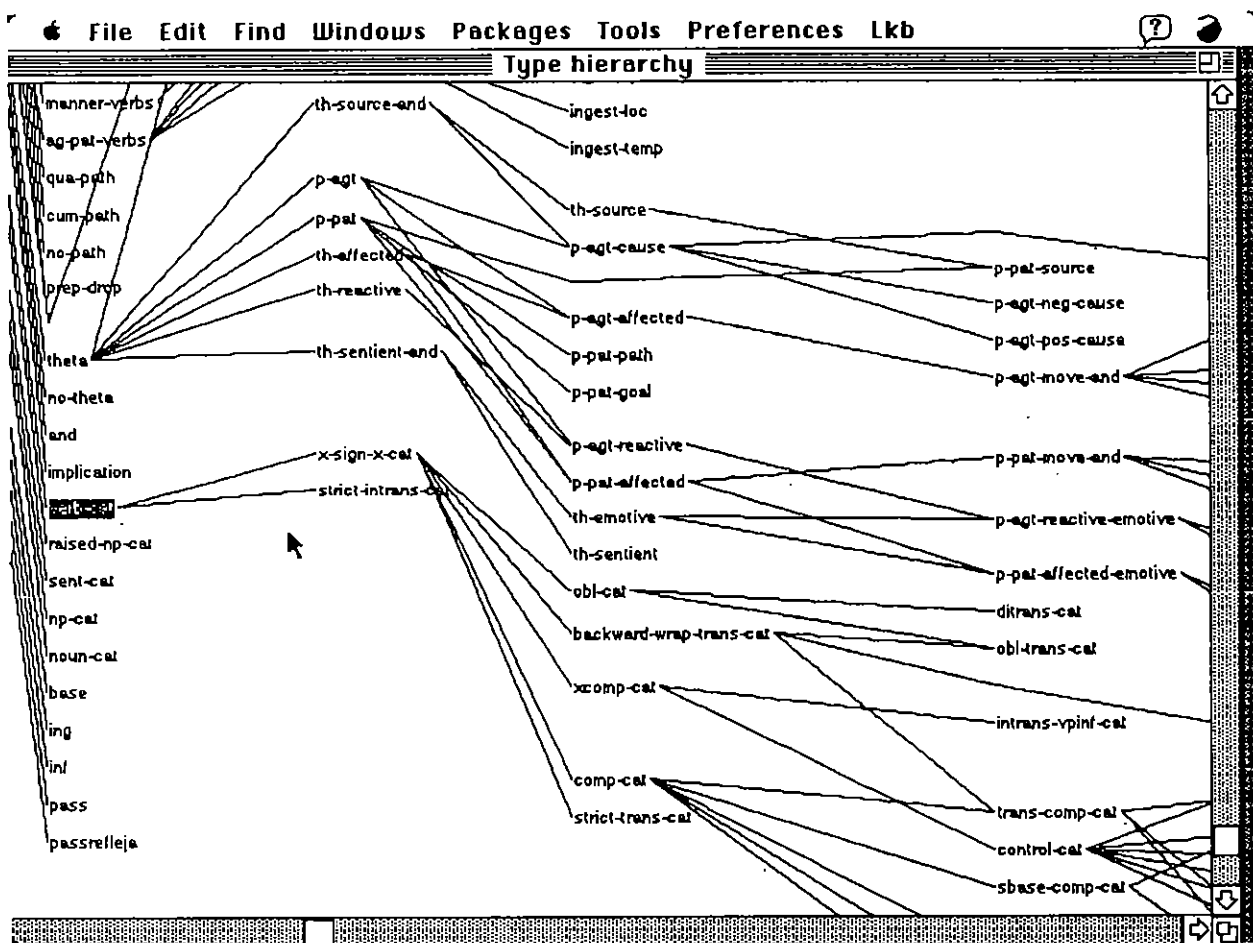


Figura 2: Clasificación verbal según la categoría

La semántica del verbo (<sem>) es una fórmula lógica formada por diversos componentes: a) un índice ontológico que aporta información referente al Aktionsart, que en el caso de los verbos siempre se trata de

eventualidades: estados, acciones, procesos, etc.; b) una serie de predicados que reflejan los distintos proto-roles temáticos de cada argumento y, por último, c) las restricciones de selección, que especifican las propiedades semánticas que deben cumplir los proto-roles ('animado', 'inanimado', 'humano', etc.).

formula (sem)

<ind> = entity

<pred> = logical-pred

<arg1> = sem.

Figura 3: Fórmula lógica

En el signo verbal se integra la información sintáctica y semántica a través de la coindexación de los argumentos subcategorizados especificados en <cat> y los argumentos de la forma lógica especificados en <sem>.

En el rasgo <rqs> se especifica también información de tipo semántico. En el caso de las entradas verbales, en la <vrqs>, se incluye información sobre la clase semántica a la que pertenece el verbo (ingestión, cocción, movimiento, percepción, dicción, etc.) e información relativa a los componentes semánticos⁷.

3.- Alternancias de subcategorización

El objetivo del presente trabajo se centra en la codificación de la información relativa a la subcategorización, concretamente la referente a las alternancias de subcategorización o "diatheses" verbales. Se entiende por alternancias de subcategorización o "diatheses", las diferentes realizaciones (alternancias) en la expresión de los argumentos de un verbo, es decir, los distintos esquemas de subcategorización en que aparecen los argumentos que combinan característicamente con un determinado verbo.

El tipo de "diatheses" verbales que se tratan a lo largo del trabajo son

⁷En el trabajo *Towards a VROS Representation* (Taulé, M., 1993) hay una descripción completa de la información contenida en la <vrqs>. (En preparación)

las siguientes: la construcción transitiva, la pasiva, la pasiva refleja, la reflexiva y la construcción intransitiva. Veámos a continuación unos ejemplos de alternancia en los esquemas de subcategorización:

(1) a. Clara come un huevo frito con patatas. (Construcción Transitiva)

b. Clara come. (Construcción Intransitiva)

c. Un huevo frito con patatas es comido por Clara. (Pasiva)

(2) a. Toni tuesta el pan. (Construcción Transitiva)

b. Se tuesta el pan. (Pasiva Refleja)

c. El pan es tostado por Toni. (Pasiva)

(3) a. El vino emborracha a Manuel. (Construcción Transitiva)

b. Manuel se emborracha. (Reflexiva)

c. Manuel se emborracha con vino. (Reflexiva)

Los verbos "comer", "tostar" y "emborrachar" de las frases (1), (2) y (3) pueden realizarse sintácticamente de forma diversa: en la primera frase, el verbo "comer" alterna la forma transitiva (1a) con la forma intransitiva (1b), además, permite también la construcción en pasiva (1c); en la segunda frase, la forma activa de (2a) alterna con la denominada pasiva refleja (2b) y la pasiva (2c)⁸; y en la frase (3), el verbo "emborrachar" es usado de forma transitiva en (3a) y reflexivamente en (3b) y (3c), la diferencia entre estas dos últimas frases ((3b) y (3c)) reside en la presencia o no del complemento preposicional.

Las alternancias que aparecen en los ejemplos mencionados no sólo reflejan cambios a nivel sintáctico, es decir, en el número y tipo de argumentos subcategorizados por el verbo, sino también suponen cambios en la expresión de los roles temáticos implicados, cambios a nivel morfológico e incluso puede comportar cambios a nivel aspectual. Por ejemplo, en la frase (1b) hay una omisión sintáctica del argumento paciente

⁸ Trataremos la pasiva y la pasiva refleja como alternancias de subcategorización.

(aunque este sea perfectamente inferible) que supone un cambio aspectual, la frase (1a) tiene una lectura de 'accomplishment' (el proceso tiene una culminación lógica) mientras que la frase (1b) es indicadora de un proceso. La frase (2) alterna una lectura causativa (2a) donde el argumento agente es explicitado, frente a una lectura no causativa (o incoativa) donde el causante de la acción no se ha expresado (2b). En las frases (3b) y (3c), el agente y el paciente es el mismo ('Manuel') y 'el vino' indica el medio con el cual el agente se emborracha, que se puede expresar (3c) o no (3b).

Las alternancias de subcategorización implican cambios a nivel morfológico, sintáctico y semántico, esto significa que la "diatheses" no debe ser tratada exclusivamente desde un punto de vista sintáctico. La "diatheses" es, en cierto modo, la interacción entre información referente al significado (roles temáticos y restricciones de selección) y la realización sintáctica de estos argumentos, como señala B. Levin (1989)⁹. La propuesta de B. Levin se basa en la idea que una buena clasificación de las alternancias gramaticales puede servir de ayuda para agrupar los significados de los verbos en clases semánticas distintas.

Toda esta información relativa a las alternancias debe ser codificada en la entrada léxica, en ella se deben reflejar todos los cambios y las relaciones que implican a nivel sintáctico y semántico. Otro aspecto es la manera de representar toda esta información que, en nuestro caso, se lleva a cabo mediante reglas léxicas.

3.1.- Las reglas léxicas

Básicamente, hay dos estrategias posibles para representar las alternancias de subcategorización en una BCL:

a) Crear tantas entradas léxicas como esquemas de subcategorización admite un determinado verbo. La adopción de esta estrategia supone, en

⁹ El presente trabajo sigue las líneas de investigación propuestas en Levin (1989) por lo que se refiere al tratamiento y clasificación de las alternancias de subcategorización.

primer lugar, un incremento del número de entradas en la BCL, la repetición innecesaria de información léxica en las entradas y, por último, no se refleja la relación entre entradas de un mismo verbo.

b) La otra estrategia posible, que aquí se propone, es la expresión de las alternancias de subcategorización a través de la formulación de reglas léxicas (Sanfilippo 1992). Mediante este procedimiento sólo se hace necesaria la creación de una única entrada verbal a partir de la cual se generan las demás posibilidades. De este modo, se reduce la cantidad de entradas léxicas y se prioriza una de ellas, la que se considera más básica o general, se elimina la redundancia de información y se capturan formalmente los distintos comportamientos sintácticos y semánticos del verbo.

Siguiendo esta propuesta, se plantea la necesidad de determinar qué esquema de alternancia es el más básico o general. Nuestra propuesta, considera, en principio, la entrada verbal que subcategoriza más argumentos (generalmente, la más informativa) como la entrada más básica.

Las reglas léxicas se definen en la BCL mediante estructuras de rasgos de la siguiente forma:

lexical-rule (rule)*
<1> = lex-sign
<0> = lex-sign.

Figura 4 : Regla Léxica

(La información que aparece entre paréntesis hace referencia al hiperónimo del tipo en cuestión en la jerarquía, es decir, el tipo 'lexical-rule' es hipónimo del tipo 'rule'.)

Las reglas léxicas consisten en un signo léxico de entrada (<1>) que da como resultado un nuevo signo léxico de salida (<0>). En el caso concreto que se estudia, el signo léxico de entrada es verbal y el resultado producido es un nuevo signo verbal.

3.2.- Representación de la 'diatheses' en la BCL

En este apartado se expone la manera en que se representa la "diatheses" verbal en la BCL, las reglas léxicas y la instanciación de éstas.

Las reglas léxicas son estructuras tipificadas y como tales forman parte de la jerarquía de tipos de la BCL. En la figura 5 se presenta una muestra de cómo éstas se organizan en la BCL.

Las reglas representadas dan cuenta de las alternancias de subcategorización que hemos denominado transitivas: uso reflexivo ('trans-reflex-alt'), pasiva refleja ('pass-refl-alt'), omisión de objeto ('indef-def-alt'), pasiva ('pass-alt'), etc.

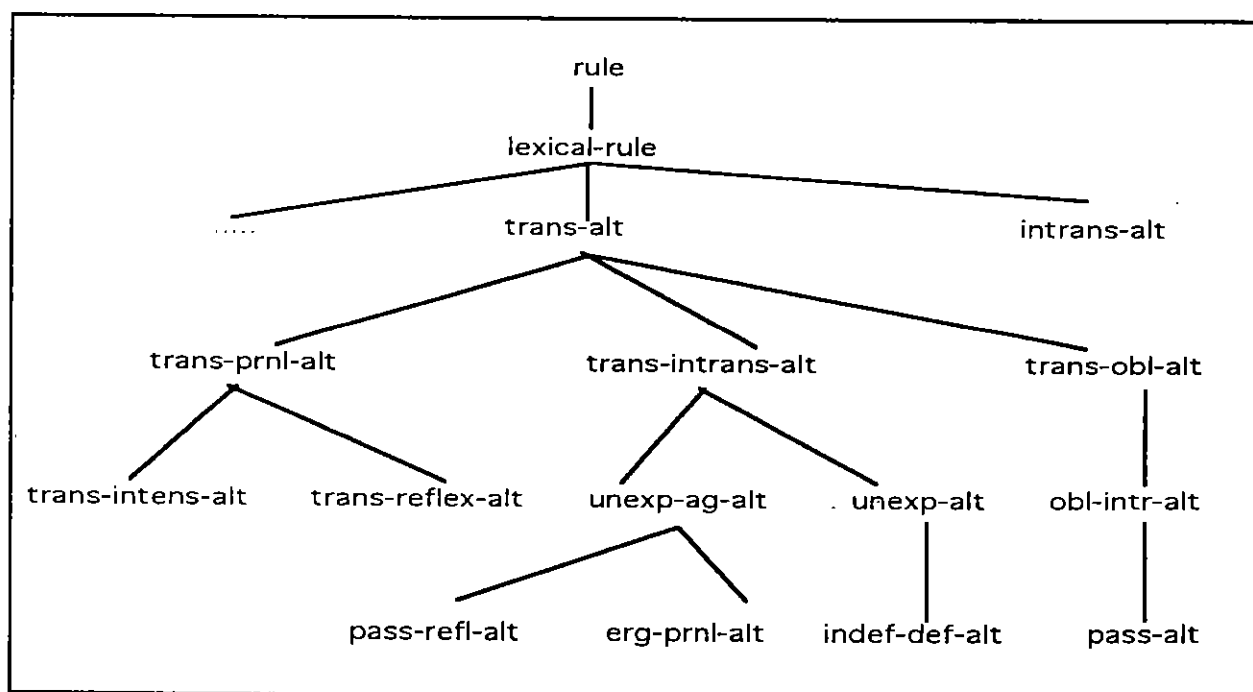


Figura 5 Jerarquía de tipos: reglas

La jerarquía de reglas se inicia con el tipo más general ('rule') que subclasifica dos tipos de reglas, las léxicas ('lexical-rule') y las gramaticales ('grammar-rule'). Un subtipo de regla léxica es el formado por las reglas de alternancia transitiva ('trans-alt') caracterizadas por tener como signo léxico de entrada un verbo transitivo ('strict-trans-sign'). Dentro de las alternancias transitivas ('trans-alt'), se distinguen distintas reglas ('trans-prnl-alt', 'trans-intrans-alt', 'trans-obl-alt', etc.) según el signo léxico de salida que generen.

El tipo de regla 'trans-prnl-alt' no implica ningún cambio de categoría, es decir, que mantiene el mismo signo léxico de salida ('strict-trans-sign') y el mismo número de argumentos subcategorizados. La regla 'trans-reflex-alt', por ejemplo, es un subtipo de esta regla que da cuenta de las construcciones que permiten el uso reflexivo (p.e.. (3b.) Manuel_i se_i emborracha vs. (3a) El vino emborracha a Manuel), este uso implica un cambio en el tipo de argumento objeto que pasa a ser un clítico pronominal así como también un cambio semántico, en el uso reflexivo siempre se produce una identificación entre el argumento agente y el paciente, son correferentes. Correferencia que no se establece en la construcción transitiva activa.

Las reglas que se incluyen en el tipo 'trans-intrans-alt' reflejan la omisión de algún argumento y, en consecuencia, generan signos de un único argumento, es decir, signos intransitivos. El tipo de cambio producido reside en el número de argumentos subcategorizados que puede implicar también cambios a nivel morfológico y semántico. Dentro de este tipo se distinguen dos grupos de reglas: las que omiten el argumento paciente 'unexp-alt' (p.e. 'indef-def-alt': (1b.) Clara come vs. (1a.) Clara come un huevo frito con patatas); y las que omiten el argumento agente 'unexp-ag-alt' (p.e. la regla de pasiva refleja 'pass-refl-alt': (2b.) Se tuesta el pan vs. (2a.) Toni tuesta el pan, que además del cambio argumental implica un cambio a nivel morfológico con la introducción del clítico pronominal "se" a la forma verbal).

Por último, el tipo 'trans-obl-alt' incluye aquellas reglas que producen como signo léxico de salida un signo oblicuo, es decir un signo que

subcategoriza un argumento preposicional. La regla de pasiva 'pass-alt' ejemplificaría este tipo de reglas (p.e.: (1c.) Un huevo frito con patatas es comido por Clara vs. (1a.) Clara come un huevo frito con patatas) que implican un cambio en la argumentación del verbo, no necesariamente en el número de argumentos subcategorizados pero sí en el tipo de argumentos implicados, además del cambio a nivel morfológico en la forma del verbo. Estas reglas suponen una reestructuración de los argumentos.

Las alternancias de subcategorización se introducen a nivel sintáctico (es decir, en <cat>) a través del rasgo <diatheses> que tiene cómo valor el tipo 'alternations'. A partir del tipo 'alternations' se ha definido una subclasificación de los diversos tipos de alternancias que se pueden encontrar. En la figura 6 se da una muestra de ésta.

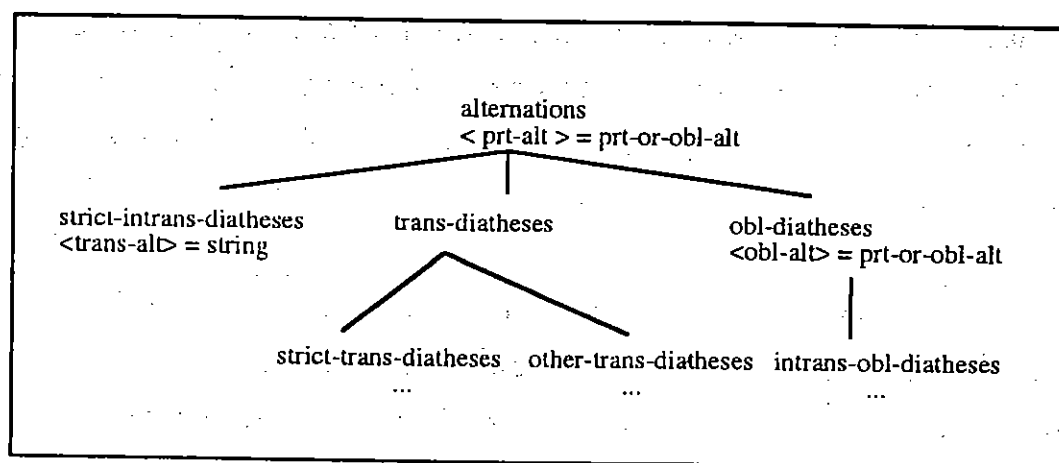


Figura 6: Jerarquía de tipos: alternancias de subcategorización

A continuación, se ejemplifica la regla léxica de pasiva refleja con su aplicación en la entrada verbal "tostar X_I_1".

En la figura 7 se presenta la entrada léxica "tostar X_I_1" como un signo verbal del tipo 'strict-trans-sign' que contiene información relativa a la ortografía (<orth>), a la sintaxi (<cat>) y a la semántica (<sem>). La información contenida en los rasgos <sense-id> y <rqs> es información relacionada con el diccionario fuente que, en este trabajo, no es relevante.

En el rasgo <cat> se especifica la categoría, 'strict-trans-cat', y los

argumentos subcategorizados por el verbo. La aproximación utilizada para expresarlo es la propuesta en la Gramática Categorial (Uszkoreit, H. 1986) que considera el verbo como una categoría compleja. El verbo se define como una estructura de rasgos que contiene un signo activo ('dir-obj-np-sign'), una dirección ('backward') y la categoría resultante ('strict-intrans-cat'). En la Gramática Categorial un verbo transitivo se expresaría como 'S/NP\NP'. En el rasgo <m-feats> (morphological-feats) se incluye la información referente a la forma verbal (<vform>), a las preposiciones regidas (<prt>), etc., y la información relativa a las alternancias de subcategorización (<diatheses>). En el caso del verbo 'tostar' el tipo de 'diatheses' es transitiva ('strict-trans-diatheses') y los dos tipos de alternancias que puede presentar se incluyen en el rasgo <trans-alt>: construcción pasiva ('pass') y pasiva refleja ('passrefl'), que permitirán la aplicación de las reglas léxicas correspondientes ('pass-alt' y 'passrefl-alt').

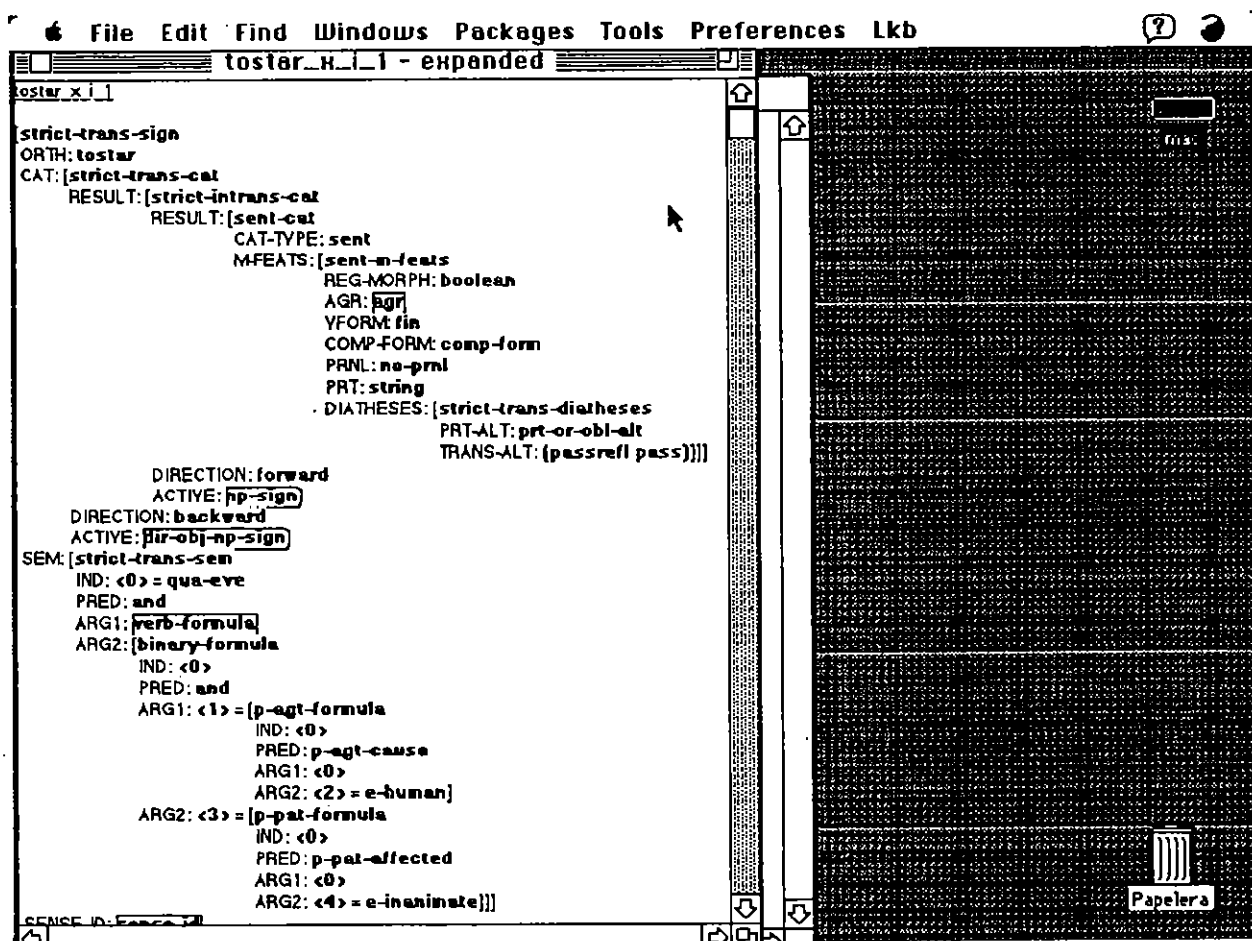


Figura 7: Entrada léxica de "tostar X_I_1"

En el rasgo <sem> se especifica la estructura argumental del verbo "tostar" mediante una fórmula lógica binaria, que indica que el verbo "tostar" en su uso transitivo es un 'accomplishment' ('qua-eve'), que el primer argumento es un agente humano indicado mediante el proto-rol 'p-agt-cause' y la restricción de selección 'e-human' y el segundo argumento un paciente con proto-rol 'p-pat-affected' rellenado con la restricción de selección 'e-inanimate'.

En la figura 8, se presenta la regla léxica de pasiva refleja ('pass-refl-alt') donde se pueden comparar los respectivos signos léxicos de entrada (<1>) y de salida (<0>) que reflejan los cambios producidos en <orth>, <cat> y <sem>.

The image shows a screenshot of a software application window titled "pass-refl-alt". The window is split into two panes. The left pane displays the input structure for the rule, and the right pane displays the output structure. Both panes show a hierarchical list of features and their values, including orthographic forms, category codes, semantic information, and morphological details.

Left Pane (Input Structure):

- pass-refl-alt
- 0: [strict-intrans-sign
- ORTH: [complex-orth
- ORTH1: se
- ORTH2: <0> = [orth]]
- CAT: [strict-intrans-cat
- RESULT: [sent-cat
- CAT-TYPE: sent
- MFEATS: [sent-m-feats
- REG-MORPH: boolean
- AGR: [agr
- PERS: <1> = person
- NUM: <2> = number]
- VFORM: passrefleja
- COMP-FORM: comp-form
- PRNL: <3> = prnl
- PRT: string
- DIATHESES: [strict-intrans-diatheSES
- PRT-ALT: prt-or-obl-alt]]
- DIRECTION: forward
- ACTIVE: hp-sign
- SEM: [strict-intrans-sem
- IND: <4> = eve
- PRED: and
- ARG1: [verb-formula
- IND: <4>
- PRED: <5> = string
- ARG1: <4>]
- ARG2: <6> = [p-agt-formula
- IND: <4>
- PRED: p-agt-affected
- ARG1: <4>
- ARG2: <7> = ob]]
- SENSE-ID: [sense-id]
- RQS: [p-pat-verbs]
- 1: [strict-trans-sign
- ORTH: <0>

Right Pane (Output Structure):

- 1: [strict-trans-sign
- ORTH: <0>
- CAT: [strict-trans-cat
- RESULT: [strict-intrans-cat
- RESULT: [sent-cat
- CAT-TYPE: sent
- MFEATS: [sent-m-feats
- REG-MORPH: boolean
- AGR: [agr
- VFORM: vform
- COMP-FORM: comp-form
- PRNL: <3>
- PRT: string
- DIATHESES: [strict-trans-diatheSES
- DIRECTION: forward
- ACTIVE: hp-sign
- DIRECTION: backward
- ACTIVE: [ir-obj-ap-sign]
- SEM: [strict-trans-sem
- IND: <4>
- PRED: and
- ARG1: [verb-formula
- IND: <4>
- PRED: <8> = string
- ARG1: <4>]
- ARG2: [binary-formula
- IND: <4>
- PRED: and
- ARG1: <9> = [p-agt-formula
- IND: <4>
- PRED: p-agt
- ARG1: <4>
- ARG2: <10> = ob]]
- ARG2: <11> = [p-pat-formula
- IND: <4>
- PRED: p-pat
- ARG1: <4>
- ARG2: <7>]]
- SENSE-ID: [sense-id]

Figura 8: Regla Léxica de Pasiva Refleja

El primer cambio que se observa es en el rasgo <orth> con la introducción del pronombre 'se', que supone una ortografía compleja en el signo resultante (<0>).

La regla de pasiva refleja convierte un signo transitivo 'strict-trans-sign' (<1>) en un signo intransitivo 'strict-intrans-sign' (<0>) en el que se ha omitido el argumento causante de la acción ('p-agt-cause') de la construcción transitiva. También se especifica para el rasgo <vform> el valor de 'passrefleja'.

Esta regla también tiene repercusiones en la estructura argumental, el argumento agente de la construcción en pasiva refleja es un 'p-agt-affected'. En la figura 9 se ven con más claridad los cambios que se producen en el rasgo <sem>.

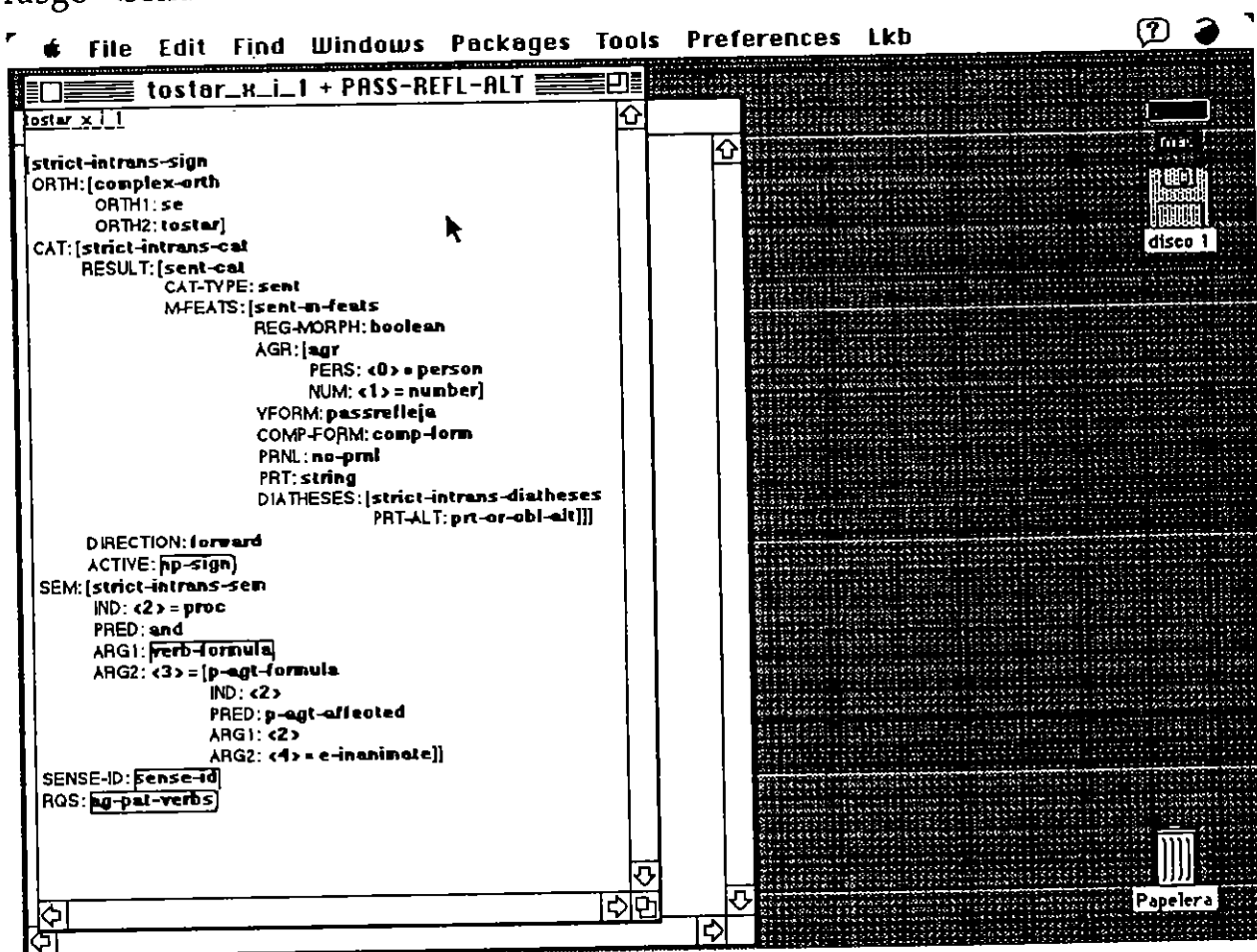


Figura 9: Aplicación de la regla de pasiva refleja sobre "tostar"

En la figura 9 se muestra la entrada léxica generada una vez aplicada la regla léxica de pasiva refleja. En la nueva entrada léxica se reflejan los diversos cambios producidos en <orth>, <cat> y <sem>. El paciente inanimado 'p-pat-affected' de la construcción pasiva es ahora, en la construcción de pasiva refleja, el agente 'p-agt-affected' inanimado. El agente humano causante de la acción de la forma transitiva 'p-agt-cause' no se expresa en la construcción de pasiva refleja.

4.- Conclusiones

En la actualidad se han desarrollado un total de seis reglas léxicas de alternancia que cubren las siguientes realizaciones de un verbo transitivo: la pasiva, la pasiva refleja, la reflexiva, el uso pronominal intensificador, el uso pronominal ergativo y las construcciones intransitivas. Estas reglas cubren, en principio, la totalidad de verbos que admiten este tipo de construcciones (las reglas son restrictivas en el sentido que sólo se aplican a aquellos verbos que admiten esas construcciones y no a todos los verbos de la Base de Conocimiento).

Las Bases de Conocimiento Léxico constituyen un recurso potente para el tratamiento de léxicos computacionales: tanto por los mecanismos de herencia que contienen como por la posibilidad de definir operaciones que permiten expresar fenómenos lingüísticos como las alternancias de subcategorización de modo elegante, evitando la redundancia de información en la codificación de ésta y permitiendo el establecimiento de relaciones y generalizaciones entre las entradas léxicas.

Las líneas futuras de investigación se centrarán en dos aspectos concretos:

- 1.- La ampliación de reglas léxicas con el fin de cubrir la totalidad de alternancias de subcategorización que permite la lengua española.

- 2.- El estudio de las alternancias en relación a las clases

semánticas de los verbos, siguiendo el planteamiento teórico propuesto por Levin, B. (1989).

Bibliografía

- Acquilex-II, The Acquisition of Lexical Knowledge, Basic Research Project 7315 ACQUILEX-II, Technical Annexe.
- Agno, A., I. Castellón, M.A. Martí, F. Ribas, G. Rigau, H. Rodríguez, M. Taulé, M.F. Verdejo, "SEISD: An Enviroment for Extraction of Semantic Information from On-line Dictionaries" en los *Proceedings of the 3rd Conference on Applied Natural Language Processing*, 1992.
- Castellón, I., Lexicografía Computacional: Adquisición de Conocimiento Léxico, Tesis Doctoral, Barcelona 1992.
- Copestake, A. "The ACQUILEX LKB: Representation Issues in Semi-Automatic Acquisition of Large Lexicons". En los *Proceedings of the 3rd Conference on Applied Natural Language Processing*, 1992.
- Dowty, D. "Thematic Proto Roles, Subject Selection and Lexical Semantic Defaults", Colloquium paper LSA, 1987.
- Dowty, D. "On the Semantic Content of the Notion 'Thematic Role', en Chierchia, G., Partee, B. & Turner, R. (eds) Property Theory, Type Theory and Natural Language. Dordrecht: D. Reidel, 1989.
- Levin, B., Towards a Lexical Organization of English Verbs, editado en la Universidad de Chigago (forthcoming), 1989.
- Pollard, C. & Sag, I. An Information-Based Approach to Syntax and Semantics: Fundamentals, Vol. I. CSLI, Stanford, 1987.
- Pollard, C. & Sag, I. Agreement, Binding and Control: Information-Based Syntax and Semantics, Vol.2. CSLI, Stanford 1992.
- Pustejovsky, J. "Generative Lexicon" en *Computational Linguistics*, 1991.
- Sanfilippo, A. "LKB Encoding of Lexical Knowledge". Aparecerá en Briscoe, T., Copestake, A., & de Paiva, V. (eds.), 1991.

- Sanfilippo, A. "Verbal Diatheses: Knowledge Acquisition, Lexicon Construction and Dictionary Compilation" (forthcoming), 1992.
- Sanfilippo, A., Briscoe, T., Copestake, A., Martí, M., Taulé, M. & Alonge, A. "Translation Equivalence and Lexicalization in the ACQUILEX LKB" *Proceedings of the 4th International Conference on Theoretical and Methodological Issues in Machine Translation*, 1992.
- Uszkoreit, H. "Categorial Unification Grammar" *Proceedings of the 11th International Conference on Computational Linguistics and the 24th Annual Meeting of the Association for Computational Linguistics*, 1986.