

Ana de Aguilar-Amat Castillo

Las colocaciones de nombre y adjetivo.
Un paso hacia una teoría léxico-semántica
de la traducción.

Tesis doctoral dirigida
por el Dr. Ramon Cerdà Massó

Departamento de Lengua Española
Facultad de Filología Española
Universidad Autónoma de Barcelona
Octubre 1993

A Ramon y Júlia

'Throughout his poetry Swinburne lays general constructions alongside each other, syntagmatically parallel collocations are a feature of verse-form and stanza-form'

J.R. Firth (1957)

'Delight, the rootless flower,
And Love, the bloomless bower:
Delight that lives an hour,
And love that lives a day...

Sin sweet beyond forgiving
And brief beyond regret.'

Swinburne,
'Before Dawn'

"... podemos estar mirando una pieza de un puzzle tres días seguidos y creer que lo sabemos todo sobre su configuración y su color, sin haber progresado lo más mínimo: sólo cuenta la posibilidad de relacionar esta pieza con otras y, en este sentido, hay algo en común entre el arte del puzzle y el arte del go: sólo las piezas que se hayan juntado cobrarán un carácter legible, cobrarán un sentido: considerada aisladamente, una pieza de un puzzle no quiere decir nada; es tan sólo pregunta imposible, reto opaco; pero no bien logramos, tras varios minutos de pruebas y errores, o en medio segundo prodigiosamente inspirado, conectarla con una de sus vecinas, desaparece, deja de existir como pieza: la intensa dificultad que precedió aquel acercamiento, y que la palabra *puzzle* -enigma- expresa también en inglés, no sólo no tiene ya razón de ser, sino que parece no haberla tenido nunca, hasta tal punto se ha hecho evidencia: las dos piezas milagrosamente reunidas ya sólo son una, a su vez fuente de error, de duda, de desazón, de espera."

Georges Perec, La vida instrucciones de uso

PREFACIO

La tesis doctoral que sigue ha sido realizada gracias al acuerdo de colaboración científica entre FUJITSU ESPAÑA, S.A. y la Universidad Politécnica de Cataluña en el área del tratamiento informático del lenguaje natural. Dicho acuerdo, convenio nº C1135, firmado el 15 de abril de 1990, permite la publicación o difusión de los resultados de la investigación científica a excepción de información de uso restringido como puedan ser listados de palabras y/o programas en código fuente. Por este motivo sólo aparecen en el texto ejemplos representativos de la teoría que se enuncia, y no los listados lexicográficos que la justifican.

El presente estudio se plantea el recuento exhaustivo de las colocaciones que forman los de las tres clases principales de palabras, Nombre y Adjetivo, trabajo que, si la ocasión lo permite, se proseguirá con el estudio de las combinaciones de Verbo y Nombre. La práctica mayoría de las lenguas poseen una diferenciación de estas tres clases de palabras en el discurso, por lo que una descripción de las combinaciones usuales en cada lengua particular puede suponer una vía de comparación y relación muy útil. En el caso de las lenguas que no poseen la categoría de adjetivo existe, sin embargo, alguna manera de expresar los conceptos de los adjetivos, ya sea a través de verbos intransitivos, como es el caso del chino (Dixon, 1982), o de algunos nombres y algunos verbos combinados con ciertas partículas, por lo que las comparaciones

entre las colocaciones o estructuras sintagmáticas de estas lenguas también podría llegar a efectuarse.

Es necesario advertir que en este estudio no se tienen en cuenta partículas que, aunque consideradas tradicionalmente adjetivos, tienen un comportamiento claramente distinto al de los llamados adjetivos descriptivos o calificativos que poseen un correlato conceptual; me refiero a los demostrativos, los posesivos, los numerales, los indefinidos, y los determinativos en general. Estas partículas carecen de representación en muchos modelos debido a que, dado su carácter subsidiario, son 'elevados', es decir, pasan a convertirse en rasgos de las categorías temáticamente relevantes o argumentos. Así, en el sistema ATLAS de traducción automática, que ha servido de plataforma a nuestro trabajo, los adjetivos determinativos no tienen una representación conceptual y son plasmados en la interlingua como arcos unarios que aportan información añadida a la información central representada por los nodos de conceptos.

Los adjetivos descriptivos o calificativos, por el contrario, se conciben como argumentos del nombre.

En la elaboración del listado de colocaciones de nombre y adjetivo que ha servido de base al siguiente estudio se ha consultado principalmente el diccionario bilingüe *COLLINS Español-Inglés Inglés-Español* (Smith, 1988) y el monolingüe de María Moliner (1966) *Diccionario de uso del español*.

Agradecimientos

Debo agradecer al jefe del proyecto de traducción automática de Fujitsu España, Jorge Vivaldi, que creyera desde el principio en el interés científico de esta investigación. Por motivos similares estoy en deuda con el director de la tesis, el doctor Ramon Cerdà Massó de la Universidad de Barcelona y director del proyecto EUROTRA. Asimismo gracias a los compañeros y profesores que se prestaron siempre a comentar cuestiones relativas a la lingüística computacional, una lista de los cuales sería demasiado larga y correría el riesgo de ser incompleta. Mencionaré, sin embargo a Kathleen McKeown del Computer Science Department en Columbia University, quien no dudó en enviarme ciertos artículos relacionados con el tema de las colocaciones, así como a Javier Gómez Guinovart de la Universidad de Vigo y a Joan Torroella de la U.A.B. por la misma razón.

Por último deseo agradecer el apoyo de mi marido, Ramon Fabregat, y de los abuelos maternos y paternos de mi pequeña Júlia, quienes tan bien me han suplido cuando ha sido necesario. Y a ella, su maravillosa sonrisa.

Barcelona, 28 de junio de 1993

PREFACIO.	V
-------------------	---

PRIMERA PARTE

COLOCABILIDAD	1
-------------------------	---

0. Introducción	1
---------------------------	---

1. El significado	7
1.1. El significado colocacional	11
1.2. El significado conceptual	14
1.3. La asociación entre palabras y conceptos	20
1.3.1. Hiperonimia e hiponimia	24
1.3.2. La falacia de la sinonimia conceptual	28
1.3.2.1. Colocaciones equivalentes	36
1.4. Pertinencia de una selección léxica	38

2. Historia del concepto de 'colocación'	43
2.1. La selección léxica por colocación	55
2.1.1. La selección prioritaria	57
2.2. Relaciones funcionales del léxico	61

3. Colocación y fijación	65
3.1. Combinaciones libres	77
3.2. Combinaciones preferentes	79
3.3. Combinaciones fijas	83

4. Niveles de coocurrencia	87
4.1. La coocurrencia sintáctica	88
4.2. La coocurrencia semántica	92
4.2.1. La elaboración de un árbol semántico	95
4.3. La coocurrencia léxica	102
4.3.1. Colocaciones y derivación	107

5. El diccionario electrónico	113
5.1. El diccionario de conceptos	117
5.2. El diccionario de palabras	121

5.3. El diccionario de colocaciones	128
---	-----

SEGUNDA PARTE

LAS COLOCACIONES DE NOMBRE Y ADJETIVO	133
---	-----

0. Introducción	133
-----------------------	-----

1. Semántica del adjetivo	136
---------------------------------	-----

2. Sintaxis del adjetivo	141
--------------------------------	-----

3. Clasificación sintáctica del nombre.	152
--	-----

3.1. Nombres derivados deverbales intransitivos (NDVI).	153
--	-----

3.2. Nombres derivados deverbales transitivos (NDVT)	154
--	-----

3.3. Nombres derivados deverbales completivos (NDVC)	156
--	-----

3.4. Nombres derivados deadjetivales (NDD)	159
--	-----

3.5. Nombres derivados funcionales (NDF)	161
--	-----

3.6. Nombres simples funcionales (NSF)	163
--	-----

3.7. Nombres simples estativos (NSE)	164
--	-----

4. Clasificación colocacional del adjetivo	168
--	-----

4.1. Adjetivos no_atributivos terminológicos (ANTT)	169
--	-----

4.2. Adjetivos no_atributivos enfáticos (ANTE)	172
--	-----

4.3. Adjetivos atributivos polares existenciales (ATPE)	174
--	-----

4.4. Adjetivos atributivos polares perfectivos (ATPP)	176
--	-----

4.5. Adjetivos atributivos polares adverbiales de acción (ATPVA)	179
---	-----

4.6. Adjetivos atributivos polares adverbiales de agente (ATPVG)	181
---	-----

4.7. Adjetivos atributivos polares adverbiales de objeto (ATPVO)	183
--	-----

4.8. Adjetivos atributivos no polares (ATNP)	185
--	-----

5. La implementación en la gramática de generación	187
--	-----

5.1. El formalismo de ATLAS y su entorno	189
--	-----

5.2. Las reglas restrictivas colocacionales	193
---	-----

6. La construcción automática de un diccionario de colocaciones	199
--	-----

7. El orden del adjetivo	203
8. Conclusiones.	217
BIBLIOGRAFÍA	220

¡Error! Marcador no definido.PRIMERA PARTE

COLOCABILIDAD

0. Introducción

Una consideración que me parece importante hacer previamente es que el propósito de esta tesis no es escribir un tratado sobre semántica general. Tampoco pretende ofrecer una nueva teoría semántica ni hacer un estudio especulativo sobre el significado. Se trata de configurar un modelo que funcione en un sistema de traducción automática dando una visión, quizá idealizada, del ámbito de las relaciones léxicas, en espera de su incorporación a una teoría más comprehensiva que defina la naturaleza de las descripciones semánticas. Esa teoría general debería ser concebida como una teoría de la traducción (Jhonson&Rosner, 1987) que proporcionara una idea global sobre la resolución de problemas cruciales como el de la ambigüedad léxica, la ambigüedad estructural (alcance conjuntivo, dependencia de los sintagmas preposicionales y estructura interna de los compuestos), y las asociaciones textuales como la anáfora y la elipsis.

El éxito de un sistema de traducción automática depende de una teoría de la traducción coherente, además de un entorno

adecuado que permita una implementación rápida y la evaluación de las teorías experimentales. Pero no se puede modificar la teoría si se carece de ella. En realidad los sistemas actuales funcionan "libres" de teoría, o, mejor dicho, tienen teorías varias embutidas ahí, los estamentos relevantes de las cuales no están explicitados o están sepultados en una masa de códigos.

Es cierto que la traducción es demasiado caótica para constituir una teoría, pero también que las teorías lingüísticas vigentes son poco realistas: trabajan sobre abstracciones del lenguaje y no sobre el lenguaje real, de manera que se convierten en algo inservible. Un sistema de traducción automática precisa de una enormidad de conocimiento lingüístico que está interrelacionado. Quizá sea aventurado hablar de una "teoría de la traducción", pero se precisa de una teoría basada en ella o que resulte del análisis de sus problemas, dado que la traducción no puede descansar ni fructificar en ninguna de las teorías actuales con todos sus méritos.

Sin embargo, aun lejos de esta utopía, es inevitable y a la vez conveniente que nos planteemos ciertas cuestiones relativas al significado y nos situemos en nuestro terreno, la lexicografía, habiendo tomado cierta perspectiva con referencia al conjunto. Es conveniente, sobre todo, porque el tema que nos ocupa entra de lleno en lo que podría llamarse "nivel semántico-léxico", área que ha sido soslayada prácticamente por la lingüística americana y, aunque en menor medida (cf. Coseriu, 1977) por los estructuralistas y la

lingüística germánica tradicional, y que, por lo tanto, carece de marco teórico propiamente dicho. Tampoco sobre cuestiones más generales como "la diversidad de formas por las cuales establecemos, en la práctica, el significado de determinadas palabras, la "circularidad" del vocabulario y la pertinencia del "contexto", se hace un completo reconocimiento teórico en la semántica tradicional" (Lyons, 1968:424).

Es en parte debido a este vacío que teorías lingüísticas como la GB (Chomsky,81), LFG (Bresnan, 1982) o HPSG (Gazdar, 1985) no han cuajado como marcos teóricos sobre los que desarrollar sistemas de traducción automática. Esto no significa, sin embargo, que una teoría de la traducción como la que pensamos que actualmente se precisa, tenga que ser verificada obligatoriamente pasando la prueba de su capacidad para ser implementada en un sistema existente. La lingüística computacional, que no es sino un aspecto englobado en lo que hoy se denomina Inteligencia Artificial, ha sufrido durante toda su existencia la ambición de quienes han pretendido obtener a toda costa un 'producto' que sirviera para acrecentar su lucro o su prestigio, y debería, a nuestro juicio, empezar a ser considerada una disciplina que trata el lenguaje como proceso cognitivo, para lo cual utiliza modelos, pero no está exclusivamente a su servicio.

Por lo general, la gente que trabaja con ordenadores suele estar acostumbrada a resolver problemas acotados, del estilo

de enigma o crucigrama, pero en la vida real los problemas no son de este tipo: "to me it seems very clear in practice that the kind of problems which have to be solved in translation are often much more like the open-ended sort of problems encountered by research scientists than like the complex-but-bounded problems printed in the puzzle corner of a Sunday paper" (Sampson, 1987:103).

El número de soluciones de un problema, en programación, es un sistema perfectamente cerrado. Está definido por la sintaxis; si sabes qué debe hacer el programa, hacer que funcione correctamente es puro razonamiento deductivo. El proceso es de conjeturas y refutaciones. Según Sampson, tentativa y error son alternativas heurísticas convenientes, pero esto no es extrapolable a otras esferas de la actividad mental. La comprensión del lenguaje es una actividad popperiana, creativa, falible, no-algorítmica, y la idea de una simulación computacional se desvanece pronto.

Ciertamente esta fue la conclusión de la primera generación de estudiosos de la traducción automática. El ALPAC (Automatic Language Processing Advisory Committee) realizó en 1966 un informe con las conclusiones sobre la primera etapa de investigaciones en el que decía que la FAHQUT (Fully Automatic High Quality Translation of Unrestricted Text) era un objetivo a largo plazo y que el trabajo en traducción automática se enfrentaba con la barrera semántica. (Es lógico, pues, pensar que la investigación técnica en este campo debe plantearse resolver

cuestiones relacionadas con la desmembración de esa barrera.)

El informe del ALPAC reconocía la aportación de la lingüística computacional a la investigación sobre el lenguaje, pero descartaba el hecho de que este campo se utilizara como método para probar teorías nuevas. Ahora bien, dado el carácter especulativo que afecta a las teorías semánticas en general y la ingente cantidad de conocimiento lingüístico que se precisa para crear un sistema experto de este tipo, nosotros opinamos que la traducción automática puede ofrecer un marco modélico para la construcción de una teoría que funcione a su vez como modelo de la realidad, y que en su intento de controlar ciertos fenómenos aporte una aclaración sobre cuáles son los problemas que se oponen al avance de un modo crucial y cuáles pueden ser relegados a la investigación posterior.

Naturalmente que un modelo que se aplique a un sistema de traducción debe ser una descripción general sobre la lengua, y esta debe intentar ser lo menos fragmentaria posible. Así, un sistema como ATLAS trata todos los fenómenos relacionados con el lenguaje escrito desde el punto de vista de la morfología, la sintaxis y la semántica. Ahora bien, el tema de la equivalencia de la traducción no deja de ser una cuestión semántica. Como suele decirse, traducir es encontrar una expresión sinónima en la otra lengua. Por eso, aunque nuestro propósito no es construir una teoría semántica general que suponga "universales" semánticos, sí está cerca de la elaboración de una teoría específica que prediga los enunciados básicos existentes en la competencia

lingüística de los hablantes de una lengua.

1. El significado

Dar una definición de lo que es el significado es algo que ha preocupado a muchos autores infructuosamente. Como dice Leech (1985:21) "un físico no se ve precisado a definir nociones como las de "tiempo", "calor", "color" o "átomo" antes de comenzar a investigar sus propiedades: las definiciones, si son necesarias, surgirán del estudio mismo". En realidad no podemos estar muy seguros de que exista algo que responda a lo que denominamos significado. Cuando tratamos de explicar el significado de una palabra, de una frase, de un fenómeno o de un conjunto de pensamientos como pueda ser una teoría, no hacemos sino paráfrasis, sustituciones, ampliaciones, restricciones o cualquier otro tipo de reorganización con el lenguaje. El significado surge de la comparación entre el enunciado inicial y el final; raramente el segundo aislado "significa" más que el primero. Esto es así porque la lengua no es más que un sistema semiológico particular, aunque, eso sí, el más complejo y extendido de los sistemas semiológicos, susceptible de convertirse en modélico. De este modo, ante la arbitrariedad de un signo, aunque sea lingüístico, siempre existe la posibilidad de enunciar un "significado" más convencional en términos de lenguaje. En este sentido, cuanto más se avance en el intento de fijar el universo de realidades humanas (mediante un diccionario de conceptos, por ejemplo), más se avanza en el dominio de la semiología y, por ende, en el de la lingüística.

Visto así, el significado de una enunciación, tal como los hablantes lo expresamos, no sería más que una **solución**, pero, en el sentido de Lakoff & Johnson (1980:186). Una 'solución' que debe ser tratada como una "metáfora QUÍMICA" y no como una "metáfora del ROMPECABEZAS":

En el momento presente la mayoría de nosotros tratamos los problemas según lo que podríamos llamar la metáfora del ROMPECABEZAS, en la que los problemas son ROMPECABEZAS para los que característicamente existe una solución correcta y una vez encontrada están solucionados para siempre. (...) Vivir mediante la metáfora QUÍMICA sería aceptar como un hecho que ningún problema desaparece para siempre. Más que dirigir las energías a resolver los problemas de una vez por todas, uno las dirigiría a encontrar los catalizadores que disolvieran sus problemas más acuciantes durante el mayor tiempo posible, y sin precipitar otros peores. La reaparición de un problema se consideraría un hecho natural, más que un fallo por parte de uno a la hora de encontrar la forma correcta de solucionarlo.

Citar aquí el trabajo de estos autores para sustentar nuestro esfuerzo en la construcción de un sistema conceptual cuando su obra critica tan rotundamente lo que ellos llaman el "mito objetivista" puede parecer contradictorio. Pero la originalidad y repercusión de su punto de vista no puede ser puesta en duda y la asumimos aquí de igual modo que un estado de derecho debería tener en cuenta pensamientos anarquistas de grandes ideólogos como Proudhon, aunque las revoluciones que sugieren no puedan ser un objetivo a corto plazo.

Según Lakoff & Johnson (1980:241), el objetivista caracteriza el significado puramente en términos de condiciones

de verdad o falsedad, cuando la verdad "no es absoluta u objetiva sino que se basa en la comprensión", porque "el significado no es nunca incorpóreo u objetivo y siempre está fundamentado en la adquisición y el uso de un sistema conceptual y las metáforas que lo estructuran".

La explicación de la metáfora convencional que dan estos autores va en contra de toda una tradición que ha dominado la cultura occidental desde los presocráticos hasta hoy, tanto en las corrientes empiristas como en las racionalistas como en la síntesis kantiana; esta corriente se ha preservado en los descendientes de los positivistas lógicos, la tradición fregeana, la tradición de Husserl, y en la lingüística en la noción de verdad de Tarski y el neorracionalismo surgido con la tradición chomskiana. Al considerar la metáfora como algo esencial a la comprensión humana y un mecanismo para crear nuevo significado y nuevas realidades en nuestras vidas, así como la base de la mayor parte de la estructura conceptual de un lenguaje natural, Lakoff & Johnson se enfrentan con la mayor parte de la tradición filosófica anglosajona que ha considerado la metáfora como un agente del subjetivismo y, en consecuencia, como subversiva en la búsqueda de la verdad absoluta.

Esta visión del significado como algo que no es nunca incorpóreo u objetivo, que siempre está fundamentado en la adquisición y el uso de un sistema conceptual y las metáforas que lo estructuran, no cuestiona, sin embargo, la existencia de ese sistema conceptual, ni tampoco la posibilidad de que

se produzca equivalencias de significado en una y otra lengua.

En ese sentido es interesante el esfuerzo por sistematizar la extensa lista de clases potenciales de metáforas enunciada por Lakoff and Johnson (1980) y su posible aplicación para la resolución de problemas de asignación de rasgos semánticos en el 'parsing' (Carbonell, 1987). Sin embargo, en generación, tarea a que nos remitimos centralmente, la cuestión es la contraria: producir lenguaje metafórico.

Cuando se trabaja en un sistema limitado técnicamente y se pretende algo tan concreto como es la traducción de un enunciado, el lingüista debe aceptar y asumir ante todo, en un esfuerzo de humildad obligado, que no existen unas condiciones de verdad y falsedad universales; no obstante, también se hace preciso confiar en que existe una realidad humana básicamente objetiva, y que las convenciones, metafóricas o no, a las que se llega mediante el lenguaje, pueden ser a su vez convenciones para una mente electrónica independientemente del proceso de codificación y decodificación. El estudio de las colocaciones pretende ser un paso desde la hipotética objetividad del sentido hacia la real subjetividad de la comprensión.

1.1. El significado colocacional

La polémica que engendra la distinción entre mundo, verdad y lenguaje es antigua. El hombre padece unas limitaciones en el intento de conocer la realidad tal cual es. Por ese motivo la imagina, la "recrea" a partir de los datos que le es posible manejar, y establece unos modelos que le sirven como herramientas de control, y luego otros modelos para controlar a los primeros, y así sucesivamente. Que esos modelos, como pueda serlo el lenguaje, no son la "realidad" es una conciencia que nunca debe perderse.

Desde esta premisa los intentos de algunos autores por descubrir el mundo a través del estudio de los mecanismos que lo expresan nos parecen equivocados, y, sin embargo, es una actitud frecuente. Casi todos los estudios topológicos que carecen de un objetivo "técnico" adolecen de ello. Por poner un ejemplo, Pérez Calvo (1986:116) trata de deducir, a través del estudio del adjetivo, que el tiempo fluye al menos tridimensionalmente, que existe una cuarta dimensión y que esto influye sobre el hombre. Según nuestro punto de vista "el tiempo" es el tema de una tesis de física o de una disciplina que aún no se ha inventado. Sí es cierto que el conocimiento del modelo ayuda en el conocimiento de la realidad que lo ha inspirado, y viceversa, pero no debe olvidarse hasta qué punto el lenguaje es una convención creada para ser útil al individuo y a la sociedad. Si queremos que una investigación sobre el lenguaje sea a su vez algo útil, debemos tomar plena conciencia de la manera en que funciona "convencionalmente", antes de concebirlo como herramienta de

conocimiento o de creación.

A este respecto afirma Leech (1985:22) que es conveniente para un estudioso de la semántica no salirse del lenguaje en el estudio del significado: "la búsqueda de una explicación de los fenómenos lingüísticos apoyándose en lo que no es lenguaje es tan vana como la tentativa de salir de una habitación que no tenga puertas ni ventanas, ya que la misma palabra 'explicación' implica un enunciado del lenguaje". La solución es, pues, conformarse con explorar "lo que hay dentro de la habitación, es decir, estudiar las relaciones que existen 'dentro' del lenguaje".

Hasta aquí compartimos plenamente el enfoque de Leech; ahora bien, dado que su postura es esencialmente mentalista, para este autor el estudio del significado que él llama "conlocativo" es uno de los "menos importantes" de entre los siete que distingue y de los que a continuación transcribimos un cuadro sinóptico:

Los siete tipos de significado de Leech (1974:45):

1. SIGNIFICADO CONCEPTUAL o sentido		Contenido lógico, cognoscitivo o denotativo
SIGNIFICADO ASOCIATIVO	2. SIGNIFICADO CONNOTATIVO	Lo que se comunica en virtud de aquello a lo que se refiere el lenguaje.
	3. SIGNIFICADO SOCIAL	Lo que se comunica sobre las circunstancias sociales del uso del lenguaje.
	4. SIGNIFICADO AFECTIVO	Lo que se comunica sobre los sentimientos y actitudes del que habla o escribe.
	5. SIGNIFICADO REFLEJO	Lo que se comunica merced a la asociación con otro sentido de la misma expresión.
	6. SIGNIFICADO CONLOCATIVO	Lo que se comunica merced a la asociación con las palabras que suelen aparecer en el entorno de otra palabra.
7. SIGNIFICADO TEMÁTICO		Lo que se comunica por la forma en que el mensaje está organizado respecto del orden y el énfasis.

Es obvio que todos los aspectos significantes en un acto comunicativo son igualmente importantes. Ahora bien, desde el punto de vista de la elaboración de un diccionario que funcione en un sistema empírico, y que suministre los datos que permitan trabajar en este campo desde una perspectiva científica y lo

menos intuitivamente posible, el efecto conlocativo es el más fácilmente perceptible: puede hallarse en los hablantes, en los documentos, en los diccionarios, y no precisa de una especulación. Incluso diré mas: el significado conlocativo es el único que puede fijarse mediante un proceso de extracción realizado con herramientas informáticas y puede ser determinado con rigor estadístico (trabajo que sería posterior a este, en el que se intenta observar la pertinencia del posible resultado para la resolución de tareas específicas).

Lyons (1968:424) extrae los siguientes principios de la obra tardía de Wittgenstein (aunque avisa de que este autor no propuso, ni intentó proponer, una teoría del "uso" de las palabras como una teoría semántica): "El único control empírico que tenemos sobre el estudio de la lengua es el "uso" de las expresiones lingüísticas en las variadísimas situaciones de la vida diaria. Expresiones como el "significado de una palabra" y "el significado de una oración (o proposición)" son peligrosamente engañosas en cuanto que nos incitan a buscar los "significados" que tienen y a identificar sus "significados" con entidades tales como objetos físicos, "conceptos" de la mente o bien "estados de cosas" en el mundo físico".

1.2. El significado conceptual

Respecto al peso del significado conceptual, denotativo o lógico, que Leech llama también 'sentido', no existe ninguna

duda de que es preciso tenerlo en cuenta en un sistema de traducción automática y en cualquier intento de establecer una comparación entre las lenguas. Aún en el caso de que se pretenda crear un sistema basado en la transferencia de árboles sintácticos, la comparación de las estructuras que se dice son equivalentes en una lengua y otra supone implícitamente la existencia de conceptos similares expresados mediante construcciones distintas. Una posibilidad para plasmar el significado conceptual es enunciarlo y asociarlo con una palabra. La información que aporta el significado denotativo constituye el diccionario de conceptos en el sistema de diccionario electrónico de ATLAS que se describe en el apartado 5.1. de la segunda parte. ATLAS, modelo de representación que ha servido de entorno experimental a este trabajo, es un sistema de traducción basado en un diccionario electrónico tripartito: un diccionario de palabras, un diccionario de coocurrencias y un diccionario de conceptos (ver el apartado 5). Cada concepto se identifica con un código alfanumérico (asociado a su respectivo 'headconcept' o enunciado en inglés) que constituye el/los valor/es CID ("Concept IDentification") de cada entrada del diccionario.

Aunque es imprescindible recoger este significado denotativo y enunciarlo para cada palabra o diferenciarlo por medio de un símbolo cualquiera, la objetividad en cuanto al grado de abstracción que se precisa para que el diccionario de conceptos dé cuenta del sistema conceptual de cada lengua de una manera coherente es discutible (ver apartado 1.3.). Cuanto

mayor es la distancia topológica entre la lengua origen y la lengua destino, menor es la biunivocidad léxica, siendo este un problema subsidiario a la ya de por sí difícil biunivocidad entre palabras (signos) y objetos conceptuales.

Sucede que el universo de las palabras y el universo de los conceptos se nutren mutuamente. La comunicación de una idea nueva o de la matización de una idea puede producir nuevo léxico o provocar ampliaciones o restricciones de significado. El uso que los hablantes hacen de los términos, asociándolos a realidades distintas a través de un mecanismo metafórico innato, constituye con el paso del tiempo un sistema de relaciones característico para cada sistema conceptual.

La tarea de crear un diccionario de conceptos y su relación con las denominaciones constituye la **onomasiología**, ciencia que parte de los conceptos para encontrar las realizaciones lingüísticas en una o más lenguas. Según esta disciplina, el concepto de "concepto" no está unido a una lengua determinada como lo estaría el concepto de "significado" o "contenido".

La onomasiología es una disciplina complementaria de la **semasiología**, de tal manera que la primera corresponde al punto de vista del emisor y la segunda al del receptor. El proceso de codificación del emisor se inicia con la "visión" del concepto y se transforma en lenguaje, mientras que el receptor recibe un mensaje que debe decodificar interpretando semánticamente

las señales hasta llegar a la "visión" del concepto.

Obsérvese que las fases de la traducción automática coinciden con estos enfoques en modo inverso, de tal manera que el proceso de análisis es semasiológico, mientras que el de síntesis es onomasiológico.

Ahora bien, la construcción lógica de pirámides conceptuales resulta muy problemática porque los conceptos son inevitablemente expresados por medio de lenguajes particulares a pesar de no estar inmanentemente ligados a ellos. Según Baldinger (1960:522), el concepto es "una abstracción obtenida a partir de la pluralidad individual de la realidad objetiva (objeto), es decir, una idea que teóricamente no está unida al cuerpo de una palabra".

La cuestión de compartición de conceptos por parte de las distintas lenguas es importante. Cuanto mayor sea la frecuencia de uso de una palabra más y mayores serán las discrepancias. Esto se debe, en último término, a la relación inversamente proporcional que existe entre la extensión y la intensión de un concepto ¹: cuanto mayor es la extensión de un concepto, menor es su intensión, y viceversa.

Obsérvese las múltiples traducciones de la palabra "fuerte"

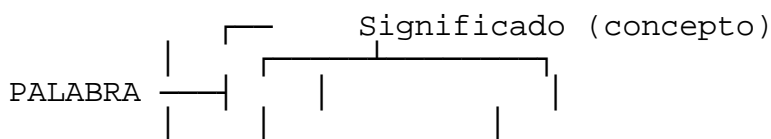
¹ La extensión de un concepto, frente a la intensión o contenido (clase de los atributos o propiedades que le caracterizan), es la clase de los objetos a los que se aplica el concepto. Por ejemplo, la extensión del concepto "flor" es el conjunto de todos los individuos a los que se aplica el predicado o contenido conceptual o intensión de "flor" (esencia, función, forma, color, aroma, etc.).

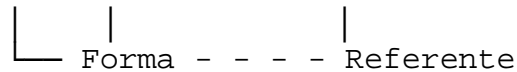
en una lengua relativamente cercana como es el inglés:

persona <i>fuerte</i>	<i>strong</i> person
ruido <i>fuerte</i>	<i>loud</i> noise
golpe <i>fuerte</i>	<i>hard</i> strike
lluvia <i>fuerte</i>	<i>heavy</i> rain
comida <i>fuerte</i>	<i>rich</i> food

¿Se debe esto a que el castellano asocia esta palabra con un concepto que posee un grado de abstracción superior -o menor intensidad-, o bien sucede que la misma palabra se asocia a distintos conceptos? Si la extensión de un concepto depende de su intensidad y hasta que no se determinan los rasgos esenciales del contenido conceptual no puede definirse su extensión, esa es la tarea que parece ser la prioritaria y que debe emprenderse a pesar de que, y parece inevitable, haya una lengua que imponga su sistema conceptual (ver apartado 1.3.).

La distinción lógica entre extensión e intensidad que usan algunos semantistas como Carnap (1956) puede enunciarse también, y así lo hace Lyons (1968:467) en términos de sentido y referencia. La gramática tradicional afirmaba que la palabra resultaba de la combinación de una determinada forma con un determinado significado. Según Lyons, "la relación entre forma y referente es indirecta, pues la forma se relaciona con su referente a través del significado intermedio (conceptual) el cual se asocia con ambos de un modo independiente" (p.418):





Ahora bien, sea como fuere que se constituyera el diccionario de conceptos ideal (o como la onomasología indicare que fuera), las asociaciones de nombre y adjetivo que aparecen en el ejemplo de la página anterior deberán constar, al menos para el inglés, en un diccionario de colocaciones. Del mismo modo deberá existir un diccionario de colocaciones del castellano que dé cuenta de fenómenos como el siguiente:

<i>bunch</i> of flowers	<i>ramo</i> de flores
<i>bunch</i> of grapes	<i>racimo</i> de uvas
<i>bunch</i> of bananas	<i>cabeza/racimo</i> de plátanos
<i>bunch</i> of keys	<i>manejo</i> de llaves
<i>bunch</i> of people	<i>grupo</i> de gente
<i>bunch</i> of hair	<i>cola/coleta</i> de pelo

o este otro:

<i>dim</i> person	persona <i>corta</i> (de luces)
<i>dim</i> light	luz <i>débil</i>
<i>dim</i> sight	visión <i>turbia/borrosa</i>
<i>dim</i> room	habitación <i>oscura/sombría</i>
<i>dim</i> outline	perfil <i>confuso</i>
<i>dim</i> opinión	opinión <i>desfavorable</i>
<i>dim</i> memory	recuerdo <i>borroso/lejano</i>
<i>dim</i> colour	color <i>apagado</i>

No cabe presumir que, aun siendo muy detallado el diccionario de conceptos, provea del suficiente número de conceptos para que cada uno esté asociado con una palabra solamente. Por el contrario, y sobretodo cuando se pretende que el diccionario de conceptos funcione en un sistema multilingüe, que no bilingüe, sucede que el número de conceptos es inferior al total de palabras

de todas las lenguas involucradas.

Así y todo, cabe considerar que el número de conceptos es presumiblemente infinito. La capacidad de los pueblos y de las personas para 'comprender' la realidad (la cual continuamente aprehenden) e imaginarla no parece tener límite. Obviamente, los sistemas de almacenaje sí tienen un límite (que por otro lado se amplía vertiginosamente de la mano del progreso tecnológico), sobre todo desde el punto de vista de la "manejabilidad" de los datos y del control de sus interrelaciones.

Pero la cota es ya lo bastante amplia para establecer, no sin esfuerzo, un diccionario con los conceptos que más usualmente se barajan en la comunicación cotidiana y asociarlos a las palabras que los sostienen al menos convencionalmente para sectores de población más o menos amplios.

1.3. La asociación entre palabras y conceptos

Como se desprende, pues, el número de palabras y el número de conceptos no es el mismo. Las palabras se asocian al significado denotativo, pero este no tiene porque ser exclusivamente la definición científica de una realidad, sino que también debe haber un enlace con el significado de uso popular, los usos metonímicos que la palabra protagoniza, las interpretaciones metafóricas o idiomáticas, etc. Por ejemplo, la palabra 'hierro' irá asociada a los siguientes conceptos (entre varios):

3bcc38 "a metallic element of atomic number 26 named
iron"
0c468c "something made of iron"
0c468f "a branding iron"
0c4691 "a soldering iron"

Y también sucede que existen conceptos que tienen asociada más de una palabra. Por ejemplo el concepto 3ce892 "a condition of being well, strong, hard" que va asociado a las palabras castellanas siguientes: *firme, determinante, reforzado, fuerte, robusto, resistente, fornido, enérgico, tenaz, vigoroso y sólido*.

Es en estos casos en que se hace patente la necesidad de un diccionario de coocurrencias que seleccione la prioridad combinatoria que existe a nivel léxico. Desde nuestro punto de vista, el significado conlocativo está infravalorado en la semántica de Leech, y este autor entra en contradicción cuando dice que "lo que hace falta, si se quiere fundamentar la semántica del modo más seguro posible, es en primer lugar un control de la forma en que se utiliza la intuición y, en segundo lugar, una exploración de los medios por los que se pueden basar más firmemente los datos intuitivos en otras clases de evidencia" (1974:107).

Ahora bien, este tipo de estructuras lexemáticas sintagmáticas que resultan de la preferencia de una lengua por la combinación de determinadas palabras están motivadas por el contenido, la historia, la existencia de otros lexemas; no son combinaciones determinadas sólo por las propiedades objetivas de la realidad extralingüística. Coseriu (1977:184)

acusa varias veces en su libro a B.Pottier de confundir el mundo real con el mundo lingüístico cuando dice que "il n'y a qu'une infime probabilité pour qu'une olive soit blanche, et une très grande pour qu'une mouette le soit" (1964:130), lo que le lleva a afirmar que "gaviota blanca" es un "hecho de lengua". Para Coseriu "mouette blanche" no es un hecho de la lengua francesa porque esta combinación no está determinada lingüísticamente, sino que se debe al conocimiento de las cosas como tales.

También Leech (1985:24) concede importancia a esta distinción y afirma que ciertas combinaciones carecen de sentido porque solo son imaginables en un mundo onírico o novelesco ("dormir sobre las puntas de los pies", por ejemplo), pero otras son contradicciones lógicas que responden a absurdos lingüísticos (por ejemplo, "dormir despierto"). Estas últimas solo significan algo si se juega a trastocar las reglas del juego lingüístico y se conciben metafóricamente.

En nuestra opinión el hecho de que se pueda establecer un juego lingüístico como el de la metáfora no es un criterio robusto para distinguir la inaceptabilidad combinatoria referida al mundo real y la inaceptabilidad referida al mundo lingüístico, puesto que metáforas las hay de todos los tipos. Sin embargo, sí que nos parece interesante establecer una diferencia entre las combinaciones que ocurren porque la realidad las dicta y las combinaciones que son producto de una estructura preferente

de una determinada lengua. Las primeras tratan de las relaciones entre los "referentes", mientras que las segundas son relaciones entre formas. Para las primeras siempre podrán imaginarse situaciones o mundos oníricos en que cualquiera de ellas sea posible, por ejemplo: "el pájaro sabio dijo al pajarito que no fuera al colegio sino que se quedara a meditar con él". Los rasgos semánticos son los que se encargarían de indicar la falsedad de esta oración, a no ser que pudiera contarse con la posibilidad de reestructurarlos para el tratamiento de discursos fabulísticos, novelescos u oníricos, lo cual, ahora mismo, es absolutamente improbable. Este es otro argumento en contra de la utilización de restricciones que utilicen solamente rasgos semánticos (ver apartado 4.2.1.).

Ahora bien, lo que nos preocupa es cómo, aun cuando no se alterara la lógica restrictiva impuesta por los rasgos semánticos, las relaciones que afectan a las formas pueden no ser correctas. Obsérvese el siguiente ejemplo:

"el cantante tomó una enfermedad genital debido a su movida vida venérea y a la promiscuidad de sus contactos de los cuáles poseyó hijos ilegales"

Se desprende del párrafo que ha habido un error en el uso de las colocaciones siguientes: "contraer una enfermedad", "enfermedad venérea", "vida sexual", "vida sexual activa", "relaciones promiscuas", "tener hijos", "hijo ilegítimo", y que la redacción correcta es la que sigue:

"el cantante contrajo una enfermedad venérea debido a su activa

vida sexual y a la promiscuidad de sus relaciones, de las cuales tuvo varios hijos ilegítimos"

Ahora bien, ¿cómo sabrá una persona hablante de otra lengua -o bien una mente electrónica- cuáles son las colocaciones adecuadas y cuáles las incorrectas si se barajan conceptos similares o idénticos?

1.3.1. Hiperonimia e hiponimia

La hiponimia también se denomina "inclusión" y puede formalizarse a través de la lógica de clases. Se dice que una palabra con mayor extensión y menor intensión incluye a otra con menor extensión y mayor intensión, por ejemplo *animal* a *elefante*, y que por lo tanto el significado de *elefante* está incluido en el de *animal*. Es decir, la inclusión incide al contrario que la intensión: si A domina B en una taxonomía, el conjunto de cosas nombradas A incluye la extensión de B, como se ha dicho, pero la intensión de B incluye la intensión de A: la intensión de *elefante* contiene todos los atributos de *animal* y otros más.

Dice Lyons (1968:466) que "el tipo de relaciones de inclusión descansa sobre la noción de referencia (por cuanto opera con clases de entidades nombradas por datos léxicos)", pero, si bien es cierto que aparentemente las relaciones hiperónimas

o hipónimas se manifiestan en el nivel léxico, estas relaciones de inclusión afectan básicamente a la organización del diccionario de conceptos. Obsérvese que, aunque los conceptos están intrínsecamente asociados a las entradas, las relaciones de inclusión no se proyectan al sintagma nominal cuando existe modificación no intensional como pueda ser un adjetivo de medida, de tal manera que un elefante es un animal, pero un *elefante pequeño* no es un *animal pequeño* (siempre que no tomemos a 'animal' como sinónimo de 'ejemplar'). Por lo tanto parece claro que las relaciones de inclusión afectan a los conceptos de manera distinta a como afectan al nivel léxico, en el cual la identidad de las palabras en combinación ejerce una influencia recíproca. Esto quiere decir que, si bien parece lógico concebir jerárquicamente el diccionario de conceptos según una taxonomía que determine relaciones de hiponimia, esta estructura difícilmente puede aplicarse al léxico de las lenguas particulares si antes no se ha descrito el contexto lingüístico de ese léxico; a saber, a) la distribución de las categorías, b) la combinatoria de los distintos tipos de subcategorías y, finalmente, c) la distribución coocurrente de los lexemas según restricciones colocacionales.

Obsérvese que el orden en que planteamos la descripción de un sistema lingüístico es inverso al que debe existir en la ejecución de los sucesivos "filtrajes" en los procesos automáticos, en los que primero se utilizarán los datos más específicos, los de coocurrencia, luego la subcategorización

tipológica y luego la distribución sintáctica y las estructuras que se deriven (con la utilización de rasgos semánticos si cabe).

Pero, si bien el diccionario de conceptos es artificial y soporta con robustez las relaciones de inclusión que le imponamos, no por eso está exento de problemas que se derivan de su asociación con las lenguas particulares.

Puede ocurrir, y ocurre, que una cultura conciba la realidad de diferente manera a como lo hace otra, a veces motivada justamente porque la realidad se le muestra con formas distintas. Se puede poner el ejemplo típico de la multiplicidad de tipos de nieve que conocen los inuit (esquimales) frente a la univocidad del criterio occidental. En este caso caben dos posibilidades:

1) Todos los conceptos inuit están en el diccionario de conceptos, de tal manera que la palabra castellana "nieve" (como la inglesa "snow") esté asociada a todos y cada uno de esos conceptos con la etiqueta "u" de "upper-concept" o "hiperónimo" (ver apartado 5.2.).

2) En el diccionario de conceptos está el concepto de nieve que goza de un mayor nivel de abstracción, y todas las palabras inuit están asociadas a este concepto con la etiqueta "s" de "sub-concept" o "hipónimo".

Obviamente, dado que el pueblo inuit no participa en la investigación de momento, y que, de hacerlo algún día, no gozaría presumiblemente del poder fáctico necesario para cambiar las decisiones ya adoptadas por países económicamente más poderosos o con más antigüedad en el proyecto o que son usuarios de lenguas

mayoritarias, la segunda opción es a todas luces la más probable.

Sea cual fuere la decisión adoptada, parece claro que las relaciones de inclusividad permiten salvar las distancias entre las estructuras conceptuales de una lengua y otra, es decir, las diferencias en el nivel de abstracción que cada lengua impone al mundo que le rodea al considerar cruciales algunas diferencias e ignorar otras.

Claro que también es posible que el problema no se deba a un desnivel de abstracción, sino que el concepto ofrecido por una cultura no tenga correlato en la otra. Por ejemplo, el concepto *0eba33* del diccionario de conceptos procede del japonés y tiene un 'head concept' que reza "the first year of the new era"; se refiere al primer año de la nueva era iniciada por la subida al poder del nuevo emperador, lo que en Japón supone empezar a contar los años desde el principio. En castellano no existe tal concepto, y la solución de asociarlo a la palabra hiperónima o super-concepto "año" resulta a todas luces insuficiente. Lo único que se puede hacer es inventar una palabra compuesta que dé una idea aproximada de la forma más sintetizada posible, que de todas formas no será algo más corto que "primer año de la nueva era imperial". Pero, afortunadamente, estos casos son los menos si se comparan con los que pueden solucionarse a través de la inclusividad.

1.3.2. La falacia de la sinonimia conceptual

La sinonimia es un término que se usa para indicar "similitud de significado". La definición más asumida del fenómeno es la que proporciona Ullmann, 1957: "sólo pueden considerarse sinónimos aquellas palabras que permiten ser reemplazadas entre sí en todo contexto dado sin el más leve cambio en el valor tanto cognitivo como emotivo". Pero estos dos requerimientos, el de reemplazabilidad en todo contexto y el de identidad cognitiva y emotiva muy raramente se cumplen. Esto es muy natural, si se considera que la lengua opera según el criterio lógico de comunicar el máximo con el mínimo esfuerzo, con lo cual resulta "caro" mantener palabras de significado absolutamente "idéntico". Es cierto que en una lengua existen palabras distintas provenientes de préstamos o que, debido a cambios semánticos o a evoluciones paralelas, tienen significados muy parecidos, pero cuando esto ocurre los hablantes suelen matizarlas o relegar algunas usándolas de distinto modo para connotar, connotarse, restringir, o en la creación de modelos o patrones sintácticos (colocaciones) que le permiten ahorrar energía en las tareas semasiológicas y onomasiológicas de codificación y decodificación del significado.

Muchos autores mantienen que no existen sinónimos reales, mientras otros se esfuerzan en distinguir tipos o grados de sinonimia en términos a menudo vagos y poco clarificativos (ver Lyons, 1968:460, "sinonimia total", "sinonimia completa", y

"sinonimia normal"). Ocurre que no es lo mismo plantearse la similitud de significado desde el punto de vista del diccionario que verla actuar en el lenguaje real. Cuando el lexicógrafo utiliza los sinónimos en las definiciones de las entradas matiza y añade todos los detalles precisos. Pero si los sinónimos no son intercambiables en los distintos contextos y situaciones el establecimiento de sus relaciones resulta un trabajo prácticamente estéril.

Pérez Calvo (1986) niega la existencia de la sinonimia en favor de las relaciones hiposémicas, según las cuales las palabras de un campo semántico son ordenables unas con respecto a otras, pero la hiponimia no es un fenómeno fácil de precisar exhaustivamente sobre todos los elementos del léxico con independencia de las relaciones de dominio que un lenguaje particular mantiene en su estructura conceptual. Palmer (1976:76) señala que, desgraciadamente, no siempre existe un término superordenado o superior (superordinate term):

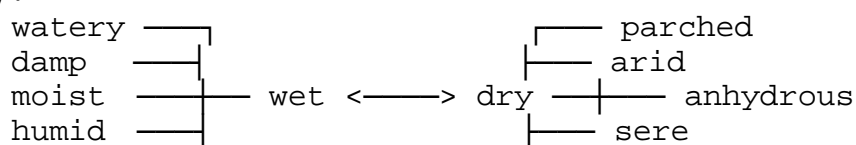
Lyons' own work led him to observe that in Classical Greek there is a superordinate term to cover a variety of professions and crafts, 'carpenter', 'doctor', 'flute player', 'helmsman', 'shoemaker', etc., but none in English. The nearest possible term is *craftsman*, but that would not include *doctor*, *flute player*, or *helmsman*. Similarly, and rather strangely, there is no superordinate term for all colour words, *red*, *blue*, *green*, *white*, etc.; the term coloured usually excludes *black* and *white* (and *grey* too), or else (used to refer to race), means 'non-white'.

Por este motivo no parece razonable optar por la hiponimia como una panacea universal, sobre todo en el nivel léxico. Las relaciones hiposémicas o hiponímicas deben estar relegadas al

nivel conceptual.

La postura de Gross&Miller (1990) nos parece quizá un punto de partida coherente entre los distintos intentos de jerarquizar el léxico, o lo que es lo mismo, en la organización de la información que deberá constar en el diccionario, porque deducen relaciones en el léxico dirigidas por oposiciones claras, es decir, por la antonimia.

Pero sucede también que, supuesta una relación sinonímica entre dos palabras, se produce a menudo una asimetría con respecto al antónimo que debiere ser común a ambos. Por ejemplo, dada una palabra como *profundo* se nos aparece claramente el antónimo de *superficial*, pero si consideramos sinónimos de esta palabra como pueden serlo *hondo*, *insondable*, *penetrante*, el mismo antónimo no funciona con la misma eficacia en todos los casos, ya que el antónimo de *insondable* parece ser *sondable*, mientras que el resto carece de antónimo si no es por su relación con *profundo*. Con respecto a esto Gross, Fisher & Miller (1989) distinguen entre **antónimos directos** y **antónimos indirectos**, de tal manera que se conciben opuestos conceptuales que no son parejas léxicas. Los grupos de sinónimos son considerados ramas asociadas por similitud semántica con un adjetivo central que relaciona al grupo con otro en contraste opuesto con el atributo de la palabra. Obsérvese el esquema proporcionado por Gross & Miller (1990:268):



soggy —┘

└— dried-up

Parece claro, pues, que el fenómeno de la sinonimia y el de la antonimia deben ser considerados a la vez, siendo determinante el segundo con respecto del primero, puesto que el hecho de compartir un antónimo produce grupos de palabras relacionadas o "sinónimos".

Por otro lado, como se recalca en el apartado 4.3., la misma palabra puede poseer distintos antónimos dependiendo del nombre con el que coloca. Retomando el ejemplo ejemplificado más arriba en torno a 'profundo', observamos:

sueño profundo/ligero
pensamiento profundo/superficial
aguas profundas/(poco profundas)
herida profunda/leve
noche profunda/clara
silencio profundo/roto

Partiendo de Palmer, F.R., 1976, concebimos cuatro tipos de sinonimia 'posible', siendo los dos primeros relativos a cuestiones de geografía lingüística o de sociolingüística más que a la semántica en sí, de tal manera que sólo el tercero y el cuarto se nos aparecen como fenómenos relativos al significado propiamente dicho:

1. Palabras que pertenecen a distintos dialectos de la lengua ubicados en áreas distintas. Es un fenómeno sincrónico que afecta mayormente al léxico rural (los nombres de las

herramientas, etc.), pero que no difiere del de las equivalencias existentes entre distintos idiomas. Simplemente, las personas usan formas distintas, pero estas no coexisten en la actuación de los mismos hablantes aunque puedan estarlo en su dominio.

2. Palabras que pertenecen a distintos registros. El problema es similar al anterior y Palmer sugiere tratarlo asimismo como si de palabras de distinto idioma se tratara. Pero en este caso los términos a menudo coexisten en la actuación de los hablantes, los cuáles pueden cambiar de uno a otro registro en la misma conversación. Un ejemplo puede ser el de "hombre", "tío", "sujeto", "tipo", "individuo", "caballero", "señor", "primo", etc.:

"Me dijeron que me esperaría un *señor* en la puerta de mi casa. Cuando bajé a la calle vi a un *hombre* andando arriba y abajo sobre la acera. Cuando el *tío* me vio se acercó y me pidió fuego. Era un *individuo* muy singular y bastante desagradable, por cierto: el *tipo* tenía la cara llena de cicatrices. Nunca había visto un *sujeto* tan malcarado. '¿Usted es el *caballero* que envía la agencia?', le pregunté."

3. El tercer grupo está constituido por palabras que se diferencian sólo por su significado emotivo o evaluativo, de tal manera que en la utilización de una u otra se pone de manifiesto la valoración subjetiva del hablante, como puede o ocurrir entre "médico" y "matasanos", por ejemplo. Es obvia también la matización positiva y negativa en la lista siguiente: "económico", "frugal", "ahorrativo", "tacaño", "agarrado".

Un fenómeno al que no alude Palmer pero que pudiere incluirse

en este grupo es el de los sinónimos que indican gradación, como puedan serlo "simpático", "chocante", "gracioso", "divertido", "cómico", "humorístico", "hilarante", "destornillante", etc."

4. En cuarto lugar incluimos los grupos cuarto y quinto de Palmer. El primero trata de palabras restringidas colocacionalmente: "tocino/mantequilla rancia", "leche agria", "huevo/fruta podrida", "vino avinagrado", "queso mohoso", "carne descompuesta", etc". Todas estas palabras, al igual que "pasado" o "en mal estado", estarían recogidas bajo el mismo concepto y Palmer dice de ellas que son "verdaderos" sinónimos, a pesar de que, por ser usadas en contextos radicalmente distintos, se contradicen con la definición tradicional de sinonimia. Este fenómeno es frecuente cuando se relaciona el léxico con "conceptos" que pretenden ser universales. Así, a una enunciación conceptual o 'headconcept' que rezara '*que resulta de guisar un alimento en agua u otro líquido*' podría referirse a los adjetivos de las siguientes colocaciones: "huevos escalfados", "patatas hervidas/guisadas", "huevo duro", "carne estofada", "pescado guisado", "verdura rehogada", "legumbre cocida", etc. Aunque lo deseable es que hubiera conceptos distintos para cuando la cocción es breve o duradera, y para cuando el líquido es, por ejemplo, aceite, no podemos prever cuán matizado estará el diccionario de conceptos, porque pudiere existir una lengua que usara una sola palabra para todos los procesos. Lo justo

sería, naturalmente que se decidiera la naturaleza y variedad de los conceptos comparando democráticamente todas las lenguas del mundo, pero esto todavía no sucede así (ver capítulo anterior).

Lo mismo ocurre con lo que constituye el grupo quinto de sinónimos de Palmer: "words that are close in meaning, or that their meanings overlap", a lo que añade un ejemplo que funciona también para el castellano: la palabra "maduro" puede referirse a una persona con el sentido de 'adulta' o con el de 'entrada en años', a una fruta, y a una situación o trabajo que va a culminarse pronto.

Visto desde la perspectiva de Leech, 1974 (ver cuadro de la página 10), y considerando que no existe un "significado" sino varios niveles del mismo, podríamos establecer los distintos sinónimos (o palabras de significado similar) en cada uno de los niveles, de manera que quedaran patentes las diferencias al mismo tiempo que las semejanzas. Así, dado un grupo de sinónimos podrían jerarquizarse sus relaciones.

Este sería un trabajo interesante y sin duda muy útil en la elaboración de un diccionario concebido para la traducción y/o la descripción del uso del idioma. Pero no podemos ocuparnos ahora de ello porque no son todos los niveles de significado los que pretendemos barajar aquí, sino que nos hemos planteado desde el principio enfocar el trabajo hacia la relación entre el significado conceptual y el colocacional como punto de partida. Así y todo cabe señalar que no habrá nunca un sinónimo en la

casilla correspondiente al significado conceptual, puesto que si dos conceptos son idénticos, entonces es que son el mismo concepto, y deben llevar el mismo código de identificación. El fenómeno de la sinonimia solo funciona a nivel léxico y pone de manifiesto claramente la distancia existente entre el universo de los conceptos y el de las palabras.

Definiremos la sinonimia, pues, como la relación entre varias palabras que van asociadas al mismo concepto sin que exista la posibilidad de que sean intercambiables salvo en aquellos casos en que no se produce absolutamente ninguna restricción colocacional o léxica, siendo imprescindible dar cuenta de esas restricciones para cada entrada léxica.

Si los matices expresados colocacionalmente en una lengua para un mismo concepto se consideran cualitativamente o cuantitativamente significativos, entonces deberá enunciarse nuevos conceptos, por ejemplo: "en mal estado cuando se refiere a la mantequilla", "en mal estado cuando se refiere a la carne", etc., pero esta delicada decisión no compete a los estudiosos de una lengua particular, sino a lo que podría llamarse "diplomacia" lingüística, esperándose de los trabajadores de la misma que decidan con precisión y al mismo tiempo con amplio margen.

1.3.2.1. Colocaciones equivalentes

No obstante puede darse el caso de que las especificaciones del diccionario de colocaciones no sean suficiente; supongamos por un momento que los conceptos que aparecen en la red, vinculados por un arco Objeto, son los siguientes:

```

3d0683      "to drop a thing"
3bdf63      "a bullet weapon which explosives by
throwing or dropping"

```

Al concepto 3d0683 van conectadas las siguientes palabras del castellano: *arrojar, dejar caer, tirar, echar, lanzar, soltar*. Al concepto 3bdf63 van conectados los siguientes nombres: *proyectil, bomba, bomba de mano, bomba atómica, granada*.

El diccionario de colocaciones podría contener las siguientes especificaciones para las palabras implicadas de estos dos conceptos:

headword1	co_rel	headword2
lanzar	oper	bomba
arrojar	oper	proyectil
soltar	oper	carga

La combinación *soltar la carga* queda excluida inmediatamente por el hecho de que *carga* no se encuentra entre las palabras conectadas al concepto 3bdf63 que aparece en la red, pero ¿como seleccionamos entre las opciones *arrojar proyectiles* o *lanzar bombas* si ambas están definidas por la misma relación de

coocurrencia y están conectadas a los mismos conceptos?

En este caso puede solucionarse. La palabra *proyectil* está asociada al concepto 3bdf63 mediante una conexión de tipo 'u' (upperconcept). Por regla general la lectura más precisa (con mayor grado de intensión) tendrá preferencia sobre su hiperónimo. Por eso, en este caso se seleccionará la colocación *lanzar bombas*.

Como se ve el sistema proporciona suficientes oportunidades para efectuar una selección correcta.

No obstante también puede ocurrir que varias palabras pertenecientes al mismo concepto puedan colocarse con el mismo nombre sin que exista ningún criterio que permita escoger una y no otra. Es lo que sucedería, por ejemplo, con el concepto 3cf56c "a state of one's character or behaviour, good-tempered and friendly"

el cual va asociado a las palabras *amigable* y *afable*, entre otras, las cuales colocan con *trato* en la misma igualdad de condiciones. Existe la posibilidad de seleccionar la palabra que tenga un más alto grado de frecuencia de aparición en los textos, o bien la que tenga un menor grado de aparición si se pretende crear un mayor efecto estilístico; tal vez seleccionar la de menor extensión para denotar mayor precisión. En todo caso siempre será posible generar un texto con cada una de las palabras equivalentes existentes, pero, dado que el fenómeno podría ser recurrente, la combinatoria de opciones podría producir

una sobregeneración poco económica, por lo que es preferible reducir las soluciones a una sola aunque sea mediante juicios arbitrarios o aleatorios, justificados por la equivalencia de significados.

1.4. Pertinencia de una selección léxica

Con respecto a las combinaciones que son "hechos de lengua", es decir, que están determinadas lingüísticamente por cada lengua particular, el diccionario de colocaciones deberá dar cuenta de ellas ayudando a la gramática a decidir cuál de las palabras asociadas a un concepto es la más pertinente en combinación con otra. Por ejemplo, supongamos que tenemos que generar una palabra que se corresponda con el concepto 0ceb74 "of a condition of power or action, strong", la cual está asociada a los siguientes adjetivos del castellano: *enérgico, vigoroso, poderoso, potente, fuerte, fornido, robusto, recio, intenso, demoledor, titánico*. Pero estos adjetivos, que pudieran considerarse sinónimos, al menos cuando se usan haciendo referencia al concepto 0ceb74, tienen diferentes usos colocacionales. Diremos *gobierno poderoso*, pero no *coche poderoso* sino *coche potente*; del mismo modo se establece una preferencia de combinación en casos como los siguientes: *esfuerzo titánico, brazo fornido, árbol robusto, sacudida enérgica, lluvia intensa, sentimiento demoledor...*

Alguno opinará que esto es solo cuestión de estilo. Pero obsérvese la distorsión de sentido que producen las combinaciones

siguientes: *gobierno titánico, coche demolidor, árbol poderoso, lluvia robusta, sacudida fornida, sentimiento potente, etc.*

Calvo Pérez (1985:48) lo enuncia en los siguientes términos cuando detalla los supuestos en que fundamenta su pragmática léxica:

"Los miembros del espacio topológico léxico son palabras abstractas, si es necesario indexicadas, cuya manifestación más genuina se delata en su "anclaje" denotativo en la Sintaxis. El paso de Sdo. entre Emisor y Receptor se deriva, entre otros supuestos, de la posición real de la palabra en el campo léxico correspondiente y tiene como referencia estructuras semánticas rectas. En consecuencia toda desviación léxica origina *virtualidad connotativa* o pérdida de semas y tiene su soporte pragmático en el postulado anterior desde cuya base E. y R., coimplicados, se apoyan en idénticos criterios para la interpretación del Sdo. oblicuo que se manifiesta frecuentemente en el discurso."

En algunas ocasiones no se trata solo de una distorsión del sentido sino que se produce una verdadera confusión o ausencia de significado; a veces esto es debido a que el diccionario de conceptos sufre el reflejo de los sistemas conceptuales de las lenguas de las personas que lo han confeccionado, en este caso mayoritariamente del japonés, por lo que se producen ciertos vacíos que debe pensarse en completar de alguna forma. Pero lo cierto es que suceden cosas como la que sigue: en dos redes (estructuras de interlingua) aparece el mismo concepto, el 3ce59a que reza "of a condition of space, boundlessly extensive" que ha sido escogido en el proceso de análisis de las oraciones japonesas como modificador de las palabras *calle* (red c.0004) y *accidente* (red c.0312). De las palabras del castellano que van asociadas a este concepto como pueden ser *ancho, amplio,*

grande, colosal, enorme, inmenso, mastodóntico, extenso, vasto, descomunal, monstruoso, desmesurado, sólo ancho cobra sentido en la asociación *calle ancha* y solamente *grande* puede ser adjetivo en la combinación *gran accidente*, aunque quizá sería más apropiado *accidente de gran magnitud* o *de primera magnitud*, con lo que *de primera/gran magnitud* sería una entrada con la categoría de adjetivo también enlazada al concepto 3ce59a.

Debe quedar claro que la responsabilidad de que se produzcan estructuras de interlingua "aceptables" desde el punto de vista de su veracidad en confrontación con la experiencia depende del proceso de análisis y la correcta asignación de los conceptos lógicos a las palabras existentes en la frase de la lengua origen, salvándose todas las cuestiones de idiomatidad (ver más adelante la frontera entre idiomatismo y colocación). Si la estructura de interlingua realiza una asignación errónea debido al trabajo deficiente del nivel semántico, la generación de una estructura correcta en la lengua destino es, claro está, imposible.

Por este motivo la correcta asociación entre conceptos y palabras de la lengua origen es el proceso que supone mayor responsabilidad semántica, mientras que en el transcurso de la generación la problemática trasciende el nivel semántico y queda relegada a las relaciones lexemáticas. Por lo tanto, y mientras la voluntad política no decida invertir en el análisis del castellano, un sistema de generación como el nuestro utiliza solamente unos cuantos rasgos semánticos en la subcategorización

de las entradas como veremos, y sin embargo realiza un gran esfuerzo en la concreción de las coocurrencias.

De todas maneras existe un completísimo árbol semántico que funciona en el análisis del japonés (ver apartado 4.2.) y que está previsto que debe funcionar para la desambiguación en el análisis del castellano. Los rasgos que produce este árbol son la base del mecanismo que se encarga de filtrar las lecturas inaceptables de las oraciones según el significado que se supone deben tener en el "mundo real" y las restricciones que existen entre ellas a ese nivel.

Sin embargo, poco pueden hacer los rasgos semánticos sin una especificación que en cada entrada indique la colocabilidad de la palabra. ¿Por qué? Pues porque si una palabra, un adjetivo, por ejemplo, está asociado a varios conceptos, como ocurriría en el caso de adjetivos que poseen mucha extensión como "alto", "bajo", "fuerte", "grande", "gordo" o "fresco" (por poner ejemplos que son especialmente polisémicos y que por lo tanto deberán constituir varias entradas diferentes) solo la información de las palabras con las que colocan -o los rasgos semánticos de esas palabras- confrontadas con las palabras existentes en el texto, podrán hacer decidir al analizador cuál es la entrada del diccionario -y el consecuente concepto asociado a ella-. Si la información colocacional es lo suficientemente rica, el número de rasgos semánticos no tiene por qué ser excesivamente elevado.

2. Historia del concepto de 'colocación'

Suele suceder en el avance lingüístico, como en cualquier otra rama del conocimiento, una vacilación en el uso de los términos cuando estos se refieren a realidades no muy bien estudiadas ni delimitadas. Pero las aplicaciones prácticas de las teorías lingüísticas exigen cierta clarificación en ese sentido y una etiquetación sistemática de los fenómenos, aunque esto no signifique, como es natural, que no pueda reorganizarse la clasificación de los micromundos cuando los criterios que se deducen de ulteriores fases de la investigación lo exijan. Así, es posible que los sistemas de control en el mundo de la semántica se orienten progresivamente en distintas direcciones. Como dice Leech (1985:19) "la ciencia no avanza como el agua que va colmando un recipiente hasta llenarlo, sino que lo hace por un proceso ininterrumpido de revisión y aclaración que lleva a una mayor claridad y profundidad de comprensión". Probablemente es cierto, como afirma Sampson (1987:104) que el proceso de conjeturas y refutaciones no es aplicable a la actividad mental en general, pero la experimentación en lingüística computacional permite observar y delimitar las lagunas existentes en las distintas teorías.

Esto es lo que está ocurriendo en el campo de la semántica léxica. Las restricciones semánticas impuestas sobre léxico marcado semánticamente tal y como fueron concebidas por Katz y Fodor (1963), son funcionales, pero "en crudo" (Sampson, 1987)

no funcionan: por mucho que subcategoricemos el verbo *comer* con el rasgo MASS (concreto) no podemos confiar en esa restricción cuando existen expresiones como "comerse las palabras". Sin embargo, desde nuestro punto de vista, esto no significa que el modelo de restricciones semánticas no sea válido, sino que existen usos idiosincrásicos distintos en cada idioma que dependen del sistema conceptual que se ha desarrollado a través de mecanismos como la metáfora, y que esos usos deben listarse en los diccionarios de colocaciones ².

Las posibilidades de combinación del léxico sobre una estructura sintáctico-semántica conocida van desde la exclusividad (fijación) a la asociación libre pasando por un status de combinatoria preferente. En esta "combinatoria preferente" las palabras se asocian unas con otras para reflejar una realidad precisa o un matiz expreso sobre esa realidad. Cada lengua ha desarrollado un tipo distinto de asociaciones en virtud de su realidad histórica, ambiental, cultural y lingüística. Este fenómeno ha sido designado de distintas formas: "solidaridades léxicas" (Coseriu, 1973), "lexías complejas" (Pottier, 1964), "combinaciones" o "combinaciones recurrentes",

² Otra posibilidad es que se disponga de un sistema que pueda identificar las metáforas a base de la comparación de atributos o de conocimiento del mundo. Por el momento parece lejana la posibilidad de beneficiarse de los avances en inteligencia artificial o de los experimentos con corpus muy limitados como los de Dan Fass (1991), y tampoco puede afirmarse que los logros en este sentido no sean sino complementarios al diccionario de colocaciones previo.

"conurrencias" o "coocurrencias", "coapariciones", "colocaciones" o "conlocaciones". Grimes (1988:167) habla de "vecindad" (neighbourhood). El término colocación fue acuñado en el entorno del contextualismo británico por J.R. Firth (1957:196). La expresión 'coocurrencia' o 'co-ocurrencia' procede de los trabajos de Zellig S. Harris (1957:283) y, junto al término de Firth, ha sido mayoritariamente utilizada hasta hoy.

La traducción al castellano del término inglés "collocation" puede despertar polémica por el hecho de existir esta palabra en el lenguaje corriente con diversos significados. Obsérvese sin embargo que el término inglés es un derivado del latín, por lo que su traducción directa (la más cómoda, por otro lado) es perfectamente aceptable, resultando más familiar que la opción 'conlocación' (la preposición latina *cum* cambia la *u* en *o* resultando *com* y *con*, pero también *co* por asimilación de la *m* ante *l* y *r* - Alemany, 1920).

Firth (1957:196) introduce la noción de "nivel colocacional", a mitad de camino entre el situacional y el gramatical, allí donde se manifiesta el significado léxico. Tal nivel de significado, aunque posea una función en los contextos concretos de situación, depende principalmente de su tendencia a coaparecer en los textos:

The statement of meaning by collocation and various collocabilities does not involve the definition of word-meaning by means of further sentences in shifted terms. Meaning by collocation is an abstraction at the syntagmatic level and is not directly concerned

with the conceptual or idea approach to the meaning of words. One of the meanings of 'night' is its collocability with 'dark', and of 'dark', of course, collocation with 'night'.

Dentro de la general distinción entre contexto lingüístico y extralingüístico, el contexto léxico puede caracterizarse como las circunstancias semánticas del uso, y se distingue del contexto gramatical (morfología y sintaxis) y de la relación del discurso con el contexto cultural:

It must be pointed out that meaning by collocation is not at all the same thing as contextual meaning, which is the functional relation of the sentence to the processes of a context of situation in the context of culture. (:195)

Las teorías de Firth sobre colocabilidad léxica tienen sentido dentro de una teoría general sobre el significado y la función social del lenguaje que no considera significativas sólo a las palabras y las frases, sino también a los fonemas y a los rasgos prosódicos de los enunciados. El significado de cada componente -paralingüístico, fonológico, gramatical, léxico, etc,- se describe en tanto que elemento que interviene en la estructura de las unidades del nivel superior. Así, desde el punto de vista de Firth, puede hablarse de 'significado fonético', de 'significado gramatical' y de 'significado semántico'. Por otro lado Firth parece coincidir con Harris en la concepción distribucional del significado, según la cual dos lexemas tendrán el mismo significado si tienen la misma

distribución a lo largo de una muestra representativa de textos.

El distribucionalismo puede ser considerado como la evolución posterior del descriptivismo americano que se impuso con la obra de L. Bloomfield Language (1933), quien parte de la definición behaviorista de significado como un esquema estímulo-respuesta que se corresponde con una forma. Bloomfield elabora un principio de descripción formal excluyendo los criterios semánticos como categorías básicas, de tal manera que sólo sea relevante el hecho de que existan diferencias entre los significados de dos formas. El distribucionalismo de Harris es su máximo heredero; afirma que la suma de todos los entornos/contextos en que puede aparecer una unidad lingüística, esto es, la suma de todas las posiciones de un elemento en relación a otros elementos es su distribución. De esta manera, podemos definir una clase N por su frecuente co-ocurrencia con otras partículas como el artículo y el adjetivo:

This suggest that morphemes can be grouped into classes in such a way that members of a class have rather similar sets of co-occurents, and each class in turn occurs with specific other classes to make a sentence structure. In structural linguistics this classification is not set up on the basis of relative similarity of co-occurents, but rather on the basis of a particular choice of diagnostic co-occurents: 'cloth' and 'paper' both occur, say, in the environment 'the () is' (i.e. after 'the' and before 'is'), where 'diminish' does not appear; we call this class N. (Harris, 1957:284).

Sin embargo Harris deduce que es virtualmente imposible

describir un lenguaje en términos de coocurrencia de morfemas (o palabras -grupos de morfemas-) a nivel individual: "it is in general impossible to obtain a complete list of co-occurents for any morpheme" (p.285), "it is impossible to survey everything that has been said or written in a language, even within a specified period" (p.292) y que este tipo de coocurrencia, aunque no puede tratarse con indiferencia, no puede usarse como elemento básico en la construcción de un sistema. (Naturalmente que Harris concibe un sistema regido por la sintaxis, no por la semántica, como enfatizamos más adelante, y por supuesto que no pensaba en la posibilidad de usar herramientas computacionales como las existentes hoy).

Harris prosigue su razonamiento afirmando que las coocurrencias individuales son esencialmente diferentes simplemente porque son un reflejo directo de las combinaciones de significado que hacen los hablantes, "and is therefore not subject at all to investigation for distributional regularities" (p.285, nota 2). Es curioso, sin embargo, que estas afirmaciones se sigan de una definición de coocurrencia según clases y construcciones que utiliza ejemplos léxicos y por lo tanto de combinaciones de significado:

For classes K, L in a construcción c , the K -co-occurrence of a particular member L_i of L is the set of members of K which occur with L_i in c : For example, in the AN construction found in English grammar, the A -co-occurrence of *hopes* (as N) includes *slight* ('slight hopes of peace') but probably not *green*. The K -co-occurrence of L_i is not necessarily the same

in two different KL constructions: the N-co-occurrents of *man* (as N_i) in N_i is a N may include *organism, beast, development, searcher*, while the N-co-occurrents of *man* in N_i 's N may include *hopes, development, imagination*, etc.

Hay que tener en cuenta, sin embargo, que el propósito del artículo de Harris es justificar la existencia de transformaciones por medio de la presencia de las mismas identidades léxicas en relación de coocurrencia:

To check whether the co-occurrences are identical in two n-class constructions (suspected of being transforms), we obtain a list of n-tuples which satisfy the first construction and ask about each one whether it also satisfies the second; and vice versa. (...)
Such checking of individual co-occurrences is not practicable without an informant, i.e. a more or less native speaker of the language. (p. 292)

El trabajo de Harris estaba orientado a crear una gramática de construcción de clases mediante un conjunto de procedimientos de descubrimiento y de los análisis distribucionales, en aras de un algebra de las transformaciones aplicable a la tipología lingüística (la comparación de diferentes tipos de lenguas y estructuras). Harris se plantea la existencia de las coocurrencias como un medio y no como un fin en sí mismo; desde la óptica de su generación, la cual estaba lejos de plantearse cuestiones de tipo semántico y que suponía que podían reducirse a la sintaxis todos los fenómenos de descripción de una lengua, es una postura coherente. Actualmente, sin embargo, el enfoque es exactamente el opuesto: la capacidad transformacional del lenguaje es la prueba de que existen relaciones léxico-semánticas

que sobreviven a la forma sintáctica que el hablante o las exigencias del discurso impongan; y este enfoque inverso se debe no sólo a que las transformaciones son difícilmente implementables y de poca utilidad en un sistema experto, sino a que esas relaciones léxico-semánticas, catalogadas convenientemente, suponen un importante punto de partida en el estudio semántico, semasiológico u onomasiológico, de la producción de sentido.

Por otro lado, y a pesar de que el distribucionalismo hizo importantes aportaciones a la lingüística estructural, en conjunto sus resultados también devienen insatisfactorios para algunos. Lyons (1977:551) observa que nunca se ha dado ninguna razón convincente para creer que la identidad de significado léxico definida distribucionalmente corresponda a lo que preteóricamente se considera como tal identidad. En el caso de que una palabra tenga varios significados como ocurre con el verbo 'turn'-girar, volver, ...- por poner el mismo ejemplo que usa Lyons, no podemos afirmar que un determinado significado ('the milk has turned' -la leche se ha cortado-) lo es porque tiene la misma distribución que 'go sour' ('the milk has gone sour' -la leche se ha vuelto agria-), si antes no hemos establecido esos significados de 'turn' de alguna manera, su identidad y su diferencia.

El problema teórico (generado por la correlación existente

entre la identidad de significado y la identidad de distribución) de si la similitud de distribución es la consecuencia y no la causa de la similitud de significado puede parecer pueril, pero es decisivo a la hora de construir un diccionario conceptual que no contemple sólo aquellos conceptos deducidos de la distribución dada en una lengua determinada, sino que esté abierto a los que proceden de otras distribuciones.

En realidad la solución a esta polémica parece encontrarse, en la práctica, en la distinción, a la que ya hemos hecho mención, entre **significado** y **concepto**. El primero estaría constituido por el significado abstraído a partir de las diferentes formas de uso contextual de la palabra y el segundo por el significado fundamental o núcleo conceptual. Existiría un tercer nivel de significación que podríamos llamar connotativo o afectivo que se desprendería de las reacciones afectivas que la palabra generara con o sin intervención de la voluntad del hablante, pero este nivel, no menos importante en cuanto al hecho comunicativo (quizá el más importante al fin y al cabo), varía de hablante a hablante, y por su difícil tratamiento objetivo queda en principio fuera del propósito de esta tesis, de ATLAS y de los programas de traducción automática en general.

El significado propiamente dicho, pues, se deriva de la contextualización de la palabra, y si entendemos por palabra aquella que posee carga semántica, nos estamos refiriendo a las colocaciones. El significado, desde este punto de vista,

es una cuestión meramente lingüística, intrínseco para cada lengua, mientras que, fuera de la relación del individuo y la sociedad con la realidad, fuera de su valoración de la realidad, estaría el mundo de los conceptos. Claro está que un diccionario de conceptos así concebido es una utopía platónica; es obvio, como ya se ha sugerido en anteriores apartados, que las lenguas mayormente implicadas o más poderosas impondrán su sistema conceptual, o, mejor dicho, su significado contextual o lingüístico. Pero no deja de ser un proyecto de gran envergadura. Cuantas más lenguas intervengan en la construcción de ese diccionario, en principio, más posibilidades hay de que el sistema conceptual que represente sea "universal".

Los problemas que genera la conexión entre el nivel conceptual y el léxico o contextual son muchos. Algunos conceptos procedentes del diccionario conceptual tienen una distribución plurilexémica en la lengua destino, con lo que será preciso incluirlos como compuestos (unidades léxicas múltiples, ver apartado 3.3.) en el diccionario de palabras, o bien generar las palabras que faltan si se trata de una colocación mediante reglas que contemplen las relaciones de coocurrencia patentes en la entrada del diccionario o en un diccionario específico (ver segunda parte).

Así, rechazar la teoría distribucional del significado no exime de admitir que hay un alto grado de interdependencia entre lexemas que coocurren colocacionalmente en los textos,

de tal manera que el significado de algunos de ellos no puede desligarse del significado de la palabra con la que colocan (por ejemplo 'mesarse los cabellos' o 'pan ácimo'). Como dice Lyons (1977:552), "en última instancia, la noción de colocabilidad es un importante correctivo a una excesiva dependencia de la noción dualista de significación". Desde nuestro punto de vista esto no es una conclusión, sino el inicio de una nueva perspectiva que concede, sin miedo a la exhaustividad de los corpus ni al proceso de acumulación de conocimiento (gracias a las nuevas herramientas computacionales), una importancia decisiva a las relaciones léxicas sintagmáticas en el sistema de la lengua.

Desde una panorámica actual, los problemas que presenta la desambiguación o la selección léxica en los sistemas de procesamiento del lenguaje natural convierten la representación de las opciones que una lengua proporciona y sus consecuencias funcionales en un enfoque central.

En el desarrollo de este trabajo vamos a utilizar la palabra **coocurrencia** para designar cualquier tipo de relación de adyacencia que sea detectada en el nivel sintáctico, semántico o léxico, mientras que el término **colocación**, como un tipo específico de coocurrencia, será utilizado para designar la relación de solidaridad léxica ³ existente entre categorías

³ Se usa aquí el término 'solidaridad léxica' en el sentido de Coseriu (1973) de que una determinada palabra funciona como rasgo diferenciador con respecto de otra con la que está sintagmáticamente relacionada.

semánticamente plenas relacionadas sintagmáticamente, a saber, el nombre, el verbo y el adjetivo, incluyendo dentro de este último grupo a los adjetivos adverbializados. Tal y como se plantea en Aguilar-Amat, 1990, el término solidaridad léxica define aquellas asociaciones entre unidades léxicas con un grado de pertenencia al conjunto de los términos transparentes mayor que el grado de pertenencia al conjunto de los términos opacos. Su significado es, por lo tanto, composicional (el significado del todo es igual a la suma de las partes), pero son características porque los hablantes las escogen siempre simultáneamente de entre el abanico de posibilidades que les ofrece su competencia. Es por ese motivo que las llamamos recurrentes. Pero, cuando estas combinaciones coocurren siempre en el marco de las mismas estructuras sintácticas, se distinguen del resto de combinaciones con el nombre de colocaciones. Así podrá hablarse de la coocurrencia entre términos de distintas colocaciones, y nos estaremos refiriendo a la relación conceptual existente entre significados asociados a categorías quizá distintas, como es el caso de "músculo agarrotado" y "agarrotamiento del músculo", entre "circunstancia atenuante" y "circunstancia que atenúa" o "palabras tiernas" y "la ternura de sus palabras", etc. (Ver apartado 4.3.).

2.1. La selección léxica por colocación

Los modernos intentos de investigación en aras de construir sistemas que permitan analizar y generar texto para, por ejemplo, traducir automáticamente, han hecho que se mida la calidad de un sistema por la calidad del 'output' o producto de salida. Si la selección léxica y sintáctica no es correcta, el sentido del texto quedará distorsionado. Por este motivo, y, a pesar de que la investigación en torno a temas relativos a la generación no ha producido escuelas coherentes de pensamiento ⁴ (al contrario de lo que ocurre en análisis ⁵, en el que se dan claramente dos corrientes: redes de transición y dominios activados léxicamente ⁶), el proceso de generación y el proceso de selección léxica inherente a ella se convierte en último responsable de la calidad de un producto de salida que no se valora sólo en cuanto a refinamiento estilístico sino, sobre todo, en cuanto a plenitud de sentido.

Está claro que los problemas que presenta el análisis son distintos a los que presenta la generación, y que mientras que

⁴ A excepción de la llamada "Meaning-Text school of linguistics" que desarrolla las ideas de Firth (cf. Mel'cuk, 1988).

⁵ Es obvio que el proceso histórico del desarrollo de un tema ligado íntimamente a sistemas empíricos demuestre una mayor inversión en el campo del análisis por ser los productos de este el punto de partida de la generación, y que las trabas halladas a lo largo de la trayectoria del análisis hayan repercutido en el avance de la investigación sobre síntesis, ni mucho menos por eso menos importante, que dió comienzo alrededor de 1974.

⁶ La expresión "dominio activado léxicamente" procede de McDonald, 1987, y se enuncia aquí haciendo referencia al análisis sintáctico dirigido por la categorización y subcategorización del léxico.

en el análisis la desambiguación es una cuestión importante, en generación la selección léxica juega un papel fundamental previamente a la realización sintáctica de la estructura de la frase: "... many of the terms in a message are surface lexical relations rather than a more abstract conceptual vocabulary; this has the obvious corollary that syntactic realization will usually occur after key words have been chosen" D.D. McDonald (1988:217).

De aquí se infiere que, una vez asolida la crucial tarea de determinar el significado conceptual que corresponde a cada palabra del texto en el proceso de análisis de la lengua origen, nos encontramos con que a menudo el concepto 'pelado' no es suficiente para producir una estructura coherente en la lengua de destino, y que el control ejercido por las restricciones y los gobiernos sintácticos, que va supeditado al léxico, no sirve para decidir la selección de la entrada léxica correspondiente al nudo.

Esto no significa, como se verá en el apartado 4, que las restricciones gramaticales impuestas según el comportamiento sintáctico de las entradas no puedan filtrar las posibilidades inconsistentes desde el punto de vista semántico en relación con la primera elección efectuada; este proceso es complementario a la selección efectuada en virtud de las relaciones léxicas significativas establecidas previamente, y también funciona cuando estas no existen.

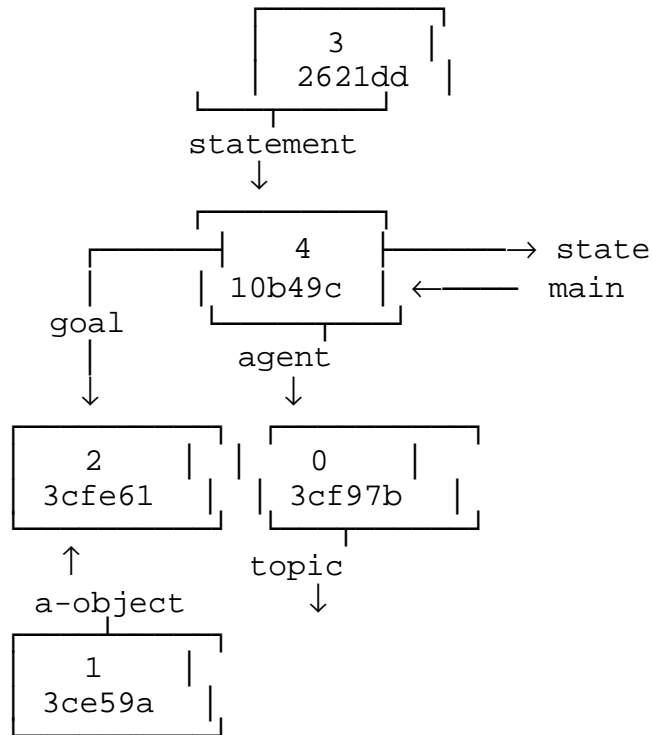
2.1.1. La selección prioritaria

¿Qué significa la primera elección efectuada? Las redes (o configuraciones abstractas conceptuales) tienen siempre marcado uno de los nodos como nodo "main". Este nodo correspondería al núcleo ('head' o 'governor') en una gramática de dependencias. Suele ser un verbo o un nombre. Si se trata de un nombre significa que la oración es de tipo existencial o atributiva, si se trata de un verbo, de él se infieren los distintos roles del resto de nodos que constituyen la oración: el agente, el objeto, el benefactivo, etc.⁷

La primera selección léxica que se produzca determinará semántica y sintácticamente el resto de las selecciones léxicas correspondientes a los otros nodos. Ahora bien, la selección léxica correspondiente al nodo principal se realizará en virtud de las relaciones colocacionales del nodo con su adyacente más inmediato en este orden: **verbo-complemento**, **nombre-modificador**, **verbo-modificador**. Estas son, por otro lado, las relaciones de coocurrencia que se prevé van a consolidarse en el diccionario de colocaciones que ha empezado a construirse por las de nombre-adjetivo, objeto de la presente tesis.

⁷ Llamamos roles a los papeles o roles temáticos de Fillmore que se usan aquí como conjunto prototípico de un paradigma ampliable según ulteriores investigaciones sobre la semántica de los sintagmas preposicionales modificadores (o complementos) del verbo.

Veamos un ejemplo. Supongamos que se pretende generar una frase del castellano correspondiente a la siguiente estructura conceptual:



En esta red cada nodo va asociado a la serie de palabras siguientes del castellano:

- 10b49c "to face a thing"
encarar, enfrentar, afrontar, dar a, hacer frente, acometer,
mirar hacia, arrostrar, estar orientado a,
orientarse a
- 3cfe61 "an entrance to a place"
entrada, acceso, vía de acceso, puerta, válvula, ingreso, admisión
- 3cf97b "a street which people or cars can use"

calle, paseo, avenida, carretera, camino, pasaje, pista

3ce59a "of a condition of space, boundlessly
extensive"

grande, mayor, amplio, ancho, colosal, inmenso, de primera
magnitud, de gran magnitud, mastodóntico,
monstruoso, oceánico, ciclópeo, macrocósmico

El diccionario de colocaciones deberá recoger la expresión "dar a la calle" (la relación se establece entre *dar* y *calle*, ya que la información acerca de la preposición regida se encuentra en la entrada léxica). Seguidamente se selecciona el modificador del nombre; en este caso el diccionario de colocaciones señala *calle mayor* y *calle ancha*, con lo que se producirán dos lecturas (o bien se selecciona simplemente la primera de ellas).

El nodo correspondiente al sujeto no está recogido como colocación, sino subcategorizado por el verbo ⁸. Su correcta selección depende en principio de la información sintáctico-semántica especificada como subcategorización en la entrada del verbo. El verbo "dar a" posee una especificación restrictiva preferente hacia nombres con rasgo semántico MASS (concreto). Por este motivo los nombres *admisión*, *ingreso*, que están categorizados semánticamente como EVENT (eventos) son rechazados por las reglas de concordancia semántica entre verbo

⁸ Esta decisión responde al hecho de que el sujeto se ha contemplado a menudo como un complemento especial o argumento externo. Frente a las ideas de Tesnière y Fillmore que lo consideran un actante más del verbo, a tradición chomskyana cree que a él corresponde una asignación temática indirecta, es decir, que no es subcategorizado por el verbo como los demás complementos sino como una categoría léxica c-dominada por el sintagma verbal.

y sujeto; lo mismo ocurre con la entrada eventual de *entrada*. Por el contrario la entrada concreta de *entrada*, así como *puerta* y *válvula* tienen el rasgo MASS. Dado que *válvula* posee dos entradas, una especificada como perteneciente al área de aplicación TEC (técnica) y la otra ANAT (anatomía), y no hay ninguna otra ocurrencia perteneciente a estos campos en el texto, la entrada de *válvula* será rechazada. Esto restringe las posibilidades a *entrada* y *puerta* - *acceso* y *vía de acceso* están categorizadas como ABSTR (abstractos). Las posibilidades resultantes son pues las siguientes:

La entrada da a la calle ancha
La entrada da a la calle mayor
La puerta da a la calle ancha
La puerta da a la calle mayor

La similitud de significado de estas oraciones indica que es posible escoger arbitrariamente (si no hay otro criterio como pueda ser la frecuencia de uso) una de ellas. De este modo vemos increíblemente reducido el número de oraciones que se hubieran obtenido si se permitiera realizar la combinatoria de todas las palabras asociadas a los conceptos de la red, evitándose por otro lado sinsentidos como "*El acceso afronta la carretera monstruosa*" o "*La válvula arrostra el oceánico pasaje*", etc.

2.2. Relaciones funcionales del léxico

En cierto modo las colocaciones no son sino un tipo de relación entre unidades de léxico de entre una diversidad de relaciones que una vez definidas pueden conformar una tipología.

En torno a este tema se produce un debate entre los "lumpers" (acumuladores) y los "splitters" (escendidores) - ver Evens (1988:8). Los primeros trabajan en modelos con un pequeño número de relaciones generales (Werner y Sowa, en el mismo volumen) y parecen tener la realidad psicológica de su parte, mientras que los segundos (Calzolari, Ahlswede & Evens, Grimes, White) crean modelos con un gran número de relaciones específicas y tienen más fama en las relaciones léxicas orientadas superficialmente. Desde el punto de vista de la psicología los "lumpers" o acumuladores estarían aliados a los mentalistas, mientras que los "splitters" o escendidores a los behavioristas.

Otros sistemas, entre los que nos incluimos, poseen familias de relaciones según las funciones taxonómicas del nombre, el verbo o el adjetivo. Ahora bien estas familias de relaciones que son más escindidas que acumuladas, tienen, como se verá para el caso de nombre y adjetivo en la segunda parte de este libro, una estructura jerárquica que las reduce a unas básicas.

El punto de partida en el momento de poner nombre a las relaciones ha sido el trabajo de Igor Mel'cuk, 1988, el cual ha definido alrededor de sesenta funciones léxicas que ha utilizado en la construcción de su ECD (Explanatory Combinatorial Dictionary).

Frente a la consideración del dominio semántico (relaciones paradigmáticas) y el dominio sintáctico (dependencia sintáctica de elementos frasales que son actantes semánticos), la asunción de un nivel léxico de análisis aporta el estudio de las posibilidades combinatorias de una palabra de tal modo que deviene una función en sentido matemático donde encontramos una función f , una palabra W_0 (argumento o "key-word") y un conjunto de palabras o frases que expresan el significado correspondiente a f . Mel'cuk da a este tipo de relaciones el nombre de **funciones léxicas** ("lexical functions" o LF), las cuales divide en dos grandes grupos: las relaciones **paradigmáticas** y las **sintagmáticas**. Las relaciones sintagmáticas o colocacionales predicen coocurrencias a nivel léxico, y las relaciones paradigmáticas expresan relaciones que no son de coocurrencia, como puedan ser la sinonimia, la antonimia, la conversión, la hiperonimia, diminutivos y aumentativos, etc.

El número de relaciones paradigmáticas identificadas por Mel'cuk es mucho mayor que el de relaciones sintagmáticas. En parte era lógico esperar que esto fuera así dado que la tradición semántica en general se ha ocupado preferentemente de este tipo de relaciones, cuando lo ha hecho. Esto se debe a que los intentos de abordar las relaciones léxico-semánticas se plantean la tarea globalmente en busca de la naturaleza universal de esas relaciones desde un punto de vista antropológico (Werner, 1988) o psicológico (Riegel, 1970). Este último punto de vista se preocupa principalmente por la organización de la memoria y concibe una

naturaleza doble en las relaciones: las **relaciones lógicas**, derivadas por abstracción de las palabras que atañen a fenómenos como la superordenación (mesa-mueble) y la coordinación (mesa-silla), y las **relaciones infralógicas** o relaciones físicas basadas en objetos denotados, eventos o cualidades que son producto de la abstracción de rasgos físicos de las entradas: partes (mesa-pata), locaciones (cebra-Africa) o sustancia (mesa-madera). Igor Mel'cuk hace también esta distinción denotando una clara preferencia hacia las relaciones lógicas frente a las infralógicas. Relaciones como, por ejemplo, la tan debatida parte-todo le parecen demasiado vagas. Pero, por otro lado, reduce la cuestión del adjetivo, aspecto central en esta tesis, a la existencia de una función léxica denominada MAGN, lo cual no deja de ser una simplificación conformista.

El trabajo que ofrecemos aquí, mediatizado por unas herramientas determinadas y enfocado a un objetivo concreto, sin embargo, debe plantearse seriamente cuál es el enfoque que le permitirá controlar el máximo de información eficaz posible sin precipitar problemas peores, y vuelvo a referirme aquí a la cita de Lakoff & Johnson de la página 6 en la Introducción. Así nosotros no vamos, en principio, a tratar de solucionar el problema de las relaciones léxicas como si se tratara de un rompecabezas, sino que simplemente nos planteamos la elaboración de un diccionario eficaz que dé cuenta de la realidad combinatoria sintagmática del léxico, con la confianza de que esto pueda funcionar como catalizador y permitir posteriormente

el abordaje de otras parcelas de la semántica (como son las relaciones paradigmáticas) y de la sintaxis (por ejemplo las transformaciones).

Desde este punto de vista concebimos la dimensión colocacional del léxico, en colaboración con la dimensión taxonómica soportada por las relaciones de hiperonimia o hiponimia del léxico y sus correspondientes conceptos, como armas eficaces para desambiguar y para efectuar la selección léxica correcta.

3. Colocación y fijación

El significado colocacional, es decir, el significado de una palabra que se obtiene gracias a la distinción efectuada por la presencia de otras palabras que aparecen en su contexto inmediato, es un tipo de significado que afecta tanto a las estructuras denominadas composicionales o transparentes, como a las no composicionales u opacas, más usualmente denominadas idiomatismos. Palmer (1976:98) afirma que "idioms involve collocation of a special kind", con lo que queda claro que el léxico se combina o 'coloca' de una manera que puede ser más o menos composicional y más o menos rígida sintácticamente hablando.

Así, ser o no ser un idiomatismo es muy a menudo un problema de grado. En Aguilar-Amat (1990:45) se sugiere que el cálculo exacto de la pertenencia de una colocación al conjunto de los términos opacos o al de los términos transparentes puede calcularse aplicando la fórmula de la distancia euclídea utilizada en la teoría matemática de los conjuntos borrosos (fuzzy sets) en virtud de encuestas de opinión o de otros criterios como puedan ser los sintácticos o los comparativos con otras lenguas. Pero el esfuerzo para determinar cuando una combinación de palabras es simple (combinaciones libres), cuando es una colocación (combinaciones preferentes) o cuando es un idiomatismo (combinaciones fijas), ya sea a través de un grado de pertenencia o de cualquier otro rasgo, deviene completamente estéril si

no se tiene en cuenta un planteamiento general del problema de la contextualización.

El tratamiento computacional de las colocaciones, que significa el recuento de las constantes de ocurrencia que se producen en los patrones sintácticos más usuales, no puede partir más que de la especificación de los contextos posibles para cada significado o concepto relativo a una palabra y sus homónimos. Este contexto puede ser una palabra, varias palabras, o un paradigma formado por los hiperónimos, hipónimos o sinónimos de una palabra, pero también puede ser una estructura invariable de carácter idiomático, como por ejemplo "castillos en el aire" (construir...) u "ojo de la cara" (costar..) o "tiempo de perros" (hacer...).

Naturalmente que dicho contexto o, como lo llama Sinclair (1991), la parte colocada, debe llevar especificaciones acerca de su determinación, del tipo de artículo que la acompaña, así como del número, género, tiempo y persona de ser necesario. Asimismo, cuando sea pertinente, deberá especificarse el eje antónimo, a través del cual podrán hallarse las palabras relacionadas o sinónimos cuando estas no estén especificadas también en la entrada.

Ahora bien, a pesar de que las colocaciones puedan ser más o menos idiomáticas, más o menos transparentes u opacas, y de que su representación lexicográfica sea parecida, el hecho es que la existencia de opacidad o idiomática en una

construcción incide en un su comportamiento sintáctico, de tal manera que ciertas transformaciones quedan automáticamente vetadas. Es lo que ocurre, por ejemplo, con *morderse la lengua*, expresión que, en su significado no literal, se resiste a la construcción de pasiva, a la adjetivación participial, a la relativización, a la pronominalización, a la modificación adjetival y nominal, a la modificación adverbial, a la determinación, a la indefinición, al plural y a la ausencia de artículo:

Juan se mordió la lengua y no dijo nada más.

*La lengua fue mordida por Juan

*La mordida lengua ...

*La lengua que se mordió Juan

*Juan se la mordió

*Juan se mordió la indiscreta lengua

*Juan se mordió la lengua de carne

*Juan se mordió la lengua intensamente

*Juan se mordió aquella lengua

*Juan se mordió una lengua

*Juan mordió lenguas

Ciertas colocaciones con usos figurativos, traslaticios... en definitiva parcialmente opacos o idiomáticos también sufren restricciones de este tipo. Obsérvese que *prestar el coche* es una construcción que admite la forma pasiva, mientras que *prestar atención* no. Ahora bien, si comparamos sus limitaciones con las que afectan a *morderse la lengua*, vemos que es bastante más flexible que esta:

Pasiva:

*La atención fue prestada al ministro por el público

Adjetivación participial:

El ministro agradeció la atención prestada

Relativización:

Nos sorprendió la atención que prestaban los niños

Pronominalización:

*El público la prestó

Modificación adjetival:

El público prestó gran atención

Modificación nominal:

*El público prestó atención de gran intensidad

Modificación adverbial:

El público prestó atención ininterrumpidamente

Determinación:

*El público prestó aquella atención

Cuantificación:

El público prestó mucha atención

Indefinición:

*El público prestó una atención

Plural:

*El público prestó atenciones

Presencia de artículo:

*El público prestó la atención

(aunque 'El público prestó la debida atención')

Por este motivo parece conveniente que las entradas que llevan especificada la parte colocada lleven además una marca de grado que permita, según una tipología previa, identificar su comportamiento sintáctico.

De este modo, y como ya queda recogido en Aguilar-Amat, 1990, se concibe una gradación que va de la transparencia a la opacidad desde el punto de vista de la composicionalidad del significado, aunque, como veremos, el criterio de composicionalidad es muy relativo. Así, diremos que una

expresión es semánticamente transparente siempre que sea divisible en constituyentes ⁹ (a pesar de que algunos de ellos puedan no ser divisibles a su vez). Una expresión transparente debería, en principio, ser susceptible de ser traducida composicionalmente a otra lengua sin mayor problema puesto que existe un significado conceptual detrás de cada una de sus partes (aunque como ya se ha visto a través de ejemplos anteriores el resultado final no es correcto si no se produce la correcta selección léxica), mientras que una expresión opaca no puede ser traducida mediante la traducción correlativa de sus partes sin alterarse absolutamente el sentido.

Sin embargo cabe notar que el criterio de composicionalidad, aunque frecuentemente asumido, adolece de cierta ingenuidad a la luz de las divergencias entre sistemas conceptuales aun muy cercanos. Moreno (1989) ofrece una definición del concepto de composicionalidad total como la poseedora de dos propiedades: la perfecta correspondencia léxica y la perfecta correspondencia estructural, es decir, si se toman las oraciones

- a) Juan habla holandés
- b) John speaks Dutch

se observa que para cada elemento terminal en a) existe un elemento

⁹ La noción de constituyente semántico parte del presupuesto de que el significado de una frase resulta de la combinación de los significados, más simples, de los componentes. Un constituyente semántico puede ser una parte de una palabra, una palabra o una secuencia de palabras. Basándonos en Cruse, 1986:35, decimos que un constituyente semántico debe ser al menos una palabra porque es el elemento más pequeño que tiene movilidad en la frase, es decir, que puede alterar su posición sin que se destruya la gramaticalidad.

correspondiente en b) que lo traduce exactamente, y las estructuras de constituyentes sintácticos en a) y b) son isomórficas.

Sin embargo, sucede que, a pesar de que la composicionalidad posea esas dos dimensiones, la léxica y la estructural, no puede establecerse ni en el nivel léxico ni en el nivel estructural de una manera sistemática. Obsérvese que en el ejemplo siguiente,

c) se habla holandés

d) Dutch spoken here (holandés hablado aquí)

donde d) es el paralelo impersonal de b), la traducción al mismo idioma deja de ser composicional. Lo mismo sucede con el ejemplo siguiente:

e) John speaks nonsense

f) Juan dice tonterías

donde e) posee una estructura muy parecida a b) pero no traduce su componente verbal igual que a).

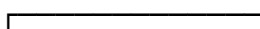
Por otro lado resulta falaz afirmar que una oración es composicional virtualmente cuando esa composicionalidad depende exclusivamente de la lengua de la oración con la que se la esté comparando. Es por este motivo que opinamos que el criterio de composicionalidad no es suficiente para sustentar una teoría del significado y debe ser usado con mucho escepticismo por los lexicógrafos. Actualmente sólo puede servir para establecer baremos comparativos entre lenguas y determinar cuál es el grado de proximidad o similitud entre ellas. Si los universales

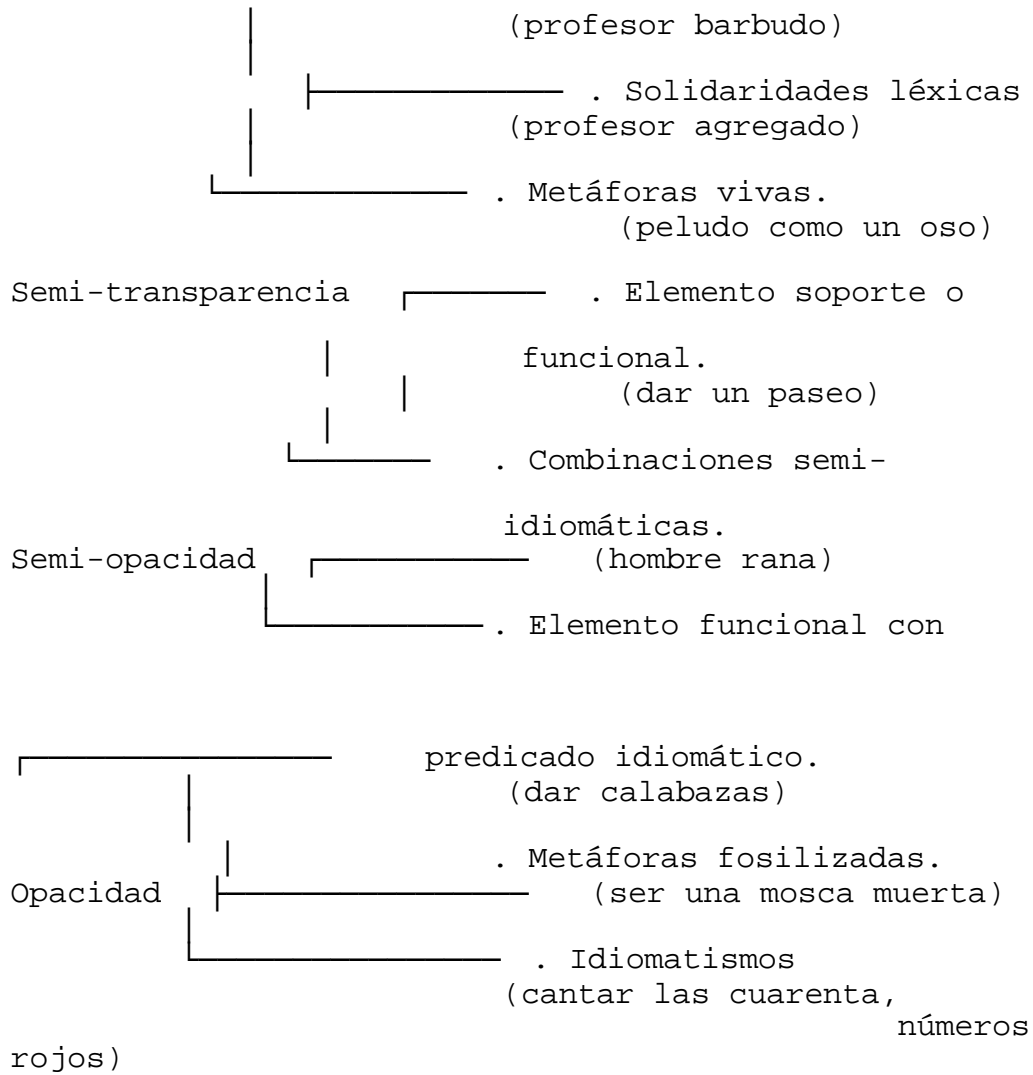
sintácticos y semánticos se fijaran algún día, la comparación con ellos podría determinar la composicionalidad de las estructuras de una lengua; o quizá, si la comunidad internacional asume la responsabilidad política de erigir un idioma (el inglés, por ejemplo) en 'medida estándar' con la que comparar los demás idiomas.

Sin embargo, la dificultad que entraña realizar una labor comparativa exhaustiva hace que el criterio de la composicionalidad se utilice intuitivamente como un punto de partida en el estudio de las relaciones del léxico entre sí.

En realidad lo que se persigue en los intentos clasificatorios es aislar las construcciones consideradas no-composicionales para que reciban un tratamiento específico, lo que a menudo aleja de una solución más profunda y duradera. Como sucede en todas las ramas de la ciencia, sin embargo, es preciso comenzar poniendo nombres o etiquetas a la realidad que se nos aparece como distinta para deducir después la similitud existente.

A este respecto en Aguilar-Amat (1990) se intenta una clasificación de las combinaciones que resumimos aquí en el siguiente esquema:

Transparencia  . Combinaciones libres.



Las combinaciones libres mantienen la transparencia de todos sus constituyentes, las solidaridades léxicas también, aunque se trate de una combinación más determinista, que es lo mismo que ocurre con las metáforas vivas. Las construcciones con un elemento funcional son semi-transparentes porque dicho elemento está vacío de significado. Ocurre algo parecido con combinaciones en que existe un elemento que sí aporta opacidad pero que no es el elemento de más peso de la expresión; la

relatividad de este "peso específico" semánticamente hablando es notable, por lo que estas construcciones pueden ser llamadas semi-transparentes o semi-opacas según se mire. Locuciones compuestas de un elemento idiomático y uno funcional son consideradas opacas pero también semi-opacas por el hecho de que se recogen en el diccionario como usos distintos de la palabra y, como se explica más adelante, subsumen el elemento funcional como una característica (subcategorización) distintiva. El resto de construcciones gozan de un alto grado de opacidad, por lo que deberán constituir unidades léxicas múltiples asociadas a un solo concepto.

La idea desarrollada en Aguilar-Amat (1990) e implementada en un sistema de transferencia, es limitar en lo posible el número de unidades léxicas múltiples, de tal manera que, dada una combinación semi-transparente o semi-opaca, el elemento funcional (vacío de significado) o el elemento idiomático (significado connotativo) aparecieran como una subcategorización de la entrada del diccionario, lo que la diferenciaría de la entrada "libre". (En un sistema conceptual además cada entrada iría conectada con un concepto distinto, obviamente). Obsérvese el siguiente ejemplo:

```
calabaza_1 => cat=n,num=sp,gen=f,sem=veget,efunc=none  
calabaza_2 => cat=n,num=p,gen=f,sem=abstr,efunc='dar'
```

Este tipo de entrada de diccionario tiene sentido por razones

de economía dada la profusión de reglas que pueden resultar de una implementación que pretenda construir reglas particularizadas de contextualización (ver apartado 3.3.).

Ahora bien, desde el punto de vista de un sistema que utiliza un diccionario conceptual como es ATLAS, los intentos clasificatorios que establecen jerarquías o gradaciones no sirven siempre para establecer qué entidades léxicas conformarán unidades múltiples y qué otras no lo harán; el diccionario de conceptos ejerce poderes fácticos en la constitución del diccionario de la lengua destino. Puede suceder que exista un concepto identificado con un número que deba ser conectado con una combinación que a los hablantes del castellano nos resulte forzada ("año primero", concepto 0eba33 "the first year of the new era"), que una combinación que nos parece hace referencia a un solo concepto, en el diccionario conceptual aparezca como dos conceptos distintos, o también que exista un concepto para cada componente y otro para la combinación de ellos como es el caso de *ángulo recto*:

0ad171 "the difference in direction between two lines or surfaces that meet, measured specially in degrees"
(ángulo)

0d420c "having a line or axis perpendicular to another line or surface"
(recto)

26d496 "angle of 90 degrees"
(ángulo recto)

Esta última actitud del diccionario de conceptos es la más corriente, y obliga a que el diccionario de la lengua destino contenga asimismo ambas opciones. Esta duplicidad que podría parecer superflua no hace sino reflejar una realidad lingüística, y proporciona una ductilidad que se hace necesaria en muchos casos, por ejemplo:

"... El triángulo rectángulo posee un ángulo recto (concepto 26d496) y otros dos ángulos (concepto 0d420c) que no son rectos (concepto 0d420c) sino agudos".

Obsérvese que de haber una anáfora el proceso de análisis se inclinará también por la lectura partida del nombre:
"...El triángulo rectángulo posee un ángulo recto y otros dos ángulos que no lo son".

En conclusión diremos que los miembros de una colocación son considerados independientemente por el diccionario de palabras y por el diccionario de conceptos, pero que, en ocasiones, cuando se trata de un tipo especial de vínculo ¹⁰, el diccionario de conceptos lo considerará una unidad, y que en esos casos el diccionario de palabras deberá tener una entrada específica asociada a ese concepto, de tal manera que podríamos hablar de nombres compuestos por idiomaticidad u opacidad y de nombres compuestos por colocación. Muchos de los términos empleados en áreas específicas pertenecerán a este último grupo ("traducción

¹⁰ Me refiero a relaciones combinatorias que obedecen a pautas objetivas de ordenación del mundo; en cuanto a las combinaciones de nombre y adjetivo se trata de la relación de coocurrencia denominal (co_rell). Ver apartado 4.3. y también la segunda parte.

automática", etc.). Esto no significa, sin embargo, que las entradas asociadas a un "concepto compuesto" (*línea tangente, ángulo obtuso, traducción automática, etc.*, deban mostrar obligatoriamente la plurilexematicidad en el "reading" ¹¹ (con guiones - *sauce_llorón* - o espacios - *sauce llorón* -), sino que la información de la palabra colocada estará en otro campo).

Para distinguir lingüísticamente un idiomatismo o compuesto de una colocación se pueden realizar una serie de pruebas sintácticas de independencia de los miembros de la combinación (ver apartado 3.2.).

¹¹ El "reading", en contraposición al "headword" que muestra el lexema o raíz de la palabra con inclusión de la afijación, es un campo de la base de datos en el que aparece la lectura de la palabra en su forma flexionada; por defecto, en el masculino del singular.

3.1. Combinaciones libres

Llamamos combinaciones libres a aquellas construcciones que se realizan sin que exista ninguna exigencia o preferencia de tipo léxico que determine sus relaciones. Por ejemplo, "gaviota blanca", "pescado frito", etc. (Naturalmente que una expresión como "gaviota frita" puede producir extrañeza por hacer referencia a una realidad no común en nuestra cultura, pero esto, como ya dejamos claro en el apartado 1.3., atañe al "mundo", que no al lenguaje).

Las combinaciones libres se forman a partir de una ausencia de restricciones en el proceso de selección léxica, aunque, claro está, se rigen por el comportamiento que dictan las reglas de combinación sintáctica y semántica. Esto es, la rección de preposición, el entorno del verbo (valencia sintáctica expresada en forma de rasgos que exigen una determinada estructura), la exigencia de un rasgo semántico determinado para alguno de los actantes, etc. Así es posible construir una expresión como "el panadero confiesa..." pero no "el pato confiesa...", porque es este un verbo que exige sujeto humano. Del mismo modo se efectúa un "chequeo" de tal manera que un verbo que subcategorice determinados complementos obligatoriamente sea rechazado como candidato cuando la estructura de la red no proporciona el entorno adecuado.

De entre las palabras asociadas a un concepto se escogerán aquellas que cumplan las exigencias sintáctico-semánticas, y si estas son varias, se dará preferencia a aquella que cumpla otras condiciones como el mayor índice de uso, la pertenencia al mismo campo semántico o área de aplicación que otras palabras existentes en el texto, o, con mayor incentivo, la pertenencia a un determinado sublenguaje al que están asociadas otras palabras del texto. Los mecanismos de control que deben dar cuenta del índice de uso y de la pertenencia a un sublenguaje, aunque previstos, todavía no están implementados en ATLAS. Como sucede a menudo en cualquier área de la investigación científica, los modelos que se crean para dar cuenta de la realidad no son completos, y, a veces, el científico piensa si su investigación no tomaría otro rumbo de contar con la solución de problemas subsidiarios o adyacentes.

Los constituyentes de una combinación libre son el material del que se componen los diccionarios, pero nunca las combinaciones en sí (a no ser en las frases de ejemplo con que suele completarse la información de la entrada).

Naturalmente que los elementos de las combinaciones libres pueden ser palabras plurilexemáticas ("la elaboración del *cabello_de_ángel* requiere mucha paciencia") y pueden alternar con combinaciones preferentes:
La *policía secreta* española infiltrada en Francia *desarticuló*
la red de los atracadores

No voy a comer *pan gomoso* ni *mantequilla rancia* habiendo un

restaurante en la esquina

Las combinaciones libres se controlan por reglas de subcategorización sintáctica y concordancia semántica según rasgos que aparecen en las entradas del diccionario de palabras.

3.2. Combinaciones preferentes

Las combinaciones preferentes o colocaciones no suelen aparecer en los diccionarios monolingües de forma marcada como ocurre en los bilingües, por la razón obvia de que estas relaciones se ponen de manifiesto en el momento que es preciso traducirlas a otra lengua. Aunque a menudo se ejemplifica el uso de una entrada por medio de su colocabilidad, esta no está tratada de modo consciente, por lo que la información se debe deducir de la información 'flotante'.

Obsérvese la definición que da María Moliner, 1966 (diccionario monolingüe del español) de la colocación "errata de imprenta":

Errata de imprenta. *Expresión de significado claro.*

Algunas veces se hace referencia a la colocabilidad del término, por ejemplo:

erizado, -a. *(aplicado particularmente al pelo).*

Pero esta actitud no es sistemática ni formalmente homogénea,

apareciendo en ocasiones en los ejemplos o como un listado para indicar referencia al final de la definición sin distinguirse de los idiomatismos:

divino,-a. ...

V. *CASTIGO divino, JUICIO divino, LETRAS divinas, Su divina MAJESTAD, PALABRA divina, divina PROVIDENCIA, SERVICIO divino, VOLUNTAD divina.*

Muchas colocaciones son usadas frecuentemente sin que el hablante tenga conciencia de estar recurriendo a relaciones prefijadas de antemano por la tradición lingüística. Esto responde al hecho de que, al contrario de lo que ocurre con las combinaciones fijas o idiomatismos, sus constituyentes suelen ir asociados a conceptos distintos. Esta característica es la que permitió clasificarlos como solidaridades léxicas y decir que pertenecían al grupo de asociaciones transparentes.

Ahora bien, desde el punto de vista de ciertos sistemas como puedan ser los que utilizan una estrategia de transferencia y tienen un diccionario bilingüe entre la lengua origen y la destino, existe una barrera difusa entre lo que se considera colocación y lo que se considera unidad plurilexémica o idiomatismo. Ambos fenómenos, evidentemente, suponen una contextualización, es decir, se detectan y se controlan a través de sus relaciones con el léxico adyacente, y es explicitando ese contexto léxico preciso, al parecer, la única manera de

proporcionar un equivalente en otra lengua a una expresión dada.

Sin embargo, una colocación supone una relación conceptual además de una relación sintagmática, mientras que un idiomatismo o una unidad léxica múltiple o plurilexemática o como quiera llamársele no mantiene ninguna relación lógica o conceptual o semántica entre los miembros que la componen. Por ejemplo, si se observa una locución como "tocinillo de cielo", resulta obvio que no existe ninguna relación lógica en la modificación por 'cielo' que sufre 'tocinillo', y, además, no existe ninguna relación entre el significado de los lexemas y la realidad a la que nombra, a saber, un dulce elaborado con huevo y azúcar. No hay más remedio que concebir esta combinación como una unidad léxica múltiple que irá asociada a un único concepto o superconcepto. Sin embargo, a pesar de que esta distinción pueda parecer clara a nivel intuitivo, en la práctica, cuando no haya una representación conceptual que nos imponga un criterio, puede resultar conflictivo establecer una barrera. Cada tipo de estructura sintagmática proporciona sistemas de reconocimiento apropiados. En el caso de las contrucciones formadas por nombre y adjetivo es clarificadora la prueba de la elisión del nombre, la pronominalización y el uso de una relativa en vez del adjetivo:

Tenemos un ángulo recto y dos que no lo son
Un ángulo que no es recto
ángulo recto = colocación

*Estoy en números rojos y en otros que no lo son

*Unos números que no son rojos
números rojos = idiomatismo

Jugaban un elefante africano y otro que no lo era
Un elefante que no era africano
elefante africano = colocación

*Le apuntarán en la lista negra y en otra que no lo es
*Una lista que no es negra
lista negra = idiomatismo

O sea, se trata de comprobar la independencia sintáctico-semántica del adjetivo con respecto del nombre. Si no se da esa independencia se trata de un idiomatismo o compuesto o unidad plurilexemática, es decir, una combinación fija.

Esta prueba también es útil para discernir colocaciones de idiomatismos de estructura NN o N_de_N:

*Salió un hombre rana y dos que no lo eran
*Un hombre que no era rana
hombre rana = idiomatismo

Vino un buque escuela y otro que no lo era
Un buque que era escuela
buque escuela = colocación

*Había queso de bola y otro que no lo era
*Queso que era de bola
queso de bola = idiomatismo

Probaron el vino de la casa y otro que no lo era
Vino que era de la casa
vino de la casa = colocación

3.3. Combinaciones fijas

En el caso de que se detecte una estructura de este tipo,

tanto la estrategia basada en la transferencia como la interlingua deberán enfrentarse al problema de representar la unión de esas palabras y, al mismo tiempo, permitir la realización de ciertas transformaciones cuando estas sean posibles (ver Aguilar-Amat, 1990:55). Naturalmente que la rigidez semántica que caracteriza estas construcciones fijas impide casi siempre que se produzcan transformaciones de una manera similar a como ocurre en el léxico libre; mayormente se trata de la imposibilidad de realizar cambios en el orden, inserciones, elisiones o pronominalizaciones, pero muy a menudo es posible incluir adverbios o cláusulas de valor adverbial, transconstruccionales, realizar nominalizaciones o dislocaciones de cualquier tipo:

me toman a menudo el pelo
¡vaya tomadura de pelo!
ya me lo han tomado tres veces, el pelo, aquí

Así como con respecto a las colocaciones, como se verá en la segunda parte de este libro, el tratamiento interlingual se manifiesta mucho más económico que en los sistemas de transferencia, el control de unidades plurilexemáticas no es tan espectacular.

La estrategia de análisis de los idiomatismos o, mejor, unidades plurilexemáticas (o "multi-word units" - MWU -) parte, naturalmente de un reconocimiento previo a través de un 'parsing'. En un sistema de transferencia basado en una gramática de

unificación, el reconocimiento de un MWU se produce a través de reglas constructoras que "crean" una unidad plurilexemática según condiciones de adyacencia y después deben eliminar la estructura original. La información morfológica de género, número, persona, tiempo y modo que pudiere haber es asumida por la nueva estructura léxica construida que aparece como tal en el diccionario del nivel superior (ver Aguilar-Amat, 1990 para una información detallada de este proceso) si el sistema posee una estructura modular con diccionarios distintos para cada nivel. El paso ulterior es la transferencia, en el nivel de interfase, de esa unidad plurilexemática de la lengua origen a un lexema de la lengua destino que contenga el mismo contenido informativo, lo cual a veces significa otra unidad plurilexemática idiomática, una no idiomática o un lexema simple.

El análisis en un sistema interlingual difiere de la transferencia principalmente en que no existe un nivel de interfase entre las lenguas implicadas, sino que se crea una estructura abstracta en forma, por ejemplo, de redes semánticas, en la que no existen palabras sino conceptos a los que se ha llegado a través de la desambiguación correcta. (Naturalmente que el proceso de desambiguación también se produce en el caso de la transferencia.) La unidad plurilexemática está asociada a un solo concepto, igual que si se tratara de un lexema simple (y es desde el concepto que deberá efectuarse la correcta selección léxica en la lengua destino). El procedimiento de parsing puede ser similar al descrito en el párrafo anterior o totalmente

diferente, pero esto no depende de las diferencias básicas entre transferencia e interlingua sino de otras circunstancias. El sistema ATLAS, por ejemplo, utiliza una **gramática de producción** (y no de unificación, como es la del modelo descrito anteriormente que procede de la estructura del sistema de transferencia de EUROTRA). En ATLAS las reglas funcionan según estructuras del tipo <CONDICIÓN>::<ACCIÓN>, donde la CONDICIÓN es del tipo de conjunciones o disjunciones de rasgos que provienen del diccionario (del diccionario de palabras y del de colocaciones) y/o de la estructura conceptual si se está en proceso de generación.

Un 'parser' de un sistema como este funciona contrastando estructuras lo más amplias posibles con el diccionario, en el cual las entradas deben estar como unidades simples y luego como unidades plurilexemáticas, si cabe, con una notación del tipo de árbol sintáctico en el que se representa la estructura de la palabra compuesta con sus palabras constituyentes según una convención de la cual la que sigue es un ejemplo:

- entrada: w#se/ /d/w#morf/ /cuenta/ /de/
- categoría: SV1
- árbol: SP1(w#se)SSY()SV1((d)/SE1(w#morf))SSY()
SN1(cuenta)SSY()SDE(de)

. Los nodos sintácticos al mismo nivel de profundidad son divididos por barras (/)

. La inserción de una palabra se realiza entre llaves {} si es obligatoria y entre paréntesis cuadrados [] si es opcional.

. Si es posible insertar un elemento opcional como pueda ser un adverbio o un sintagma adverbial se indica la posición vacía con doble barra (//), y lo mismo para indicar que la posición de dos constituyentes es intercambiable.

4. Niveles de coocurrencia

Cuando Firth (1957:195) hablaba de nivel colocacional se refería a lo que nosotros preferimos llamar registros colocacionales, a saber, colocaciones usadas con preferencia en una determinada clase social o en un sublenguaje:

Just as phonetic, phonological, and grammatical forms well established and habitual in any close social group provide a basis for the mutual expectancies of words and sentences at those levels, and also the sharing of these common features, so also the study of the usual collocations of a particular literary form or genre or of a particular author makes possible a clearly defined and precisely stated contribution to what I have termed the spectrum of descriptive linguistics, which handles and states meaning by dispersing it in a range of techniques working at a series of levels.

Pero, de momento, las consideraciones sociales quedan fuera de nuestro estudio, así como la caracterización estilística de un determinado autor. El trabajo con sublenguajes es una opción que puede ser interesante cuando la aprehensión de ciertos fenómenos del idioma contemplado en su totalidad se escapa al control de las herramientas o los métodos de que se dispone. Pero este no es el enfoque escogido aquí, sino que, por el contrario se pretende dar cuenta del mayor grado de estandarización posible de la lengua.

Cuando aquí hablamos de niveles de coocurrencia es porque nos parece importante distinguir la esencia de los mecanismos de control que se manifiestan a través de las reglas de la gramática y de la subcategorización de las entradas en el diccionario

de palabras (un sistema como el nuestro consta de un 'diccionario de conceptos', un diccionario de colocaciones' y un 'diccionario de palabras', ver apartado 5). Los distintos niveles de coocurrencia, sintáctica, semántica y léxica ¹², a los que nos referiremos encuentran sus fronteras definitorias cuando se trabaja en un sistema empírico. Los fenómenos de coocurrencia sintáctica y semántica se acotan en el diccionario de palabras, la coocurrencia léxica en el diccionario de colocaciones.

4.1. La coocurrencia sintáctica

La idea de coocurrencia sintáctica se refiere a la sintaxis vista como una disposición en sucesión cronológica de las unidades lingüísticas. Da cuenta de las relaciones sintagmáticas existentes entre los elementos de una frase y asigna valores categoriales a los paradigmas. Además permite establecer una clasificación de los verbos y de los nombres predicativos según su "valencia sintáctica" o, dicho de otra manera, según su capacidad para combinarse con diferentes actantes o complementos ¹³. Una catalogación de los distintos tipos de complementos como,

¹² La coocurrencia como fenómeno general afecta también a los niveles fonológico y morfológico, pero su comentario queda fuera de los temas relacionados con el significado conceptual que nos interesa matizar aquí.

¹³ Usamos aquí el término 'actante' tomado por Tesnière de la gramática alemana de los años 50 y 60, pero damos por válida cualquier otra apreciación: argumentos, objetos, etc. y modificadores o adjuntos u otros para los complementos no subcategorizados por el

por ejemplo, los papeles temáticos de Fillmore, permite controlar, previa subcategorización de cada verbo, la oportunidad de una selección léxica, la función del actante (en el caso de los verbos ergativos, por ejemplo, el sujeto no será agente sino objeto), la generación de las preposiciones correctas, etc. Así, por ejemplo, los llamados verbos de movimiento tendrán un 'organigrama' o 'entorno' ("frame") en el que constarán cuatro actantes (o, si se prefiere, dos actantes y dos complementos): sujeto, objeto, origen y destino ("el camión transporta las naranjas desde Valencia hasta la frontera").

Naturalmente que se puede plantear la cuestión de hasta qué punto son posibles estas explicaciones estructurales sin considerar y tomar como base las relaciones semánticas. En realidad hasta hace poco más de quince años el planteamiento era exactamente el opuesto, y se opinaba que la sintaxis o la estructura formal del lenguaje lo era prácticamente todo, siendo el significado nada más que "un término global que abarca todos los aspectos del lenguaje sobre los que sabemos poco" (N. Chomsky 1957:93). La semántica había quedado totalmente abandonada en manos de los filósofos y los antropólogos. Sin embargo, actualmente la semántica se considera una "disciplina integrante de la lingüística" (Leech 1985:14) y su estudio es "el punto de partida más fructífero y apasionante". De todas formas no

verbo. No pretendemos clarificar nada en ese sentido.

se puede caer en el polo opuesto y estudiar el "contenido" del lenguaje sin hacer referencia al plano formal de la expresión. Es inconcebible una teoría del significado aislada de la sintaxis, porque no podemos identificar los enunciados básicos de una lengua sin identificar las estructuras a cuyos significados se refieren los enunciados básicos, y este proceso es en parte anterior al trasvase de significado, y en parte posterior a la selección léxica de los nodos dominantes o núcleos que lideran ese trasvase de significado.

La coocurrencia sintáctica está controlada en el diccionario de palabras sobre todo por la información que se recoge en las entradas verbales (ver el apartado 5), las cuales especifican cuales son los roles temáticos que irán asociados a los complementos de este verbo, cual es la función sintáctica que ejercerán y a qué categoría sintáctica se vincula ese complemento. Asimismo se señalan las preposiciones pertinentes para los oblicuos regidos y el orden en que deberán generarse. El orden indicado es el que se considera preferente y se sigue en la generación, pero deberá tenerse en cuenta que, en fase de análisis, este orden puede estar alterado en el texto original.

Nombres, adjetivos y adverbios también sostienen información que atañe a su realización sintáctica, pero, por ahora, en menor medida.

Los nombres subcategorizan rasgos que indican individualidad, colectividad, continuidad (no contables) o discontinuidad (contables), que afectan a la estructura

morfemática (generación del singular o del plural y presencia o ausencia de artículos), pero que también proporcionan información sobre el tipo de complemento (caso de los colectivos). El control de los complementos exige asimismo que los nombres que son EVENT o STATE se clasifiquen como predicativos. Los nombres predicativos son deverbales cuyos complementos con "de" o "por" tienen rol de agente, objeto, etc., lo que permite diferenciar estos de los genitivos y otros sintagmas preposicionales complementos del nombre.

Los adjetivos tienen información sobre la manera en como forman el grado (sintáctico, morfológico) y si son superlativos léxicos. Este es un primer paso en la clasificación del adjetivo, pero, como se verá, la información colocacional es decisiva en este aspecto.

Los adverbios están subcategorizados asimismo por rasgos (posición temporal, modal, manera, actos de habla, etc.) que les proporcionan cierta capacidad para ser restringidos o restringir ciertas estructuras sintácticas.

4.2. La coocurrencia semántica

Con nivel de coocurrencia semántica nos referimos aquí

al control de significado que afecta a la oportunidad del acto comunicativo según las exigencias de la lógica del mundo real; es decir, es el nivel que controla las "concordancias" semánticas. Ya hablamos de la distinción entre semántica del mundo real y semántica lingüística, lexemática o colocacional, y ya dijimos que, en el proceso de generación, la subcategorización semántica está supeditada a la selección léxica y sólo tiene valor cuando no existe especificaciones a ese nivel.

La coocurrencia semántica se construye a través de rasgos y de restricciones. Los rasgos semánticos que aparecen como subcategorización en los nombres sirven, además de para ser objeto de restricción por parte del verbo, para distinguir las entradas homónimas relacionadas con conceptos distintos y ayudan en el proceso de enlace. Asimismo colaboran a controlar fenómenos relacionados con la determinación y la generación del artículo.

Es importante aclarar que los rasgos y las restricciones semánticas que juegan un papel tan importante en el proceso de análisis para desambiguar oraciones en las que aparece léxico polisémico no son básicas en el momento de efectuar la selección léxica en el proceso de síntesis. Véase un ejemplo de cada caso:

1) Supongamos una frase en la que aparece la palabra "radio", la cual es a la vez el nombre de una institución, de un aparato electrodoméstico, de un elemento químico, de una parte de la rueda de una bicicleta y de la mitad del diámetro de una circunferencia. Si la frase que se pretende analizar ofrece

los suficientes datos como puedan ser el género del nombre, el control de un verbo subcategorizado semánticamente o la especificación de coocurrencia de los adjetivos presentes u otros ¹⁴, la frase podrá desambiguarse y se escogerán los conceptos apropiados al recto sentido. Es decir, si la frase es algo así como: "estoy oyendo la radio" se rechazarán las entradas con género masculino, y, además, dado que el verbo 'oir' está subcategorizado de manera que sólo admite objetos con el rasgo HUMAN (humano), AUTOM (automatismos) o EVENT (eventos), la lectura de 'radio' como institución (LEGAL), quedará rechazada. Lo mismo ocurrirá si en la frase que se pretende analizar el nombre va modificado: "el radio está roto", porque en ese caso el adjetivo lleva en la entrada la información de que sólo modifica a nombres subcategorizados como MASS (concretos) o BODYP (parte del cuerpo) o AUTOM (automatismos), con lo que el nombre 'radio' en el sentido abstracto de "mitad del diámetro" quedará rechazado.

Naturalmente que la sutileza del criterio con el que se hayan escogido los rasgos semánticos y se hayan determinado las relaciones entre ellos y los verbos es determinante en este aspecto (ver apartado 4.2.1. para objeciones).

Todo esto en el supuesto de que "oir la radio" no aparezca como colocación en el diccionario, con lo que el proceso de filtro de las entradas se simplificaría bastante. En realidad,

¹⁴ Cada entrada léxica posee un campo específico en que se detalla las palabras o los grupos de palabras con los que coloca, tal y como se verá al detallar la forma del diccionario de palabras.

como se verá, un buen recuento de las relaciones léxicas sintagmáticas es suficiente en la mayoría de los casos, pudiéndose optar por un espectro reducido de rasgos semánticos no identificado con una representación del mundo sino sólo útil en el control de ciertos fenómenos sintáctico-semánticos o de concordancia semántica, mejor dicho.

2) Pero ahora supongamos que pretendemos generar esas frases que ya hemos desambiguado y de las que obtuvimos sendas representaciones interlinguales.

En este punto se trata de escoger la palabra más relevante que cada caso exige según la competencia de los hablantes de la lengua destino; aquí los rasgos semánticos son redundantes. Véase lo que ocurre con los mismos ejemplos:

Las oraciones "el radio está roto" y "el radio mide cinco centímetros" se generarán a partir de unas redes o estructuras interlinguales en las que el número de concepto asignado a "radio" en cada caso es distinto. Es decir, en el proceso de generación, así como en el de producción de sentido por parte del hablante desde un punto de vista onomasiológico, la polisemia propiamente dicha no existe. Lo conflictivo en este caso, como ya hemos venido diciendo, es la selección de la palabra adecuada en función de su distribución y su colocabilidad cuando las entradas asociadas al concepto son varias.

4.2.1. La elaboración de un árbol semántico

Debo avisar al lector de que este apartado constituye una digresión. Sin embargo, me permito incluirlo aquí, con perdón previo hacia quien se lo salte, porque me parece que puede dar una idea de como la autora ha llegado a la conclusión de que la subcategorización semántica (sin la información colocacional) es sólo una burda urdimbre que, ciertamente, fundamenta una estructura conceptual básica, pero que difícilmente puede servir para controlar el universo de realizaciones metafóricas que se desencadenan a partir de ella para cada lengua particular. Esta imposibilidad es la que ha diezmado los intentos históricos de control por rasgos semánticos del léxico que partían de árboles semánticos que se construían por inducción a partir de una intuición lingüística más o menos enciclopédica.

La otra opción, que ingenua (pero no infructuosamente) concebimos, era la de construir un árbol semántico a partir de un procedimiento deductivo que tuviera en cuenta el análisis progresivo de las relaciones de solidaridad léxica que se ponen de manifiesto al contrastar un par de lenguas particulares (el castellano y el inglés, por ejemplo), sin concebir en absoluto, de momento, que ese árbol semántico pudiese servir para ninguna otra lengua. Así, (léase con conmiseración) en un lejanísimo futuro, los árboles semánticos deducidos de los pares de lenguas contrastados podrían ser comparados a su vez y elaborar un sistema de primitivos semánticos, si no útil, al menos justificado.

El siguiente ejemplo puede dar una idea del caos al que es fácil llegar yendo por este camino:

Supongamos que el ordenador debe traducir el verbo "provocar". El diccionario COLLINS (Smith, 1988) recoge las siguientes traducciones:

1. provocar: to provoke
2. provocar: to rouse, to stir up
3. provocar: to tempt, to entice

Naturalmente que la primera es la traducción más general, la segunda se utiliza en los contextos en que provocar se interpreta como "enojar" y la tercera cuando significa "tentar".

(Es posible que el empleo indiscriminado de la primera lectura bastara para dar una burda pero inteligible traducción en todos los contextos en que pudiere aparecer la palabra "provocar". Pero esto no es siempre así, por motivos lingüísticos o extralingüísticos. En todo caso es obvio que determinados contextos preferirán la segunda o bien la tercera lectura de este verbo.)

¿Cómo puede decidir el ordenador cuál es la oportuna para cada caso? Lo usual es que cada una de las entradas correspondientes a las diferentes lecturas subcategorice la información que permita discriminar pertinentemente según los rasgos semánticos que posean los argumentos regidos por ellas. Por ejemplo:

provocar_2: sema_del_OD=+PSI_PROCESS

Con lo que tenemos que la lectura 2 se caracteriza por regir un objeto directo cuyo sema es de tipo "proceso psicológico",

lo cual daría lugar a la correcta selección en el caso de:

(provocar_2):
provocar indignación/cólera - to rouse sb. to fury
provocar lástima/pena - to rouse sb. to pity
provocar risa - to make sb. laugh

Obsérvese, sin embargo, que la traducción apropiada para "provocar risa" es diferente a la de los otros dos casos, siendo difícil diferenciar ese proceso psicológico de los otros y utilizar por lo tanto un nuevo rasgo semántico. Esto nos forzaría a reconsiderar el rasgo +PSI_PROCESS y ha dividirlo en +PSI_STATE y +PSI_EVENT (evento psicológico), siendo este último el que encajaría con "risa". Naturalmente que esto provocaría la aparición en el diccionario de una entrada más, correspondiente a este verbo:

provocar_2_a:sema_del_OD=+PSI_STATE
provocar_2_b:sema_del_OD=+PSI_EVENT

Ahora bien, además de estas lecturas el diccionario indica otros usos del verbo "provocar":

Lo que sería la entrada 3 de "provocar" se enuncia como sigue:

provocar (cambios,etc) - to bring about/to lead to

Esto se refiere a:
provocar una subida de precios
provocar un reajuste de precios
provocar una reorganización de personal
provocar una toma de decisiones, etc.

Así que esta lectura del verbo debiere estar subcategorizada del siguiente modo:

provocar_3:sema_del_OD=+EVENT

Ahora bien, si el sistema de rasgos con el que se pretende trabajar es un árbol jerárquico, el rasgo EVENT será superior al rasgo PSI_EVENT (que permitía restringir 'provocar risa'), con lo que la entrada 2_b quedaría inmediatamente absorbida por la entrada 3 . Esto daría pie a la enunciación de un principio de no-dominancia:

un árbol semántico debe estar concebido de tal manera que un nudo excluya la posibilidad de marcar positivamente una palabra que es susceptible de ser marcada por un nudo dominado por él.

Lo cual va en contra de la esencia misma de un árbol semántico, por lo que debemos deducir que, o bien el conjunto de rasgos semánticos no debe constituir un árbol, o bien deben marcarse las entradas de manera que se dé cuenta de las excepciones que se producen en las restricciones:

provocar_3:sema_del_OD=+EVENT,-PSI_EVENT.

En caso de optarse por esta segunda opción, es obvio que se produciría una segunda exigencia o principio de adyacencia:

todas las entradas susceptibles de ser marcadas por un rasgo quedarán automáticamente marcadas también con el rasgo negativo correspondiente a los rasgos adyacentes a ese

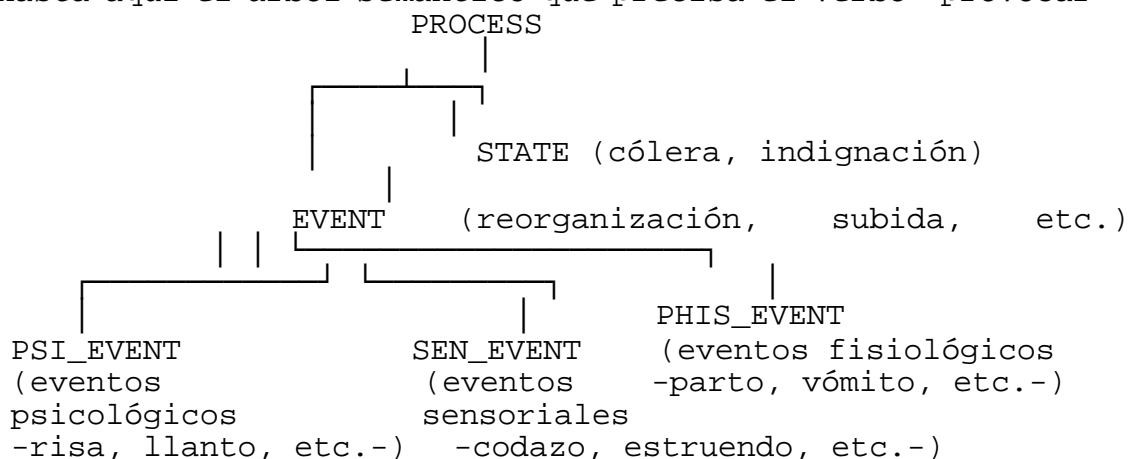
rasgo

Siguiendo por este camino, deberemos concebir aún otro rasgo para dar cuenta de otra lectura de "provocar" que, si bien entraría dentro de la subcategorización de la entrada_3,

tiene sin embargo una traducción distinta:

provocar_5:sema_del_OD:+PHIS_EVENT
provocar un parto - to induce birth
provocar el vómito - to induce vomit

Tanto "parto" como "vómito" deberían ser subcategorizados como eventos fisiológicos (+PHIS_EVENT). Obsérvese como quedaría hasta aquí el árbol semántico que precisa el verbo 'provocar':



En otro orden de cosas, y prosiguiendo con la información del diccionario, observamos que La lectura 3 de "provocar" se caracteriza por el tratamiento dativo del verbo, de tal manera que el sujeto deberá ser no-humano:

provocar_3: sema_del_SUJ=-HUMAN
no me provoca la idea - the idea doesn't appeal to me
me provoca ir - I feel like to go
me provoca un helado - I would like an ice-cream

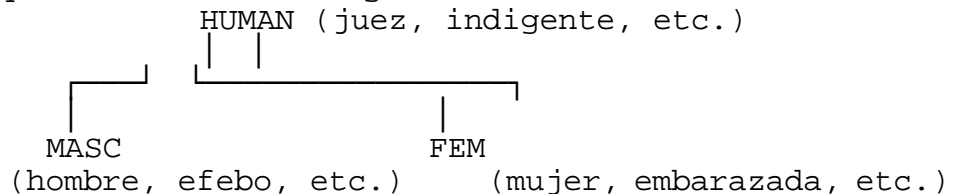
el mar provoca a bañarse - the sea tempts one to bathe

El enunciado **principio de adyacencia** permitirá que todos los rasgos hermanos del rasgo +HUMAN vayan marcados como -HUMAN, pero se precisa también que vayan marcados negativamente todos los rasgos dominados por rasgos marcados como -HUMAN. Esto significa que debe volver a enunciarse el principio de adyacencia en los siguientes términos:
todas las entradas susceptibles de ser marcadas positivamente por un rasgo quedarán automáticamente marcadas con el rasgo negativo correspondiente a los rasgos adyacentes a ese rasgo y a los rasgos dominados por ellos

Pero sigamos. Lo que sería la lectura 5 de "provocar" se refiere al hecho de provocar a alguien para despertar su instinto sexual. Esta lectura aparece en el diccionario COLLINS restringida a la realización de un objeto "mujer" mientras que "provocar" cuando se trata de un hombre se interpreta en relación al instinto de pelea.

provocar_6:
provocar a una mujer - to stir/to stimulate a woman
provocar a un hombre - to rouse a man

Esto supondría que el rasgo HUMAN debería dividirse en masculino y femenino de la siguiente manera:



Existe otra lectura de "provocar" según su traducción al inglés, que da cuenta de la aparición de un objeto directo también de tipo evento pero con matices que podrían ser de tipo aspectual o de modo de la acción: puntualidad, incoatividad, causalidad...

provocar_7:	
provocar una explosión	to cause an explosion
provocar un fuego	to start a fire
provocar una revolución	to spark off a revolution

Esta información no puede ser contemplada en el mismo árbol porque es de otra naturaleza. Las diferencias matizadas en el inglés provocarían distintas entradas con restricciones aspectuales distintas, tema en el que no vamos a entrar aquí.

Por último cabe señalar la existencia de otras entradas que dan cuenta de la lectura dativa de provocar (me provoca un café -I would like a coffee, I do fancy a coffee...-, no me provoca estudiar hoy -I'm not in the mood for ...-, etc.) que vamos a dejar de lado.

Me parece que lo anteriormente expuesto da una idea de la imposibilidad de restringir por medio de rasgos semánticos todas las relaciones colocacionales del léxico. El estudio exhaustivo de cada entrada exige un árbol que puede no coincidir con el árbol considerado en una entrada anterior, de tal manera que se hace difícil imaginar una jerarquía universal.

De aquí resulta obvia la conveniencia de un diccionario

sistematizado de colocaciones que permita filtrar el léxico y dejarlo 'desnudo' para un tratamiento semántico basado en unos pocos rasgos de amplio espectro.

4.3. La coocurrencia léxica

Supongamos que en una red aparece el concepto 3ce7d8 "to express one's opinion" que va enlazado con los verbos *comentar*, *manifestar*, *exponer*, *presentar*, *demostrar*, *expresar*, *decir*, *mostrar*, *hablar*. Suponemos que algunos de estos verbos están en el diccionario de coocurrencias de la forma siguiente:

headword1	co_rel	headword2
manifestar	oper1	opinión
exponer	oper1	teoría
mostrar	oper1	postura
demostrar	oper1	sentimiento

Suponiendo también que exista una red donde haya un concepto relacionado mediante un arco OBJECT con el nodo correspondiente al concepto verbal, y que ese concepto tenga entre las palabras asociadas a él una con el headword2 (por ejemplo *sentimiento*), se escogerá el verbo relacionado (*demostrar*) y se rechazará el resto. La relación de coocurrencia ("cooccurrence relation" o CO-REL) indica el tipo de relación sintagmática entre los "headwords" o palabras clave y es importante en la implementación

(ver la segunda parte de este libro).

La información sobre coocurrencia léxica tal como aparece en el diccionario de colocaciones previsto para la generación se usa para efectuar este tipo de discriminaciones por medio de la relación de coocurrencia, estando esta determinada según la distribución sintáctica de las colocaciones, es decir: existe un grupo de relaciones característico para la combinación de nombre y adjetivo, otro para la combinación de verbo y nombre, etc. Estas relaciones de coocurrencia establecen pares de colocaciones distinguiendo diferencias funcionales entre ambos miembros; la colocación de verbo+nombre, por ejemplo, utilizará unas etiquetas que indicarán matizaciones de la idea temática de objeto, y, en relación con ello, el tipo de preposición que pueda ser preciso seleccionar. Estas relaciones sub-temáticas dependientes del léxico pueden ser muy útiles también de cara a controlar los procesos transformacionales a los que las coocurrencias se ven sometidas en la estructura superficial.

Este es un asunto que se ha rozado en algún momento del presente texto (apartado 2) como derivación de la teoría de Harris, pero que parece pertinente centrar en este apartado.

Es conveniente delimitar claramente el alcance de los términos que se han venido utilizando por los estudiosos de manera más o menos frecuente en torno a las relaciones léxicas sintagmáticas, como son el de función léxica, el de coocurrencia y el de colocación:

Función léxica:

Relación existente entre las palabras. Pueden ser sintagmáticas

(coocurrentes) o paradigmáticas (no coocurrentes):
antonimia, sinonimia, hiperonimia).

Ej.: Relación sintagmática:
caballo/relinchar
peinar/cabello
ir/-lugar-

Relación paradigmática:
antonimia: fracaso/éxito
sinonimia: éxito/triunfo
hiperonimia: éxito/resultado

Coocurrencia:

Relación sintagmática profunda entre dos lexemas independientemente de su realización sintagmática superficial.

Ej.: a) plegar/silla
b) recargar/mechero
c) alto-bajo/voz
d) ligero-profundo/sueño

Se produce un paso de la coocurrencia a la colocación regido a través de reglas de transformación. Puede ocurrir que el número de transformaciones sea nulo, caso en que coocurrencia y colocación coinciden: cabello/rubio, (*rubiar el cabello, *la rubiosidad del cabello)

Colocación:

Realización sintagmática superficial de la relación de coocurrencia entre dos lexemas. (Contrastar con ejemplos anteriores).

a') plegar la silla (V+N)
silla plegable (N+A)
silla plegada (N+Part)

b') recargar el mechero (V+N)
mechero recargable (N+A)
mechero recargado (N+Part)

c') voz alta-baja (N+A)
hablar alto-bajo (V+Ad)

d') sueño ligero-profundo (N+A)
dormir ligeramente-profundamente (V+Ad)

la profundidad del sueño (N+N)

Las colocaciones se clasifican por medio de las relaciones de coocurrencia.

Relación de coocurrencia:

Relación existente entre los miembros de una colocación a nivel léxico. Se trata de un subtipo de la relación sintagmática expresada en la colocación.

Ej. Relación sintagmática N+A, subtipos:

co_rel1terminológica:
instrumento musical
científico nuclear

co_rel2 enfática:
perfecto imbécil
loco rematado

co_rel3existencial:
rostro agraciado
color claro

co_rel4 perfectivos:
local abarrotado
lugar acordonado

co_rel5activa:
conducta delictiva
comportamiento instintivo

co_rel6agentiva:
coche rápido
bailarina ágil

co_rel7objetiva:
problema fácil
amor imposible

co_rel8no_polar:
cerveza alemana
pelo blanco

(Las relaciones de coocurrencia para las colocaciones N+A están explicadas en la segunda parte de esta tesis).

Catalogar las relaciones colocacionales según estas etiquetas permite controlar fenómenos sintácticos como la gradación, o morfológicos como la adverbialización en -mente, así como utilizar filtros en el proceso de desambiguación del análisis. Pero, además, es preciso plantearse si la relación temática que se pone de manifiesto en una relación superficial como pueda ser una colocación es la misma o puede, en la práctica, relacionarse con la relación existente en otra relación superficial. Es decir:

1. ¿Es útil determinar las posibles transformaciones colocacionales de las coocurrencias?

2. ¿Es posible identificar y controlar las relaciones que darían pie a las transformaciones?

La respuesta a la primera pregunta es "sí", puesto que permitiría restringir el número de colocaciones a reseñar en el diccionario, siendo posible elaborar reglas que adaptaran la realización de la selección léxica a cada caso.

La segunda pregunta exige una respuesta meditada. La realización sintagmática o colocacional de las relaciones de coocurrencia depende en gran parte de la idiosincrasia del proceso de formación de palabras, pero sucede que los cambios semánticos afectan a menudo de forma distinta al radical y a su derivado.

4.3.1. Colocaciones y derivación

Si la lengua tuviera un comportamiento regular, verbos, nombres, adjetivos y adverbios estarían relacionados por sus radicales, siendo suficiente con quitar el sufijo formativo de la palabra (o los sufijos, en los derivados secundarios) para hallar dicho radical, la parte "general y abstracta" (Alemany, 1920:1). Pero sucede que esa parte abstracta puede tener un equivalente de otra categoría con distinto radical. Obsérvese, por ejemplo, la colocación 'cabello rubio'. En su caso no existe, por razones históricas, un verbo que sea 'rubiar', y no se producen transformaciones colocacionales. Pero, dado que actualmente se practica tal proceso cosmético, existe la colocación "aclarar el pelo" o "aclarar el cabello", con lo que se pone de manifiesto una relación entre categorías con radicales distintos. Claro que esto sucede porque existe otra colocación equivalente a la anterior, léase "cabello claro", que es quien da pie a la transformación en la colocación verbal "aclarar el pelo". De esta manera, y como ya se sugirió al hablar de la antonimia (ver apartado 1.3.2.) tendremos colocaciones que funcionan como dicotomías cuyas partes son eje de un conjunto relacionado por su antonimia indirecta con el otro miembro:

cabello claro	<----->	cabello oscuro
cabello rubio		cabello castaño
cabello dorado		cabello moreno
		cabello negro

Esto quiere decir que, dada una palabra, por ejemplo, un adjetivo, y una vez identificado su significado por indicación del nombre con el que coloca, debería especificarse su relación con un par dínamo eje, relación que le proveería de antónimo (directo o indirecto) y de posibilidades transformacionales si él mismo careciera de ellas (entendiendo por posibilidades transformacionales la explicitación de la relación de su radical con el de otras categorías).

Ahorabien, esto sólo podría suceder en ciertas combinaciones de nombre y adjetivo, aquellas denominadas "polares" (co_rel3 a co_rel7), puesto que, como se verá, las relaciones de coocurrencia se conciben precisamente en virtud de una clasificación que prevea el comportamiento sintáctico de las colocaciones. Pero, a la vez, la existencia de una colocación equivalente con una polaridad clara puede ayudar a la hora de asignar las etiquetas de las relaciones de coocurrencia; es decir, obsérvese la diferencia entre el ejemplo anterior de "pelo rubio" y el siguiente de "pelo blanco". La colocación "pelo (o cabello) blanco" es una construcción atributiva que pudiere en principio parecer polar, puesto que tradicionalmente se atribuye a blanco el antónimo 'negro', pero que no funciona polarmente en la colocación con 'pelo': *una persona de 'pelo blanco'* (indicador de la edad o de alguna enfermedad) no es sinónimo de *una persona de 'pelo claro'*, y por lo mismo no es antónimo de *una persona de 'pelo oscuro' o 'pelo castaño' o*

'*pelo negro*'. Este es el motivo de que esta colocación se etiquete como *co_rel8* (ver cuadro anterior), y no como *co_rel3* que sería la correspondiente a '*pelo rubio*'.

El hecho de que una coocurrencia carezca de transformaciones colocacionales o de distintos "patterns of cooccurrence" como indica Sinclair (1991:112) no está solamente relacionado con el fenómeno de la polaridad. Los adjetivos terminológicos, también llamados "pertainyms" por Gross&Miller (1990:268), identificables básicamente porque no son atributivos, también tienen esta característica:

- pedra preciosa
- *preciar la piedra
- *pedra preciada
- *preciosidad de la piedra, etc..

Pero también sucede, y esto es lo que pretendíamos mostrar aquí, que la relación morfológica entre los radicales de los lexemas no siempre determina una relación semántica, y no funciona regularmente en la producción de colocaciones. Si recuperamos el ejemplo del concepto 0ceb74 "of a condition of power or action, strong", el cual está asociado a los siguientes adjetivos del castellano: *enérgico, vigoroso, poderoso, potente, fuerte, fornido, robusto, recio, intenso, demoledor, titánico*, observamos la existencia de colocaciones como "coche (AUTOM) potente" y "gobierno (HUMAN) poderoso", pero nos damos cuenta de que los derivados de "potente", los formados con el prefijo negativo *in-* y el de antelación, prioridad o encarecimiento

pre- no colocan con automatismos sino con humanos, contradiciendo

una posible regularidad semántica del radical:

gobierno poderoso	coche potente
*gobierno potente	*coche poderoso
gobierno prepotente	*coche impotente
gobierno impotente	*coche prepotente

Es decir, las palabras formadas por derivación de un mismo radical no están siempre relacionadas de modo regular con el mismo comportamiento colocacional. Como dice Alcoba (1988:2), "mientras en sintaxis se puede decir que las oraciones existentes son un subconjunto de las oraciones posibles, en morfología no se puede decir lo mismo, que las palabras existentes sean un subconjunto de las palabras posibles o regulares constituidas de acuerdo con determinadas reglas morfológicas."

Ahora bien, la hipótesis lexicista no puede desestimar las propiedades de los nombres derivados, ya que no es fácil, como dice Chomsky (1967:144) "ver como un enfoque transformacionalista de los nombradores derivados podría dar razón del hecho de que tanto las estructuras en que aparecen como su propia estructura interna y, con frecuencia, sus propiedades morfológicas sean las de los sintagmas nominales. (...) El problema reside en que las evidentes generalizaciones relativas a las propiedades de las formas básicas y derivadas, en lo que a distribución se refiere, sólo podían expresarse dentro de aquel marco apoyándose en transformaciones gramaticales:

no había otra manera de expresar el hecho de que los contextos en que *refuse* aparece como verbo y aquellos en que *refusal* aparece como nombre guarden una estrecha relación entre sí. Sin embargo, cuando se separa el lexicón del componente categorial de la base y se analizan sus rúbricas apoyándose en rasgos contextuales, esta dificultad se desvanece: podemos introducir *refuse* en el lexicón como pieza dotada de ciertos rasgos selectivos y de subcategorización estricta fijos, pieza que estará exenta de los rasgos categoriales [nombre] y [verbo]; entonces, unas reglas morfológicas en gran parte especiales, peculiares sólo de la pieza del caso, determinarán la forma fonológica de *refuse*, *destroy*, etc., cuando estas piezas aparezcan en posición de nombres; y el hecho de que, ya sean nombres o verbos, *refuse* lleve como complemento de un sintagma nominal o una oración completiva reducida, mientras que *destroy* sólo admita un sintagma nominal como complemento, se expresa por la estructura de rasgos de la rúbrica léxica "neutral", (de igual modo que sucede con las propiedades selectivas)."

Proponemos pues, con Chomsky (1967:145), como hipótesis provisional, que "una gran cantidad de las piezas aparecen en el lexicón provistas de rasgos selectivos y de subcategorización estricta fijos, pero que permiten elegir en lo que respecta a rasgos vinculados a las categorías léxicas de nombre, de verbo y adjetivo; y la rúbrica léxica podrá especificar que los rasgos semánticos dependen en parte de la elección que se haga entre

aquellos rasgos categoriales. Tal es, por lo demás, la situación típica en el lexicón: en general, en las rúbricas léxicas entran en juego ciertas condiciones booleanas entre rasgos, que expresan dependencias condicionales de diversos géneros, y en la medida en que sean regularidades, podrán expresarse mediante reglas de redundancia del lexicón."

Pero, a lo que íbamos, ¿qué sucede con las irregularidades? Según Chomsky (1967:143, nota 10) el alcance de las subregularidades que realmente hay se ha exagerado mucho en los trabajos que adoptan la postura transformacionista. Pero eso no exime de considerarlas. La opción más adecuada, desde nuestro punto de vista, consiste en mantener la información categorial, a despecho de la hipótesis provisional enunciada más arriba, y subcategorizar cada entrada con información concerniente a su relación "genealógica" con lexemas de otra categoría, estén relacionados morfológicamente de un modo regular o irregular, o incluso cuando la relación no sea morfológica.

5. El diccionario electrónico

La traducción automática, como ocurre con muchos otros sistemas, constituye un proceso de acumulación de conocimiento que permite seleccionar una información de salida correcta en el menor tiempo posible. En este caso el tipo de conocimiento que debe ser procesado es el lenguaje natural, por lo que los diccionarios constituyen la fuente de información primordial.

Ahora bien, un diccionario que va a ser empleado por un ordenador debe estar concebido de una manera absolutamente sistemática, puesto que las órdenes que regirán su utilización serán interpretadas rígidamente. Los diccionarios creados para ser usados por humanos utilizan muy a menudo ejemplos y anotaciones que presuponen un conocimiento del mundo por parte del lector. Naturalmente esto no tiene sentido cuando se trata de una mente electrónica, al menos de momento. Lo importante, por lo tanto, es decidir cuál es la información que va a subcategorizar cada entrada del diccionario, cuál es la información que va a precisarse en los análisis que permitan dar cuenta de la estructura de la frase para que el sentido permanezca inalterable en una representación abstracta.

La representación abstracta de la frase, sea cual sea el método de traducción utilizado, es una herramienta indispensable para salvar las diferencias lingüísticas que existen entre las lenguas particulares. Tanto los sistemas llamados de transferencia, los cuales poseen un diccionario muy parecido a un diccionario bilingüe en el que las entradas de una y otra lengua están directamente relacionadas, como los sistemas basados en un sistema conceptual, la precisan. Esta abstracción es fruto de una serie de procesos en que las idiosincrasias sintácticas de cada lengua se convierten a una forma canónica establecida por convención.

El diccionario electrónico en el que se basa el sistema ATLAS recibe el nombre de Diccionario Electrónico del EDR (Electronic Dictionary Research) y ha sido creado por el *Japan Electronic Dictionary Research Institute, LTD* con sede en Tokyo, con la aportación del Estado y de varias empresas privadas.

El diccionario del EDR no está concebido como una mera versión mecánica de un diccionario convencional diseñado para uso humano, sino que pretende aportar toda la información necesaria para el procesamiento del lenguaje natural, incluyendo aquella que normalmente se da por sobreentendida o se explica mediante un ejemplo. Pretende, además, ser independiente de cualquier aplicación específica o algoritmo.

El diccionario está dividido en tres: el diccionario de palabras, el diccionario de conceptos y el diccionario de coocurrencias.

La información superficial está contenida en el diccionario de palabras. El diccionario de palabras está diseñado de manera que cada entrada está conectada con un concepto del diccionario de conceptos mediante un número de identificación, de tal manera que el proceso semántico principal se realiza sobre conocimiento conceptual, y, a su vez, está conectado con otros diccionarios de otras lenguas por medio de los conceptos, de manera que estos actúan a modo de una interfase.

El diccionario de conceptos contenía 400.000 conceptos en abril de 1990. La información contenida en este diccionario es independiente de las lenguas particulares. Los autores del

diccionario de conceptos definen un concepto como la clase de imágenes consistente en los atributos y componentes que pueden ser reconocidos independientemente del contexto o situación en el que la palabra es usada. Por ejemplo, el concepto *c.silla* es un conjunto de imágenes de varias sillas representadas por la palabra *silla*, y también el conjunto de imágenes creado por el grupo de atributos que describe a una silla. De esta forma un concepto viene a ser tanto una definición extensional como intensional de la palabra.

Un concepto que aparece sin ninguna limitación especial es un conjunto que a su vez funciona como un subconjunto de sí mismo. Por ejemplo, el concepto *c.pájaro*, que forma un conjunto definido por varios atributos, puede descomponerse en subconjuntos que dependen de cada atributo. Así *pájaro* puede ser visto como *un animal que vuela* y también como *un animal comestible*, lo que supondrá dos conceptos (conjuntos) distintos que son subconjuntos de *c.pájaro*. Un concepto incluye todos los atributos prescritos por las relaciones con otros conceptos, a saber, *pájaro* como agente de *volar*, *pájaro* como objeto de *comer*, etc. El criterio para distinguir conceptos no es distinto del que se usa en los diccionarios convencionales.

El diccionario de coocurrencias provee de información semántica sintagmática, que permite discriminar parte del léxico como candidato a relacionarse con otra parte del léxico. Esta información podría ser ampliada con datos semánticos

paradigmáticos que permitieran establecer otro tipo de relaciones entre el léxico y salvar distancias graves como la de la anáfora, la metonimia o la metáfora.

Naturalmente, si bien estos tres diccionarios están contruidos independientemente, el hecho de que estén conectados entre sí a través de los 'headwords' y de los CIDs hace que sean un realidad un único gran diccionario que podría, si se precisara, tener el formato de una base de datos única con información altamente compleja.

5.1. El diccionario de conceptos

En ciertos pasajes de este libro se ha señalado la dificultad que entraña la elaboración de un diccionario de conceptos universal que dé cuenta de los sistemas conceptuales existentes en todas las lenguas de una manera, digamos, democrática. De hecho la viabilidad de un diccionario de este tipo es cuestionable. Sin embargo, este trabajo está concebido sobre el supuesto de que, al contrario de lo que ocurre con las palabras, muchos conceptos son compartidos por muchas lenguas del mundo, y que, lo que es más importante, las relaciones existentes entre ellos son universalizables. De esta manera parece razonable suponer que también las relaciones existentes entre los distintos sistemas conceptuales puedan ser definidas a través de relaciones semánticas como parte-todo o super-sub (hiponimia e hiperonimia), o relaciones casuales como agente, escenario, condición, etc. Así, el diccionario de conceptos se divide en descripciones de conceptos y clasificaciones de conceptos según sus relaciones.

El diccionario de conceptos descansa sobre la base del modelo de dependencia conceptual de Schank (1975), influenciado a su vez por la gramática de casos de Fillmore, la cual, a mediados de los años sesenta, intentó solucionar los problemas de las gramáticas basadas en la teoría estándar de Chomsky. Pero, si bien las gramáticas de casos tienen una dependencia excesiva de las lenguas particulares, el modelo de dependencia conceptual persigue un tipo de abstracción superior mediante una estructura

de redes de relaciones semánticas en torno a una acción (concepto verbal). De esta forma la representación semántica adquiere una forma canónica, es decir, las distintas modalidades de ofrecer información se representan de la misma manera. En este sentido el modelo de dependencia conceptual de Schank supone una representación "reduccionista" de las valencias casuales.

En el diccionario de conceptos encontramos símbolos y "headconcepts", que son identificadores de los símbolos, y que poseen una forma a modo de cortas paráfrasis definitorias de los conceptos. Los símbolos son series alfanuméricas que responden a una clasificación lógica de la que aquí se muestra un ejemplo:

1121E Paper/ sheet made from vegetable fiber << 1121 Human artifacts/ inanimate things made by humans << 112 inanimate objects/ physical objects without life existing in space << 11 Physical objects/ things that exist in physical space and whose existence can be verified physically << 1 'Mono' concepts/ concrete, perceptible things, or things whose existence can be recognized through thought

Cada una de las lecturas figurativas o de significado traslaticio de una palabra supone un nuevo concepto, por ejemplo, "avinagrado" cuando se aplica a "vino" es un concepto y cuando se aplica a "carácter" es otro distinto, o "baño" cuando se refiere a la acción de sumergirse en un líquido o a la capa

de un metal sobre otro (baño de plata).

Una palabra puede expresar muchos conceptos diferentes. Por norma general, cuanto más corriente sea la palabra más conceptos representa. Identificar cada uno de estos conceptos es la tarea primordial en el análisis de una oración. Esto implica, naturalmente, un análisis semántico. Sin embargo, el análisis semántico que pueden realizar las gramáticas de casos aun cuando incorporan las restricciones seleccionales propuestas por Katz&Fodor (1963) no es siempre exitoso. Los problemas que se derivan de este procedimiento son los siguientes:

1. No se ha fijado el suficiente número de casos, y los existentes carecen a menudo de bases sólidas, fenómeno observable cuando se pretende establecer el entorno casual de cada verbo individual.

2. No se ha fijado un número apropiado de rasgos semánticos, y los existentes no permiten el buen funcionamiento de las restricciones seleccionales.

Teóricamente estos problemas deberían solucionarse en el diccionario de conceptos gracias a las relaciones de hiperonimia e hiponimia (u otras como parte-todo, equivalencia, etc.), pero esto significaría arrastrar hasta el final del análisis un buen número de objetos erróneos, lo cual no parece lo más acertado. Por otro lado, los mismos autores del diccionario reconocen que, a pesar de haber analizado un gran volumen de textos para

describir las relaciones entre conceptos, no parece posible dar cuenta de todas las relaciones existentes sin que el volumen del diccionario aumente de un modo tal que lo convierta en inefectivo.

La propuesta de este trabajo es utilizar un diccionario de colocaciones en el proceso de selección léxica de la generación, pero resulta además que, a la vista de los problemas existentes en el proceso de análisis, la información colocacional se muestra crucial para colaborar en las tareas de desambiguación. Los rasgos semánticos de Katz&Fodor o la información casual son procedimientos válidos, pero precisan de la matización colocacional que cada lengua particular ejerce a través de sus restricciones seleccionales.

5.2. El diccionario de palabras

El diccionario de palabras se encarga de reseñar la información gramatical del léxico de una lengua particular.

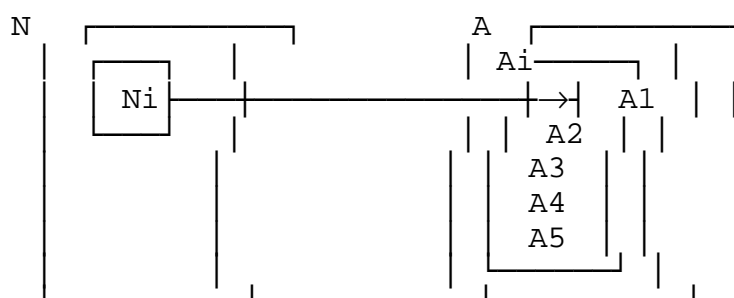
De este modo cada entrada del diccionario debe llevar información morfológica que permitirá establecer concordancias sujeto-verbo y nombre-adjetivo e información sintáctica, que permita identificar la palabra como perteneciente a una categoría y su relación funcional o temática con otros miembros de la frase.

Asimismo es preciso que las entradas posean información relativa al área de aplicación o al sublenguaje al que pertenecen, lo cual permitirá desambiguar entre entradas homónimas cuando se detecte que el léxico del texto pertenece mayoritariamente a un área determinada.

La existencia de rasgos semánticos en el diccionario de palabras puede servir para desambiguar oraciones y para escoger la entrada correspondiente al concepto adecuado para cada ocasión, pero las dificultades que aparecen a la hora de establecer el conjunto o el árbol de rasgos que debe regir las restricciones seleccionales hace pensar en esta solución como secundaria. Estas restricciones seleccionales, como ya dijimos, deberán ser protagonizadas prioritariamente por restricciones colocacionales. En la asunción de este hecho se plantea la necesidad de organizar el diccionario de palabras de una manera consecuente y distinta a la concepción clásica de los diccionarios

tradicionales.

Volviendo a la definición de coocurrencia de un término con una clase de Harris (página 49), observamos que este autor concibe las relaciones entre las clases de léxico N y A como una relación del subconjunto N_i sobre A_i :



Sin embargo, a la hora de construir un diccionario esta relación no siempre se ha racionalizado. La cuestión es la siguiente: si en la entrada de un nombre en el diccionario indicáramos la lista de adjetivos que pueden acompañarle (como por otro lado suele suceder en los grandes diccionarios -ver, por ejemplo, la entrada de "onda" en el María Moliner que produce "onda corta", "onda electromagnética", "onda estacionaria", "onda larga", "onda normal", "onda ultracorta", "onda ultrasonora", etc.-) sería imposible asociar un concepto distinto a cada uno de los significados que se crean por colocación con el adjetivo; ciertamente tendría sentido crear siete u ocho entradas distintas para "onda" si estas se conciben como compuestos por colocación. Pero también resulta que la palabra sola, sin colocar, es el común denominador de las colocaciones,

es decir, irá asociada a un solo concepto con un grado de abstracción mayor (o mayor extensión si se prefiere). Así pues, deberá concebirse una entrada para cada compuesto por colocación y otra para la palabra aislada, y al adjetivo como un ente camaleónico con pluralidad semántica que se multiplicará tantas veces como cambie de significado en virtud de su acompañante nominal, indicado como colocación en la entrada.

De este modo se concibe la información reflejada en el diccionario como la aplicación de A_i sobre N_i , o en todo caso, como una relación recíproca.

Según Allerton (1984:23), quien realiza una distinción entre tres tipos de modificación o "adaptación" de una palabra (que él llama "tailoring", "reshaping" y "characterization"), "nouns seem to have the most fixed meanings of other word classes, whereas prepositions come at the other end of the scale being most amenable to 'tailoring'. (...) In semantic 'tailoring' we have observed specialization in the meaning of one word under the influence of a neighbouring individual word (with a more fixed meaning)". Por todo esto son los adjetivos quienes sufren mayormente el "desdoblamiento de personalidad" que supone la creación de una nueva entrada asociada a un concepto diferente, y es bajo cada una de esas entradas donde se anotan los nombres o los rasgos semánticos de los nombres con los que ese adjetivo coloca.

Es importante que la distribución de la información en el diccionario no responda a un juicio arbitrario. La

producción de entradas se realiza en virtud de los nuevos significados que surgen debido a la modificación o adaptación que sufre una palabra por vecindad de otra, especificándose cuál es la palabra así influyente (información colocacional), de tal forma que vemos como semánticamente es el nombre el que modifica al adjetivo, aunque gramaticalmente se exprese al revés.

Allerton, (1984) describe así el fenómeno:

... the traditional notion of modification refers to the relative contribution to the overall meaning of a phrase made by the component words. In other words, whereas semantic 'tailoring' is a relationship between two word(-meaning)s, modification is a relationship between words and the phrase they are part of.

In a particular phrase like *good pianists* we say that *good* modifies *pianists*, and this means that, in constructing the meaning of the whole phrase, we take 'pianists' as our semantic basis, and modify it, with the notion of 'good' to arrive at the complex meaning 'good pianists'. When, on the other hand, we say that *pianist* 'tailors' *good*, compared with its general range of meaning, is narrowed under the influence of *pianist*.

La cuestión es la siguiente: ¿Cómo debe plasmarse la información colocacional en el diccionario de palabras para que sea útil en la desambiguación de palabras polisémicas? ¿Debe un nombre llevar información sobre los adjetivos con los que coloca y viceversa para los adjetivos? ¿No duplica esto la información en el diccionario?

Sinclair (1991:116) divide las colocaciones en tres grupos basándose en el hecho de que las palabras implicadas en una colocación tienen diferentes índices de frecuencia de aparición, con lo que la colocación tiene un valor distinto en la descripción

de cada una de las dos palabras:

"downwards collocations": cuando el nodo (palabra estudiada) tiene una mayor frecuencia que la palabra 'colocada'.

"upwards collocations": cuando el nodo tiene menor frecuencia que la palabra 'colocada'.

"neutral collocations": cuando la frecuencia es similar en ambos miembros de la colocación.

Sinclair afirma lo siguiente: "There appears to be a systematic difference between upward and downward collocation. Upward collocation, of course, is the weaker pattern in statistical terms, and the words tend to be elements of grammatical frames, or superordinates. Downward collocation by contrast gives us a semantic analysis of a word". Esto podría sugerir una ordenación del diccionario según índices de frecuencia. Pero existe una dificultad en este sentido, y es que el diccionario de frecuencias debería elaborarse teniendo en cuenta la existencia de palabras homónimas que pertenecen a conceptos distintos, por lo que un procedimiento automático de establecimiento de frecuencias sobre un corpus de textos no es lo más apropiado. Tal procedimiento debería ir supeditado a distinciones estructurales o colocacionales que permitieran distinguir la frecuencia de aparición de 'abierto' en relación a 'ventana', la frecuencia de 'abierto' en relación a 'carácter', y 'abierto' en relación a 'herida', además de la frecuencia general que permitiría ordenar otras cuestiones, por ejemplo, de carácter fonético.

La estrategia que vamos a seguir es la siguiente: Suponiendo una colocación de N y A en la que A es un adjetivo terminológico, se constituye una nueva entrada para N con la especificación del adjetivo colocado. Si, dada una colocación N y A, A es un adjetivo atributivo, se constituye una nueva entrada para cada nuevo concepto del adjetivo generado por los nombres con los que coloca (ver los criterios para la clasificación de adjetivos en el apartado 4 de la segunda parte).

En la entrada de un adjetivo atributivo se indica cuáles son los nombres con los que coloca y que le hacen 'adaptarse' semánticamente, su antónimo directo, su antónimo indirecto y/o sinónimos, de haberlos, para cada caso. Por ejemplo:

blanco	nc = vino
	ant1 = tinto
	ant2 = oscuro
	syn1 =
	syn2 = claro
blanco	nc = voz
	ant1 = adulto
	ant2 = grave
	syn1 = infantil;
	syn2 = agudo
blanco	nc = pan;harina
	ant1 = moreno
	ant2 = integral
	syn1 =
	syn2 = normal
blanco	nc = arroz
	ant1 = oscuro
	ant2 =
	syn1 = hervido
	syn2 =

También es posible, indicar la clase (rasgo) semántico del nombre, de tal manera que por la abstracción del rasgo pueda deducirse la coocurrencia de otros lexemas no especificados en la entrada:

```
abierto    nc = PLACE (lugar,campo,mar,etc.)
           ant = cerrado;acotado
           sin = descubierto;ilimitado
```

De esta manera los rasgos semánticos no constituyen reglas de selección por sí mismos, pero participan en las restricciones colocacionales.

5.3. El diccionario de colocaciones

El diccionario de colocaciones para la generación consiste simplemente en una relación entre 'headwords', siendo estas las unidades que restan a la extracción de los morfemas de género y número. Las relaciones de coocurrencia deben ser lo suficientemente sutiles para dar cuenta de las diferencias en los casos en que una palabra coloca de diversas formas; pero además, la elección hecha del concepto determina ya un pequeño conjunto de palabras de tal manera que, una vez determinada una relación entre dos unidades, resulta sencillo establecer una comparación con ese grupo y escoger la que encaja (ver segunda parte).

Smadja&McKeown (1991:229) distinguen dos tipos de colocaciones, las que llaman "*compound collocations*" que se caracterizan por ser secuencias fijas de palabras, y las "*predicative collocations*", más flexibles, que afectan a combinaciones que pueden ir separadas por otras palabras (representado *word[]word*) y que pueden aparecer en cualquier orden dentro de la frase; "but words in such pairs often are not interchangeable with other words in similar semantic classes".

En el apartado 3 de la primera parte nosotros hemos hablado de los **compuestos por colocación** refiriéndonos a aquellas combinaciones que pueden aparecer como conceptos únicos en el

diccionario de conceptos; aunque nuestro criterio no ha tenido en cuenta aspectos de rigidez sintáctica como los que han motivado la distinción de Smadja&McKeown (importante sobretodo para el análisis), los objetos designados coinciden en este caso tanto como las denominaciones utilizadas. En lo que a las colocaciones de N+A se refiere, nuestros compuestos por colocación son estructuras no atributivas y no predicativas que son caracterizadas por la relación de coocurrencia *co_rell* (terminológicos). La rigidez sintáctica de estas colocaciones es una de sus características principales. En aras de una unificación terminológica que atenúe la complejidad llamaremos desde ahora a estas colocaciones *no_atributivas* **colocaciones compuestas**.

El resto de colocaciones observadas, a las que nosotros hemos llamado **atributivas**, son las "predicative collocations" de estos autores. En este sentido nosotros hemos partido de Gross&Miller (1990) quien divide los adjetivos en "atributivos" (ascriptive) y "no atributivos" (non ascriptive), siendo los atributivos divididos a su vez en "predicativos" y "no predicativos". Los primeros incluyen los adjetivos que son solamente predicativos (hambriento, ahíto, etc.), que en castellano se construyen con el verbo 'estar', y también aquellos adjetivos que permiten la transformación en una estructura predicativa ¹⁵.

¹⁵ Según algunos autores como Luján (1980) y Vendler (1968) todas las combinaciones de nombre y adjetivo tienen una estructura profunda del tipo *N es A* o bien un nombre y una oración subordinada de relativo

Los segundos se refieren a colocaciones de adjetivos como 'mínimo' (mínimo indispensable, *lo indispensable es mínimo) , (loco rematado, *el loco es rematado), 'perfecto' (perfecto idiota, *el idiota es perfecto), 'diverso' (diversa índole, *la índole es diversa) o 'mero' (un mero capricho, *el capricho es mero), los cuales no forman compuestos por colocación pero tampoco permiten la forma predicativa.

(Como se verá en la segunda parte de este trabajo, la dependencia semántica del adjetivo con respecto del nombre es tal, que no puede afirmarse que un adjetivo pertenezca a una u otra clase si no se tiene en cuenta su contexto. Es por ese motivo que nosotros no ofrecemos sólo una clasificación del adjetivo, una división en clases susceptible de incorporarse en forma de rasgos en la subcategorización de las entradas del diccionario, sino una descripción de las relaciones que se establecen entre ellos y los nombres, patentes en las combinaciones recurrentes de la lengua.)

Cuando Smadja&McKeown hablan de "predicative collocations" se preocupan por el hecho de que haya palabras mediando la supuesta adyacencia de la relación sintagmática. Naturalmente que esto es un problema para la identificación de las colocaciones en el proceso de extracción automática a través de corpus de texto.

restrictiva (especificativa).

abstracta y no de la frase superficial, al menos en generación.

En el análisis no podrá desambiguarse antes de establecido un primer objeto que muestre las relaciones sintagmáticas, haciéndose en ese momento por rechazo de los objetos sin relaciones de coocurrencia patentes en el diccionario. Esto supondría la eliminación de los objetos no pertinentes fruto de una sobregeneración ocasionada por pluralidad semántica. También, en la gramática de análisis, podrían establecerse reglas que tuvieran en cuenta la existencia de elementos medianeros si fuera necesario, pero lo más práctico parece ser identificar, mediante su relación de coocurrencia o mediante una marca accesoria, las colocaciones que son indisociables. De esta manera el resto, marcado por omisión, puede ser tratado en una estrategia general que pruebe a establecer un encaje saltándose una, dos o tres palabras sucesivamente.

Por otro lado hay que tener en cuenta que si el diccionario de colocaciones pretende utilizarse para la desambiguación en el análisis de oraciones, deberá incorporar información sobre el número de concepto de las entradas que en él se relacionan, porque de otra manera no podrían identificarse los miembros de una colocación cuando hubiera varias entradas homónimas. En realidad, dado que en muchas ocasiones la información colocacional ayuda a precisar el número de entradas que deben efectuarse (ver apartado 5.2.), lo lógico es que tal información acompañe a cada entrada de forma sistematizada, de tal manera que exista un único diccionario.

SEGUNDA PARTE

LAS COLOCACIONES DE NOMBRE Y ADJETIVO

0. Introducción

Cuando se observa el panorama de los intentos clasificatorios del adjetivo uno se encuentra con que los estudiosos han tratado de describir y etiquetar ciertos fenómenos atendiendo solamente a la naturaleza del adjetivo como un ente categorial independiente. Pero ya hemos señalado que el adjetivo es una categoría mediatizada, modificada, por el nombre al que acompaña y que la naturaleza de esas relaciones es nuestro objetivo prioritario.

Una clasificación prototípica del adjetivo que tome en cuenta tanto su formación (sufijación o prefijación) como su sintaxis (prenominal, postnominal o ambos, predicación con 'ser' o 'estar') como su semántica (medidas, colores, nacionalidades, etc.) puede aventurarse tras la observancia de las constantes relacionales, pero siempre irá supeditada a ellas, y sólo será útil para controlar la gramaticalidad de las combinaciones libres cuando ninguna relación de coocurrencia haya sido detectada.

Véase un ejemplo. Un adjetivo como "verde" es un adjetivo

de color, pero denota estado en "fruta verde" (unripe fruit), actitud en "viejo verde" (randy man), tema en "chiste verde" (blue joke), período en "años verdes" (teenage years). Cada una de estas lecturas constituirá una entrada, por lo que será preciso considerar la naturaleza del nombre que le acompaña para desambiguar entre ellas, y antes de aplicar la regla general para combinaciones libres que dice que los adjetivos de color (tipol) sólo combinan con nombres no argumentales (ver apartado 2) y funcionales no derivados.

Suponiendo que al adjetivo en cuestión sólo le correspondiera una entrada, como en el caso de "varonil" (virile), el análisis no supondría problemas de ambigüedad, pero en la generación deberá decidirse cuando debe ser "varonil" (voz varonil), cuando "viril" (miembro viril) y cuando "masculino" (sexo masculino), sin que ninguna diferencia de campo o rasgo entre ellos permita resolver esa cuestión.

Por esta razón concebimos un diccionario de colocaciones que relacione nombres y adjetivos. Como ya quedó profusamente explicado en la primera parte de este libro, dado que en la generación la red semántica (o estructura abstracta) de la frase proporciona un concepto distinto para cada caso nominal (voz, miembro, sexo, vestimenta, etc.) se escogerá la opción léxica que se corresponda con cada uno según el diccionario de colocaciones, de manera que no se produzcan combinaciones no útiles como *voz viril, *miembro masculino o *sexo varonil.

Ahora bien, además de establecer un vínculo, hemos definido

de qué tipo de vínculo se trata a través de las etiquetas correspondientes a las relaciones de coocurrencia; de esta manera se dispone de información que permite discriminar las distintas relaciones léxicas que posea una palabra y catalogarlas en virtud de posibles transformaciones, modificaciones de grado, predicación copulativa o de relativo, predicación con verbo morfológicamente explícito o bien implícito, etc. La clasificación que seguirá, concebida en virtud de las relaciones y que se ha nutrido de intentos no relacionales como el de Vendler (1968), Siegel (1980) y, sobre todo, Gross&Miller (1990) y Tamba-Mecz (1980), nos parece la más útil para un sistema interlinguar como el nuestro, siendo quizá, este sistema, un intento de emular fielmente el proceso humano de producción del lenguaje.

1. Semántica del adjetivo

La nomenclatura empleada en los estudios clasificatorios existentes es poco manejable, como se verá. Los mismos fenómenos a menudo llevan nombres distintos, de manera que se solapan o interfieren unos con otros y, viceversa, se utilizan palabras idénticas o muy parecidas para nombrar realidades distintas.

Así, Vendler (1968) habla de **extensibilidad** y de adjetivos "transferibles" para referirse al fenómeno según el cual la adjetivación de un nombre conduce a la correspondiente adjetivación de su hiperónimo:

mosca verde > insecto verde > animal verde

y califica de no_extensibles a otros adjetivos, como los de medida, por ejemplo, y los que llama de "predicación relativa" (p. 94), es decir, aquellos que no están adscritos directamente al nombre sino sólo a sus actividades:

elefante pequeño > *animal pequeño (medida)
rey débil > *hombre débil (predicación relativa)

Naturalmente existen adjetivos que pueden pertenecer tanto al primer grupo (extensibles) como al segundo (el orden del adjetivo influye en la desambiguación para el castellano, ver apartado 7.):

cocinero bueno > hombre bueno / *hombre bueno

Pero Siegel (1980) se refiere al mismo fenómeno matizándolo con el término de **intersectividad**, produciéndose esta en el caso de que haya "extensibilidad" (Vendler) o sustitución posible por un nombre co-extensional, y lo relaciona con la cuestión de la **restrictividad** (especificación) y no_restrictividad (explicación o aposición):

*Un *jugador* retirado es una *persona* retirada
Un *jugador* muerto es una *persona* muerta

Así, los adjetivos serán restrictivos cuando sean no_intersectivos (como *retirado*), término procedente de la

semántica de Lyons (1968), y *no_restrictivos* cuando haya intersektividad (como *muerto*). Los primeros son, además, considerados **modificadores** o atributivos, mientras que los segundos son **predicativos**.

Por otro lado Pérez Calvo (1986) habla de **categorematicidad** y **sincategorematicidad** para referirse al mismo fenómeno de la extensibilidad, y aun añade el concepto de **proximidad**, afirmando que los adjetivos de color (categoremáticos) están más "próximos" al nombre que los de tamaño (sincategoremáticos). Pero, en realidad, la aportación de Pérez Calvo es dividir los adjetivos en dos grupos atribuyéndoles la condición de **antónimos** y **no antónimos**. El primer grupo afecta a los adjetivos que forman dobles como alto/bajo, bueno/malo, mientras que el segundo afecta a adjetivos como "rojo", "peruano" o "terrestre". Vendler (1968) considera la existencia de "pares", pero sólo en vinculación con el grupo de los adjetivos de medida (A2), el segundo de los nueve tipos de adjetivos que distingue.

Por su parte Siegel (1980) considera tres grupos de adjetivos: los intersectivos, los *no_intersectivos* y los de medida. Pero usa además la distinción entre adjetivos **intensionales** (intersectivos) y **extensionales** (*no_intersectivos*), lo que se suma a una concepción especial de la relatividad que ofrece este autor. La tradición distingue entre adjetivos **absolutos**, que aportan información definitoria o esencial (máquina lavaplatos), y **relativos** que distinguen al nombre por comparación con la misma cualidad de otros entes (máquina ruidosa). Pero

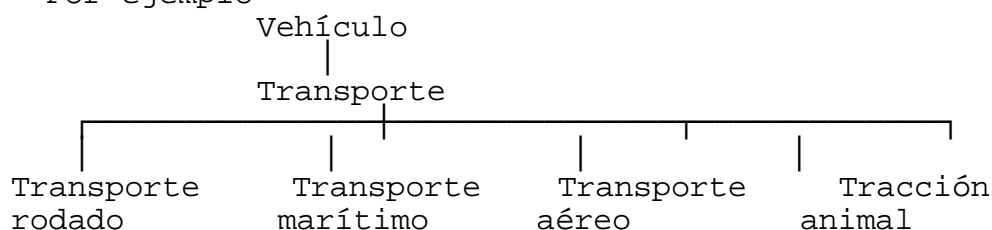
para Siegel los adjetivos de medida no encajan dentro de ninguna de estas dos clases por lo que crea para ellos un grupo aparte. Los adjetivos de medida fallan el test de extensionalidad (sustitución por un co-extensional) pero no porque sus significados sean relativos a los de los nombres sino por motivos de indeterminación. Esto significa que estos adjetivos no son complementos de nombre propiamente dichos y no le caben al autor en su "teoría doble" del adjetivo - algunos adjetivos son a la vez predicativos (intensionales, absolutos, intersectivos, no_restrictivos) y modificadores (extensionales, relativos, no_intersectivos, restrictivos).

Todo esto es muy oscuro y a nuestro parecer poco útil. Siegel confunde varios fenómenos al identificarlos.

En contra de este autor nosotros pensamos que el problema de la intersektividad y/o extensibilidad no depende del adjetivo, sino del hecho de que en ocasiones adjetiva al nombre como ente y en ocasiones adjetiva la acción que este realiza, es decir, depende del rol temático que desempeña o, mejor dicho, de su relación sintagmática. Cuando el adjetivo se refiere no al ente sino a sus hechos es porque el nombre es **funcional** (término acuñado por Vendler), es decir, existe participación agentiva en un proceso verbal. Lógicamente, en estos casos puede producirse cierta ambigüedad dada la polisemia de adjetivos de amplio espectro como 'bueno', pero esto no ocurrirá cuando el nombre no sea funcional.

Las cuestiones de intensionalidad y extensionalidad no afectan directamente al adjetivo sino a través de su relación con el nombre. Cuando Siegel habla de co-extensionalidad parece referirse a lo mismo que Palmer (1976:76) llama co-hiponimia como se deduce del ejemplo (ver página anterior), pero las relaciones de hiponimia no son fáciles de determinar en el nivel léxico, tal y como se explicó en el apartado 1.3.2. de la primera parte, y, en muchas ocasiones, la adjetivación lo que provoca es una mayor profundidad en el árbol de relaciones.

Por ejemplo:



En realidad parece difícil establecer una clasificación del adjetivo si no es atendiendo a razones que, aunque sean básicamente semánticas, tengan un reflejo en un comportamiento sintáctico regular. A este respecto parece ser mucho más lúcida la distinción de Gross&Miller (1990) que divide los adjetivos en dos grupos básicos: *ascriptive adjectives*, cuya relación semántica básica es la antonimia directa o indirecta, por lo que constituyen grupos de palabras relacionadas entre sí, y los *nonascriptive adjectives* o *pertainyms* que carecen de antónimos y cuyas propiedades sintácticas y semánticas son una mezcla de las de los adjetivos y los nombres usados como complementos

de otros nombres, por lo que *"rather than attempt to integrate them into either structure, we maintain a separate listing of nonascriptive adjectives with pointers to the corresponding nouns"*.

2. Sintaxis del adjetivo

La identidad categorial del adjetivo ha sido muy discutida. Se le ha asociado al sustantivo debido a su origen morfológico, de tal manera que la R.A.E., Alcina&Blecua (1975), hablan de "nombre sustantivo" y "nombre adjetivo"; sin embargo, otra corriente procedente de la gramática de Port Royal (1660), desemboca en el generativismo con la aportación al adjetivo de un componente verbal.

Sobre lo que no hay duda es que el adjetivo es una categoría central del lenguaje. Según Sechehaye (1926) citado por Pérez Calvo (1986), el hombre concibe el mundo *dramáticamente* a través del verbo y *estáticamente* en una actitud contemplativa. El adjetivo propicia la visión diferencial del mundo presentado analogías y diferencias.

Desde nuestro punto de vista, y partiendo de esta concepción dualista del lenguaje con un eje basado en el díptico acción/contemplación o evento/estado o, mejor dicho, **evento/actante**, el adjetivo sería una categoría de acercamiento de los miembros de la dicotomía. Es decir, cumple una función doble de asociar acciones a las cosas (verbos a los nombres) y cosas a las acciones (nombres a los verbos), a la vez que se ocupa también de establecer relaciones de los entes entre sí o de los eventos entre sí.

En el primer grupo (el de asociación de acciones a las cosas) encontramos adjetivos con sufijos -ante o -ble, -ero,

-a, -ado (entre otros), como por ejemplo en *mesa plegable*, *alarido espeluznante*, *mano justiciera*, *voz trémula*, *plan trazado*, en los que el adjetivo refleja un evento presente o pasado (caso del participio ¹⁶) o la potencia de un evento. Son parafraseables en la estructura **N que V**, donde V es el verbo explícito en la adjetivación. Así: una mesa que puede plegarse, un alarido que espeluzna, una mano (metonimia por persona) que ajusticia, una voz que tiembla, un plan que alguien trazó.

El segundo (asociación de cosas a las acciones) contiene adjetivos con sufijos en -al como los de *relación conyugal*, *despliegue policial*, *pago quincenal*, *conflicto racial* donde se observa la unión de un nombre derivado deverbal predicativo y un adjetivo que se corresponde con alguno de los complementos de la acción: cónyuges (agente), policía (agente), quincena (complemento temporal), y razas (agente). Esta relación es parafraseable en la estructura **N de N** donde el complemento se analiza como un argumento del verbo explícito morfológicamente o bien implícito en la nominalización.

La relación de los entes entre sí queda manifiesta en las combinaciones *reserva zoológica* (reserva y zoología), *aparato eléctrico* (aparato y electricidad), *material escolar* (material y escuela); son también parafraseables en la estructura **N de N** (reserva de zoología -animales-, aparato de electricidad, material de escuela), pero aquí el complemento no es un argumento,

¹⁶ Respecto a la distinción entre participio y adjetivo ver Abaitua, (1991).

sino un modificador.

Por último, la relación de eventos es patente en *prueba eliminatoria, afirmación determinante, conducta escandalosa, pregunta provocativa*. Son parafraseables también en la estructura **N que V**, pero donde N es un deverbal predicativo, de manera que los actantes de los verbos implícitos o morfológicamente explícitos en el nombre y en el adjetivo pueden no coincidir:
prueba que elimina -> Ni prueba a Nj y Ni elimina a Nj
pero,

conducta escandalosa
conducta que escandaliza -> Ni se conduce y Nj se escandaliza.

Pérez Calvo (1986), desde el punto de vista de su pragmática-léxica, defiende que en la estructura del léxico no existen categorías derivadas. Tanto sustantivos como adjetivos o verbos son, desde su punto de vista, "primitivos" en ese nivel, lo que llama "categorías puras". Así se opone a la gramática generativa-transformacional que hace énfasis en la identidad resultativa de la categoría adjetivo, distinguiendo entre adjetivos primitivos (María es guapa), adjetivos primitivos que remiten a constructos léxicos diferentes (una triste mujer/una mujer triste), y adjetivos derivados. Entre estos últimos señalan la existencia de adjetivos derivados cuya relación lexemática carece de identidad de significado entre el adjetivo y la categoría anterior, y pone Pérez Calvo como ejemplo a *ordenador portátil*

(que puede ser *portado), y adjetivos con identidad de significado morfológico (abarcable, respetable, etc.).

Pero no es fácil establecer la línea divisoria entre las relaciones lexicológicas con identidad de significado y sin identidad de significado. Ocurre que en castellano ya no existe "portar", pero existió en un estrato léxico de un pasado no muy lejano dejando su huella en verbos como 'trans-portar' y otros derivados como 'trans-porta-ble', 'porte', 'portamonedas', 'portador', 'porteador', 'portaderas', 'portavoz' e incluso 'portarse', 'portante' o 'portantillo' con significados derivados. También resulta que pervive en otras lenguas románicas como el catalán (portar -> llevar), el francés (porter), etc.

Un criterio podría ser el de considerar como derivados solamente aquellos adjetivos que coexisten con su antecesor en el mismo estrato diacrónico y sincrónico de la lengua, y desechar el resto, tal y como lo haría un estudiante no nativo, pero este no deja de ser un rasero un tanto artificial.

Por otro lado cabe decir que, si bien la transformación categorial y la derivación son fenómenos reales y con unas consecuencias bien marcadas, se hace difícil aceptar una clasificación del adjetivo que, basada sólo en ellos, no considere cuestiones de semántica léxica que aparecen cuando se tiene en cuenta la función específica del adjetivo en cada caso. Véanse sino los siguientes ejemplos y las paráfrasis que les corresponden:

- a) circunstancia favorable
(circunstancia que favorece a Ni)
- b) detalle despreciable
(detalle que es despreciado por Ni/que carece de 'precio')
- c) bienes deleznable
(bienes que 'deleznan' (resbalan) /bienes que no duran)
- d) estado deplorable
(estado de Ni que hace 'plorar' -llorar- a Nj/Nj llora a causa del estado de Ni)
- e) cantidad considerable
(cantidad que es considerada por Ni)
- f) trato afable
(trato de Ni que permite a Nj 'hablar' -hablar-/Nj habla gracias al trato de Ni)
- g) agua potable
(agua que puede ser bebida por Ni)

Si consideramos la función sintáctica de cada caso observamos que en a) el nombre es sujeto, en b) es sujeto paciente en una interpretación y a la vez sujeto de un verbo existencial, en c) sujeto de un verbo existencial; en d) el nombre es sujeto de un verbo causativo, siendo en realidad el núcleo del complemento causal de la segunda interpretación, con un sujeto existencial en el complemento. En e) el nombre es sujeto paciente otra vez, pero ahora claramente objeto; f) se asemeja a d), y, por último, g) es como e).

Así llegamos a la conclusión de que la catalogación del léxico según su morfología es útil en una clasificación paradigmática del adjetivo que distinga clases semánticas según la información aportada por los afijos implicados (sufijos, prefijos o conjunción de ambos), trabajo que, efectuado para

cada categoría de forma sistemática, resultaría complementario del conocido de Alemany Bolufer (1920). La utilidad de esta clasificación morfológica en el control de transformaciones sintácticas sobre coocurrencias idénticas en colocaciones distintas es patente pero, por desgracia, no funciona siempre (véase apartado 4.3.1.).

Sin embargo tal clasificación resulta estéril en el estudio de las relaciones sintagmáticas que determinan las relaciones coocurrentes de nombre y adjetivo, cuestión que, como hemos señalado desde el principio, centra nuestro interés en aras de una descripción que permita optimizar un diccionario destinado al procesamiento automático de esos datos.

En relación con dichas relaciones sintagmáticas Vendler (1968) establece una clasificación en nueve tipos de adjetivos y, previamente, una clasificación del nombre, que resumimos aquí para que pueda cotejarse con nuestra propuesta (apartados 3 y 4).

Según Vendler los nombres se dividen en tres grupos: derivados deverbales (e_nominals y a_d_nominals), originales y Nm o nombres de medida.

Los nombres originales a su vez se dividen en funcionales, parcialmente funcionales y no funcionales, refiriéndose al hecho de que puedan o no realizar una acción. No nos detendremos en la explicación de cada uno porque esta aparecerá matizada en la clasificación por nosotros dispuesta, y explicada más adelante

(apartado 3). Los distintos tipos de adjetivos se estructuran según se relacionen con uno u otro tipo de nombre.

Cuando un nombre es predicativo afecta al tipo de adjetivo que puede acompañarle (ver la clasificación tipológica del adjetivo más adelante). De tal manera que el adjetivo sirve para desambiguar el tipo de nombre y viceversa. Los adjetivos de tipo1 (cuadrado), por ejemplo, sólo podrán asociarse a nombres puros no funcionales (mesa) (ver próximo apartado), dado que no son adjetivos que puedan adverbializarse, es decir, que no pueden modificar a una acción.

Del mismo modo los adjetivos de tipo3 (amable), pueden adverbializarse y modificar a nombres deverbales, puesto que están adscritos a las actividades del sujeto, y no al sujeto mismo. Así:

nombre puro no funcional + A1
flor roja (es *rojamente)
nombre deverbal + A3
entrada lenta -de Pepe- (entra lentamente)

y no

*flor justa
*entrada roja

Ahora bien, este fenómeno precisa de matizaciones de las que se dará cuenta más adelante; observar la excepción:

nombre deverbal + A3:
bailarina grácil (que baila grácilmente)
nombre puro no funcional + A3:
flor grácil ((de talle) que se mueve grácilmente)

Además, existe casos de adjetivos que pertenecerían tanto a A1 como a A3, puesto que están adscritos tanto al sujeto como a su actividad (ver apartado 1).

En cuanto al adjetivo Vendler establece la siguiente clasificación:

- A1** - ejemplo: flor roja
 - estructura profunda:
 - NA <- N es A (la flor es *roja*)
 - proceden de una relativa restrictiva (N que es A)
 - adjetivan a nombres originales no funcionales
 - son transferibles (manzana roja > fruta roja)
 - adscritos al sujeto: atributivos

- A2**- ejemplo: casa grande
 - estructura profunda:
 - NA <- N que es A para N (casa que es *grande* para [ser] una casa),
o bien
 - NA <- N cuyo Nm es A (casa cuyo tamaño es *grande*)
 - adjetivan a nombres de todo tipo
 - constituyen pares (grande/pequeño)
 - no tienen derivativos adverbiales
 - no son transferibles

- A3** - ejemplo: huida rápida
 - estructura profunda:
 - NA <- N que (e[V+]) D (N que huye *rápidamente*)
 - N es el sujeto del verbo explícito en la nominalización.
 - adjetivan a nombres deverbales, originales funcionales o parcialmente funcionales
 - tienen derivativos adverbiales
 - no son transferibles
 - adscritos a las actividades del sujeto: predicativos

- A4**- ejemplo: problema fácil
 - estructura profunda:
 - NA <- N que es A de [V-] (un problema que es fácil de solucionar [por mí])
 - adjetivan nombres que son el objeto de un verbo transitivo
 - frente a A3, admiten combinaciones con nombres simples: una mesa lenta (de montar)

A5 - ejemplo: hombre ansioso [por]
- la estructura profunda para el castellano sería:
N qu está A [prep V] (un hombre que está ansioso por triunfar)

- rigen preposición
- hay control de sujeto (correferencia entre N y sujeto de V)
- no se puede pasivizar: *to join the class was eager of him
(en castellano a veces sí: triunfar era deseado por él) ¹⁸

A6- ejemplo: hombre prudente
- estructura:
NA <- N es A [cuando/al V] (el hombre es prudente al ir), o
bien,
NA <- qu N [V] es A (que el hombre vaya es prudente)
- se implica una oración sustantiva de tipo predicativo
- a diferencia de A3 rechaza la nominalización agentiva de la
acción: *la ida de Juan fue prudente

A7 - ejemplo: posible, necesario, beneficioso...
- estructura:
qu N [V] es A (que vayamos es necesario) ¹⁹

A8 y **A9** son subgrupos de A7 que cumplen sólo algunas y no todas las transformaciones de su matriz.

La clasificación de Z. Vendler puede ser útil desde el punto de vista del análisis en el campo de la traducción automática, sobre todo después de concebir ciertas modificaciones que nos

¹⁷ Nota del autor.

¹⁸ Nota del autor.

¹⁹ Esta estructura se parece mucho a la segunda opción de la anterior. Por otro lado Vendler no considera la combinación de este grupo de adjetivos con un nombre, siendo posible en castellano: un mal necesario, un asunto importante, un negocio beneficioso, etc.

parece la mejoran con vistas al castellano, pero, en el caso de la generación en un sistema interlinguar, no sirve prácticamente para nada.

Naturalmente que, estando todos los nombres y los adjetivos marcados convenientemente, en el control del análisis se puede usar como filtro, es decir, para dar preferencia a ciertas combinaciones en menoscabo de otras. Proporciona, por lo tanto, un criterio para desambiguar tanto estructuras de nombre y adjetivos polisémicos, como complementos con 'de', que serán argumentos si el nombre es predicativo (ver 3.1. y 3.2.). Pero tampoco es un procedimiento efectivo en todos los casos. Los usos figurativos, metafóricos o metonímicos necesitarán estar contemplados aparte.

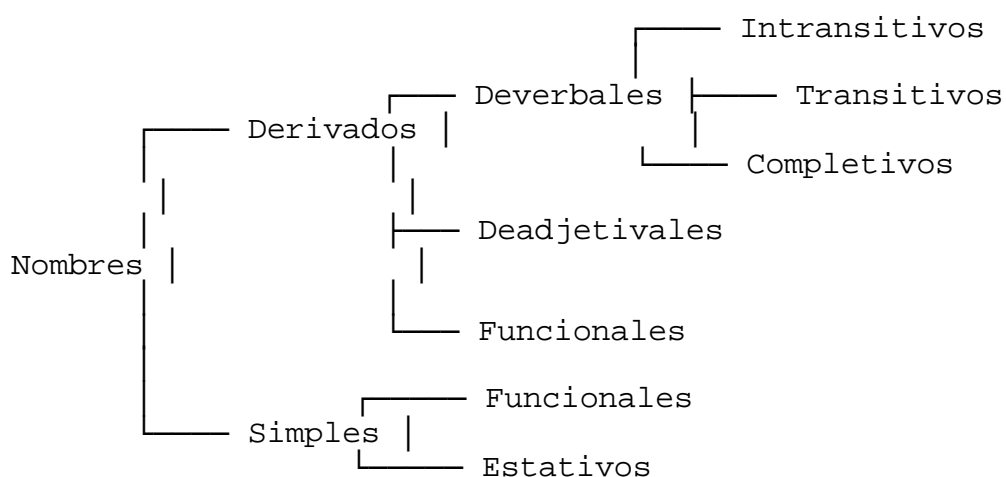
Pero desde el punto de vista de la generación la clasificación general de Vendler deviene estéril porque no ayuda para nada en el proceso de selección léxica que se efectúa a partir de los conceptos de la red (interlingua) o para escoger la traducción correcta en el nivel de interfase (transferencia). 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 aplica a *habla*.

No vale la pena insistir en la necesidad de un diccionario de colocaciones, pero sí en el hecho de que las distintas colocaciones de una palabra pueden distinguirse mediante etiquetas relacionales distintas. A este efecto hemos concebido una clasificación que, como la de Vendler, tiene en cuenta que la naturaleza argumental (predicativo) o no argumental de un nombre determina cierto tipo de relaciones temáticas con el adjetivo, mientras que los nombres derivados no argumentales (derivados de los complementos del verbo) y los nombres simples (puros) condicionan relaciones de otro tipo.

3. 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 deverbales intransitivos
NDVT	Nombres derivados deverbales transitivos
NDVC	Nombres derivados deverbales completivos
NDD	Nombres derivados deadjetivales
NDF	Nombres derivados funcionales

NSF	Nombres simples funcionales
NSE	Nombres simples estativos

3.1. Nombres derivados deverbales intransitivos (NDVI).

Esta denominación agrupa aquellos nombres que proceden de verbos que poseen un sólo argumento, independientemente de que este 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 empresarios* (los empresarios se unen, NDVI) y *la unión de las piezas por el zapatero* (el zapatero une las piezas, NDVT).

3.2. 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 *la acusación del fiscal de prevaricación al ministro

el cambio de opinión de la chica le sorprendió
*el cambio de la chica de opinión le sorprendió

Cuando la preposición regida no es 'de' y no existe un tercer complemento, la rigidez es menor:

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 por el trailer, el traslado de los muebles de una casa a la otra por la empresa, el recorrido de la flecha desde el arco hasta el objetivo, etc.*

3.3. Nombres derivados de verbales 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 ²⁰.

Cabe quizá aclarar que llamamos argumentos o complementos a los que pertenecen al *patrón verbal*, y adjuntos a los demás, coincidiendo, en el caso de sintagmas preposicionales, con la antigua denominación de complementos regidos fuertemente y complementos regidos débilmente. Cuando consideremos el conjunto de todos los complementos y adjuntos en la definición de un verbo hablaremos del "frame" o *entorno verbal*.

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) ²¹; 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

²⁰ La modificación nominal a través de complementos con 'de' precisa de una clasificación relacional similar a la que se ofrece más adelante en este libro para el adjetivo.

²¹ Los criterios de clasificación según rasgos semánticos se encuentran en el documento interno titulado Rasgos semánticos para nombres, en el cual se indican los criterios de delimitación. "Ruido" se considera un evento de tipo sensorial, así como "luz" o "grito".

los adjuntos del verbo ²².

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)

²² El ejemplo es de Santiago Alcoba de la Universidad Autónoma de Barcelona.

Lo cual 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 poseerán un homófono resultativo los nombres derivados deverbales que sean transitivos. A veces la lengua genera una forma irregular que se corresponde con el objeto directo del verbo, como sería el caso de *producto*.

Los nombres derivados deverbales completivos resultativos deberán ir subcaracterizados con una marca distintiva, porque a efectos de coocurrencia tienen un comportamiento similar a los nombres simples existenciales.

3.4. 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 (subjettivos) 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.

3.5. 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, ya sea debido a la incoherencia de sus reglas de formación tanto como a cambios semánticos, y que, de ser deseada la plasmación de las relaciones, estas deberán ser explicitadas de otra forma, con la enunciación del verbo en forma de rasgo, por ejemplo (nom_agt_verb='hablar'). Esta actitud es una opción hacia la reseña de las excepciones.

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 propiamente lingüística con un objetivo de control determinado, a saber, el de relacionar los adjetivos con los nombres, como se verá.

Así y todo, 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.

3.6. 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, ora, confiesa, etc.
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, por ejemplo:
madre amorosa => que (trata) amorosamente a su hijo
científico riguroso => que (trabaja) rigurosamente

Los nombres funcionales, ya sean derivados o simples, son susceptibles de ser adjetivados también por adjetivos no adverbializables, de tal manera que puede crearse cierta ambigüedad (de la que ya se habló al tratar la semántica del adjetivo) cuando el adjetivo es tal que puede modificar tanto al nombre como a la acción que realiza, como ocurre en el caso

de *buen jefe* o *jefe bueno*.

3.7. 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. También es cierto que en ocasiones estos nombres que llamamos 'simples', son derivados de otras categorías latinas, pero si tuviéramos en cuenta consideraciones a este nivel nos saldríamos de nuestro registro. La calificación de 'estativo' ha sido escogida por oposición a 'funcional'.

Ejemplos de nombres simples estativos son: *árbol, mofeta, lápiz, libro, tecla, nota, mueble, atril, manta, rosa, revista, consultorio, casa, fideo, 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 NDVC resultativo; 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 y se comporta como simple.

Además, este es un nombre que resulta de la nominalización de un adjetivo (mueble -> un bien mueble o movable), por lo que en todo caso debiere ser un NDD. Hay que distinguir, sin embargo, entre los NDD morfológicos 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 recordar que los nombres derivados deverbales completivos resultativos 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**:

árbol robusto => *árbol que es robustamente

Naturalmente que el lexicógrafo se encuentra a menudo con casos dudosos. Tómese el ejemplo de "luz", o de otros nombres eventivos sensoriales como "ruido" o "voz". Se podría decir que implicitan verbos como "iluminar" y "sonar", y que permiten transformaciones adverbiales del tipo:

luz intensa => luz que ilumina intensamente
ruido fuerte => ruido que 'suena' fuertemente

por lo que deberían ser catalogados como nombres simples funcionales, pero, por otro lado, estas transformaciones adverbiales son tautológicas, siendo "iluminar" a "luz", a "sol", a "lámpara" tan inherentes como "crecer", "comer" y "beber" era a "planta", y dijimos que las acciones de este tipo no afectaban a terceros y no eran valorables.

Estos nombres se combinan, según Vendler (1968), 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 pudiera servir para facilitar la tarea de establecer un cuadro de relaciones que permitiera, al igual que hacen las relaciones sintácticas, 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:

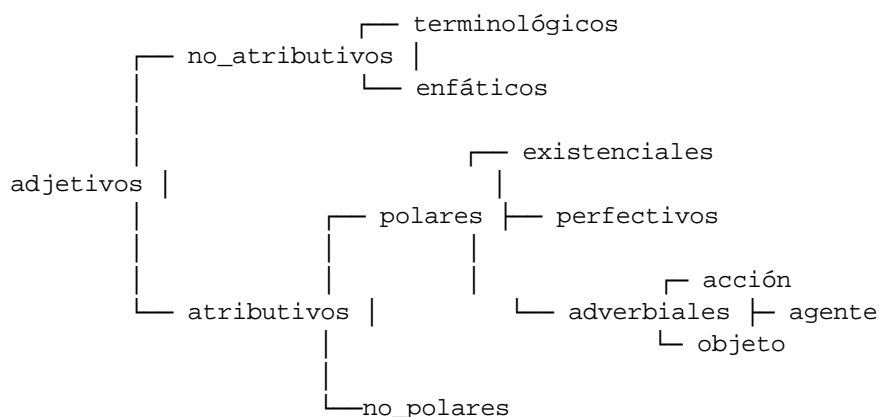
flor oval -> *planta oval

Ahora bien existen más objeciones a una teoría general

de combinaciones de N y A. Como ya se sugirió 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 estas se fosilizan, deben estar caracterizadas en cada caso, y un diccionario de colocaciones puede ser muy útil al respecto.

4. Clasificación colocacional del adjetivo

La clasificación que hemos concebido para el adjetivo, en virtud de los distintos papeles que juega con respecto al nombre al que está asociado, es la siguiente:



Lo que resulta en los siguientes ocho tipos:

- ANNT Adjetivos no_atributivos terminológicos
- ANTE Adjetivos no_atributivos enfáticos
- ATPE Adjetivos atributivos polares existenciales
- ATPP Adjetivos atributivos polares perfectivos
- ATPVA Adjetivos atributivos polares adverbiales de acción
- ATPVG Adjetivos atributivos polares adverbiales de agente
- ATPVO Adjetivos atributivos polares adverbiales de objeto
- ATNP Adjetivos atributivos no_polares

4.1. Adjetivos no_atributivos terminológicos (ANTT)

Los adjetivos de este grupo aparecen en muchas combinaciones con un comportamiento similar. Gross&Miller (1990) llaman a estos adjetivos "pertainyms", nombre que inventan recogiendo la definición de Levi (1978) en la que se les define como 'of, pertaining to, or associated with (some noun)'; anteriormente Tamba-Mecz (1980) los identifica y los denomina "relacionales".

Son adjetivos de este tipo los de las combinaciones siguientes: *movimiento artístico, fiesta nacional, instrumento musical, investigación policial, acta notarial, relación marital, actividad económica, ingeniero electrónico, etc.*

Sintácticamente se caracterizan porque no son predicativos, es decir, que la construcción N es A es incorrecta o muy forzada, y, en general, no usada:

- * este movimiento es artístico
- * la fiesta es nacional
- * el instrumento es musical
- * etc.

Estos adjetivos pueden relacionarse con cualquier tipo de nombre de los anteriormente reseñados.

Juegan un papel similar al de un nombre en posición de modificador, y, morfológicamente, suelen ser adjetivos

denominales; de esta manera en la colocación se combinan dos términos, y no un término y una propiedad (o atributo). Por ese motivo los hemos denominado terminológicos; es preciso señalar, sin embargo, que no debe haber una predisposición a identificar este grupo sólo con aquellos adjetivos que resultan de dividir el léxico en especialidades según la lógica científica (a saber la química, la biología, etc). Aunque estos últimos están, naturalmente, incluidos en este grupo, el hecho de que se parta de un criterio sintáctico-semántico hace que las consideraciones extralingüísticas no sean decisorias. Esto no significa que no pueda asociarse un área de aplicación a cada entrada posteriormente.

Otras de sus características sintácticas, ya señaladas por Gross&Miller (1990) es que estos adjetivos no son susceptibles de experimentar gradación (*ingeniero muy electrónico), no pueden ser nominalizados (*la electronicidad del ingeniero) y no poseen antónimos directos, aunque en inglés (y en castellano por su influencia) se suelen combinar con 'no' (no_atributivo).

No pueden combinar con otros adjetivos, si son atributivos, en conjunción copulativa (*el ingeniero electrónico y eficiente) ni cuando se intenta insertarlos (*el ingeniero eficiente electrónico), pero sí pueden combinarse con ellos si se colocan en posición prenominal o al final de la combinación: *un eficiente ingeniero electrónico, un ingeniero electrónico eficiente.*

Las colocaciones de adjetivos y nombres de este tipo suelen abundar en los glosarios y han constituido a menudo motivo de discusión entre los estudiosos que los han considerado unidades léxicas múltiples y los han tratado del mismo modo que a los idiomatismos, unidades plurilexemáticas idiosincrásicas u opacas, cuando, en realidad, forman pares de conceptos bien definidos en todas las lenguas con hermandad científica. A este respecto una traducción a través de conceptos con la ayuda de un diccionario de colocaciones que escogiera la palabra conveniente, en caso de haber más de una, parece una idea más que razonable ²³.

Gross&Miller (1990:269) conciben como una complicación el hecho de que algunos adjetivos puedan ser usados a la vez como atributivos y como terminológicos, como es el caso de *persona nerviosa* y *alteración nerviosa*, *mueble económico* y *crisis económica*, *persona vital* y *órgano vital*, *suceso cómico* y *actor cómico*. Obsérvese que las colocaciones atributivas admiten gradación (*persona muy nerviosa*, *mueble muy económico*, *persona muy vital*, *suceso muy cómico*), pero no las otras.

Desde nuestro punto de vista esto no constituye un problema dado que no tratamos de hacer una clasificación del adjetivo

²³ De hecho un sistema basado en la transferencia como es el de EUROTRA se ha planteado también el trasvase de información terminológica a través de códigos asociados a los términos, lo cual es un procedimiento prácticamente idéntico al tratamiento interlinguar.

sino de su comportamiento relacional. Esto significa que las etiquetas que indican la naturaleza de la relación de coocurrencia entre un nombre y un adjetivo serán distintas en un caso y en el otro, y esto bastará para controlar las realizaciones colocacionales y sintácticas de esa combinación, al menos en generación. Luego existe la posibilidad de que la camaleonidad del adjetivo no sea tratada por subcategorización sino controlada por contexto. Aunque esto influye en la configuración de la gramática de generación, no lo hace necesariamente en el diccionario. Si se pretende que este sirva también para el análisis, las distintas realizaciones de un adjetivo deberán ser entradas distintas.

4.2. Adjetivos no_atributivos enfáticos (ANTE)

Este grupo de adjetivos se relaciona con el nombre de una manera un tanto especial. Suelen preferir la posición prenominal, y, a menudo, adjetivan a nombres que son nominalizaciones de adjetivos (no derivados, como ya se explicó en el apartado 3.7.). Se trata de adjetivos como *perfecto* en 'un perfecto idiota', *rematado* en 'un loco rematado', *mero* en 'un mero capricho', *indispensable* en 'lo mínimo indispensable', *completo* en 'un completo error'.

La denominación de 'enfáticos' no responde a consideraciones sobre la entonación, aunque dada la enfatización semántica no

se descarta que puedan ser mas proclives que otros adjetivos a sufrir una acentuación tonal.

Estos adjetivos combinan a menudo con nombres de adjetivales, ya sean derivados morfológicos u homófonos, pero también con nombres funcionales y existenciales, siempre con un matiz valorativo-superlativo o peyorativo: un *simple* soldado, un *insignificante* insecto, etc.

Al igual que el grupo de los terminológicos no pueden ir en posición predicativa (*el idiota es perfecto, *el loco está rematado, *el capricho es mero, etc.), y, como ellos, no admiten gradación y no poseen antónimo directo ni indirecto. Pero, así como aquí los hemos considerado no_atributivos, Gross&Miller (1990:269) los consideran atributivos. Ciertamente que estos adjetivos responden a un juicio valorativo, pero no nos parece eso suficiente para considerarlos atributivos en el sentido de que 'ascribe to their head nouns values of attributes' tal y como ellos mismos definen en la introducción.

Estos adjetivos eran destinados a un cajón de sastre no caracterizado por Vendler (1968).

4.3. Adjetivos atributivos polares existenciales (ATPE)

Los adjetivos de este grupo son denominados atributivos pero no como contraposición a los predicativos, sino todo lo contrario, porque tienen la capacidad de convertirse en tales, o, como dice Luján (1980:17), porque "cualesquiera que sea su posición derivan de adjetivos predicativos en la estructura profunda". Llamamos estructura predicativa aquí a la de complementación nominal que se forma con uno de los verbos copulativos 'ser' o 'estar', por contraposición a las estructuras de complementación verbal en torno a verbos eventivos. Sin embargo también existe quien denomina predicativa a la estructura eventiva y atributiva a la copulativa. No vamos a entrar en el problema de la nomenclatura que es más pedagógico que teórico.

Estos adjetivos acompañan principalmente a nombres simples existenciales, a nombres funcionales derivados o simples, y a nombres deverbales completivos resultativos.

Son, por ejemplo: *defecto feo, ojos vivarachos, carácter serio, mujer bonita, brisa suave, etc.*

Llamamos polares a este tipo de adjetivos coincidiendo con la visión ya comentada de diversos autores según la cual estos adjetivos se caracterizan por estar organizados en grupos de antónimos (recuérdese que los antónimos no siempre son directos) y esto organiza el léxico en grupos de palabras relacionadas (ver apartado 1.3.2. de la primera parte).

Debe tenerse en cuenta que aquí no tratamos la polaridad

como una propiedad del adjetivo sino en su relación sintagmática. Recuperando el ejemplo de la página 109, apartado 4.3.1 de la primera parte, el adjetivo 'blanco', cuyo antónimo es presumiblemente 'negro' en términos generales, quedaría excluido del grupo de relación cuyo eje es *claro-oscuro* en relación a cabello, porque las connotaciones de 'pelo blanco' van más allá de esta dicotomía ("blanco" se opone tanto a "rubio" como a "moreno"). Por este motivo lo catalogábamos como un adjetivo atributivo no-polar.

Asimismo ocurre que los antónimos varían según el nombre al que van asociados; como ya se especificó en el apartado 5.2. de la primera parte, página 126, 'voz blanca', 'ropa blanca', 'vino blanco' o 'pan blanco' tienen antónimos diferentes.

Llamamos 'existencial' a este tipo de adjetivación (y a la explicada a continuación con el verbo 'estar') en oposición a 'adverbial', ya que esta no es una adjetivación plausible de adverbializarse, o, lo que es lo mismo, no califica a acciones, ni a actantes, sino a entes pasivos.

Puesto que se trata de adjetivos polares o relativos, son susceptibles de sufrir gradación. Así: *habitación (muy) amplia, tela (muy) fina, libro (muy) grueso, tobogán (muy) alto*, etc.

El hecho de que algunos de estos adjetivos puedan enunciarse con 'estar' (*Juan está gordo*) no confiere carácter de eventualidad (verbal) a la combinación, a diferencia de lo que ocurre con los ATPP que se exponen en el próximo apartado (*el lugar está*

acordonado). En estos casos la atribución se tiñe con un matiz aspectual y podría enunciarse también con el verbo 'ser' seguido de un indicador temporal: *Juan es gordo -actualmente-*, lo que no ocurre con los ATPP (**el perro es sediento -actualmente-*).

Pero la ambivalencia del uso de 'ser' y 'estar' es compleja y no afecta sólo a los adjetivos, por lo que el tema nos alejaría en exceso del campo de estudio.

4.4. Adjetivos atributivos polares perfectivos (ATPP)

Estos adjetivos acompañan generalmente a nombres simples existenciales, a funcionales derivados o simples y a deverbales completivos resultativos.

Al igual que el tipo anterior de adjetivación, se trata de adjetivos polares que pueden formar parte de estructuras predicativas, aunque, en este caso, las prefieren, y no pueden parafrasearse con el verbo 'ser'. Les hemos asignado el calificativo de perfectivos porque conllevan un proceso y un resultado. Así, por ejemplo:

el lugar está acordonado
el lápiz está afilado
los comensales están satisfechos
el perro está sediento
el canario está hambriento
la sopa está caliente
el local está abarrotado
el cielo está nublado
el lector está suscrito
el trabajador está afiliado

En los ejemplos se aprecia claramente dos grupos de adjetivos, los participiales y los denominales. Los denominales se corresponden con nombres experimentales (hambre, sed, calor...). La transformación de estas estructuras en colocaciones del tipo V+N no será posible en el caso de que no exista ningún verbo asociado con el adjetivo, como ocurre en el caso de algunos de estos denominales, pero no por eso deja de existir un proceso experimental a menudo expresable mediante un verbo soporte:

estar hambriento	<	tener/sentir hambre
estar apenado	<	tener/sentir pena

Ahora bien, es preciso distinguir entre las colocaciones ATPP denominales en las que el nombre es sujeto experimentante como en *perro sediento* o *canario hambriento*, y las colocaciones ATPP en que no lo es, como en *sopa caliente*, lo cual tiene cierta incidencia en el tema de la posición del adjetivo (ver apartado 7).

En cuanto a los participiales hay que recordar que, además de su papel en la construcción de las formas verbales compuestas y en la voz pasiva, los participios cumplen una función adjetiva. Tal y como señala Abaitua (1990:39), la mejor evidencia de la transformación en adjetivo la aportan los verbos con participio irregular: bendecido/bendito, imprimido/impreso, sujetado/sujeto, freído/frito, etc., donde la forma irregular suele emplearse como adjetivo.

Pero aun en su uso como adjetivos, muchos participios conservan la posibilidad de transformarse en una estructura verbal transitiva:

acordonar el lugar
afilarse el lápiz
satisfacer a los comensales
abarrotar un local

o intransitiva (correspondiéndose esta última con verbos pronominales):

suscribirse el lector
abonarse el socio
afiliarse el trabajador
nublarse el cielo

Ahora bien, no siempre que un adjetivo sea identificado como participial se debe deducir automáticamente que existe una estructura verbal coocurrente; véase, por ejemplo, el caso de *decisión acertada*, en que la colocación **acertar una decisión* no es usada por los hablantes.

Ocurre, pues, que la especialización del adjetivo le aleja semánticamente del verbo original: *el informe es correcto* ≠ *el informe está corregido*. En este caso, como en los ejemplos no participiales siguientes,

terreno arbolado
vestimenta abigarrada
pelo leonado
voz almibarada

el verbo que se precisa en la estructura copulativa es 'ser', y por ese motivo son considerados ATPE y no ATPP.

Cuando el adjetivo puede emparentarse con una estructura verbal se acerca a las relaciones adverbiales (explicadas en próximos apartados) por la relación temática que le une al nombre, pero obsérvese que no son adjetivaciones de acciones ni de actantes, de manera que no son adverbializables.

La polaridad de este tipo de adjetivos puede manifestarse léxicamente (abarrotado/vacío, nublado/despejado) o bien mediante adverbios de cantidad o negación. La existencia de polaridad indica, como ya es sabido, la posibilidad de que exista gradación.

4.5. Adjetivos atributivos polares adverbiales de acción (ATPVA)

La adjetivación de este tipo modifica la acción expresada en el nombre, y no al nombre mismo, por lo que se combina con nombres deverbales predicativos.

Pueden tomar la forma de una estructura predicativa. Por defecto toman el verbo 'ser' como cópula. También son susceptibles a la gradación.

Distinguimos el hecho de que sean 'de acción' porque el adjetivo califica a la acción misma y sólo indirectamente al agente o al objeto de esa acción. Es el caso de: *conducta ejemplar, comportamiento reprochable, trato amigable, charla animada, movimiento brusco, acción brutal, suceso accidental, respuesta*

exacta, objeción principal, mirada atenta, disparo certero, propuesta atrayente, decisión precipitada, juicio aventurado, comentario acertado, etc.

El hecho de que el núcleo nominal de la colocación sea un nombre predicativo implica que puede existir un complemento del nombre o sintagma nominal con 'de' que se realizará como agente, y en caso de ser un verbo transitivo, un segundo complemento con rol de objeto: *la conducta ejemplar del alumno, la atrayente propuesta de Pepe de comprar bombones, etc.*

Obsérvese la capacidad de esta estructura de transformarse en la estructura verbal **N V Adv**:

N se conduce ejemplarmente
N se comporta reprochablemente
N trata amigablemente a
N charla animadamente con
N se mueve bruscamente
N actúa brutalmente
N sucede accidentalmente
etc.

Estos adjetivos son denominados polares porque puede concebirse un antónimo directo o indirecto aunque este a veces resulte de la negación adverbial del adjetivo o de una de las

palabras relacionadas por antonimia indirecta:
charla animada / charla lacónica
acción brutal / acción no violenta
suceso accidental / suceso premeditado
movimiento brusco / movimiento suave
trato amigable / trato arisco
conducta ejemplar / conducta reprobable
mirada atenta / mirada distraída
disparo certero / disparo errado
etc.

4.6. Adjetivos atributivos polares adverbiales de agente (ATPVG)

Este tipo de adjetivación califica al agente de una acción, y por lo tanto acompaña a nombres funcionales ya sean derivados o simples, pero también a nombres existenciales o completivos resultativos a los que se les atribuye propiedades, ya sean de tipo medicinal o de otro tipo.

Así, esta clase de adjetivos se divide en dos grupos: en el primero la acción está implícita o morfológicamente explícita en el nombre (nombres funcionales deverbales), mientras que en el segundo lo está en el adjetivo.

Algunos ejemplos del primer grupo son: *secretaria eficaz, nadador rápido, seguidor asiduo, cazador cauto, esposa fiel, conductor prudente, persona puntual, etc.*

Este grupo responde a la transformación en la estructura **N que V Adv**, donde V está implícito o morfológicamente explícito en el nombre:

secretaria que [trabaja] eficazmente
nadador que nada rápidamente
seguidor que sigue asiduamente
cazador que caza cautamente
esposa que [se conduce] fielmente, etc.

Ejemplos del segundo grupo son: *circunstancia atenuante, medicamento analgésico, producto alimenticio, dolor enloquecedor, comentario alusivo, libro ameno, programa educativo, escote provocativo, público asistente, etc.*

Este segundo grupo permite la transformación en **N que V**,

donde V está implícito en el adjetivo:

circunstancia que atenúa
medicamento que calma (de *an-* y *álgos* -dolor)
producto que alimenta
dolor que enloquece
mujer que atrae, etc.

La relación morfológica existente entre el nombre o el adjetivo y el verbo correspondiente no siempre es fácil de determinar por motivos históricos, como ya se explicó en anteriores apartados; pero esto no la niega, sino que pide una explicitación de las relaciones a través de rasgos si se pretende algún tipo de control transformacional como es el paso de **NA** a la estructura **N que V** o **N que V Adv.**

La polaridad de estos adjetivos puede basarse en la negación morfológica: *cazador incauto, secretaria ineficaz...* pero este no es un comportamiento sistemático, sino que suele ser mucho más frecuente utilizar otros adjetivos como *torpe, incompetente*, etc. pero también según restricciones de colocación: *un torpe cazador, una secretaria incompetente, un dolor soportable, una circunstancia agravante*, etc. A veces existen cuestiones pragmáticas que provocan la inexistencia de un antónimo, como sería el caso de 'medicamento analgésico' -> *medicamento doloroso, pero sí existe la posibilidad de expresar el contrario por medio de los adverbios 'poco' o 'nada' y, por anglicismo, con el 'no': *un medicamento nada/poco/no analgésico*.

Como ya se ha comentado en otras ocasiones, el hecho de

que exista polaridad supone la posibilidad de gradación, aunque la capacidad de formar el superlativo sintáctico con 'muy' no implica que pueda formarse el morfológico con el sufijo -ísimo, lo cual va indicado en forma de rasgo para cada entrada del adjetivo.

4.7. Adjetivos atributivos polares adverbiales de objeto (ATPVO)

Las combinaciones de este tipo se caracterizan por el hecho de que el nombre de la colocación es el objeto de una acción implícita o morfológicamente explícita en el adjetivo.

Son combinaciones del tipo: *problema fácil, silla plegable, mesa abatible, sonido audible, paraguas plegable, mujer accesible, bien deleznable, material fungible, estado deplorable, detalle despreciable, condición indispensable, etc.*

El sujeto agente de la acción es recuperable a través de la preposición 'por' o 'para': un *problema fácil* para mí (yo soluciono el problema fácilmente), una *mesa plegable* por mí (yo pliego la mesa)... Aunque la paráfrasis suena un poco forzada, es evidente que existe tal sujeto quien 'soluciona fácilmente el problema' o 'pliega la mesa' o 'accede a la mujer', y que el rol del nombre es objetual.

Ahora bien puede ocurrir que el rol del nombre no sea el de objeto directo, sino el de objeto oblicuo, como sería el caso de *precio asequible* (yo accedo al precio), o bien puede

ocurrir que se trate de un verbo dativo, como en el caso de *comida apetecible* (la comida me apetece), verbos que en otras lenguas tienen una estructura objetual clara ("I like/I feel like the food", "J'aime le manger"). Esto hace que este sea el grupo de adjetivos más controvertible.

En general se combinan con nombres existenciales, funcionales simples o derivados y completivos resultativos.

Obsérvese que estas construcciones se asemejan temáticamente a las ATPP que se construían con el verbo 'estar', (el lugar está acordonado => alguien ha acordonado el lugar) pero en aquel caso el aspecto de la colocación era terminativo, mientras que aquí es ingresivo (o "pre-ingresivo").

La polaridad se refleja, como en el grupo anterior, a través del prefijo de negación (sonido **inaudible**), pero también a través de otros lexemas o de la ayuda de adverbios: *mujer accesible/estrecha, comida apetecible/asquerosa o nada apetecible, problema fácil/difícil, paraguas plegable/no plegable o normal, etc.*

Las construcciones de objeto admiten gradación, aunque, como sucede con todos los adjetivos que hemos denominado adverbiales, esta no intensifica la atribución sobre el nombre, sino el modo en que el sujeto realiza la acción sobre el objeto: un precio muy asequible (-> un precio al que alguien accede muy fácilmente), un problema muy fácil (-> un problema que alguien soluciona muy fácilmente), etc.

4.8. Adjetivos atributivos no polares (ATNP)

Este grupo de adjetivos se caracteriza por no poseer polaridad. Acompañan a nombres existenciales, funcionales en su vertiente no funcional (no agentiva) y completivos resultativos, pero nunca van junto a adverbiales de acción. Estos adjetivos son atributos del nombre, no de sus acciones, por lo que no admiten la adverbialización en *-mente*.

Se trata de combinaciones como: *bandera roja, cerveza alemana, porcelana china, emmental danés, bandera tricolor, color amarillo, sacerdote anglicano, monje benedictino, etc.*

Estos adjetivos son predicativos (se pueden parafrasear según la forma **N es A**).

Aunque los colores posean contrarios en la escala cromática según la física óptica, lingüísticamente estos adjetivos no poseen antónimos, porque decir de algo que no es rojo no implica necesariamente que sea verde, ni que no es azul que es amarillo. En este sentido los adjetivos atributivos no polares se parecen a los terminológicos. Como ellos carecen de gradación lógica, de tal manera que no parece adecuado generar un sintagma como "muy rojo" o "muy alemán". Pero, sin embargo, el análisis debe prever la posibilidad de hallar en los textos adjetivos de este tipo con gradación sintáctica, ya que los hablantes usan esas construcciones con la intención de significar 'rojo muy intenso' y 'carácter alemán muy acusado'. Este tipo de gradación se diferencia de la gradación por polaridad puesto

que no atañe a la atribución aportada por el adjetivo, sino al nombre, quien es "totalmente", "completamente", "en todos sus aspectos", rojo o alemán.

Estos los aspectos de la gradación y la polaridad distingue a este grupo de los ATPE. En realidad este tipo de combinaciones es menos proclive a formar colocaciones, y, cuando lo hace, a menudo se trata de combinaciones en las que el adjetivo no es 'inherente' (Quirck, R. et al., 1985), es decir, no caracteriza al referente del nombre directamente o bien expresa connotaciones extralingüísticas: lista negra, números rojos, viejo verde, cabeza cuadrada, etc.

5. La implementación en la gramática de generación

El proceso para realizar una implementación que tenga en cuenta los fenómenos observados en la investigación y que se atenga a los criterios definidos consiste en la enunciación de conclusiones y su posterior formalización.

Esto significa que las reglas de control enunciadas en la gramática deberán empezar por comparar las palabras existentes (ya sea en el texto -análisis- o en relación con un concepto -generación por interlingua- o una palabra de la lengua origen -generación por transferencia-, con las palabras relacionadas sintagmáticamente en el diccionario de colocaciones que le atañe (por ejemplo el de nombre y adjetivo).

Si las palabras implicadas tienen explícita una relación morfológica o semántica con palabras de otra categoría o de la misma, deberá rastrearse también la posibilidad de que exista otra relación colocacional manifiesta que permita desambiguar entre las distintas lecturas de una palabra (análisis) o las distintas palabras de una lectura (generación).

Algunas propiedades morfológicas de las palabras deben estar contempladas en el diccionario, a saber, la posibilidad de que un adjetivo se adverbialice con el sufijo -mente o la de que se forme un superlativo con -ísimo. Naturalmente que deberán ser contemplados los casos en que la asunción de estos sufijos afecta a la raíz de la palabra, como sucede en el caso

de 'amable', 'antiguo' o 'joven':

amable/s	-> ambabil-ísimo
antigu/os,as	-> antiqu-ísimo
joven/jóven/es	-> jovenc-ísimo

En estos casos se dobla o se triplica la raíz según convenga.

Pero otras cuestiones como son la gradación sintáctica y las transformaciones estructurales podrán ser tratadas mediante reglas generales en la gramática de generación siempre que se haya identificado una estructura bimembre como una colocación, tal y como ya se ha comentado anteriormente. De todas formas, aunque es conveniente que la implementación sea inteligente y económica al máximo, en muchas ocasiones esto se consigue diseñando convenientemente el diccionario y la información de las entradas. También hay que considerar las ventajas que puede aportar la uniformidad de cara a futuros cambios o mejoras. Lo más importante, sin embargo, es dejar abierta la posibilidad de que una generalización sea omitida, dado que en las lenguas las excepciones son corrientes.

Este tipo de decisiones a tomar es frecuente en la práctica de la lingüística computacional. La buena armonía entre la gramática y el diccionario es primordial para obtener un resultado satisfactorio (calidad y rapidez), de tal manera que el reparto de funciones no puede ser excluyente. Uno de los mayores errores de los implementadores, a nuestro juicio, ha sido relegar el trabajo lexicográfico a un servilismo totalmente injustificado.

En los próximos apartados se ejemplifica la implementación de las restricciones seleccionales colocacionales en la gramática de generación de ATLAS. Las reglas permiten escoger la palabra marcada con una etiqueta que la identifica colocacionalmente de entre el grupo de palabras asociadas a un concepto. Aunque una misma palabra puede estar relacionada varias veces a través de la misma relación de coocurrencia, nunca será con el mismo concepto.

5.1. El formalismo de ATLAS y su entorno

ATLAS es un sistema diseñado para la traducción multilingüe que trata de salvar la enorme distancia existente entre las lenguas europeas y las lenguas asiáticas. Para ello se basa en el procedimiento denominado interlingua, el cual se apoya en la antigua aspiración estructuralista de lograr una organización del "significado" válida para todas las lenguas²⁴. La representación conceptual tiene la forma de una red semántica, un grafo direccionado compuesto por nodos y arcos, en el que los nodos no están constituidos por palabras sino por conceptos, ya sean del tipo 'ente' como del tipo 'evento' o 'predicado'. Las relaciones entre dos nodos se llaman binarias, mientras que los atributos se expresan como arcos unarios que se refieren

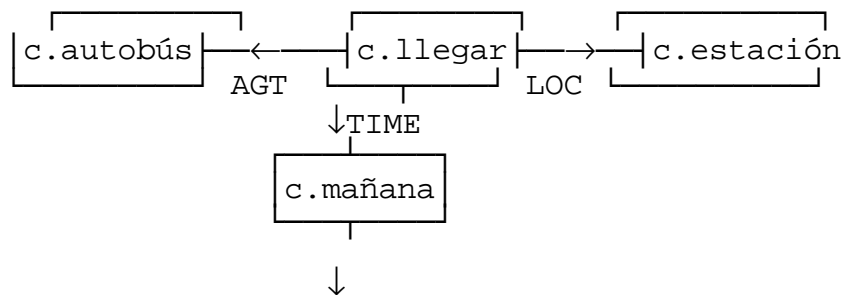
²⁴ Para una descripción reflexiva sobre la interlingua ver Abaitua, J., Soler, J. & Vivaldi, J., 1989.

a un único nodo. Esto es:

(<relation>, <concept 1>, <concept 2>)
(<attribute>, <concept 1>)

Las relaciones binarias son relaciones temáticas del tipo AGENTE, OBJETO, INSTRUMENTO, ORIGEN, DESTINO, LUGAR, MODO, etc., mientras que los atributos se refieren a cuestiones de tiempo (PASADO, FUTURO), aspecto (FINALIZADO, PROGRESIVO), número (TODOS, VARIOS), definición (GENERICO, ESPECIFICO), etc.

La expresión gráfica de estas relaciones facilita su interpretación. Obsérvese el ejemplo correspondiente a la oración "el autobús ha llegado a la estación esta mañana":



THIS

El análisis sintáctico que ha dado pie a esta estructura abstracta o red semántica consiste en la identificación categorial de las palabras implicadas en la frase, la identificación del sujeto, el predicado y sus complementos.

El procedimiento consiste en aplicar reglas gramaticales a las entradas extraídas del diccionario después de efectuado un análisis morfológico correcto. El análisis está basado en

una gramática de estructura de frase en la forma de reglas libres de contexto que permite un análisis determinista.

La interpretación de las reglas de reescritura de acuerdo con la información que provee el diccionario puede, sin embargo, producir más de un objeto debido a ambigüedades de tipo estructural.

Es lo que ocurre en las construcciones del tipo de "comí los fideos con un tenedor", donde se producen dos objetos, uno en el que "con un tenedor" es un modificador de "fideos" y otro en el que es un adjunto del verbo. Este problema debe resolverse, la mayoría de las veces, mediante análisis semántico. Las personas que concibieron ATLAS previeron que el análisis semántico se efectuara en el nivel conceptual, de tal manera que las relaciones existentes en la clasificación conceptual pudieran deshacer estas ambigüedades (comer -> instrumento -> tenedor). Sin embargo, como ya se mencionó al hablar del diccionario de conceptos, parece difícil concebir un diccionario donde todas las relaciones conceptuales existentes estén reseñadas.

Desde nuestro punto de vista la crucial tarea del análisis semántico debería efectuarse en el nivel léxico mediante la explotación de la información colocacional. Eso permitiría efectuar un análisis adecuado al sistema conceptual de cada lengua particular, puesto que el léxico de esa lengua protagonizaría las restricciones seleccionales, y, además, evitaría arrastrar hasta el último nivel la sobregeneración de objetos erróneos, que, en las oraciones largas con muchos

sintagmas preposicionales (de las que suelen estar llenos los textos técnicos), puede ser ingente.

La gramática de ATLAS se basa en la producción, como ya dijimos, que no en la unificación. Esto significa que la información (los rasgos) compatibles no se combinan para obtener un objeto nuevo, sino que la información nueva debe producirse de acuerdo con los rasgos de que se dispone y las concordancias que se establezcan. La gramática consiste en una estructura modular que el implementador puede organizar como desee y variar si los fenómenos nuevos que se van implementando lo requieren. Lo más útil, sin embargo, es concebir una estructura modular lo suficientemente especializada para que cada fenómeno posea un tratamiento según una trayectoria independiente que no provoque interacciones problemáticas. Naturalmente que los módulos superiores, o de acceso primario, deben ser lo suficientemente generales para determinar el camino correcto a seguir a través de la red modular. El orden en que estén establecidos los módulos no es importante, y tampoco lo es el orden de las reglas, dado que recogen información de tipo declarativo.

El lenguaje de usuario que se utiliza en la gramática de ATLAS está basado en reglas del tipo <CONDICIÓN> :: <ACCIÓN>. La condición supone conjunciones o disjunciones de rasgos que provienen tanto de la estructura conceptual como de la información recogida en el diccionario como de la gramática misma, por asignaciones secundarias realizadas a través de las mismas reglas.

En nuestro caso concreto la parte de la condición debe referirse a la identidad categorial y a la relación de coocurrencia que relaciona la palabra con otra con la que coloca.

La acción puede consistir en la indicación de cambio de módulo, en el envío de un rasgo a un nodo anterior o posterior, en la fijación temporal o permanente de un rasgo en el nodo, en el salto a través de un determinado arco o en la generación de la palabra.

Es importante señalar la existencia de un mecanismo llamado LOOK-AHEAD que permite recorrer toda la red semántica antes del proceso de generación para apreciar los aspectos estructurales que determinarán las acciones que se efectúen.

5.2. Las reglas restrictivas colocacionales

La utilización de las etiquetas de coocurrencia en el control de la selección léxica en la gramática de generación es distinta y bastante más crucial que en la gramática de análisis, donde la confrontación con un diccionario de colocaciones no es sino una herramienta más, aunque a veces decisiva, en las tareas de desambiguación. En primer lugar nos plantearemos la implementación en la fase de generación, donde, bien el concepto (caso de la interlingua), bien la entrada semánticamente pertinente correspondiente a la palabra de la lengua origen

(caso de la transferencia), ya han sido seleccionados. Naturalmente que las consideraciones que se expresen aquí para la técnica de transferencia son sólo suposiciones, mientras que las que se ajustan al marco de ATLAS, sistema de interlingua y punto de referencia de este trabajo, son de aplicación directa, y por eso ocupan un lugar central.²⁵

En una estructura abstracta del tipo 'red semántica', el grupo de reglas que condicionaría el salto del nodo del nombre al nodo del adjetivo si hubiera una relación de coocurrencia entre ellos debería ser del siguiente tipo, donde + indica conjunción, | indica disyunción y ~ negación:

NOMBRE -----

```

/ NOUN + ANTT / I.mod: (ANTT) / G / * ;
/ NOUN + ANTE / I.mod: (ANTE) / G / * ;
/ NOUN + ATPE / I.mod: (ATPE) / G / * ;
/ NOUN + ATPP / I.mod: (ATPP) / G / * ;
/ NOUN + ATPVA / I.mod: (ATPVA) / G / * ;
/ NOUN + ATPVG / I.mod: (ATPVG) / G / * ;
/ NOUN + ATPVO / I.mod: (ATPVO) / G / * ;
/ NOUN + ATNP / I.mod: (ATNP) / G / * ;

/ NOUN + ~ANTT + ~ANTE + ~ATPE + ~ATPP + ~ATPVA + ~ATPVG +
  ~ATPVO + ~ATNP / I.mod / G / * ;

```

ADJETIVO -----

```

/ ADJ + ANTT / O.mod: (ANTT) / G / * ;
/ ADJ + ANTE / O.mod: (ANTE) / G / * ;
/ ADJ + ATPE / O.mod: (ATPE) / G / * ;
/ ADJ + ATPP / O.mod: (ATPP) / G / * ;

```

²⁵ La distinción entre los sistemas de transferencia y de interlingua no es más que una distinción estratégica que no varía sustancialmente el problema fenomenológico lingüístico a tratar ni las posibles soluciones a nivel teórico.

```

/ ADJ + ATPVA / O.mod: (ATPVA) / G / * ;
/ ADJ + ATPVG / O.mod: (ATPVG) / G / * ;
/ ADJ + ATPVO / O.mod: (ATPVO) / G / * ;
/ ADJ + ATNP / O.mod: (ATNP) / G / * ;

/ ADJ + ~ANTT + ~ANTE + ~ATPE + ~ATPP + ~ATPVA + ~ATPVG + ~ATPVO
+ ~ATNP / O.mod / G / * ;

```

En una notación simplificada se leería así:

NOMBRE -----

```

[<CONDICIÓN:
  Categoría: NOMBRE
  Relación de coocurrencia: X
  Tipo de arco: INPUT
  Condición del nodo opuesto: X>
  ::
<ACCIÓN:
  Tipo de acción: GENERAR
  Mensaje: >]

```

ADJETIVO -----

```

[<CONDICIÓN:
  Categoría: ADJETIVO
  Relación de coocurrencia: X
  Tipo de arco: OUTPUT
  Condición del nodo opuesto: X>
  ::
<ACCIÓN:
  Tipo de acción: GENERAR
  Mensaje: >]

```

La regla es aplicable de modo general a todas las categorías:

```

[<CONDICIÓN:
  Categoría: CAT
  Relación de coocurrencia: REL
  Tipo de arco: OUTPUT|INPUT
  Condición del nodo opuesto: REL>
  ::
<ACCIÓN:

```

Tipo de acción: GENERAR
Mensaje: >]

El generador cuenta con un mecanismo que le permite barajar todas las fichas (entradas) asociadas con el nodo que se indica como opuesto hasta encontrar la condición requerida, en este caso el rasgo indicador de determinada relación colocacional.

Obsérvese que en el caso de que el nombre no posea ninguna relación de coocurrencia se omite el condicionante de la etiqueta, y se genera sin otra condición. En realidad son muchas otras las condiciones que podrían aplicarse, a saber, una comprobación de la coocurrencia del hiperónimo (1) o del hipónimo (2) del nombre con el adjetivo, la coocurrencia en caso de pertenencia a un conjunto de palabras relacionadas con un eje de antónimos (3), o la coocurrencia de la base si se trata de un derivado o del derivado si se trata de una base (4), así como de la palabra que esté ocupando el lugar de la base en el caso de irregularidades por radical culto (urbano/ciudad) u otros (orador/hablar):

1. [<CONDICIÓN:
 Categoría: CAT
 Relación de coocurrencia: NONE
 Headword del hiperónimo: SUP-HEAD
 Relación de coocurrencia de SUP-HEAD: REL
 Tipo de arco: OUTPUT|INPUT
 Condición del nodo opuesto: REL>
 ::
 <ACCIÓN:

Tipo de acción: GENERAR
Mensaje: >]

2.[<CONDICIÓN:

Categoría: CAT
Relación de coocurrencia: NONE
Headword del hiperónimo: SUP-HEAD
Relación de coocurrencia de SUP-HEAD: NONE
Headword del hipónimo: SUB-HEAD
Relación del coocurrencia de SUB-HEAD:REL
Tipo de arco: OUTPUT|INPUT
Condición del nodo opuesto: REL>

::

<ACCIÓN:

Tipo de acción: GENERAR
Mensaje: >]

3.[<CONDICIÓN:

Categoría: CAT
Relación de coocurrencia: NONE
Headword del hiperónimo: SUP-HEAD
Relación de coocurrencia de SUP-HEAD: NONE
Headword del hipónimo: SUB-HEAD
Relación del coocurrencia de SUB-HEAD:NONE
Headword del sinónimo: SYN-HEAD
Relación de coocurrencia de SYN-HEAD:REL
Tipo de arco: OUTPUT|INPUT
Condición del nodo opuesto: REL>

::

<ACCIÓN:

Tipo de acción: GENERAR
Mensaje: >]

4.[<CONDICIÓN:

Categoría: CAT
Relación de coocurrencia: NONE
Headword del hiperónimo: SUP-HEAD
Relación de coocurrencia de SUP-HEAD: NONE
Headword del hipónimo: SUB-HEAD
Relación del coocurrencia de SUB-HEAD: NONE
Headword del sinónimo: SYN-HEAD
Relación de coocurrencia de SYN-HEAD: NONE
Headword de la base: BASE-HEAD
Relación de coocurrencia de BASE-HEAD: REL
Tipo de arco: OUTPUT|INPUT
Condición del nodo opuesto: REL>

::

<ACCIÓN:

Tipo de acción: GENERAR

Asimismo podría comprobarse la identidad de las áreas de aplicación o los campos semánticos en que se hubieran catalogado las entradas relacionadas, o bien comprobar la identidad del área de la palabra con la del texto completo de haberse definido este al principio.

Finalmente también cabría la posibilidad, antes de escoger una entrada arbitrariamente o de tomarlas todas, de seleccionar aquellas con un índice de frecuencia de aparición más elevado; sin embargo esto deriva en el problema subsidiario del establecimiento de las distintas frecuencias de los distintos homónimos en virtud de sus discrepancias conceptuales. De esta manera, es fácil deducir que el establecimiento de un índice de frecuencia para cada entrada no puede ser hecho sino después de elaborado un diccionario donde las entradas se distingan, entre otras cosas, por su colocación, de manera que el recuento informatizado pueda utilizar esa información para distinguir unas entradas de otras en los casos de pluralidad semántica.

Naturalmente que las sucesivas comprobaciones de coocurrencia a través de los diccionarios de colocaciones podrán y deberán efectuarse para el nodo opuesto en el caso de que fallen para el antepuesto, a saber, la comprobación de las relaciones de coocurrencia del hiperónimo, el hipónimo, el sinónimo y la base.

6. La construcción automática de un diccionario de colocaciones

La elaboración manual del diccionario de colocaciones es, obviamente, costosa, y es lógico plantearse la extracción automática de los pares de palabras a partir de un corpus seleccionado de textos, porque ya hemos visto que, al menos en el proceso de selección léxica en un sistema interlinguar, un diccionario de estructuras bimembres es muy útil, aun cuando no se varíe la estructura del diccionario de palabras.

Posterior a la informatización de datos mediante la utilización de un escáner y de un reconocedor óptico de caracteres (OCR) ²⁶, un punto de partida importante podría ser la identificación previa de la categoría a la que pertenecen las palabras del texto, lo cual no parece difícil de lograr tras una comparación con el diccionario de palabras de ATLAS o con cualquiera de los que actualmente existen, puesto que casi todos ellos incluyen esta información.

Naturalmente que existe el problema de los homónimos, pero, dado que lo que se va a cotejar es la vecindad de cada palabra, parece que, en gran parte, esta colaboraría en la desambiguación de la homonimia léxica (variantes de la misma categoría con distinto significado). Pero esta solución es un pez que se muerde

²⁶ La fiabilidad de reconocimiento de la tecnología OCR es entre un 98 y un 99,9 por ciento.

la cola, porque para poder desambiguar los distintos significados por medio de su colocación, sería preciso que el diccionario de palabras llevara la información colocacional de cada variante o acepción, lo cual si bien no excluye los homónimos léxicos del recuento automático, impide su diferenciación.

En cuanto a los homónimos sintácticos (variantes idénticas pertenecientes a distintas categorías), un proceso de análisis morfológico y sintáctico podría determinar la categoría léxica según la estructura sintáctica de la oración y la categoría del sintagma.

Lo mismo ocurre con los homónimos morfológicos (variantes idénticas que pertenecen a categorías gramaticales distintas, como por ejemplo pasado y presente, primera y tercera persona, etc.).

El análisis sintáctico sería asimismo de utilidad cuando existieran palabras insertas en la colocación, como puedan ser, para el caso de los adjetivos, adverbios (*baraja descaradamente marcada*) o cuantificadores (*bibliografía muy extensa, curva poco cerrada, producto altamente contaminante*), modificadores nominales (*acto de presentación inaugural*), verbos copulativos (*la caída fue aparatosa, la carne está tierna*), etc.

Claro está que este tipo de situaciones puede complicarse por cuestiones de derivación (*el hijo luchó para demostrar su legitimidad -hijo legítimo*), o de pronominalización, elisión, anáfora... temas que las gramáticas suelen resolver con dificultad.

El principal problema podría ser la resolución de la ambigüedad estructural, puesto que la duda respecto a la paternidad de los sintagmas preposicionales imposibilitaría la identificación de las colocaciones del tipo **N de N** y de los entornos verbales, por lo que un estudio profundo sobre la complementación nominal resulta urgente.

Así y todo, un programa que busque y extraiga automáticamente los pares de palabras de categorías establecidas (nombre y adjetivo o verbo y nombre, por ejemplo) parece posible, tanto si su posición es colindante en el texto como si después de cierto análisis se produce una adyacencia del léxico en el árbol sintáctico.

Dicho programa puede informar además de la posición prenominal o postnominal en el caso del adjetivo, y según se conciba el algoritmo, de la prefijación y/o sufijación, así como de la existencia de determinadas preposiciones en el caso de las regencias verbales o nominales, la definitud o indefinitud, etc.

La cobertura de un diccionario construido por este método depende de la selección de los textos a cotejar. Estos pueden ser escritos o también orales (en cuyo caso se precisará de una transcripción o de un analizador del habla), pueden ser textos científicos y técnicos, artísticos, periodísticos, que reproduzcan registros cultos o el lenguaje vulgar, etc., siempre, parece lo más indicado, que reproduzcan el idioma de una época

determinada, y cuya distancia sincrónica o dialectal no sea excesiva.

La obtención de vocablos por muestreo tiene un antecedente en trabajos como el estudio estadístico sobre el Corpus del Español Mejicano Contemporáneo (CEMC) cuyo propósito era la elaboración de diccionarios para la evolución didáctica desde el lenguaje más básico. Ham Chande (1989) define la muestra como "una parte de un todo con afán científico de inferencia válida sobre el todo". La estratificación de la muestra divide la lengua en tres niveles: lengua culta, lengua sub-culta, lengua no-estándar. El tamaño absoluto o relativo de los distintos niveles responde a una ponderación subjetiva inevitablemente: la consideración del valor relativo de cada nivel y género respecto al idioma. Una de estas consideraciones es que algunos géneros son de mayor riqueza léxica que otros. A menor tamaño mayor extensión de vocabulario. Así la literatura es léxicamente más rica que el periodismo, y los textos técnicos que los dialectales, etc.

La investigación en traducción automática ha sostenido siempre en su trasfondo la utopía de traducir textos científicos para hacer más fácil el trasvase de tecnología de los países avanzados a los más necesitados. Parece lógico, pues, que siguiendo con esta tradición, el diccionario de colocaciones partiera de una muestra que contemplara este tipo de textos. Una división por disciplinas en la muestra conllevaría, automáticamente, un conjunto clasificado de diccionarios de

colocaciones.

7. El orden del adjetivo

El orden del adjetivo es importante en el proceso de análisis porque, en el caso de adjetivos con posición fija (información que aparece en la entrada en forma de rasgo), la posición antepuesta o pospuesta al nombre permitirá desambiguar entre entradas homónimas. Pero, sobre todo, en el proceso de generación es importante que el adjetivo se genere en la posición más natural y que más colabore a la recta comprensión e interpretación de la frase.

La solución que aportamos a este problema, vinculada con el estudio anterior, no pretende suplir la extensa investigación que este debatido tema exige. Son varios los aspectos que no vamos a tener en cuenta aquí, como son, factores enfatizantes, acentos contrastivos o intenciones exclamativas que provocan la anteposición del adjetivo independientemente de su función.

Por otro lado habría que considerar, en un estudio detallado de este tema, cómo los complementos preposicionales del nombre o del adjetivo influyen en la posición del adjetivo, así como la existencia de regencias. La presencia de modificadores adverbiales, de gradación o comparativos, parece obligar al adjetivo a ir pospuesto, por ejemplo. Asimismo podría tener cierta influencia la función del sintagma nominal en que va incluido el adjetivo, su topicalización, etc.

Desestimamos también consideraciones sobre el llamado "epíthetum ornans", anteposición del adjetivo usada frecuentemente por los literatos. Nosotros no vamos a ocuparnos de la lengua literaria por parecernos este un tema alejado de la traducción automática de momento. Pero cabe señalar que es a las claras un error identificar *epíteto* con adjetivo antepuesto, dado que, como dice Gili Gaya (1943), el epíteto también puede seguir al sustantivo.

Sería asimismo interesante completar los trabajos existentes con un estudio de las variaciones rítmicas, la entonación, la longitud de los grupos fónicos y la agrupación silábica que se producen respectivamente en los casos de adjetivo pospuesto y antepuesto, pero ello queda muy alejado del enfoque léxico-semántico que nos ocupa.

Algunos autores (Salvador Fernández, 1951 y la R.A.E., 1973) señalan que en muchas ocasiones el sustantivo y el adjetivo forman agrupaciones indisolubles, con un orden determinado e invariable: pena capital, cielo raso, alta mar, fuego fatuo, idea fija... Lo que coincide con el concepto de colocación defendido en este trabajo.

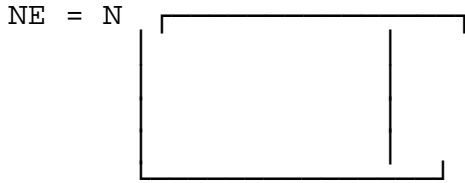
Tradicionalmente, sin embargo, la polémica se ha centrado en dos tipos de fenómenos (Luján, 1980): por un lado la existencia de unos pocos adjetivos (Alcina&Blecua, 1975 proporcionan una lista bastante completa) que cambian de significado según su posición sea antepuesta o pospuesta al nombre, y por otro, los nombres que no alteran su interpretación según su posición pero

a los que, en virtud de aquella, se les reconoce una función distinta: restrictiva o especificativa si es pospuesta, apositiva o explicativa si es antepuesta.

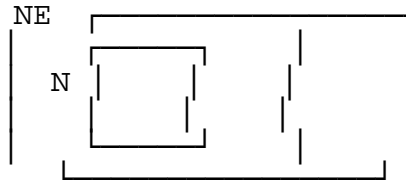
Sin embargo, tanto G.Sobejano (1956) como M.Luján (1980) han observado ya que el adjetivo explicativo puede colocarse delante y detrás del sustantivo, y que, además, también el adjetivo antepuesto puede ser restrictivo. Se afirma, además, que toda estructura que incluya un nombre y un adjetivo es, en su estructura profunda, una oración de relativo restrictiva (Luján, 1980 y Vendler, 1968), por lo que parece que en una representación abstracta el resultado semántico de la restricción y la no restricción deba ser exactamente el mismo. El caso es que cuando se diga 'abrigo gris' se comunique la existencia de un abrigo de ese color, independientemente de que existan otros abrigos de otros colores (especificación) o no haya ningún otro, información extralingüística que no altera el contenido del mensaje.

No pretendo restar relevancia a la existencia de la distinción entre aposición y restricción; en una representación booleana se ve claramente como, dado el conjunto NE de niños enfermos, en el caso de la especificación (*los niños que están enfermos no van al colegio*) el conjunto N (niños) es idéntico al conjunto NE, mientras que en el caso de la explicación (*los niños, que están enfermos, no van al colegio*) el conjunto N es un subconjunto incluido en el conjunto NE:

Especificación:



Explicación:



Pero este es un problema que depende de la extensión del significado que se le dé a la palabra "niños", es decir, al hecho de usar la palabra como atributo (contexto intensional) de Pepe, María y Laura, o si se utiliza como genérico (contexto extensional) de todos los niños. De esta manera podría afirmarse que no estamos tratando dos modos de adjetivar, sino que existe un modo único, la especificación, que puede ser universal o no universal. E incluso, como dice Cruse (1977) la especificidad podría dividirse en 'neutral', de uso más frecuente, y 'remarcada', de uso menos frecuente. Pero estas cuestiones, mientras no se relacionen con aspectos concretos como la definición, la determinación, el plural o el aspecto, atañen mayormente a la filosofía formal, y no parece que la implementación en un sistema empírico pueda beneficiarse de ellas de momento.

Por otro lado, obsérvese que la oración subordinada de relativo explicativa en *los niños, que están enfermos, no van al colegio*, al transformarse en una estructura adjetival resulta *los niños, enfermos, no van al colegio*, y no **los enfermos niños no van al colegio*.

La diferencia de función del adjetivo según su posición es un fenómeno lingüístico del que no ha sido posible señalar "directrices claras" (Lapesa, 1975), en parte porque las causas de la preferencia de posición del adjetivo se aparecen como variadas, y, en parte, porque las explicaciones aducidas son poco o nada sistematizables. Penadés (1988:23) señala que "la carencia de una teoría lingüística explícita y coherente en que fundamentar cualquier explicación de estos hechos [...] lleva a los gramáticos a hablar, por ejemplo, de la significación del sustantivo, de la diferencia estilística entre el adjetivo antepuesto y el adjetivo pospuesto y de los factores lógicos que rigen la colocación ²⁷ de la unidad estudiada, además de asignar un sentido lógico o metafórico a un adjetivo, en virtud de una posición considerada para ellos normal o desviante, pero todo ello se hace sin haber establecido previamente qué se entiende por significación, y por estilística, y sin haber aclarado cuáles son esos factores lógicos que regulan un fenómeno de naturaleza tan alejada de la lógica e inserto, más bien, en el conjunto de los fenómenos correspondientes al ámbito de la cultura. Por otra parte, la propuesta de la gramática tradicional sobre este aspecto del adjetivo constituye más su descripción que su explicación, [...] siendo aquélla en modo alguno satisfactoria".

²⁷ Penadés no utiliza el término 'colocación' para referirse a la solidaridad léxica, sino a la posición del adjetivo con respecto al nombre.

Se dice que el adjetivo antepuesto tiene carácter subjetivo o afectivo frente al pospuesto, que tiene carácter objetivo (Gili Gaya, 1943), pero que la preferencia es meramente una cuestión de estilo: el adjetivo antepuesto da al texto carácter sintético, y el pospuesto analítico. También se afirma que los adjetivos pospuestos indican una voluntad descriptiva o de análisis (Fernández Ramírez, 1951), mientras que los antepuestos se traducen en una voluntad de carácter estimativo. También que la posición postnominal distingue intelectualmente (Lapesa, 1975), mientras que la prenominal indica cualidad subjetivamente valorada. Sobejano (1956) sostiene que la posposición es la norma y la anteposición una contranorma motivada por razones de afectividad, impresionismo, motivos métricos o eufónicos, elegancia, afectación, gusto literario.

Todas estas son argumentaciones que rozan la verdad pero no proporcionan suficientes mecanismos para ejercer un control sobre el fenómeno. Tratan de hacer una descripción general de un problema que tiene una motivación parcial, que depende de una clasificación previa en virtud de la esencia del adjetivo y del nombre implicados. La anteposición o la posposición no es sólo una potencialidad del adjetivo en sí mismo, sino que también depende de su colocación, como se observa en el siguiente ejemplo:

buena chica / chica buena
buena parte / parte buena

donde el comportamiento de *buena* en el primer caso (mismo significado) es diferente al comportamiento en el segundo caso (distinto significado), lo cual se debe a que se trata de dos lecturas de 'parte': una cuantitativa (una buena parte = una gran cantidad) y la otra cualitativa (la parte buena = el lado/trozo bueno).

Señalábamos en el apartado 1 que la ambigüedad que presentan ciertas lenguas en relación con el problema semántico de la intersektividad y la no intersektividad ("a beautiful dancer" puede significar "a beautiful person" o "beautiful as a dancer") el castellano puede resolverlo mediante la posición del adjetivo, de tal manera que en el caso de:

un buen jefe
un jefe bueno

el primero califica sólo a las acciones relativas al nombre funcional, o sea que no es intersectivo, y el segundo sí lo es. (Recuérdese que los casos de ambigüedad sólo se atribuían a nombres funcionales).

Esto hace pensar que existe una implicación más importante de lo que parece entre la posición del adjetivo y las relaciones temáticas que se desprenden de la colocación.

Si contemplamos uno de los casos que se suele citar para mostrar el cambio de significado del adjetivo en posición prenominal, el de

a) una triste mujer (mujer insignificante)

b) una mujer triste (mujer apenada)

observamos que en a) la 'tristeza' no es un atributo inherente de la mujer, mientras que sí lo es en el caso de b). Es decir, en a) la mujer provoca 'tristeza', de tal manera que la condición de 'triste' pertenece al observador, al hablante, jugando la mujer el papel de ente valorable, un papel con matiz de **objeto**. Por el contrario, en el caso b) la mujer posee claramente carácter de **sujeto-experimentante** de la condición expresada en el adjetivo. De esta manera, la posición llamada subjetiva (prenominal) confiere al nombre objetividad, mientras que la posición llamada objetiva (postnominal) concibe al nombre como al sujeto de la relación sintagmática.

Desde nuestro punto de vista los adjetivos que no cambian de significado según su posición con respecto al nombre se regirían por fenómenos similares, de tal manera que cabe concebir el cambio de significado de algunos adjetivos en la anteposición como una fosilización semántica de ciertos casos desprendidos de una tendencia general. Esta confiere al adjetivo una doble dimensión calificativa, de tal modo que la lengua refleja dos vertientes psicológicas que se producen en el ámbito de la comunicación. Así la distinción, que no es de significado, entre *buen jefe* y *jefe bueno*, la cual vimos más arriba. Obsérvese también el caso de *pequeña casa* y *casa pequeña*, donde sólo en *casa pequeña* la casa es el sujeto poseedor del atributo. Y el caso de *débil voz* y *voz débil*, donde *débil voz* indica una

percepción débil ('se oyó una *débil* voz al otro lado del teléfono'), mientras que *voz débil* indica una emisión débil ('el enfermo hablaba con *voz débil*'); nótese que esto justifica el que sea admisible una construcción como 'débil grito', pero que 'grito débil' suponga cierta paradoja.

Ahora bien, dado que en la relación N+A no existe siempre una condición eventual o verbal implícita, resulta confuso utilizar los términos de objeto y sujeto en relación con ciertas colocaciones. Por otro lado lo usual es contraponer objeto a agente y sujeto a predicado, mientras que hemos reconocido que, a pesar de que toda atribución es en realidad una predicación, al menos existe una clase de adjetivos que nunca son predicativos (y tampoco atributivos), como es el caso de los terminológicos o relacionales. No sería, sin embargo, la primera vez que se considerara una relación de transitividad entre la noción sustantiva y la adjetiva. Tamba-Mecz (1980) utiliza esta idea para dividir a los adjetivos en calificativos y relacionales, aduciendo de los primeros que poseen transitividad intrínseca y de los segundos que la poseen extrínseca, lo cual puede explicitarse mediante las perífrasis 'el que tiene' (calificativos) y 'que concierne' (relacionales). La autora de este modo se avanza a Quirk (1985) en la diferenciación entre las relaciones de inherencia y las de no inherencia, aunque este autor no restringirá esta propiedad a ningún tipo de adjetivo. Quirk (1985) tiende además a identificar las relaciones de no

inherencia con las de carácter figurativo.

Con respecto a la posición del adjetivo estas consideraciones nos parecen útiles, porque a la luz de la clasificación colocacional que hemos establecido, observamos que los adjetivos que se corresponden claramente con la perífrasis 'el que tiene' y el 'que concierne', es decir, los ATNP y los ANTT, son de posición pospuesta obligada en un 95%. Esto significaría que la transitividad, ya sea intrínseca o extrínseca, inherente o no inherente, provoca posposición, mientras que la valoración, la cuantificación y calificaciones similares, en las que no se da transitividad (por bien que el nombre sea el objeto valorado, no hay tránsito o flujo del nombre hacia el adjetivo) preferirán la anteposición.

Así, los colores o las nacionalidades (adjetivos atributivos no_polares, ATNP) no admiten la posición prenominal, es decir, la posición de la subjetivización:

un coche rojo
*un rojo coche

un turista alemán
*un alemán turista

Tampoco los adjetivos no_atributivos terminológicos (ANTT) pueden anteponer el adjetivo al sustantivo:

una música instrumental
*una instrumental música

una banda magnética
*una magnética banda

En los casos en que la relación es del tipo ATNP o ANTT

y el orden es prenominal por exigencias de la colocación, deberá constar en la entrada. Es el caso de *alta mar, alto voltaje, alta tensión...*, que son fácilmente reconocibles.

Por el contrario, los adjetivos no_atributivos enfáticos (ANTE) deberán generarse en posición prenominal, porque su carácter es claramente subjetivo o valorativo: *un perfecto fracaso, un completo éxito*, etc. Pero también existen excepciones a esta regla como son *un loco rematado, un triunfo redondo*, las cuáles también deberán llevar la marca de posición.

Los restantes cinco grupos de adjetivos son más problemáticos, porque en ocasiones prefieren la posposición y en ocasiones la anteposición. Una solución provisional podría ser optar por la generación del adjetivo pospuesto al nombre salvo indicación expresa. Está por ver si la correcta generación de la posición adjetival en el caso de los grupos restantes puede ir regida por reglas generales; naturalmente que esto supone un estudio exhaustivo de colocaciones generadas correcta o incorrectamente que derivaría en una subclasificación de las clases definidas como se verá.

En las líneas siguientes aventuramos ciertos criterios que controlarían la generación del adjetivo antepuesto.

Los adjetivos atributivos polares existenciales (ATPE), los cuales colocan usualmente con nombres simples existenciales,

nombres deverbales completivos resultativos y nombres funcionales en su vertiente objetual (que no agentiva), serán generados después del nombre, porque concebimos que existe transitividad en la relación, es decir, que el nombre es el sujeto de la atribución.

Los adjetivos atributivos polares perfectivos (ATPP), cuando no van en la posición predicativa natural, pueden anteponerse al nombre si son participiales. En este caso, sin embargo, no se trata de una ausencia de transitividad, sino que parece que, dado que el nombre es sujeto paciente en la estructura verbal equivalente (el lugar está acordonado < el lugar ha sido acordonado), la transitividad se pueda expresar

pasiva o activamente. Así:

atravesó a codazos el *local abarrotado/abarrotado local*

el *abarrotado local/local abarrotado* olía a humanidad

Sin embargo, no pueden anteponerse los adjetivos de colocaciones ATPP que comparten coocurrencia con estructuras verbales intransitivas inergativas (afiliarse, suscribirse)

o inacusativas (morirse) como era, por ejemplo:

el trabajador está afiliado

> *trabajador afiliado*

*afiliado trabajador

el lector está suscrito

> *lector suscrito*

*suscrito lector

el perro está muerto

> *perro muerto*

*muerto perro

Pueden anteponerse cuando son sujetos experimentales (el *hambriento canario* devoró el alpiste), pero suelen ir pospuestos en el caso de ser ATPP's denominales cuando el nombre no es agentivo:

la sopa caliente le dió fuerzas
*la caliente sopa le dió fuerzas

Cuando las colocaciones son del tipo atributivo polar adverbial de acción (ATPVA), el adjetivo muy a menudo califica a los efectos de esa acción. Sin embargo, así como los nombres funcionales poseían cierta ambigüedad porque podían ser calificados por su función al mismo tiempo que como entes, los nombres de acciones carecen de esa ambigüedad. Es por ese motivo que, aun cuando pueden llevar el adjetivo pospuesto, tenderemos, a falta de indicaciones expresas, a generar la posición prenominal. Así: *aparatosa caída, vigoroso salto, violento golpe, brusco viraje, brutal procedimiento, triste actuación, etc.*

En el caso de las colocaciones atributivas polares adverbiales de agente (ATPVG), cuando el verbo está implícito en el nombre (nombre funcional) o es un nombre deverbal, puede optarse por anteponer el adjetivo: *fiel esposa, rápido nadador, cauto cazador...*, mientras que se pospondrá siempre que se trate de colocaciones del segundo grupo en las que el verbo está relacionado con el adjetivo: *circunstancia atenuante, producto alimenticio, dolor enloquecedor, público asistente, etc.*

Por último, las colocaciones atributivas polares adverbiales de objeto (ATPVO) llevarán siempre el adjetivo pospuesto. Se produce una transitividad clara entre los miembros de la colocación, y, aunque el nombre tenga papel objetual, la construcción lo eleva a la categoría de sujeto paciente de la atribución. Recuérdese que estas colocaciones, como las ATPP pero con aspecto ingresivo, pueden identificarse con estructuras en voz pasiva. Algunos ejemplos son: *teoría comprobable*, *material combustible*, *abrigo reversible*, etc.

8. Conclusiones

En un esfuerzo de abstracción se establecen las siguientes conclusiones:

1. Las colocaciones de una lengua particular determinan el significado en el nivel léxico, constituyendo el complemento y la realización del significado conceptual, el cual es imprescindible fijar para lograr una correcta comprensión, emisión y/o traducción (apartado 1 de la primera parte).

2. Las colocaciones constituyen un aspecto básico de las mediatizaciones lingüísticas que una lengua particular impone a su propio sistema de reglas (apartados 2 y 3 de la primera parte).

3. Un diccionario de colocaciones puede estar concebido de dos maneras (apartados 4 y 5 de la primera parte):

- a) relacionado con un diccionario conceptual
- b) no relacionado con un diccionario conceptual

En el segundo caso el significado conceptual se especificará a partir de relaciones hiperonímicas y del eje antónimo, el cual determinará un conjunto de palabras relacionadas o sinónimos (apartado 1.3.2. de la primera parte).

4. En el campo de la traducción automática, tanto el análisis (desambiguación) como la generación (selección léxica) se benefician del control impuesto por restricciones colocacionales, ya sean

directas o indirectas, a través de las relaciones hiperónimas, sinónimas, de área de aplicación, etc. (apartados 1.4. y 2.1. de la primera parte y apartado 5. de la segunda parte).

Las relaciones derivativas contempladas con objeto de establecer restricciones colocacionales deben ser especificadas individualmente (apartado 4.3.1. de la primera parte).

5. En las colocaciones de nombre y adjetivo se observa cierta adaptación conceptual del adjetivo. Esta naturaleza camaleónica del adjetivo es controlada a través de una clasificación que tiene en cuenta las distintas relaciones temáticas que se desprenden de su variado comportamiento colocacional (apartados 1, 2 y 4 de la segunda parte). La clasificación también ayuda en la generación del orden adecuado de la colocación (apartado 7 de la segunda parte).

Las repercusiones del trabajo en el campo de la lingüística son de diversa índole. En el terreno de la lingüística computacional facilita la reflexión necesaria para la elaboración de un diccionario que colabora en la resolución de problemas hasta ahora irresolubles como la desambiguación y la selección léxica, con claras ventajas en el campo de la traducción automática y en la inteligencia artificial en general.

Con respecto a la lexicografía supone un avance en la concreción de temas largamente intuitivos pero poco meditados, y una propuesta en aras de la construcción de un diccionario electrónico sistemático que tenga en cuenta la diversidad de las aplicaciones a que pueda

destinarse, y, al mismo tiempo, no deje de ser un producto en sí mismo.

En cuanto a la lingüística teórica se sugiere que se ponga al servicio del diccionario así como el diccionario ha estado a su servicio, y que se aúne el esfuerzo para lograr una aproximación de los universales lingüísticos a la realidad particularizada de las lenguas. Esta aproximación daría lugar a avances en la aplicación de técnicas concretas.

De esta forma, de la investigación se deduce una propuesta de trabajo obvia: la elaboración de varios diccionarios de colocaciones que contemplen las combinaciones Nombre + Adjetivo, Verbo + Nombre y, si cabe, Verbo + Adverbio y Verbo + Adjetivo. Estos diccionarios deberán estar interrelacionados de manera que constituyan uno sólo y, además, poseerán una arquitectura que permita su inmediato acoplamiento a un diccionario general de la lengua.

BIBLIOGRAFÍA

ABAITUA, Joseba, 1990, *Cuestiones de léxico y aspecto. Un análisis formal del participio* en Boletín de la SEPLN (Sociedad Española para el Procesamiento del Lenguaje Natural), nº 9, 1991

ABAITUA, J., SOLER, J. & VIVALDI, J., 1989, *Problemática de la interlingua como estrategia de traducción automática*, Boletín de la SEPLN (Sociedad Española para el Procesamiento del Lenguaje Natural), nº 7, pp. 157-166, UPC, Barcelona, 1989

ABEILLÉ, Anne & SCHABES, Yves, *Parsing Idioms in Lexicalized TAGs*, Fourth Conference of the European Chapter of the Association for Computational Linguistics, 10-12 Abril, Univ. de Manchester 1989, Institute of Science and Technology, pp. 1-9

AGUILAR-AMAT CASTILLO, A., 1990, Tratamiento computacional de las unidades léxicas múltiples en el sistema de traducción automática EUROTRA, Proyecto de Investigación para el Programa de Doctorado de Informática y Lingüística de la Universidad Autónoma de Barcelona (Biblioteca de Letras), B., 1990

AISO, H., ISODA, M., KAMIBAYASHI, N., MATSUNAGA, Y., 1986, *Model for Lexical Knowledge Base* en Proc. of COLING-86, Bonn, 1986, pp. 451-453

AKKERMAN, E., MASEREEUW, P., MEIJS, W., 1985, Designing a computerized lexicon for linguistic purposes, ed. Rodopi (Ascot Report No 1), Amsterdam, 1985

ALBANO, R., CUMMING, S., SONDEHEIMER, N., 1990, *How to Realize a Concept: Lexical Selection and the Conceptual Network in Text Generation* en Machine Translation n. 5, Kluwer Academic Publishers, 1990, pp. 57-78

ALCINA FRANCH, J. & BLECUA P., J.M., 1975, Gramática Española, ed. Ariel, B., 1989

ALCOBA, S., 1988, *Categoría léxica de las palabras compuestas en Verba*, nº 15, pp.1-30, 1988

ALEMANY BOLUFER, J., 1920, Tratado de la formación de palabras en la lengua castellana. La derivación y la composición: Estudio de los sufijos y prefijos empleados en una y otra, Librería General de Victoriano Suárez, Madrid, 1920

ALLERTON, D.J., 1983, *Three (or four) levels of word cooccurrence restriction* en LINGUA, nº 63, pp. 17-40, Elsevier Science Publishers B.V. (North-Holland), 1984

ALPAC, 1966, Language and Machines; Computers in Translation and

Linguistics, Washington, DC: Publication 1416, National Academy of Sciences.

BÁEZ, V. & PENADÉS, I., *Diccionario informatizado de construcciones oracionales y el proyecto Esquemas Sintáctico-semánticos del Español* en LEA (Lingüística Española Actual), XII, 1990, p. 108

BAHNS, J., 1993, *Lexical collocations: a contrastive view*, en ELT Journal, vol. 47/1, January 1993, Oxford University Press, 1993

BALDINGER, 1966, *Sémantique et structure conceptuelle (Le concept 'se souvenir')* en Cahiers de lexicologie, 8.

BAR-HILLEL, Jehoshua, 1955, *Language and information* en Machine Translation of languages, editado por W.N.Locke & A.D.Booth, Technology Press of the Mass. Inst. of Tech., N.Y., London 1955, pp.183-193

BAR-HILLEL, Y., 1967, *Dictionaries and meaning rules* en Foundations of Language, n° 3, 1967, pp. 409-414

BARTON, G.E., BERWICK, R.C., RISTAD, 1987, E.S., Computational complexity and Natural Language, MIT Press, 1987

BENSON, E., BENSON, M., ILSON, R., 1986, The BBI Combinatory Dictionary of English, John Benjamins Publishing Company, 1986

BIERWISCH, M., 1967, *Some semantics universals of german adjectivals* en Foundations of Language, n° 3, pp. 1-36

BLOOMFIELD, L., 1933, Language, ed. Holt, Rinehart & Winston, New York, 1933

BRESNAN, Joan & KAPLAN, Ronald y otros, 1982, *The mental representation of grammatical relations*, MIT Press series on Cognitive Theory and Mental Representation, J. Bresnan, L. Gleitman, Samuel Jay Keyser, editores. The Massachusetts Institute of Technology, 1982

BUCHMANN, B., 1984, *Early history of machine translation* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

CACCIARI, C. & TABOSI, P., 1988, *The Comprehension of Idioms* en Journal of Memory and Language n. 27, Academic Press, 1988, pp. 668-683

CALZOLARI, N., 1988, *The dictionary and the thesaurus can be combined* en Relational models of the lexicon, ed. M.W. Evens, Cambridge University Press, N.Y., 1988

CAMPPIA, P., CIRAVEGNA, F., COLOGNESE, A., 1992, *Knowledge extraction from texts by sintesi* en Proc. of COLING-92, Nantes, 1992, pp. 1244-1247

CARBONELL, J.G. & TOMITA, M., 1986, *Another stride towards knowledge-based machine translation* en Proc. of COLING'86, University of Bonn, 1986

CARBONELL, J.G. & TOMITA, M., 1987, *Knowledge-based machine translation, the CMU approach* en Machine Translation (theoretical and methodological issues), ed. S. Nirenburg, Cambridge University Press, 1987, pp. 68-89

CARNAP, R., 1958, Introduction to semantics and formalization of logic, Harvard University Press, USA, 1975.

CASARES, Julio, 1969, Introducción a la lexicografía moderna RFE-Anejo LII, Madrid, CS de IC, 1969

CHAFFIN, R., 1988, *The nature of semantic relations: a comparison of two approaches* en Relational models of the lexicon, ed. M. W. Evens, Cambridge University Press, N.Y., 1988, pp. 289-334

CHOMSKY, N., 1957, Las estructuras sintácticas, trad. cast. de C.P. Otero, Siglo XXI, México, 1974

CHOMSKY, N., 1964, Aspectos de la teoría de la sintaxis, versión esp. de C.P. Otero, ed. Aguilar, 1976

CHOMSKY, N., 1967, *Remarks on Nominalization*, en Readings in English Transformational Grammar, R.A. Jacobs & P.S. Rosenbaum, Ginn & Company, Waltham, Mass., 1970

CHOMSKY, Noam, 1981, Lectures on Government and Binding, Foris Publications Holland, Dordrecht, 1988

CHOMSKY, Noam, 1988, Language and problems of knowledge, The MIT Press, 1988 Massachusetts.

CHURCH, K.W. & HANKS, P., 1990, *Word Association Norms, Mutual Information, and Lexicography* en Computational Linguistics, vol. 16, n. 1, March 1990, pp. 22-29

CHURCHLAND, P.M. & SMITH CHURCHLAND, P., *¿Podría pensar una máquina?* en Investigación y Ciencia, ed. española de Scientific American, Marzo, 1990.

COLLINS, V.H., 1969, A book of english proverbs, Greenwood Press, Connecticut (USA), 1974

COSERIU, E., 1964, *Structure lexical et enseignement du vocabulaire* en Actes du premier colloque international de linguistique appliquée, Nancy, 1966, pp. 175-217

COSERIU, E., 1977, Principios de semántica estructural, ed. Gredos, M. 1991, trad. cast. de Marcos Martínez

CRUSE, D.A., 1975, *The pragmatics of lexical specificity* en Journal

- of Linguistics, vol. 13, Great Britain, 1977, pp. 153-165
- CRUSE, D.A., 1986, Lexical Semantics, Cambridge Textbooks in Linguistics, Cambridge University Press 1986
- CULLINGFORD, R.E., ONYSHKEVYCH, B.A., 1987, *An experiment in lexicon-driven machine translation* en Machine Translation (theoretical and methodological issues), ed. S. Nirenburg, Cambridge University Press, 1987, pp. 278-301
- DAGAN, I., RACKOW, U., SCHAWALL, U., *Automatic Translation of Noun Compounds* en Proc. of COLING-92, Nantes, Aug. 23-28, 1992, pp. 1249-1253
- DAHL, V., 1981, *Translating Spanish into Logic through Logic*, en American Journal of Computational Linguistics, vol. 7, n. 3, July-September 1981
- DAHLGREN, K., McDOWELL, J., STABLER, E.P., 1989, *Knowledge Representation for Commonsense Reasoning* en Computational Linguistics, vol. 15, n. 3, September 1989
- DANLOS, L., 1985, The linguistic basis of text generation, Cambridge University Press, G.B., 1987
- DE ROECK, A., 1984, *Linguistic theory and early machine translation* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987
- DIK, S.C., The theory of Functional Grammar, 1989, Foris Publications, Holland, 1989
- DIXON, R.M.W., 1982, Where have all the adjectives gone?, *Janua Linguarum, Series Maior 107*, Indiana University, Mouton Publishers, Berlin, 1982
- DOWTY, D.R., WALL, R.E., TETERS, S., Introduction to Montague Semantics, Reidel Publishing Company, Holland 1981
- DUTOIT, D., 1992, *A set-Theoretic Approach to Lexical Semantics* en Proc. of COLING-92, Nantes, 1992, pp. 983-987
- DYER, M.G. & ZERNICK, U., 1986, *Disambiguation and Language Acquisition through the phrasal lexicon* en Proc. of COLING'86, University of Bonn, 1986
- ECO, U., 1984, Semiótica y filosofía del lenguaje, ed. Lumen, Barcelona 1990
- EVENS, M.W., 1988, Relational models of the lexicon, Cambridge University Press, G.B., 1988
- FASS, D., 1991, *met*: A method for Discriminating Metonymy and Metaphor by Computer* en Computational Linguistics, vol. 17, n. 1, March 1991,

pp.49-90

FASS, D. & WILKS, Y., 1983, *Preference Semantics, Ill-Formedness, and Metaphor* en American Journal of Computational Linguistics, vol. 9, n. 3-4, July-December, 1983, pp. 178-187

FERNÁNDEZ ALONSO, M.R., 1982, Contribución al estudio del adjetivo calificativo atributivo en español, Editorial de la Universidad Complutense de Madrid, Madrid, 1983

FERNÁNDEZ LEBORANS, M.J., 1977, Campo semántico y connotación, ed. Planeta (Cupsa Editorial), Madrid, 1977

FERNÁNDEZ RAMÍREZ, S., 1951, Gramática española. Los sonidos, el nombre y el pronombre, Revista de Occidente, Madrid, 1951

FILIPENKO, M.V., RAKHILINA, E.V., PADUCHEVA, E.V., 1992, *Semantic dictionary viewed as a lexical database* en Proc. of COLING-92, Nantes, Aug. 23-28, 1992, pp. 1295-1299

FILLMORE, Ch., 1968, 'The case for case', en Universals in Linguistic Theory, Bach, E. & Harms, R., Holt, Rinehart and Winston, New York, 1968, pp. 1-90

FIRTH, J.R., 1957, Papers in linguistics 1934-1957, Oxford University Press, N.Y., 1969

FODOR, J.D., 1977, Semántica, ed. Cátedra, M., 1985 (traducción al castellano de F. Aliaga García)

FRASER, Bruce, 1970, Idioms within a transformational Grammar en Foundations of Language, n. 6, 1970, pp. 22-42

FRAWLEY, W., 1988 Relational models and metascience en Relational models of the lexicon, ed. M. W. Evens, Cambridge University Press, N.Y., 1988, pp. 335-374

FUENTES RODRIGUEZ, Catalina, 1987, Enlaces Oracionales, ed. Alfar, Sevilla 1987

GATES, D., MCCARDELL, R., NIRENBURG, S., NYBERG, E., 1989, Generation, en Machine Translation, n.4, Kluwer Academic Publishers, 1989, pp.149-168

GAZDAR, G., EWAN, K., PULLUM, G., & SAG, I., 1985, Generalized Phrase Structure Grammar, Cambridge. Mass., Harvard University Press, 1985.

GOLAN, I., LAPPIN, S., & RIMON, M., An active bilingual lexicon for machine translation en Proc. of COLING-88, vol. I, Budapest, 1988, p. 205

GRACIA, Olga & MELERO, Maite, 1990, Construcciones de Verbo Soporte, Actas del XX Congreso de la Sociedad Española de Lingüística, Tomo II, pp. 653-667, Gredos, M. 1990

GREGORY, M., 1980, *The recent work of M.A.K. Halliday: 'Language as Social Semiotic'*, en Applied Linguistics, vol.1, pp. 74-81, Oxford University Press, 1980

GRIMES, J.E., 1988, *Information dependencies in lexical subentries* en Relational models of the lexicon, ed. M. W. Evens, Cambridge University Press, N.Y., 1988, pp. 167-182

GRISHMAN, Ralph, 1986, Computational linguistics, Cambridge University Press, 1986

GROSS, Maurice, 1982, *Une classification des phrases 'figées' du français* en Revue Québécoise de linguistique, 11.2, 1982

GROSS, Maurice, 1984, *Les noms traceurs* en Cahiers de Lexicologie, 44, 1984

GROSS, Gaston & CHAURAND, Jacques, Typologie des noms composés, Section du comité national du C.N.R.S. 42, Universidad Paris XIII

GROSS, D. & MILLER, K., 1990, *Adjectives in WordNet* en International Journal of Lexicography, vol.3, n.4, Oxford University Press, 1990

GROSS, D., FISHER, U. & Miller, G.A. (1989), *The Organization of Adjectival Meanings*, en Journal of Memory and Language, 28, 1989, pp. 92-106

GRUBER, J.S., 1967, *Topicalization in child language* en Foundations of Language, n° 3, pp. 37-65

GUTIERREZ ORDOÑEZ, Salvador, Variaciones sobre la atribución, Univ. de León, 1986

HAHN, U. & REIMER, U., *TOPIC Essentials* en Proc. of COLING-86, Bonn, 1986, pp. 497-503

HAM CHANDE, Roberto, 1989, *Del 1 al 100 en lexicografía* en Jornadas del Colegio de Méjico, Invest. Ling. Lexicográfica, 1989

HAMADA, T. & SHIRAI, K., 1986, *Linguistic knowledge extraction from real language behaviour* en Proc. of COLING'86, University of Bonn, 1986

HARRIS, Z.S., 1957, *Co-occurrence and transformation in linguistic structure*, en Language, vol.33, n. 2, pp. 283-340

HERBST, T., *Adjective Complementation: A valency Approach to Making EFL Dictionaries* en Applied Linguistics, vol. 5, n. 1, Oxford University Press, 1984, pp. 1-11

IDE, N. & VÉRONIS, J., 1992, *A feature-based model for lexical databases* en Proc. of COLING-92, Nantes, 1992, pp. 588-595

ISABELLE, P., 1984, *Machine Translation at the Taum Group* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

JOHNSON, R. & ROSNER, M., 1984, *Machine Translation and software tools* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

KAMP, J.A.W., 1975, *Two theories about adjectives* en Formal Semantics of Natural Language, ed. by E.L. Keenan, Cambridge University Press, 1975

KATZ, B. & LEVIN, B., 1988, *Exploiting lexical regularities in designing natural language systems*, en Proc. of COLING-88, vol. I, Budapest, 1988, p. 316

KATZ, J.J., 1967, *Recent issues in semantic theory* en Foundations of Language, n° 3, pp. 124-194

KATZ, J.J. & FODOR, J.D., 1963, *The structure of a semantic theory* Language, 39, pp.17-210

KATZ, Jerrold J., 1966, La realidad subyacente del lenguaje y su valor filosófico, ed. Alianza Universidad, M., 1975

KEENAN, E.L., 1973, Formal Semantics of Natural Language, Cambridge University Press, 1975

KEMPSON, Ruth, 1982, Teoría semántica, ed. Teide, B., 1982

KIRSCHNER, Zdeněk, 1988, *Traditional means in machine translation* en Proc. of COLING-88, vol. I, Budapest, 1988, p. 328

KITTREDGE, R. & TUTIN, A., 1992, *Lexical choice in context: generating procedural texts* en Proc. of COLING-92, Nantes, Aug. 23-28, 1992

LAKOFF, G. & JOHNSON, M., 1980, Metáforas de la vida cotidiana, ed. Cátedra, M., 1986, (traducción al castellano de C. González Marín)

LAPESA, R., 1975, *La colocación del adjetivo atributivo en español*, en Homenaje a la memoria de Don Antonio Rodríguez-Moñino, 1910-1970, Castalia, Madrid 1975

LARSON, M.L., 1989, La traducción basada en el significado, ed Eudeba, Buenos Aires, 1989

LAUSBERG, Heinrich, 1966, Manual de retórica literaria, ed. Gredos, M., 1966

LEECH, G., 1974, Semántica, ed. Alianza Universidad, M. 1985, (trad. cast. de J. L. Tato)

LEVI, J.N., 1978, The syntax and semantics of complex nominals, New York: Academic Press, 1978.

LEWIS, D., 1976, *General Semantics*, en Montague Grammar, ed. by Barbara H. Partee, Academic Press, N.Y., 1976

LYONS, J., 1968, Introducción en la lingüística teórica, ed. Teide, B., 1979 (traducción castellana de R. Cerdà Massó)

LYONS, J., 1977, Semántica, ed. Teide, B., 1989 (traducción castellana de R. Cerdà Massó)

LUJÁN, M., Sintaxis y semántica del adjetivo, ed. Cátedra, Madrid, 1980

MAAS, Dieter, 1989, *Implementation of phraseological units*, en The dictionary coding manual, editado por Pius ten Hacken, Eurotra Reference Manual 6.1., diciembre 1989

MAKKAI, A., 1972, Idiom Structure in English, ed. MOUTON, Paris, 1972

MANN, W.C., PARIS, C.L. & SWARTOUT, W.R., 1991, Natural Language Generation in Artificial Intelligence and Computational Linguistics, Kluwer Academic Publishers, Massachusetts, 1991

MARTÍN VIDE, C., 1985, *De semiótica matemática: sobre el modo de significar del discurso matemático* en Actas del Congreso de Lenguajes Naturales y Lenguajes Formales, Barcelona, Octubre de 1985

MARTINET, André, 1969, La lingüística, ed. Anagrama, Barcelona, 1975

MATSUMOTO, Y., NAGAO, M., UTSURO, T., *Lexical Knowledge Acquisition from Bilingual Corpora* en Proc. of COLING-92, Nantes, 1992

McARTHUR, T., 1981, Longman Lexicon of Contemporary English, ed. Longman Group, 1989

McCLELLAND, James & Rumelhart, David E., 1986, Parallel Distributed Processing. Explorations in the Microstructure of Cognition, Institute of Cognitive Science, Univ. of California, MIT Press, Massachusetts, 1986

McDONALD, D.D., 1987, *Natural language generation: complexities and techniques* en Machine Translation (theoretical and methodological issues), ed. S. Nirenburg, Cambridge University Press, 1987, pp. 192-223

McDONALD, D.D., 1991, *On the Place of Words in the Generation Process* en Natural Language Generation in Artificial Intelligence and Computational Linguistics, ed. C.L. Paris, W.R. Swartout & W.C. Mann, Kluwer Academic Publishers, 1991

McKEOWN, K.R., 1985, Text Generation, Cambridge University Press, 1985

McKEOWN, K.R. & SMADJA, F., 1991, *Using collocations for language generation* en *Comput. Intell.*, 7, pp. 229-239, Canada, 1991

MEYA, M., 1989, *Lexicalizaciones y expresiones idiomáticas. Su procesamiento*, en *SPLEN*, UPC, nº 7, enero de 1989, pp. 106-113

MELBY, Alan K., 1986, *An early approach to lexical transfer* en *Proc. of COLING-86*, Univ. of Bonn, 1986, pp. 104-106

MEL'CUK, I. & ZHOLKOVSKY, A., 1988, *The explanatory combinatorial dictionary* en *Relational models of the lexicon*, ed. M. W. Evens, Cambridge University Press, N.Y., 1988

MILLER, G.A., 1990, *Nouns in WordNet: A lexical Inheritance system* en *International Journal of Lexicography*, vol.3, n.4, pp. 245-263, Oxford University Press, 1990

MONTAGUE, R., 1974, *Ensayos de filosofía formal*, ed. Alianza Universidad, MADRID 1977

MONTREYNAUD, Florence, PIERRON, Agnès et SUZZONI, François, 1984, *Dictionnaire de proverbes et dictons*, Les usuels du Robert, Paris 1984

MORENO, J.C., 1989, *Semi-compositionality and idioms in machine translation: the functional word strategy*, EUROTRA ESPAÑA, Madrid. Documento interno.

NEWMeyer, F.J., 1973, *The regularity of idiom behaviour* en *LINGUA*, vol. 34, n. 4, November 1974, pp. 327-342

NEWMeyer, F.J., 1980, *Linguistic Theory in America (The first Quarter-Century of transformational generative grammar)*, Academic Press, N.Y., 1980

NIRENBURG, S., 1987, *Knowledge and choices in machine translation* en *Machine Translation (theoretical and methodological issues)*, ed. S. Nirenburg, Cambridge University Press, 1987, pp. 1-21

NIRENBURG, S., 1989, *Knowledge-based machine translation*, en *Machine Translation*, n.4, Kluwer Academic Publishers, 1989, pp. 5-24

NIRENBURG, S. & LEVIN, L., 1989, *Knowledge Representation Support* en *Machine Translation*, n.4, Kluwer Academic Publishers, pp. 25-52

NIRENBURG, S. & NIRENBURG, I., 1990, *A framework for Lexical Selection in Natural Language Generation*, en *Proc. of COLING-90*, Budapest, 1990, pp. 471-475

NIRENBURG, S., RASKIN, V., TUCKER, A.B., 1986, *On Knowledge-based Machine Translation* en *Proc. of COLING-86*, University of Bonn, 1986

NIRENBURG, S., RASKIN, V., TUCKER, A.B., 1987, *The structure of*

interlingua in TRANSLATOR en Machine Translation (theoretical and methodological issues), ed. S. Nirenburg, Cambridge University Press, 1987, pp. 90-110

NYBERG 3rd, E., McCARDELL, R., GATES, D., & NIRENBURG, S., 1989, *Generation en Machine Translation n.4*, Kluwer Academic Publishers, pp. 149-168

NYBERG, E.H. III & MITAMURA, T., *Hierarchical lexical structure and interpretive mapping in machine translation*, en Proc. of COLING-92, Nantes, Aug. 23-28, 1992, pp. 1254-1259

PADUCHEVA, E.V. & RAKHILINA, E.V., *Predicting co-occurrence restrictions by using semantic classifications in the lexicon* en Proc. of COLING-90

PALMER, F.R., 1976, Semantics, a new outline, Cambridge University Press, G.B., 1976

PELLAT-MASSO, Luisa, 1989, *Une description formelle des expressions figées de l'espagnol* en Mémoire de DEA de linguistique informatique, Université Paris VIII, 1989

PENADÉS MARTÍNEZ, I., 1988, Perspectivas de análisis para el estudio del adjetivo calificativo en español, Servicio de Publicaciones de la Universidad de Cádiz, 1988

PÉREZ CALVO, J., 1986, Adjetivos puros: estructura léxica y topología, Anejo nº 2 de la revista CUADERNOS DE FILOLOGIA UNIVERSIDAD DE VALENCIA

POLLARD, 1991, Topics in Constraint-Based Syntactic Theory, The third European Summer School in Language, Logic and Information, Saarbrücken, 1991

POTTIER, B., 1964, *Vers une sémantique moderne* en Travaux de linguistique et de littérature, 2, 1

QUIRK, R., GREENBAUM, S., LEECH, G., SVARTVIK, J., 1985, A Comprehensive Grammar of the English Language, Longman Group Limited, N.Y., 1985

R.A.E., 1973, Esbozo de una nueva gramática de la lengua española, Espasa Calpe, M. 1989

RASKIN, V., 1987, *Linguistics and natural language processing en Machine Translation (theoretical and methodological issues)*, ed. S. Nirenburg, Cambridge University Press, 1987, pp. 42-58

REY, Alain et CHANTREAU, Sophie, 1987, Dictionnaire des expressions et locutions, Les Usuels du Robert, Paris, 1987

RIEGEL, K., 1970, *The language acquisition process: a reinterpretation of related research findings*, en Theory and research in Life-Span Development Psychology, Goulet & Baltes eds., New York

Academic Press, 1970, pp. 357-399.

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

SAMPSON, G., 1984, *Machine Translation: A non conformist view of the state of the art* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

SCHANK, R., 1975, Conceptual Information Processing, North Holland, Amsterdam, 1975

SEARLE, J. R., 1990, *¿Es la mente un programa informático?* en Investigación y Ciencia, Ed. española de Scientific American, Marzo de 1990

SEIDL, Jennifer, McMORDIE, W., 1978, English Idioms, Oxford University Press, Oxford, 1978

SELLS, P., 1985, Teorías sintácticas actuales, ed. Teide, Barcelona, 1989, trad. cast. de R.Cerdà Massó

SERRANO, S., 1975, Elementos de lingüística matemática, ed. Anagrama, Barcelona, 1988

SERRANO, S., 1980, Signos, lengua y cultura, ed Anagrama, Barcelona, 1981

SHANN, Patrick, *Machine Translation: A problem of linguistic engineering or of cognitive modelling* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

SIEGEL, M.E.A., 1976, Capturing the adjective, ed. by J.Hankamer, Harvard University, Garland Publishing, N.Y., 1980

SINCLAIR, J., 1991, Corpus, Concordance, Collocation, Oxford University Press, 1991

SMADJA, F.A., 1989, *Lexical Co-ocurrence: The Missing Link* en Literary and Linguistic Computing, vol. 4, n. 3, Oxford University Press, 1989

SMADJA, F.A., 1992, *How to compile a bilingual collocational lexicon automatically*, en Proceedings of the AAI workshop on statistically based NLP techniques, Columbia University, 1992

SMADJA, F.A., 1993, *Retrieving Collocations from Text: Xtract*, Computational Linguistics, vol. 19. 143-178

SOBEJANO, G., 1956, El epíteto en la lírica española, ed. Gredos, Madrid, 1970 (2ª ed. revisada).

SONDHEIMER, N., CUMMING, S. & ALBANO, R., 1990, *How to realize a concept: lexical selection and the conceptual network in text*

generation en Machine translation, Kluwer Academic Publishers, 1990, pp. 57-78

SOWA, J., 1988, *Using a lexicon of canonical graphs in a semantic interpreter* en Relational models of the lexicon, ed. M. W. Evens, Cambridge University Press, N.Y., 1988

STEFFENSEN, M., 1986, *Register, Cohesion, and Cross-cultural Reading Comprehension* en Applied Linguistics, vol. 7, n. 1, Oxford University Press, 1986, pp. 71-85

STOCK, O., *Parsing with Flexibility, Dynamic Strategies, and Idioms in Mind* en Computational Linguistics, vol. 15, n. 1, March 1989, pp. 1-18

TAMBA-MECZ, I., 1980, *Sur quelques propriétés de l'adjectif de relation* en Travaux de Linguistique et de Littérature, Strasbourg, 1980, pp 119-132

TARTEE, B.H., 1976, Montague Grammar, Academic Press, N.Y. 1976

TUCKER, A.B., 1987, *Current strategies in machine translation research and development* en Machine Translation (theoretical and methodological issues), ed. S. Nirenburg, Cambridge University Press, 1987, pp. 22-41

UCHIDA, H., HAYASHI, T. & KUSHIMA, H., 1985, ATLAS: Automatic Translation System, Fujitsu Scientific and Technical Journal 21, 1985, pp. 317-329

ULLMANN, S., 1975, Natural and Conventional Signs, PdR Press Publications in SEMANTICS 1, The Peter De Ridder Press, Lisse/Netherlands, 1975

VALLÉS, Montserrat & LAG, Jesús, Función distribucional y función contrastiva de ser y estar, tesis doctoral dirigida por R. Cerdà Massó, Univ. de Barcelona, 1974

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

WARWICK, S., 1984, *An overview of post-ALPAC developments* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

WEINER, E.J., 1984, *A knowledge Representation Approach to Understanding Metaphors* en Computational Linguistics, vol. 10, n. 1, January-March, 1984

WEINREICH, Uriel, 1969, *Problems in the Analysis of idioms* en Substance and Structure of Language, ed. Jaan Puhvel, Univ. of California Press, 1969

WERNER, O., 1988, *How to teach a network: minimal design features for a cultural acquisition device or C-KAD* en Relational models

of the lexicon, ed. M. W. Evens, Cambridge University Press, N.Y., 1988, pp. 139-141

WHERLI, E., 1984, *Recent developments in theoretical linguistics and implications for machine translation* en Machine Translation Today, ed. by M. King, Edinburgh University Press, 1987

WHITE, J.S., 1988, *Determination of lexical-semantic relations for multi-lingual terminology structures* en Relational models of the lexicon, ed. M. W. Evens, Cambridge University Press, N.Y., 1988, pp. 183-198

WINOGRAD, T., 1983, Language as a cognitive Process, Volume 1: Syntax, Stanford University, USA, 1983

YAGUELLO, M., 1981, Alicia en el país del lenguaje, ed. Mascarón, Madrid, 1983

ZERNIK, U., 1990, *Lexical Acquisition: Where is the Semantics* en Machine Translation, vol. 5, n. 2, June 1990

ZULOAGA, Alberto, 1975, *La fijación fraseológica* en Thesaurus, BICC, XXX, 2, 1975, pp. 225-248

ZULOAGA, Alberto, 1975, *Estudios generativo-transformativistas de las expresiones idiomáticas*, en Thesaurus, BICC, XXX, 1, 1975, pp. 1-48

Ampliación bibliográfica

AGUILAR-AMAT CASTILLO, A., 1993, *La estructura sintáctica del nombre*, boletín nº 14 de la SPLN, IX Congreso, Barcelona, 1994 (en prensa)

KATZ, J. & POSTAL, P., 1964, An Integrated Theory of Linguistic Description, MIT Press, Cambridge, MA

JACKENDOFF, R., 1993, *X-Bar Semantics*, Semantics and the Lexicon, ed. J. Pustejovsky, Kluwer Academic Publishers, Dordrecht, 1993

KATZ, J. & FODOR, J., 1964, *The Structure of a Semantic Theory*, en The Structure of Language: Readings in the Philosophy of Language, J.A. Fodor y J.J. Katz, eds., Prentice Hall, Englewood Cliffs, NJ, 1964, pp. 479-518

PUSTEJOVSKY, J., 1991, The generative Lexicon, Computational Linguistics 17 (4)

PUSTEJOVSKY, J., 1993 *Type Coercion and Lexical Selection*, en Semantics and the Lexicon, ed. J. Pustejovsky, Kluwer Academic Publishers, Dordrecht, 1993, pp. 73-96

RUSSELL, B., 1940, An inquiry into meaning and truth, Allen & Unwin,

London, 1940

SMITH, C., BERMEJO, M. & CHANG-RODRIGUEZ, E., 1988, COLLINS Diccionario Español-Inglés Inglés-Español, Ed. Grijalbo, Barcelona, 1990

SOWA, J.F., 1993, *Lexical Structures and Conceptual Structure, Semantics and the Lexicon*, ed. J. Pustejovsky, Kluwer Academic Publishers, Dordrecht, 1993