

Glosas para las colocaciones en el Diccionario de Colocaciones del Español*

MARGARITA ALONSO RAMOS
Universidade da Coruña

1. INTRODUCCIÓN

El presente trabajo tiene como objeto el establecimiento de pautas para la descripción lexicográfica de las colocaciones, más en particular, para su descripción semántica. Se trata de encontrar la mejor formulación para representar lexicográficamente el sentido de colocaciones verbales como *granjearse la antipatía*, *despertar la admiración* o colocaciones adjetivas como *fervente admiración*, *ganas locas*, etc.¹ Como es sabido, hasta ahora las colocaciones no han recibido una descripción rigurosa en los diccionarios actuales del español: unas veces aparecen tratadas como expresiones idiomáticas, otras, aparecen dispersas entre los ejemplos de una acepción dada, etc. (*vid.* entre otros, Bargalló *et al.* 1997-1998, Castillo Carballo 2002). Por esta razón, hemos emprendido la tarea de elaborar una base de datos colocacional, centrada por el momento en los nombres de sentimiento en español (Alonso Ramos 2003a, 2004).

Nuestro proyecto del *Diccionario de Colocaciones del Español* (de ahora en adelante, *DiCE*) se enmarca dentro de la Lexicología explicativa y combinatoria (LEC, Mel'čuk *et al.* 1995), que es el componente léxico de la Teoría Sentido-Texto (Mel'čuk

* Quisiera aprovechar la oportunidad para agradecer a Leo Wanner y a Begoña Sanromán la lectura de las diferentes versiones de este texto. Mi agradecimiento también se dirige a Igor Mel'čuk por recordarme que todo se ha dicho ya hace cincuenta años. Este trabajo se enmarca en el proyecto de investigación HUM2005-08052-C02-02, financiado por el Ministerio de Educación y Ciencia.

¹ El término *colocación* no es interpretado del mismo modo por los distintos investigadores. Simplificando, hay dos interpretaciones: una como combinación frecuente de palabras, en el sentido estadístico (Sinclair 1991) y otra como una combinación en la que una palabra exige la presencia de otra para expresar un sentido dado, en el sentido más lexicográfico (Hausmann 1979). Aquí nos limitaremos a la segunda interpretación.

1997, entre otros). En este marco teórico se dispone de una herramienta que permite codificar las relaciones colocacionales: las *funciones léxicas* (vid. Mel'čuk 1996, entre otros). Una función léxica (FL) modeliza la relación existente entre dos unidades léxicas en donde una de ellas, llamada *base* de la colocación, controla la elección léxica de la otra, llamada *colocativo*. Así, por ejemplo, la FL Magn codifica la relación existente entre los siguientes pares de nombres y adjetivos: *honda pena*, *terrible vergüenza*, *ganas locas* y *fervente admiración*. En estas colocaciones, el colocativo expresa la intensificación de la base.

En el *DiCE*, así como en todos los diccionarios enmarcados en la LEC,² codificamos las colocaciones de cada uno de los lemas por medio de las FFLL. Con ello describimos tanto semántica como sintácticamente la colocación dado que una FL codifica simultáneamente información semántica e información sintáctica. A diferencia de otros diccionarios de colocaciones publicados como el BBI (Benson *et al.* 1986), el LTP (Hill y Lewis 1997) y el OCD (Crowther *et al.* 2002) en donde las colocaciones son simplemente agrupadas semánticamente por la proximidad del significado de los colocativos y clasificadas por la clase de palabras del colocativo, en el *DiCE* las colocaciones son objeto de una descripción más detallada, tanto desde un punto de vista semántico como sintáctico.

Ahora bien, si el lenguaje formal constituido por las FFLL aporta sistematicidad y rigor a la descripción de la colocación, también eleva el grado de dificultad para el manejo del diccionario a los usuarios no familiarizados con este formalismo. Esta es una de las razones por las que en el *DiCE* proponemos también una codificación en metalengua natural que glose o parafrasee el sentido de la FL, siguiendo la práctica iniciada en el proyecto en curso *Lexique actif du français* (LAF, Polguère 2000). Así, por ejemplo, la glosa *intenso* acompañará a la FL Magn en las entradas correspondientes; o la glosa *empezar a sentir* será la paráfrasis en metalengua natural de la FL *IncepOper₁* de todo nombre de sentimiento.³ La glosa no es más que una somera indicación del significado del colocativo. Así, bajo la glosa *intenso* agrupamos diferentes adjetivos como *cerval*, *atroz*, *visceral*, *fuerte*, etc. que, en combinación con MIEDO, cumplen aproximadamente el mismo papel, lo que no quiere decir que tengan estrictamente el mismo significado. Como veremos, existen varias inter-

² Vid. para el ruso Mel'čuk y Zholkovsky (1984) y para el francés, los cuatro volúmenes del *Dictionnaire explicatif et combinatoire du français contemporain* (DEC, Mel'čuk *et al.* 1984-1999), así como el proyecto *Lexique actif du français*.

³ A lo largo del trabajo irán apareciendo diferentes combinaciones de FFLL, sin más explicación que la glosa a su lado, ambas con la fuente Courier New. Puesto que ese trabajo se concentra más en las glosas que en las FFLL, creemos que el lector puede seguir el hilo de la argumentación sin entrar en el detalle del formalismo. El lector interesado en una presentación de las FFLL en español puede consultar Alonso Ramos (1993).

pretaciones de *glosa*, según nos centremos en glosar el nombre de la FL o el colocativo; es decir, podemos parafrasear el sentido modelizado por el nombre de la FL o podemos parafrasear el colocativo. Mostraremos que para la correcta formulación de las glosas intervienen diversos factores que impiden la correspondencia automática entre una FL y una glosa.

Por tanto, en ese trabajo nos proponemos establecer una serie de criterios para la formulación de las glosas. Se trata de ahondar en la teorización de las paráfrasis de las FFL atendiendo simultáneamente a dos frentes. Uno, con finalidad esencialmente práctica, apunta a buscar la estandarización y la homogeneización de las glosas. Hay que evitar la inflación de glosas porque se pierde justamente la generalización ofrecida por la FL. Otro, con una finalidad más teórica aunque también con reflejos en la práctica, se dirige a reflexionar sobre la naturaleza teórica de lo que estamos llamando *glosa* de una función léxica.

Antes de entrar de lleno en la descripción semántica de las colocaciones, necesitamos introducir brevemente el *DiCE*, tal y como aparece en su versión en Internet. En la siguiente sección presentaremos brevemente la estructura de un artículo lexicográfico y las posibilidades de consulta. A continuación, expondremos la noción de FL, intentando separarla de su «envoltorio» formal. Queremos subrayar que no ponemos en duda que el sistema de las FFL ofrece la respuesta a cómo describir lexicográficamente las colocaciones, pero sí mostraremos que su envoltorio ha oscurecido esta respuesta. Pasaremos enseguida a centrarnos en la naturaleza teórica de la noción de glosa de una colocación, esbozando los criterios para una buena formulación. Asimismo, mostraremos el test de formulación de glosas distribuido a profesionales nativos y a aprendices de español y discutiremos los resultados. El último apartado estará dedicado a detallar cuál es el estado actual de nuestro proyecto lexicográfico y cuáles son nuestras futuras líneas de investigación.

2. PRESENTACIÓN DEL *DiCE* EN INTERNET

Podemos señalar tres características que hacen del *DiCE* un producto específico:

1) A diferencia de otros recursos lexicográficos que son una réplica electrónica de la versión en papel, el *DiCE* fue concebido desde el principio como una base de datos electrónica. Esta característica le permite proporcionar más información y sobre todo, un acceso más fácil a ella.

2) El proceso de redacción del *DiCE* está esencialmente basada en corpus. Cada colocación está atestiguada con varios ejemplos extraídos del corpus, la mayoría del *Corpus de referencia del español actual* (CREA). Puesto que cada registro de la base de datos tiene un campo con ejemplos, el diccionario en sí mismo contiene un

corpus de colocaciones, corpus que puede ser separado del resto de la información incluida en el diccionario.

3) El *DiCE* lleva asociado un módulo didáctico que permite explotar las facilidades de un recurso electrónico y el gran conjunto de ejemplos (*vid.* Alonso Ramos 2005 y 2006).

Existe una «demo» o prototipo de lo que puede ser un diccionario de colocaciones del español desde 2004.⁴ Así, en las siguientes direcciones electrónicas, <<http://www.dicesp.com>> o <<http://dicesp.cesga.es>>, se pueden consultar diez nombres de sentimiento, con sus diferentes sentidos: *admiración, alegría, amistad, cariño, celos, dolor, enemistad, gana, orgullo y sospecha*. Aunque la nomenclatura del *DiCE* se limita, por el momento, al campo semántico de los nombres de sentimiento, incluimos todas las acepciones del lema o palabra polisémica. Nuestra unidad lexicográfica es la *unidad léxica (UL)*, es decir, una palabra tomada en una sola acepción determinada y provista de todas las informaciones que especifican su comportamiento cuando se utiliza en esa acepción. Así, al nombre *pena* le corresponderán diferentes entradas en donde se encontrarán las colocaciones relacionadas con la UL que refiere al 'sentimiento', así como las vinculadas a la UL con el sentido 'sanción'.

En el *DiCE*, se accede inicialmente a la información por la base de la colocación.⁵ Como en cualquier otro diccionario en línea o en CD, a la izquierda aparece la lista de lemas. Una vez que se selecciona uno de ellos con el ratón, se despliega la lista de UL asociadas a ese lema. Para cada una de las UL se recoge la siguiente información (*vid.* fig. 1):

- a) La *etiqueta semántica*, que representa su significado genérico.⁶
- b) La *forma proposicional*, en la que aparecen los participantes de la situación designada por el nombre. Por ejemplo, en la situación designada por AMOR I.1a, hay siempre dos participantes: *amor de alguien por alguien*; el que siente el amor y el objeto del amor, que será también un individuo (*amor por sus hijos, el amor del padre al hijo, su amor hacia María, etc.*).

⁴ Los inicios del *DiCE* se remontan a 1999. Desde entonces, ha recibido financiación de diferentes proyectos de investigación: BFF2002-04226-C03-01 (Ministerio de Ciencia y Tecnología y Feder), PGDIT02PXIB30501PR y PGDITPXIC101401PN (Xunta de Galicia). Actualmente, trabajamos en el marco del proyecto HUM2005-08052-C02-02 (Ministerio de Educación y Ciencia).

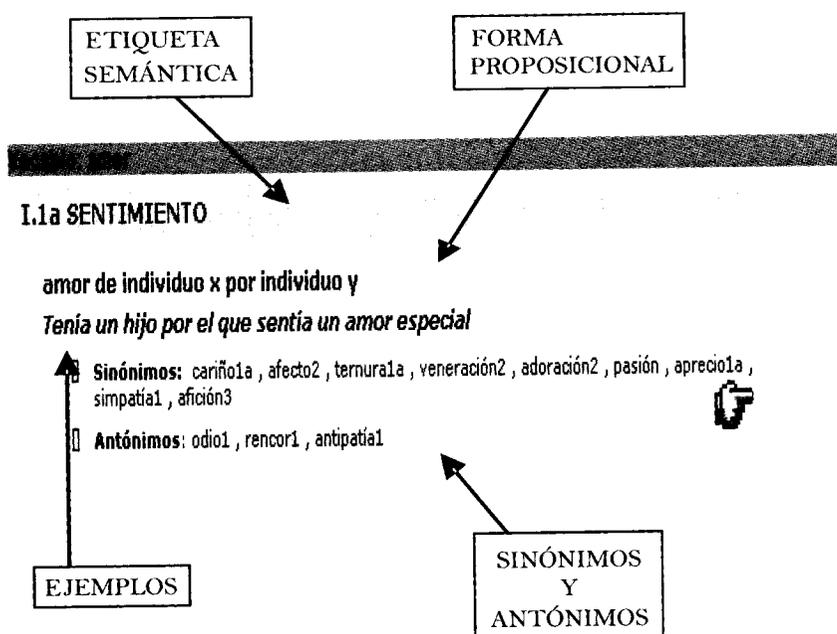
⁵ Aunque también es posible acceder por el colocalivo, la información está introducida con una orientación de producción o codificación.

⁶ Sobre la noción de etiqueta semántica, *vid.* Milićević (1997) y Polguère (2003b).

- c) Ejemplos reales con la indicación de la fuente. Por defecto, la fuente es el CREA, si bien en algunos casos hemos recurrido al corpus Lexesp (<<http://clie.fil.ub.es>>), al Corpus del español desarrollado por M. Davies (<<http://www.corpusdelespanol.org>>), a la Web o a las principales obras lexicográficas del español.
- d) Los (cuasi-)sinónimos y (cuasi-)antónimos, que ayudarán al usuario a seleccionar el sentido.

FIGURA 1.

Vista de la zona semántica del DiCE



A partir de cada UL se puede escoger entre cinco tipos de consulta:

1) Atributos de los participantes: bajo esta rúbrica se agrupan los atributos que refieren a los participantes de la UL en cuestión. Algunos de los atributos se combinan con la UL, pero otros no. Así, por ejemplo, en la entrada para ADMIRACIÓN encontraremos aquí tanto *digno de admiración* como *admirable* ya que son dos modos de referirnos a uno de los participantes de la situación 'admiración', aquello que nos despierta admiración.

2) UL+Adjetivo: aquí encontraremos adjetivos colocativos que se combinen con la UL, tanto antepuestos como pospuestos o en posición atributiva; por ejemplo: el *aburrimento es notable* o *notable antipatía* o *aversión notable*.

3) Verbo+UL: agrupa los verbos colocativos con los que la UL funciona como complemento directo o como complemento de régimen preposicional, del tipo *sentir*, *despertar*, *mostrar antipatía* o *gozar de respeto*.

4) UL+Verbo: agrupa los verbos colocativos con los que la UL funciona como sujeto; por ejemplo, *entrar*, *dar ganas*. Obsérvese que el verbo concuerda con el nombre: *me entraron ganas de salir*.

5) Nombre de UL: se incluyen aquí nombres colocativos que preceden al lema introducidos por la preposición *de*, como por ejemplo, *atisbo de esperanza*, *acceso de cólera* o *mirada de odio*.

De este modo, el usuario puede acceder a las colocaciones de cada UL, seleccionándolas en primer lugar por la clase de palabra del colocativo. Una vez que el usuario ha elegido una de estas combinaciones, se desplegará una lista de glosas asociadas a las FFLL de la entrada en cuestión. Así, por ejemplo, si el usuario quiere escoger un adjetivo que se combina con la UL *alegría* 1a, seleccionará ALEGRÍA+Adjetivo. Ahí aparece una lista de adjetivos, cada uno de los cuales aparece precedido de una glosa que apunta a describir someramente su significado; por ejemplo, *intensa* (que le llevará a *desbordante*, *enorme*, *gran*, *impagable*, *indecible*, *inmensa*, etc.), *más intensa de lo conveniente* (*descontrolada*, *desmesurada*), *compartida por muchos* (*generalizada*), *que no se interrumpe* (*constante*), *que no dura* (*breve*, *efímera*, *pasajera*), *causada por un buen motivo* (*sana*), *causada por un mal ajeno* (*alevosa*, *maligna*) y así un largo etcétera.

El modo interactivo permite a los usuarios que conozcan la herramienta de las FFLL consultar cuál es la FL que describe cada colocación. Así, en el caso de la colocación *incondicional admiración*, la información completa se ofrece de la siguiente manera:

intensa FL MAGN

incondicional Orson Welles siempre profesó una incondicional admiración hacia John Ford.

Además de la glosa y de los ejemplos, en el caso de los nombres y en el de los verbos que se combinan con el lema como complemento, ofrecemos también una sucinta indicación del *régimen* de la colocación. Así, por ejemplo, la notación [ART ~]

indica que el lema debe llevar un artículo u otro determinante equivalente en la colocación; o se indica cuál es la preposición que introduce un actante dado en compañía de un verbo colocativo. Así en el caso de *le quitaron las ganas a Juan*, el verbo *quitar* aparece seguido de este régimen: [ART ~ a X].

El *DiCE* es un diccionario concebido especialmente para la producción o codificación. El modo de acceder a la información contenida en él refleja la estrategia de la producción: si un usuario quiere saber qué otro verbo puede utilizar para decir que *tiene esperanzas*, lo encontrará a partir del nombre ESPERANZA y seleccionando Verbo+ESPERANZA, podrá escoger, por ejemplo, *abrigar*. Ahora bien, es posible también hacer una consulta inversa y encontrar a partir de *abrigar* con qué otros nombres se combina este verbo y cuál es su glosa. Si se pulsa sobre este verbo, se lanza la consulta inversa y encontraremos las colocaciones formadas por los siguientes nombres: *confianza, desconfianza, deseo, hostilidad, ilusión, inquietud, miedo, odio, recelo, rencor, resquemor, sospecha y temor*.

Como vemos, en el *DiCE* no se concede a las FFLL mucha visibilidad, a pesar del papel crucial que desempeñan en el desarrollo de cada uno de los artículos lexicográficos. A continuación nos centraremos en la noción de FL intentando desvincularlas de su notación formal. Como ha señalado Polguère (2003a), la noción de FL debe ser considerada independientemente de su formalización actual, pues esta última no es apropiada como soporte para el aprendizaje de la noción y para su utilización.

3. NOCIÓN DE FUNCIÓN LÉXICA

La noción de FL desempeña un papel primordial en el marco teórico de la Lexicología explicativa y combinatoria cuyos resultados se plasman en los diferentes volúmenes del *Dictionnaire explicatif et combinatoire du français contemporain* (DEC, Mel'čuk *et al.* 1984/ 1999), en el proyecto en curso del *Lexique actif du français* (LAF, Polguère 2000) y ahora también en español, en el *DiCE* (Alonso Ramos y Sanromán 2000). Existen ya varias presentaciones de la noción de FL.⁷ Aquí nos gustaría poner de relieve la noción misma de FL.

3.1. Cuestiones preliminares

Grosso modo, una FL es un contenido semántico asociado a un esquema sintáctico, tal que su expresión depende de la UL a la que esta FL se aplica. En otras

⁷ Las primeras presentaciones se remontan a los años 60. *Vid.* Žolkovskij y Mel'čuk (1965, 1967). Mel'čuk (1982), Mel'čuk *et al.* (1995: 125-152), Mel'čuk (1996), Wanner (1996), así como Mel'čuk (en este volumen).

palabras, una FL corresponde a un significado general y abstracto que puede ser expresado léxicamente en una gran variedad de modos, según la UL con la que este significado es expresado. En términos de FFLL, una colocación formada por L_1 (= base) y L_2 (= colocativo) se presenta como $f(L_1) = \{L_2\}$, en donde L_1 es la *palabra llave* de la correspondiente FL y L_2 , su *valor* que es también una UL o un conjunto de UL cuasi-sinónimas. Así, por ejemplo, la FL Magn asocia la UL GANA(S) con un conjunto de adjetivos como *terribles*, *enormes*, *locas*, etc. que expresan el sentido 'intenso'. La misma FL aplicada al nombre CARIÑO nos devuelve otro conjunto de adjetivos: *grande*, *intenso*, *profundo*.

Creemos que las FFLL tienen ya una larga historia y que han demostrado ser la herramienta que modeliza la relación colocacional entre una base y un colocativo. Obsérvese que la característica principal de una colocación es la dependencia funcional del colocativo en relación con la base; es decir, a distintas bases, distintos colocativos para expresar aproximadamente el mismo sentido. Parece evidente entonces que para modelizar este tipo de relación, la noción de función se presente como la más adecuada.

Sin querer ser exhaustivos, podemos especificar tres razones por las que las FFLL nos parecen la herramienta más adecuada para representar las colocaciones o, más en general, las elecciones controladas léxicamente. En primer lugar, la herramienta de las FFLL permite poner de manifiesto el vínculo conceptual que existe entre las relaciones como la sinonimia, la antonimia, etc., que funcionan en el nivel paradigmático y las colocaciones, que funcionan en el nivel sintagmático. Como Mel'čuk (1996: 38) ha subrayado, ambos tipos de relaciones son de una misma naturaleza lógica y pueden ser representadas por medio del mismo lenguaje. Tomemos un ejemplo. Si un hablante quiere hablar de 'causar alegría', puede optar por una UL que está en relación sintagmática con *alegría* como *dar*, o por un derivado como *alegrar*. La paráfrasis entre *dar alegría* y *alegrar* es una muestra de que ambas relaciones pueden ser codificadas por medio del mismo aparato conceptual (en términos de FFLL, CausFunc_1):

- (1) a. *Pocas cosas me daban más alegría que las colaboraciones de Borges.*
 b. *Pocas cosas me alegraban más que las colaboraciones de Borges.*

Si queremos registrar en un diccionario las distintas maneras de expresar 'causar alegría', debemos incluir tanto la colocación verbal como el verbo relacionado paradigmáticamente. La FL permite, por tanto, codificar tanto uno como otro, independientemente de su relación paradigmática o sintagmática con la palabra llave.

En segundo lugar, las FFLL constituyen un sistema estructurado de relaciones léxicas que permite al lexicógrafo abordar una descripción exhaustiva y sistemática de las posibles colocaciones controladas por una UL dada. Por tanto, las FFLL forman una especie de «rejilla» de análisis que permite prever, según el sentido de la

palabra llave, qué FFLL debería tener. Así, por ejemplo, al redactar la entrada de un nombre predicativo, el lexicógrafo sabe que debe buscar un posible verbo de apoyo (codificado por la FL Oper; p. ej. *dar un paseo, soltar un estornudo*); si se trata de un nombre que designe un artefacto, debe proporcionar el colocativo verbal que designe su utilización (codificado por Real; p.ej. *conducir un coche*); si el significado del lema incluye algún componente semántico graduable, habrá que consignar los colocativos que expresan la intensificación (codificados por Magn; p.ej. *ganas locas, profundo cariño*). De esta manera, cuando el lexicógrafo acude al corpus en busca de colocaciones de una UL dada, tiene en mente ya una serie de FFLL, que le facilitarán la búsqueda.⁸

En tercer lugar, una FL posibilita la descripción completa de una colocación puesto que codifica simultáneamente información semántica e información sintáctica. Las FFLL permiten describir que, por ejemplo, *tener respeto* [por alguien] y *tener el respeto* [de alguien] son dos colocaciones verbales, próximas semánticamente pero en absoluto sinónimas: la primera equivale semánticamente al verbo asociado morfológicamente *respetar*, mientras que la segunda es próxima a la versión pasiva *ser respetado*. Las diátesis de estas dos colocaciones son distintas: en la primera, la expresión del experimentador de 'respeto' desempeña el papel de sujeto del verbo *tener*, mientras que en la segunda, el sujeto gramatical expresa el objeto del respeto. Igualmente, las FFLL facilitan la distinción entre dos colocaciones verbales como *me dan ganas* y *me da vergüenza*, desde el punto de vista semántico y desde el punto de vista sintáctico: para la primera, la FL codifica el contenido semántico 'empezar en alguien' y la relación sintáctica «palabra llave-sujeto del colocativo verbal»; en cambio, para la segunda, el contenido semántico codificado por la FL es 'causar en alguien' y la relación sintáctica «palabra llave-complemento del colocativo verbal».

Acabamos de mostrar algunas razones que justifican la elección de las FFLL como herramienta de descripción de las colocaciones. Ahora bien, las FFLL, o mejor, el lenguaje de notación de las FFLL presenta varios problemas. Desde un punto de vista formal, las FFLL son demasiado vagas y poco explícitas para ser tratadas computacionalmente. La combinación de las FFLL simples en FFLL complejas y configuraciones de FFLL, de un lado, y los distintos ámbitos de los índices actanciales, de otro, dificultan en gran medida la inteligibilidad de las FFLL (*vid.* Alonso Ramos

⁸ Es necesario señalar que con frecuencia se encuentran en el corpus colocaciones para las que no hay FFLL estándar, es decir que no se aplican a muchas palabras llave ni tienen muchos valores como es el caso de las FFLL estándar. Se trata de casos como *compartir la alegría*: no hay una FL estándar que sirva para codificar el sentido 'tener un sentimiento por simpatía con otro que también lo tiene'. Sin embargo, el colocativo *compartir* debe ser también registrado en el diccionario.

2003a para los distintos problemas que presenta la codificación tradicional). En Alonso Ramos (1993), trazamos los primeros rasgos que deben tenerse en cuenta en la elaboración de la gramática de las FFLL. Posteriormente, Kahane y Polguère (2001) han desarrollado lo que llaman la *codificación explícita*, en donde representan separadamente la información semántica y la información sintáctica codificadas en la FL, haciendo más explícita esta información y más tratable computacionalmente. Desde el punto de vista del usuario, las FFLL son demasiado formales y oscuras para un usuario que prefiere, sin duda, una codificación en lengua natural, camino por el que hemos optado en el *LAF* (Polguère 2000) y en el *DiCE* con la formulación de las glosas. Con todo, nos gustaría subrayar aquí que independientemente del mayor o menor acierto con el lenguaje de notación con el que se codifiquen las FFLL, la noción de FL es inherentemente válida y corresponde a un hecho lingüístico universal, como veremos a continuación.

3.2. Las FFLL como universal lingüístico

Recientemente, Polguère (2003a: 118) ha incidido en la idea, ya formulada previamente por Mel'čuk en otras ocasiones (por ejemplo, Mel'čuk 1996: 89) de que las FFLL no son una ficción teórica desarrollada en el marco de la Teoría Sentido-Texto, sino que corresponden a un hecho lingüístico universal y que constituyen una parte esencial del conocimiento lingüístico. Su argumentación se basa en que al igual que todo hablante sabe, consciente o inconscientemente, que la sinonimia existe, que toda UL está potencialmente vinculada a otras cuyo sentido es aproximadamente equivalente, también sabe intuitivamente que hay elecciones léxicas condicionadas; es decir, un hablante sabe que hay algo extraño en una combinación como *hacer un paseo*, aunque no lo hay en *hacer un viaje*; igualmente, le chocará oír *hondas ganas*, aunque aceptará sin problemas *honda pena*. Un hablante de español tiene cognitivamente presente que a la UL PENA se le puede aplicar la FL Magn, desde el momento en que sabe que no puede intercambiar los adjetivos de *honda pena* y *ganas locas*, a pesar de que ambos significan ahí aproximadamente lo mismo. Al igual que un hablante reconoce como una violación de su lengua *el silla*, aunque no sepa conscientemente que existe una noción como el género gramatical de los nombres, asimismo rechazará *hacer un paseo*, aunque no esté informado de la noción de FL que codifica los verbos de apoyo. Por lo tanto, aprender una lengua consiste también en aprender que las FFLL existen. Al decir que una FL dada existe, por ejemplo Magn, queremos decir que existe como noción en sí, independientemente de los pares de UL vinculados por esa relación. Si al aprender una lengua, somos conscientes de que Magn existe, pondremos más atención a la hora de escoger un adjetivo intensificador porque sabremos que está léxicamente condicionado por el nombre del que se predica. Asimismo, si somos conscientes de que la FL Oper existe en las lenguas, sabremos que no basta aprender la traducción de un nombre predicativo en la lengua que

estudiamos: al aprender la UL francesa *promenade*, debo aprender también cuál es su verbo de apoyo, *faire* y no *donner*, a diferencia del español.

Las FFLL forman parte del conocimiento lingüístico de todo hablante y constituyen un universal lingüístico. Esta idea expresada explícitamente por Polguère (2003a) apunta a tratar cada FL como una *meta unidad léxica*, que existe en la competencia lingüística de los hablantes de toda lengua. Así, al igual que en toda lengua hay sinónimos, es decir UL relacionadas por la sinonimia, hay también «Oper», es decir verbos de apoyo seleccionados por un nombre predicativo; hay «Magn», es decir, modificadores de nombres o de verbos expresando la intensificación, etc.

Como vemos, estamos entrando en la naturaleza profunda de las FFLL, su estatuto: ¿qué tipo de entidad es una FL? Para poder contestar a esta pregunta, debemos poner de relieve antes la triple naturaleza de las FFLL.

3.3. Triple naturaleza de las FFLL

Cuando se habla de FFLL, no siempre se utilizan los términos con el mismo sentido. Igualmente, se es poco riguroso al utilizar las notaciones y no siempre se sabe a qué nos referimos con la notación $f(L)$, es decir una FL dada aplicada a una palabra llave L. Esta notación se utiliza, en ocasiones, para referirse al hecho de aplicar una FL f a la palabra llave L o al resultado de la aplicación.⁹ Así, se puede decir que se accede al conjunto de colocativos formado por *mortal*, *fuerte*, *encarnizado*, etc. por medio de la aplicación de la FL Magn a la UL *odio* (cuya notación sería $Magn(odio)$); pero también se puede decir que $Magn(odio)$ representa el conjunto formado por *mortal*, *fuerte*, *encarnizado*, etc.

No se explicita claramente cuál es el objeto modelizado por una FL: la colocación, el colocativo o la meta unidad léxica, como hemos visto en la subsección anterior. La falta de claridad en el uso de los términos y las notaciones vinculadas a las FFLL no es casual, sino que está en clara relación con su triple naturaleza (*vid.* Wanner y Alonso Ramos 2005). Analicemos cada una de las facetas de esta triple naturaleza.

En primer lugar, una FL f es una entidad en sí, que forma parte del léxico de todas las lenguas. En la bibliografía de la TST, es usual referirse a las FFLL como UL profundas (*vid.* Mel'čuk 1988). En ese sentido, una FL f opera en el mismo nivel de representación que cualquier otra UL plena. Así, una FL f ocupa un nodo en la representación sintáctica profunda de una oración dada, al igual que cualquier otra UL

⁹ Por ejemplo, Polguère (2003a: 125) utiliza en el mismo texto la palabra *application* y la notación $f(L)$ en dos sentidos. En primer lugar, con el sentido de 'hecho de aplicar una función': «[...] une FL f est une entité en soi. C'est une information codée dans la mémoire du locuteur et qui lui sert à accéder à un petit ensemble de lexies données, retourné par l'*application* $f(L)$ ». En segundo lugar, con el sentido de 'resultado de la aplicación': «[...] accès à des ensembles spécifiques de lexies ou expressions linguistiques (les *applications* $f(L)$)».

libremente escogida.¹⁰ La diferencia entre esta última y una FL es que la FL es una UL generalizada, en el sentido de que todavía no se tiene la aplicación de la función y por lo tanto, no se sabe cómo se va a realizar una FL dada. Así, si en una representación sintáctica profunda, tenemos el nodo *Magn* modificando la UL *odio*, todavía no sabemos si lo que se va a generar en el nivel superficial será *mortal* o *fuerte*, por ejemplo.

En segundo lugar, una FL es, como su nombre indica, una función f que se aplica a una UL L . Es durante el paso del nivel sintáctico profundo al nivel sintáctico superficial cuando se lleva a cabo la interpretación funcional de la notación $f(L)$, ya que es en el nivel sintáctico superficial donde la FL f se lexicalizará por una UL colocativa dada. Por ejemplo, *Magn(odio)* representa el conjunto de colocativos que en sintaxis superficial pueden funcionar como modificadores de la UL *odio*, expresando su intensificación, como *mortal*, *fuerte*, entre otros. Por lo tanto, la notación $f(L)$ representa el resultado de la aplicación, el valor de la función.

En tercer lugar, una FL representa una colocación, una relación entre una base L y un colocativo L' , que forma parte del valor de $f(L)$. Por ejemplo, *Magn(odio)* modeliza la relación entre una base *odio* y el colocativo *mortal*, que es un elemento del conjunto representado por *Magn(L)*. Dado que el par $(L, L' \in \text{Magn}(L))$ representa la colocación, se puede decir que la FL *Magn* modeliza la colocación.

Para responder por tanto a la pregunta de qué tipo de entidad es una FL, debemos concretar a qué interpretación nos estamos refiriendo: a su interpretación como UL generalizada (o meta UL), a su interpretación como resultado de aplicar la función, por lo tanto al conjunto de colocativos o a su interpretación como relación entre una base y un colocativo.

Dado que nuestro interés se centra aquí en las glosas de las colocaciones, vamos a centrarnos en el contenido semántico de las FFLL. En tanto que UL generalizadas, la mayoría de las FFLL tienen un sentido vago. Como Kahane y Polguère (2001) han señalado, por ejemplo, la FL *Magn* es, en sí, una abstracción de la intensificación, pero en cuanto la «traducimos» a una paráfrasis lingüística le damos un colorido semántico que ella misma no tiene. La intensificación, en estado puro, nunca es expresada como tal ya que el valor de la FL siempre aporta su añadido particular. Desde este punto de vista, la FL *Magn* tiene un sentido vago que puede ser descrito por un conjunto alternativo de expresiones lingüísticas como 'intenso', 'muy/mucho' o 'en alto grado'. Sin embargo, la aplicación de esta FL a una UL dada, como por ejemplo

¹⁰ En la TST se distinguen dos niveles sintácticos, uno profundo y otro superficial. En el nivel de representación sintáctica profunda, más orientado hacia la semántica, los nodos están ocupados por UL plenas o por FFLL. Así, la colocación *odio mortal* estará representada en ese nivel por el siguiente subárbol: ODIO—ATTR—*Magn*. En cambio, en el nivel de representación sintáctica superficial, ya aparece como nodo el colocativo específico, que es un elemento del valor *Magn(odio)*: ODIO—modif—MORTAL.

Magn(*sospecha*), sí puede recibir una paráfrasis bastante precisa. En el primer caso, siempre que optemos por una de las glosas por defecto de Magn, estamos ya sesgando la interpretación de la FL: no es exactamente lo mismo el sentido ‘intenso’ que el sentido ‘en alto grado’, por ejemplo.¹¹ En el segundo caso, la glosa de uno de los valores no es necesariamente válida para todos los demás, dado que los valores de la aplicación de FL *f* a la UL L no son necesariamente sinónimos. Entre otras razones, los valores pueden diferir semánticamente porque pueden ser el resultado de aplicar la FL a diferentes componentes semánticos de la UL L. Así, por ejemplo, una *vehemente sospecha* y una *sospecha grave* son ambas colocaciones modelizadas por Magn, aunque no podamos decir que significan lo mismo.¹² En el primer caso, se intensifica el componente semántico de ‘creencia’ incluido en ‘sospecha’, mientras que en el segundo, se intensifica el componente negativo del hecho sospechado: una sospecha es grave cuando aquello de lo que se sospecha es considerado muy malo, muy negativo moral o socialmente, como un crimen. Así, podemos tener *la grave sospecha de que María ha cometido un crimen*, pero no *la grave sospecha de que María acaba de llegar* (salvo que el hecho de que María llegue sea interpretado como algo negativo).

Como veremos a continuación, las glosas de las colocaciones pueden ser interpretadas de diferentes maneras, debido en gran medida a la propia naturaleza de la noción de FL. De esta manera, una glosa puede representar: 1) el contenido semántico de la FL *f* como UL profunda; 2) el contenido semántico de los valores de la aplicación de la FL *f* a la UL L; o 3) el contenido semántico de la colocación al completo.

4. NOCIÓN DE GLOSA

La noción de glosa no ha recibido, hasta el momento, ninguna definición explícita, aunque su práctica ya tiene cierta tradición tanto en el proyecto *LAF* como en el *DiCE*. Aunque quizás el concepto ‘glosa’ no esté todavía lo suficientemente delimitado, es conveniente formular ya una definición de la noción correspondiente para poder avanzar en su profundización. Por *glosa* de una FL, entendemos una fórmula en

¹¹ Si no buscamos precisión semántica y no siempre tenemos que buscarla, la mejor forma de codificar el contenido semántico de la FL es por medio de su propio símbolo. Esta es la opción por la que han optado Kahane y Polguère (2001) en donde el nombre de la FL seguida de su estructura argumental entre corchetes es la codificación del contenido semántico. Así, Magn[#] representa el contenido semántico de la FL Magn que toma # como palabra llave. La traducción en metalengua natural de Magn[#] necesariamente nos obliga a optar entre sentidos específicos como ‘# es intenso’ o ‘# en alto grado’ o ‘muy #’.

¹² La FL puede ser enriquecida con subíndices para indicar cuál es el componente semántico de la palabra llave afectado por la intensificación. Así, *vehemente* sería codificado por Magn_{[‘creencia’]}} mientras que *grave* sería codificado como Magn_{[‘malo’]}}.

metalengua natural controlada que permite identificar el sentido de la expresión lingüística modelizada por la FL. Dado que las colocaciones son esencialmente un fenómeno de codificación, las glosas no deben ser formuladas como una paráfrasis que analice el sentido de la colocación, sino como un elemento que sirva para identificar el sentido que se quiere expresar. Para poder analizar en profundidad el sentido de una colocación, necesitaríamos formular explícitamente la definición de la base y examinar cómo el sentido aportado por el colocativo interactúa con el sentido de la base, qué componentes semánticos de la base son afectados, etc. Aunque este procedimiento es factible, la descripción semántica en detalle no es la finalidad de un diccionario de codificación de colocaciones. Al igual que lo que ocurre con las FFLL, el objetivo de las glosas no es perseguir la precisión semántica: su deber es apuntar al conjunto de unidades léxicas que expresan el sentido identificado por la FL (*vid.* Mel'čuk 1996: 80).

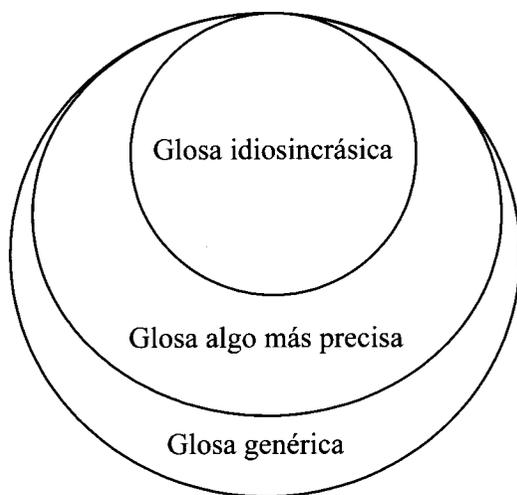
La glosa también pertenece a un metalenguaje; es decir, cuando decimos que la glosa de *Magn* es *intenso*, tomamos este adjetivo del español en su sentido más básico. La glosa *intenso* o cualquier otra pretende representar el sentido de la FL *Magn* o del colocativo representado por ella, pero en sí misma, la glosa no es un sentido del español, sino del meta-español controlado para identificar los sentidos de las colocaciones. Por esta razón, rechazamos glosas como por ejemplo *grave* para la FL *Magn* de los nombres etiquetados semánticamente como 'enfermedad' (*vid.* en cambio, Popović 2003: 14). Desde nuestro punto de vista, *grave* no puede ser propuesto como glosa ya que el sentido de este adjetivo impide que pueda pertenecer a una metalengua: se trata de un colocativo por excelencia que cambia su significado dependiendo de la base (cf. *sospecha grave* vs. *enfermedad grave*). Cuando es seleccionado por nombres que designan enfermedad, intensifica el componente semántico 'consecuencias perjudiciales' que posee el sentido 'enfermedad'. Así, *una enfermedad grave* será aquella cuyas consecuencias son **mu**y perjudiciales. Obsérvese que, por tanto, la pretendida glosa *grave* no explica el sentido de *Magn* si no se combina con la definición de 'enfermedad'.¹³

Aunque acabemos de fijar la noción de glosa como una fórmula que identifica el sentido de la expresión lingüística que modeliza la FL, hemos visto que el objeto modelizado por las FFLL puede ser interpretado de diferentes maneras. A continuación, analizamos cómo la triple naturaleza de las FFLL coincide con tres posibles modos de entender las glosas.

¹³ Al ofrecer *grave* como paráfrasis de *Magn* ('enfermedad'), Popović está optando por la glosa como valor por defecto de la FL, lo que en sí es válido pero el valor por defecto tiene que ser poco marcado idiomáticamente para poder tener alguna utilidad como glosa de la FL.

Las colocaciones pueden recibir tres tipos distintos de glosa, dependiendo del grado de abstracción que podamos o que queramos atribuir a la glosa. Hay colocativos que admiten una glosa que puede ser interpretada como la traducción en metalengua natural del nombre de la FL. Así, por ejemplo, el colocativo de *honda pena* puede ser glosado por *en alto grado*, que sería una traducción de la FL Magn en metalengua natural. En cambio, *sospecha vehemente*, que sería codificada por la misma FL ya no admite fácilmente esa glosa, sino algo como *en la que X está muy seguro de que Y es cierto*, en donde se describe el sentido de la relación del colocativo con la base. Entre ambos extremos, podemos encontrar también colocativos que pueden ser glosados por el valor por defecto de la FL. Por ejemplo, el colocativo *desvanecerse* codificado por la FL compleja *IncepPredMinus(miedo)* puede glosarse por *disminuir*. Tenemos entonces glosas para la FL en tanto que UL generalizada, glosas para el valor de la FL y glosas para la relación del colocativo con la base, es decir, para la colocación al completo. Los tres tipos de glosas están jerarquizados en grados de abstracción. Así, las glosas más abstractas, vinculadas a la FL en tanto que UL generalizada pueden ser siempre reformuladas como la glosa más concreta de la colocación. Por ejemplo, es posible glosar el colocativo de *honda pena* como *que hace sufrir mucho a X*. La dirección contraria, de lo específico a lo general, no parece satisfacer siempre el grado de inteligibilidad: ¿hasta qué punto es comprensible *sospecha en alto grado*? Podríamos representar las glosas gráficamente por círculos concéntricos en donde el círculo central representa las glosas más específicas o idiosincrásicas y el último, a las glosas más genéricas.

FIGURA 2.
Tres tipos de glosas



El primer tipo de glosa se relaciona con la FL como UL profunda y consiste en una paráfrasis en metalengua natural de la fórmula de la FL. La llamaremos *paráfrasis de la fórmula FL*. Así, por ejemplo, la FL compleja IncepPredMinus puede recibir siempre la siguiente glosa: *empezar a ser menos*, que es la traducción en metalengua natural del nombre de la FL. Independientemente de la semántica de las palabras llave a las que esta FL se aplique, la glosa siempre será válida, como se puede observar en las siguientes colocaciones: *el miedo se desvanece*, *las dudas se disipan* o *su vitalidad ha mermado*.

El segundo tipo de glosa tiene más en cuenta la interpretación de la FL como aplicación de la función. Se trata del *valor por defecto de una FL*. En ocasiones, este valor puede ser escogido independientemente de la palabra llave. Por ejemplo, como acabamos de ver, el valor por defecto para IncepPredMinus es *disminuir*. En otras ocasiones, el valor por defecto de una FL depende de la naturaleza semántica de la palabra llave. Así, no tendremos la misma glosa para la misma FL y el mismo valor, pero distinta palabra llave si estas son de distinta naturaleza semántica; por ejemplo, para *miedo atroz* y *crimen atroz*, aunque en ambos casos sea Magn la FL que las represente, la glosa debe ser distinta: para el primer caso, sirve *intenso* como valor por defecto de Magn aplicado a los nombres de sentimiento. Sin embargo, para el segundo, *crimen intenso* no sería interpretable.

Para casos más idiosincrásicos como el de *crimen atroz*, se puede proponer un tercer tipo de glosa: la *pseudo-definición de la relación del colocativo con la base*. Así, en lugar de escoger un valor comodín, el menos marcado idiomáticamente, como glosa de una función, se ofrece una definición abreviada del colocativo en la colocación; por ejemplo, *atroz* en *crimen atroz* sería glosado como *que impresiona mucho*.

Los tres tipos de glosa tienen sus ventajas y sus inconvenientes. Así, las paráfrasis de la fórmula FL presentan la ventaja de que son más cercanas a la representación semántica de la colocación y más adaptadas para un posible sistema de generación automática, pero tienen el inconveniente de que son menos legibles para un usuario humano. En cambio, los valores por defecto son más elegantes estilísticamente lo que los hace más legibles para el usuario (nativo, no necesariamente para el aprendiz). Por último, las pseudo-definiciones tienen a su favor la precisión semántica pero ésta trae consigo pérdida de generalización, con lo que le resta adaptación a los sistemas automáticos. Independientemente de los pros y los contras dependientes de la utilización final de las glosas, es necesario decir que, en ocasiones, los propios hechos de la lengua, es decir, las propias colocaciones piden en unos casos más un tipo de glosa que otro. Por ejemplo, un colocativo como *grave*, dicho de *sospecha* sólo puede ser descrito semánticamente por una pseudo-definición y no admite ni un valor por defecto ni la paráfrasis de la fórmula de la FL. Obsérvese que *sospecha intensa* o *sospecha en alto grado* no describen el significado de *grave sospecha*: a

pesar de que se exprese una intensificación, el componente semántico sobre el que recae la intensificación no es el componente genérico de 'sospecha', por lo que rechaza la paráfrasis 'sospecha intensa'. Por lo tanto, si se quiere identificar el sentido de la colocación en metalengua natural, no hay más solución que ofrecer una definición del colocativo en la colocación. En otras ocasiones, en cambio, el valor por defecto puede funcionar sin ninguna objeción como glosa. Es el caso de *Magn(pena)*, que siempre puede ser glosada como *intensa pena*.

Sea cual sea el tipo de glosa por el que optemos, es necesario establecer una serie de requisitos para su correcta formulación.

5. LA FORMULACIÓN DE LAS GLOSAS

Comenzaremos por estipular los requisitos que debe cumplir una glosa. A continuación, propondremos una lista de factores que determinan la formulación de la glosa. Las dos últimas subsecciones se centrarán en el aspecto práctico: primero, ejemplificando la elección de un tipo de glosa según los distintos factores para colocaciones en concreto y segundo, validando diferentes glosas por medio de tests aplicados a profesionales y a hablantes nativos cualificados.

5.1. *Requisitos*

Estipulamos los siguientes requisitos que deben cumplir las glosas.

1) Adaptación a su finalidad. Si la finalidad de las glosas es didáctica y orientada hacia estudiantes de español, se pondrá más énfasis en la claridad de las glosas que en su precisión. Por ejemplo, una colocación como *despertar simpatía* [en alguien] puede ser glosada bien como causar que alguien sienta simpatía, bien con una fórmula más cercana a la FL utilizando variables actanciales como *Y causa que X sienta simpatía hacia Y*. Si optamos por la finalidad didáctica, evitaremos las variables actanciales como X e Y y se tendrá en cuenta el nivel de dificultad del léxico. Así, por ejemplo, para la glosa de *Able₁* se rechazará *propenso a* por considerar este adjetivo de mayor nivel de dificultad que la glosa que *tiende a*. Si lo que perseguimos es un metalenguaje más científico y más cercano de las fórmulas de FLL, la glosa será poco elegante estilísticamente, pero quizás más precisa.

2) Corrección idiomática. La formulación no debe violar el español. Aunque la glosa forma parte de un metalenguaje, se trata siempre de un meta-español controlado para describir expresiones colocativas del español. No aceptaremos, por tanto, ejecutar como glosa de *Oper₁(golpe)* = *dar* (Morante Vallejo 2004). A pesar de que se

pueden ejecutar acciones, no cualquier acción puede ser ejecutada: *ejecutar un golpe no respeta las normas del español y no puede ser propuesto como glosa.

3) Adecuación al patrón sintáctico de la colocación. En lo posible, la glosa debe reflejar el patrón sintáctico de la colocación, como lo hacen las FFLL. Según esto, una colocación verbal en donde la base funcione como sujeto y otra en donde funcione como complemento, no deberían tener la misma glosa, a pesar de expresar el mismo sentido. Esto es lo que ocurre con las colocaciones codificadas por las FFLL vacías $Oper_1$ y $Func_1$. Por ejemplo, *tener miedo* en *Todos tenían miedo* y *reinar el miedo* en *Reinaba el miedo entre ellos*, serán glosados respectivamente como *sentir miedo* y *el miedo existe*. Como vemos, este requisito de que la glosa sea paralela con el patrón sintáctico de la colocación resta naturalidad a la glosa y contribuye a su multiplicación. Por ejemplo, siguiendo estrictamente el patrón sintáctico, deberíamos proponer glosas diferentes para las colocaciones equivalentes semánticamente pero en donde los actantes funcionan como distintos complementos. Por ejemplo, *proporcionar alegría* [a alguien – O.Ind.] frente a *llenar de alegría* [a alguien – O.D.]. En cambio, si sólo tenemos en cuenta el criterio semántico, la siguiente glosa podría ser válida para ambas: *causar alegría en alguien*. El hecho de que el experimentador en una colocación funcione como complemento indirecto, mientras que en la otra funcione como complemento directo se describe por los subíndices actanciales de la FL y también por el esquema de régimen.

4) Nivel de precisión semántica. Con vistas a evitar una inflación de glosas y buscando ganar cierta generalización, es aconsejable también valorar hasta dónde pretendemos que abarque la glosa; es decir, si formulamos glosas muy precisas y orientadas más al valor de la FL que a la FL, no se ganaría ninguna generalización y se llegaría al punto de que todas las glosas serían diferentes. Por ejemplo, según Popović (2003: 59), en el *LAF* se propone la glosa (que nosotros traducimos al español) *ser cliente en un restaurante* para describir la colocación que designa la utilización de un establecimiento como *restaurant* (*sortir au restaurant*); mientras que la glosa para la utilización de *clínica* (*se trouver dans une clinique*) es *ser paciente de una clínica*. Popović indica que debido al diferente significado de las bases, a pesar de que compartan la misma etiqueta semántica ‘establecimiento’, las glosas tienen que ser distintas. Ahora bien, también podemos plantearnos si no es posible ganar alguna generalización. Efectivamente, también en español, un *paciente* es el que utiliza los servicios de una clínica o un hospital, mientras que un *cliente* es el que utiliza los servicios de un establecimiento en donde compra o consume. Obsérvese que para describir semánticamente los nombres *paciente* o *cliente* dentro de las entradas respectivas de *clínica* y de *restaurant*, necesitamos mencionar el sentido ‘utilizar’, como nosotros aquí lo acabamos de hacer. Por lo tanto, parece que ‘ser

paciente' es más complejo semánticamente que 'utilizar los servicios': es decir 'paciente' incluye semánticamente 'utilizar los servicios', pero no a la inversa.¹⁴ Si optamos por la glosa más simple semánticamente, podemos ganar una generalización y ofrecer la glosa utilizar los servicios tanto para *clínica* como para *restaurante* o cualquier otro establecimiento, dado que en definitiva, un establecimiento es siempre un lugar que ofrece servicios a personas que los utilizan.

5.2. Factores que determinan la formulación de la glosa

Como ya hemos visto, en la formulación de una glosa intervienen diferentes factores o parámetros. Puesto que toda FL consta de tres elementos, analizaremos los diferentes factores desde tres perspectivas: factores que conciernen a la FL en sí, factores que conciernen a la palabra llave de la FL o base de la colocación y factores que atañen al valor de la FL o al colocativo. Empezaremos por los que conciernen a la FL.

5.2.1. Factores que conciernen a la FL

1. Glosa estándar para la FL

Algunas FLL pueden ser traducidas más fácilmente que otras a una metalengua natural. Así, no es posible asociar una glosa en metalengua a una FL vacía como $Oper_1$ sin saber cuál es la naturaleza semántica de la palabra llave. Sin embargo, otras FLL siempre tienen asociada una glosa, como es el caso de *Manif* con la glosa *manifestarse*. En este caso, diremos que la FL tiene una *glosa estándar*.

2. Glosa dependiente de la etiqueta semántica de la palabra llave

Algunas FLL tienen asociada una glosa dependiente de la etiqueta semántica de la palabra llave. Este es el caso de $Oper_1$ y A_1 , por ejemplo. Si se aplican a nombres de sentimiento, la glosa será, respectivamente, *sentir* y *que siente*. Así, $Oper_1$ (*alegría*) = *experimentar* sería glosado como *sentir* y A_1 (*alegría*) = *alegre*, como *que siente*. Así, si el diccionario va siendo redactado por campos semánticos, se puede ir proponiendo una glosa prototípica para cada FL según la etiqueta semántica de la palabra llave. Por ejemplo, para la glosa de $Oper_1$ aplicado a nombres de actos de habla como *crítica*, *acusación* o *reproche* podríamos proponer *hacer que* es el valor por defecto de esta FL con los nombres de este campo semántico; como glosa de $Oper_1$ para nombres que designen sanción como *pena*, *sanción* o *castigo*, se puede proponer *poner*, etc. Pero también es verdad que no podemos esperar que todo $Oper_1$ de un nombre pueda tener una glosa diferente a su valor. Así, habrá casos

¹⁴ Glosar *sortir au restaurant* como 'être client dans un restaurant' es como glosar *conducir un coche* por 'ser conductor de un coche'.

como *dar un paseo* para los que la glosa *hacer* produciría una violación de las normas del español, a pesar de que 'paseo' designa una acción. Por lo tanto, hay que bloquear la herencia por defecto: se trata de evitar que $Oper_1(paseo)$ se glose como *hacer*, bloqueando la herencia a través de la etiqueta semántica de la palabra llave, y ofrecer como glosa, en este caso, el valor de la FL.

Otras FFLL no tienen asociada una glosa dependiente de la etiqueta semántica de la palabra llave. Por ejemplo, el contenido semántico de *Real* no es fácilmente traducible a metalengua natural. En las descripciones teóricas de esta FL se le suele asociar la siguiente paráfrasis: 'hacer con lo designado por la palabra llave aquello para lo que está destinado o lo que se supone que se debe hacer'. Ahora bien, no es siempre fácil determinar qué es aquello para lo que está destinado lo designado por la palabra llave, como ocurre en el caso de los nombres de sentimiento. Hay nombres como *deseo* cuyo objetivo es claro: la realización o el cumplimiento de ese deseo. Las colocaciones *saciar*, *satisfacer*, *cumplir* o *realizar un deseo* son fácilmente codificadas por la FL *Real*. Sin embargo, otros nombres como *cariño* o *cólera* designan un sentimiento que se realiza cuando su manifestación se vuelca sobre el referente del segundo actante (actante que funciona como Objeto, no como Causa). Así, por ejemplo, la colocación *dar cariño* debe ser codificada como $Real_1$ ya que *dar cariño* no es lo mismo que *sentir cariño* ($Oper_1$). Desde este punto de vista, un sentimiento como el cariño está destinado a ser demostrado con gestos o con palabras dirigidas hacia el objeto de cariño. Asimismo, el destino de un sentimiento como la cólera es manifestarla contra el objeto de cólera, de tal manera que al hacerlo, uno se libera del sentimiento. Por ese motivo, las colocaciones *descargar* o *desahogar* serán el valor de $Real_1(cólera)$. En ambos casos, la glosa de esta FL podría ser demostrar ~ hacia/contra Y, pero en este caso la glosa de *Real* se confunde con la de otra FL $Caus_1Manif$.¹⁵

3. Glosa según la estructura actancial

Algunas FFLL hacen referencia a un actante de la palabra llave. Evidentemente, la glosa de $Oper_1$ no puede ser la misma que la de $Oper_2$, ya que intervienen actantes distintos en la colocación. Así, para los sentimientos, la glosa de la primera FL es *sentir*, pero para $Oper_2$, no se puede proponer ciegamente una glosa sin atender al papel semántico del actante Y. Si este actante es un Objeto, la glosa será ser objeto de, pero si se trata de una Causa, tendremos ser la causa de.

¹⁵ En el LAF han optado por una formulación mixta $Real_1$ -manifestation para colocaciones como *manifestar la haine* y tratan solo como $Real_1$ la colocación *déchaîner la haine*, cuya glosa adaptada al español encajaría perfectamente para la colocación española: actuar en relación con Y bajo la influencia del odio.

Así, dependiendo del papel semántico del actante que interviene en una colocación, la misma FL con el mismo índice actancial puede recibir diferentes glosas (Kahane y Polguère 2001). Por ejemplo, entre los nombres de sentimiento, los hay que tienen tres actantes, como *rencor de X hacia Y por Z*, y los hay de dos, como *miedo de X por Y*. Obsérvese que el segundo actante de *rencor* designa el blanco, la víctima que sufre el rencor de alguien. Sin embargo, el segundo actante de *miedo* designa más bien la causa que motiva el miedo. Por lo tanto, la glosa de $Oper_2$ aplicada al nombre *rencor* no será la misma que aplicada a *miedo*, ya que el papel desempeñado por el segundo actante de los nombres no es igual: en el primer caso, será el objeto, la víctima del sentimiento, mientras que en el segundo, se trata de la causa del miedo.

5.2.2. Factores concernientes a la palabra llave de la FL

Con respecto a la palabra llave o la base de la colocación, los factores a tener en cuenta para formular una glosa de la colocación son:

1. Glosa según la etiqueta semántica

La naturaleza semántica de la base de la colocación es determinante para la formulación de la glosa. A propósito de *atroz*, ya hemos visto cómo un colocativo debe ser glosado de un modo diferente, dependiendo de la naturaleza semántica de la base. Popović (2003: 47) ha señalado que la influencia de la etiqueta semántica en la redacción de las glosas no es la misma según el contenido semántico de la FL. Parece bastante evidente que existe más correlación entre la etiqueta de la palabra llave y la glosa de una FL vacía como $Oper$ que con una FL plena como *Manif*. Así, por ejemplo, la glosa de un verbo de apoyo seleccionado por un nombre de sentimiento no puede ser la misma que la glosa para un verbo de apoyo seleccionado por un nombre de acción (*albergar un temor* vs. *dar un paseo*).

2. Glosa según los componentes semánticos afectados

Para la elección de la glosa no sólo intervienen la etiqueta semántica de la palabra llave sino también otros componentes semánticos que configuran su sentido. Esto hace que dos valores de la misma FL aplicada a la misma palabra llave no reciban la misma glosa. Así, *dormir a pierna suelta* y *dormir como un lirón* son ambas codificadas por la FL *Magn*, pero la intensificación recae sobre distintos componentes semánticos de 'dormir'. En el primer caso, se intensifica el componente semántico 'reposar' incluido en el sentido 'dormir', mientras que en el segundo, se intensifica la duración del sueño.¹⁶ Así, de una persona que haya tenido pesadillas durante un sueño

¹⁶ Como ya vimos en la nota 12, para indicar la diferencia, se añaden subíndices a la FL. La notación sería: $Magn_{[reposar]}$ y $Magn_{[tiempo]}$.

que ha durado mucho, podríamos decir que *ha dormido como un lirón*, pero no que *ha dormido a pierna suelta* porque se supone que no habrá descansado. Así, en este caso, la glosa sería con completa despreocupación, mientras que *como un lirón* sería glosado como muchas horas.

5.3. Factores concernientes al valor de la FL

Con respecto al valor de la FL o al colocativo, los factores a tener en cuenta para formular una glosa de la colocación son:

1. Glosa según la productividad del colocativo

Cuando se trata de un colocativo muy productivo como *intenso*, que funciona como un valor de Magn de muchas palabras llave, la glosa entendida como valor por defecto coincidirá con él. En cambio, un colocativo especialmente idiomático como *cerval*, que sólo se combina con *miedo*, nunca coincidirá con una glosa precisamente por ser altamente idiomático.

2. Glosa según el grado de representatividad del colocativo

En algunos casos, el sentido del colocativo expresa principalmente el sentido codificado por la FL, pero en otros, tiene algún añadido semántico. Así, según sea el sentido específico del colocativo, se podrá representar mejor o peor por la glosa de la FL. Por ejemplo, la FL Bon(*alegría*) codifica valores que pueden ser separados en dos grupos: de un lado, *sana, pura* y de otro, *dulce, tranquila, sencilla*. Los primeros adjetivos expresan una valoración positiva del sentido 'alegría': se trata de una alegría causada por un buen motivo. En cambio, el segundo grupo de adjetivos aporta un añadido semántico: no se modifica sólo el sentido 'alegría' sino que se añade una indicación sobre el comportamiento del experimentador. Así, en *una sencilla alegría*, el experimentador está tranquilo y no da muestras de excitación. En consecuencia, la glosa estándar de Bon será más adecuada para los colocativos que no aportan un añadido semántico. Diremos que *sana* es un buen representante de la FL Bon(*alegría*), mientras que el colocativo *sencilla* es peor representante en el sentido de que aporta un añadido semántico específico.

3. Glosa según la autonomía semántica del colocativo

La naturaleza del sentido del colocativo desencadena un tipo u otro de glosa. Hay colocativos con más peso semántico que otros; es decir que tienen mayor autonomía con respecto a la base y constituyen una entrada lexicográfica con su propia definición (Alonso Ramos y Mantha 1996, Alonso Ramos 2003b). Así, frente a *dominar*, en por ejemplo, *Le domina la envidia* que ni siquiera figura en muchos diccionarios de la lengua, el verbo *obedecer* es más autónomo y su sentido es concebible sin ir

necesariamente emparejado con una base que lo seleccione.¹⁷ De esta manera, a la hora de glosarlo en una colocación como *obedecer una orden*, este colocativo es más fácilmente definible que *dominar* seleccionado por *envidia*.

5.4. Puesta en práctica

Como acabamos de ver, la formulación de una glosa para describir la aportación del colocativo al sentido de la colocación no es tarea fácil. Pensamos que los factores que deben primar a la hora de elegir la mejor formulación de una glosa son los que conciernen a la FL. A continuación, mostramos cómo se pueden combinar los distintos factores y cuál es al que le asignamos mayor peso:

- 1) ¿Tiene asignada la FL una glosa bien estándar (como Bon), bien dependiente de la etiqueta semántica (como Magn)?
- 2) ¿El colocativo es un buen representante de la FL en el sentido de que incide sobre el componente genérico de la palabra llave y no aporta un añadido semántico específico (*sana alegría*)?
- 3) ¿El colocativo es una UL autónoma semánticamente como UL (*obedecer con orden*)?
- 4) ¿Buscamos la precisión semántica o buscamos la generalización?

Según sean las distintas respuestas a estas cuatro preguntas, optaremos por la glosa estándar de la FL o por una glosa del colocativo como valor por defecto o por la glosa de la relación del colocativo con la base. Pongamos en práctica estos criterios:

Si se trata de glosar la colocación *sana alegría*, puesto que la FL Bon tiene asignada una glosa estándar y el colocativo es un buen representante de Bon, optaremos por la glosa de la FL: buena. En el caso de *honda pena*, puesto que la FL Magn tiene asignada una glosa dependiente de la etiqueta semántica y el colocativo es un buen representante de Magn, optaremos por la glosa del colocativo como valor por defecto: intensa.

Para glosar la colocación *sencilla alegría*, la FL Bon tiene asignada una glosa estándar, pero el colocativo no es un buen representante ya que tiene un añadido semántico específico. Si buscamos la precisión semántica, optaremos por la glosa de la colocación: X no estando excitado, buena. Si pretendemos la generalización, elegiremos como en el caso anterior la glosa de la FL: buena.

¹⁷ En algunos diccionarios actuales del español, muchos de los colocativos no aparecen descritos semánticamente ni en la entrada de la base ni en la entrada de la colocación. Así, por ejemplo, en el *Lema*, el verbo *dominar*, muy frecuente con los nombres de sentimiento, no recibe ninguna descripción. No ocurre lo mismo en el *DUE* en donde se define como «ser tenido por alguien cierto estado de ánimo, sentimiento, pasión o vicio en tal medida que no puede sustraerse a ellos».

Si se trata de glosar *expresar alegría*, puesto que la FL Caus₁Manif tiene una glosa estándar y el colocativo es autónomo semánticamente, optaremos por la glosa de la FL: manifestar. Si el colocativo no es autónomo semánticamente, como en el caso de *reina la alegría* (Func₀(*alegría*)), optaremos también por la glosa de la FL: existir.

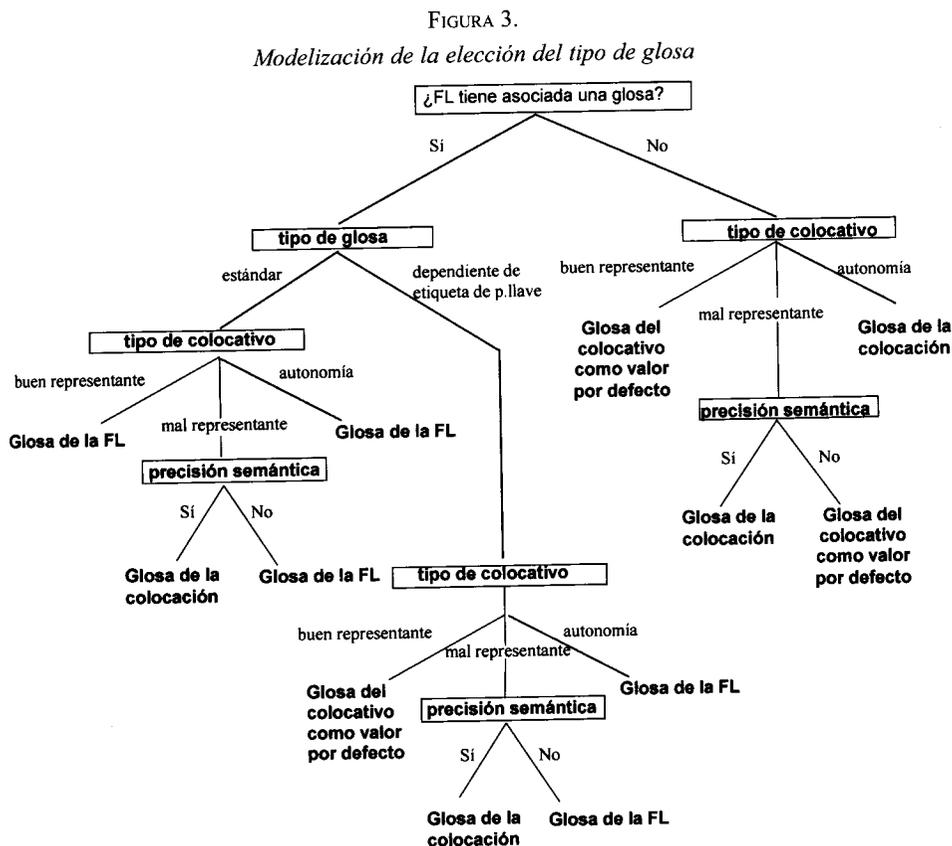
Examinemos ahora los casos en que la FL no tiene una glosa estándar ni dependiente de la etiqueta semántica de la palabra llave. Si se trata de un colocativo que es un buen representante de la FL, como es el caso de *cumplir un deseo*, la glosa será la del colocativo como valor por defecto: realizar. Si el colocativo no es un buen representante de la FL ya que aporta un añadido semántico específico, como es el caso de *descargar la cólera*, tendremos dos opciones: bien elegiremos la glosa del colocativo como valor por defecto si lo que buscamos es la generalización: demostrar; bien optamos por la glosa del colocativo en la colocación si buscamos más precisión semántica: hacer víctima a alguien de la ~. Cuando el colocativo es una UL autónoma semánticamente como *obedecer*, se elegirá la glosa de la colocación. Así, *obedecer una orden* sería codificada por la FL Real₃(*orden*) con la siguiente glosa: hacer lo que X le manda. Lo mismo ocurre si la UL no es autónoma semánticamente, como *dominar* en *la alegría domina*.¹⁸ Optaremos por la glosa de la colocación, es decir por la relación del colocativo en interacción con la base: tan intensa que determina su manera de actuar.

Dado que la etiqueta semántica de la palabra llave desempeña un papel primordial, como criterio operativo tendremos: a misma FL con misma palabra llave, misma glosa semántica; es decir, se procura glosar igual los distintos valores de la misma FL. Si algún elemento del valor de la FL incide sobre un componente semántico específico de la palabra llave o aporta algún añadido semántico, debemos valorar si queremos optar por la generalización o por la precisión semántica.

En conclusión, primará siempre la glosa estándar de la FL. Si la FL no tiene asignada una glosa estándar, se optará por la glosa del colocativo, bien tomándolo como valor por defecto, bien definiéndolo en el contexto de la colocación. El factor de la autonomía semántica del colocativo está directamente vinculado al hecho de si la FL tiene o no una glosa estándar. Así, como hemos visto, colocativos más o menos autónomos serán glosados por la glosa de la FL o recibirán una pseudo-definición, según sea el tipo de FL que los codifique.

Los diferentes caminos que se pueden seguir para elegir la glosa de una colocación pueden ser modelizados en el siguiente árbol de decisión (*vid.* fig. 3.):

¹⁸ El ejemplo que figura en el *DiCE* es el siguiente: «Una alegría desconocida te dominaba, la juventud volvía a correr por tus venas, todo era aún posible».



5.5. Validación de las glosas por medio de tests

Para poder validar nuestra manera de operar en el *DiCE* y verificar si estos criterios son operativos, hemos querido ponerlos a prueba por medio de dos tests, uno destinado a profesionales de Lexicografía y conocedores del fenómeno colocativo (test 1), y otro destinado a hablantes nativos cualificados (test 2). En el primero, la tarea consistía en elegir la glosa que a juicio del colaborador describiera mejor la colocación. Por ejemplo, ante el ejemplo siguiente *Ese profesor le inspira una profunda admiración*, debían elegir entre las glosas siguientes:

1. causa admiración en alguien
2. Y causa que X sienta admiración por Y
3. provoca en alguien admiración
4. Y causa que la admiración de X por Y exista

El test constaba de 20 ejemplos en donde figuraba una colocación con un nombre de sentimiento. El resultado muestra gran desacuerdo entre los colaboradores. De los 20 ejemplos, sólo en dos casos hay coincidencia total, es decir que todos optan por la misma glosa. Así, la glosa propuesta para la FL CausFunc₁, que codifica el ejemplo anterior, es mayoritariamente la (1), que podría ser considerada el valor por defecto. Sin embargo, en el caso de la glosa para la FL Caus₂Func₂, que codifica *ganarse la admiración* de alguien, no hay unanimidad: algunos optan por causar admiración hacia sí; otros por la glosa más técnica Y causa que la admiración de X se dirija a Y; y todavía, otros por un colocativo, provocar, que es sinónimo de *despertar admiración*, pero no de *ganarse la admiración*. Obsérvese que en el primer caso, son las cualidades de alguien las que hacen que otros sientan admiración por esa persona. Sin embargo, en el segundo, la persona tiene que hacer algo para que otros sientan admiración por ella; es algo buscado.

El test dirigido a nativos no estaba destinado a valorar su competencia léxica sino la adecuación del DiCE entre glosas y colocativos. Por esta razón, fue pasado a profesionales o estudiantes de Filología, pero no especialmente avezados en el concepto de colocación. El test requería dos tareas. En la primera, se trata de elegir el colocativo que se adecue mejor a la glosa. Por ejemplo, «Si decimos que mi llegada *causó que la alegría existiera* en María, entonces es que

1. despertó alegría en ella
2. infundió alegría en ella
3. le quitó la alegría

La segunda tarea consistía en rellenar el hueco con el colocativo que se ajuste más a la glosa. Así, por ejemplo, «busca el verbo que exprese el sentido 'causa admiración en alguien' en *El profesor le— una profunda admiración*».

Los resultados son mucho mejores en la tarea 1 que en la 2. Es curioso darse cuenta de que incluso al hablante nativo le supone mucho menos esfuerzo reconocer una UL que se ajuste a un sentido que encontrar esa UL. A pesar de lo reducido del test, no parece apreciarse una correlación entre el resultado con la glosa. Por ejemplo, el colocativo *entrar* en la colocación *entrar ganas* siempre es escogido con éxito independientemente de la glosa: una paralela al patrón sintáctico y próxima a la fórmula de la FL las ganas de algo han aparecido en alguien, pero menos natural; y otra, más natural pero más lejana de la fórmula FL empezar a sentir ganas de algo. Es necesario señalar que el grado de éxito o de error en los ejercicios sobre colocaciones tiene que ser matizado en muchas ocasiones. El criterio para determinar si una colocación es correcta o incorrecta es bastante escurridizo, especialmente con colocaciones de un nivel de lengua avanzado. Así, por ejemplo, algunos hablantes nativos eligieron el colocativo *calmar* para expresar el sentido

'causar que el enfado sea menor', cuando el colocativo previsto como correcto era *apacar*. En otro nivel de lengua, en cambio, un hablante nativo, evidentemente, no dudará entre *hacer* y *dar* al señalar el colocativo seleccionados por *paseo*.

No hay duda de que para la buena formulación de glosas en un diccionario, hay que tener en cuenta si el diccionario está destinado a hablantes nativos o a aprendices de español. Aunque todavía hace falta idear tests más complejos y pasárselos a más personas, creemos que es posible ya afirmar que un hablante nativo prefiere la glosa que hemos llamado *valor por defecto*. Por ejemplo, preferirá encontrar como descripción semántica de la colocación *paliar el desasosiego* la glosa disminuir que la *paráfrasis de la fórmula FL* causar que el desasosiego sea menor. Sin embargo, esa preferencia no tiene por qué coincidir con la de un aprendiz de español. Es probable que un aprendiz desconozca el valor por defecto, pero pueda entender la paráfrasis dado que se trata de un metalenguaje muy simple semánticamente.¹⁹

6. CONCLUSIONES

En el momento actual, el *DiCE* consta de 3600 relaciones codificadas por FFLL, de las cuales sólo una parte tiene asignada una glosa. El *DiCE* se va construyendo introduciendo primero las FFLL y buscando, después, las glosas adecuadas. Sin duda, este procedimiento es debido a que el metalenguaje de las FFLL está más rodado que el de las glosas, que todavía tienen un estatuto vago, pero nos pone de manifiesto la diferente manera de concebir las glosas.

A modo de reflexión final, nos gustaría llamar la atención sobre la circularidad existente entre las FFLL, las colocaciones y las glosas. Veámoslo con un ejemplo. La representación del colocativo *mortal* por la fórmula *Magn(odio)* se comprende si se le asocia su paráfrasis lingüística *intenso*. Por lo tanto, estamos traduciendo una expresión lingüística por otra fórmula que, a su vez, recibe una traducción lingüística, que es lo que hemos venido llamando *glosas*. Esto nos lleva a plantear las glosas como la traducción en metalengua natural del contenido semántico de la FL, de tal modo que una vez ya codificada la colocación por medio de una FL, podemos reinterpretarla en metalengua natural. En este sentido, la glosa es la paráfrasis en metalengua natural del nombre de la FL. Así, la glosa en «meta-español» de *Magn* será *intenso*, en «meta-francés» será *intense*, en «meta-inglés»...

¹⁹ En mi experiencia al consultar el *LAF*, puedo decir que en ocasiones, la glosa de la FL, entendida como valor por defecto, me suscita más dificultad que el propio valor de la FL. Este es el caso de la glosa 'concrétiser' para la FL *Real₁(hostilité) = faire preuve de, manifester, montrer* [ART ~].

Sin embargo, es posible también concebir la glosa como un puente entre la colocación y la FL. El proceso natural que sigue un lexicógrafo a la hora de codificar colocaciones en un diccionario es formular una paráfrasis de su significado. Así, por ejemplo, ante la colocación *perder la admiración* [de alguien], el lexicógrafo la parafraseará por 'ya no ser objeto de la admiración' o por la equivalente 'dejar de recibir la admiración'. Estas paráfrasis son directamente traducibles por FFLL: 'ya no' o 'dejar de' es el sentido de la FL *Fin* y 'recibir' o 'ser objeto de' es la traducción de la FL *Oper₂*. Por lo tanto, la FL que codifica el colcativo *perder* en la entrada de *admiración* será *FinOper₂*.

Estas paráfrasis son las que nosotros proponemos como glosas. En algunos casos, se corresponden directamente a la traducción del nombre de la FL, mientras que en otros casos, la glosa está más próxima a la definición del colcativo. La formulación de las glosas tiene que ser rodada con más datos y probada con usuarios del *DiCE* para poder afinar las técnicas de homogeneización y generalización, pero creemos que estamos en el buen camino, aunque este sea largo.

REFERENCIAS BIBLIOGRÁFICAS

- ALONSO RAMOS, Margarita (1993): *Las Funciones Léxicas en el Modelo Lexicográfico de I. Mel'èuk*, tesis doctoral, Madrid, Universidad Nacional de Educación a Distancia.
- (2003a): «Hacia un Diccionario de Colocaciones del español y su codificación», en M. A. Martí *et al.*, eds., *Lexicografía computacional y semántica*, Barcelona, Edicions de l'Universitat de Barcelona, pp. 11-34.
- (2003b): «La nature des colcatifs: leur statut en tant qu'unités lexicales», en F. Grossmann y A. Tutin, eds., *Les collocations : analyse et traitement, Travaux et Recherches en Linguistique appliquée*, Amsterdam, Editions De Werelt, pp. 45-60.
- (2004): «Elaboración del *Diccionario de colocaciones del español* y sus aplicaciones», en P. Bataner y J. de Cesaris, eds., *De Lexicographia. Actes del I Symposium Internacional de Lexicografía*, Barcelona, IULA y Edicions Petició, pp. 149-162.
- (2005): «Semantic Description of Collocations in a Lexical Database», en F. Kiefer *et al.*, eds., *Papers in Computational Lexicography COMPLEX 2005*, Budapest, Linguistics Institute and Hungarian Academy of Sciences, pp. 17-27.
- (2006): «Towards a dynamic way of learning collocations in a second language», *Proceedings of EURALEX 2006*, Turín, vol. II, pp. 909-921.

- y Suzanne MANTHA (1996): «Description lexicographique des collocatifs dans un *Dictionnaire explicatif et combinatoire*: articles de dictionnaire autonomes?», en A. Clas, P. Thoiron y H. Béjoint, eds., *Lexicomatique et Dictionnaires*, A.U.P.E.L.F-FMA, Montréal-Beyrouth, pp. 233-253.
- y Begoña SANROMÁN (2000): «Construcción de una base de datos de colocaciones léxicas», *Revista de la Sociedad Española de Procesamiento del Lenguaje natural*, 24, pp. 97-98.
- BARGALLÓ ESCRIVÁ, María, José CARAMÉS DÍAZ, Verónica FERRANDO ARAMO y José A. MORENO VILLANUEVA (1997-1998): «El tratamiento de los elementos lexicalizados en la lexicografía española monolingüe», *Revista de lexicografía*, 4, pp. 49-65.
- BATTANER, M. Paz, dir. (2001): *Lema, Diccionario de Lengua Española*, VOX. Barcelona, Spes.
- BENSON MORTON, Evelyn BENSON, Robert ILSON (1986): *The BBI Combinatory Dictionary of English: A Guide to Word Combinations*, Amsterdam, John Benjamins [= BBI].
- CASTILLO CARBALLO, M^a Auxiliadora (2002): «El aspecto colocacional en la producción lexicográfica», en I. Pérez Pascual y M. Campos Souto, eds., *Cuestiones de Lexicografía*, *Linguas e Lingüística* 2, Lugo, Tris Tram, pp. 97-105.
- CROWTHER, Jonathan, Sheila DIGNEN y Diana LEA (eds.) (2002): *Oxford Collocations Dictionary for Students of English*, Oxford, Oxford University Press.
- HAUSMANN, Franz Josef (1979): «Un dictionnaire des collocations est-il possible?», *Travaux de Linguistique et Littérature de l'Université de Strasbourg*, 17, 1, pp. 187-195.
- HILL, Jimmie y Michael LEWIS, eds. (1997): *LTP Dictionary of Selected Collocations*, London, LTP.
- KAHANE, Silvain y A. POLGUÈRE (2001): «Formal Foundation of Lexical Function», *Proceedings of Collocation: Computational Extraction, Analysis and Exploitation*, 39th Annual Meeting and 10th Conference of the European Chapter of the Association for Computational Linguistics, Toulouse, pp. 8-15.
- MEL'ČUK, Igor (1982): «Lexical Functions in Lexicographic Description», en *Proceedings of the VIIIth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, UCB, pp. 427-444.
- (1988): *Dependency Syntax: Theory and Practice*, Albany, N.Y., The SUNY Press.
- (1996): «Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon», en L. Wanner, ed., *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam-Philadelphia, John Benjamins, pp. 37-102.
- (1997): *Vers une linguistique Sens-Texte. Leçon inaugurale*, Paris, Collège de France.
- (en este volumen), «Colocaciones en el Diccionario».
- , André CLAS y Alain POLGUÈRE (1995): *Introduction à la lexicologie explicative et combinatoire*, Bruxelles, Duculot.

- MEL'ČUK, Igor, *et al.* (1984-1999): *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*, I-IV, Montréal, Les Presses de l'Université de Montréal.
- y Alexander ZHOLKOVSKY (1984): *Explanatory Combinatorial Dictionary of Modern Russian. Semantico-syntactic Studies of Russian Vocabulary*, Vienna, Wiener Slavistischer Almanach.
- MILIĆEVIĆ, Jasmina (1997): *Étiquettes sémantiques dans un dictionnaire formalisé du type Dictionnaire Explicatif et Combinatoire*, Memoria de licenciatura, Montréal, Université de Montréal.
- MORANTE VALLEJO, Roser (2004): *Conocimiento lingüístico y recursos computacionales para una aplicación de aprendizaje de vocabulario asistido por ordenador: VOCABLE*, tesis doctoral, Barcelona, Universitat de Barcelona.
- POLGUÈRE, Alain (2000): «Towards a Theoretically-Motivated General Public Dictionary of Semantic Derivations and Collocations for French», *Proceedings of the Ninth EURALEX International Congress, volume II*. Stuttgart, Universität Stuttgart, pp. 517-527.
- (2003a): «Collocations et fonctions lexicales: pour un modèle d'apprentissage», en F. Grossmann y A. Tutin, eds., *Les collocations : analyse et traitement, Travaux et Recherches en Linguistique appliquée*, Amsterdam, Editions De Werelt, pp. 117-133.
- (2003b): «Étiquetage sémantique des lexies dans la base de données DiCo». *T.A.L.*, 44, 2, pp. 39-68.
- POPOVIĆ, Stéfan (2003): *Paraphrasage des liens de fonctions lexicales*, Memoria de licenciatura, Montréal, Université de Montréal.
- SINCLAIR, John (1991): *Corpus, Concordance, Collocation*, Oxford, Oxford University Press.
- WANNER, L., ed. (1996): *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam-Philadelphia, John Benjamins.
- WANNER, L. y M. ALONSO RAMOS (2005): «What type of entity is a Lexical Function?», en J. Apresjan y L. Iomdin, eds., *East West Encounter: Second International Conference in Meaning-Text Theory*, Moscú, Slavic Culture Language Publishing House, pp. 518-528.
- ŽOLKOVSKIJ, Aleksandr e Igor MEL'ČUK (1965): «O vozmožnom metode i instrumentax semantičeskogo sinteza» [Sobre un método y algunos instrumentos de la síntesis semántica], I: *Naučno-texničeskaja informacija*, 5, pp. 23-28.
- e Igor MEL'ČUK (1967): «O semantičeskom sinteze», *Problemy kibernetiki* [Sobre la síntesis semántica], v. 19, pp. 177-238 [Traducción francesa : *T.A. Informations*, 1970, 2, pp. 1-85].