

## El papel de la información semántica en el aprendizaje de las colocaciones léxicas<sup>1</sup>

**Begoña Sanromán Vilas**

Universidad de Helsinki  
PL 59 (Unioninkatu 38 B)  
00014 Helsingin yliopisto  
SUOMI/ FINLANDIA  
begona.sanroman@helsinki.fi

### Abstract

According to our test results, advanced learners of Spanish still make mistakes in the use of collocations. For this reason, this paper insists on the need to create tools in order to facilitate learning collocations more effectively. The study of some English collocation dictionaries reveals that they lack semantic information that can help the user to select the suitable collocate from the lists offered. From our viewpoint, semantic information in this type of dictionaries should not be an option, but a priority, because general dictionaries normally do not include it. We introduce here, as a proposal, the *Diccionario de colocaciones del español* (DiCE), a currently ongoing project. Collocations in DiCE are described by means of lexical functions and glosses. Lexical functions provide syntactic and semantic information about collocations, however, as a formal lexicographical tool, they are difficult for particular users. Glosses, on the contrary, paraphrase the meaning of collocates in a basic Spanish.

### 1 Introducción

Uno de los retos más difíciles a los que se enfrenta un aprendiz de una lengua extranjera es el fenómeno léxico de las colocaciones. El hecho de que una palabra, pongamos por caso *enemigo*, exija la presencia de otra, por ejemplo *acérrimo*, para expresar un sentido particular, aquí 'en alto grado', es una propiedad universal del léxico. Sin embargo, tales combinaciones no siempre pueden traducirse literalmente entre las lenguas. Así, un *enemigo acérrimo* es un enemigo "de sangre" en finés (*verivihollinen*) y un *fumador es empedernido* en español, pero es "pesado" en inglés (*heavy smoker*).

Tradicionalmente en la enseñanza de lenguas, el léxico ha sido tratado casi exclusivamente en su faceta de listas de vocabulario que el estudiante debía memorizar. La rentabilidad de este esfuerzo es muy baja si no se aprovecha, al mismo tiempo, para enseñar cómo se combinan con otras las nuevas palabras. En la actualidad, la situación está cambiando. Em-

---

<sup>1</sup> Este trabajo está subvencionado por el Ministerio de Educación y Ciencia (España) a través del proyecto HUM2005-08052-C02-02/FILO.

pieza a tomarse conciencia de la necesidad de crear herramientas de trabajo que faciliten el aprendizaje de las colocaciones a los hablantes no nativos (Lewis 2000).

Una herramienta indispensable en este proceso es el diccionario, a partir del cual se puede elaborar material didáctico para la enseñanza de una lengua (Alonso Ramos 2002). En español, sin embargo, todavía no disponemos de un diccionario especializado en colocaciones.<sup>2</sup> Por esta razón, desde hace unos años venimos trabajando en el proyecto del *Diccionario de colocaciones del español* (DiCE)<sup>3</sup> (Alonso Ramos 2000, 2003; Alonso Ramos y Sanromán 2002) siguiendo el modelo de la Lexicología explicativa y combinatoria (Mel'čuk et al. 1995 y Mel'čuk et al. 1984-1999).

En este trabajo partimos de la idea de que la información semántica no puede ser facultativa sino esencial y prioritaria en los diccionarios de colocaciones. Sin ella, el usuario, especialmente si se trata de un aprendiz de una lengua extranjera, no puede producir ni tampoco interpretar una colocación. Nuestro objetivo será mostrar cómo consignamos la información semántica de las colocaciones en el DiCE. Dentro de nuestro marco teórico, las colocaciones son descritas por medio de las funciones léxicas (FFLL) (Wanner 1996, Mel'čuk 1998), herramienta lexicográfica formal que permite clasificar las colocaciones de acuerdo con su sentido y/o su función sintáctica. Somos conscientes de que el manejo de las FFLL supone un esfuerzo añadido para el usuario particular, por ello nuestra propuesta actual es introducir una glosa que explique el significado de la colocación (Alonso Ramos 2003 y 2005).

## 2 La competencia colocacional en aprendices avanzados de español

El término "competencia colocacional" (Hill 2002, Alonso Ramos 2003) da cuenta del dominio que tienen los hablantes en la producción y comprensión de colocaciones. En nuestra experiencia con hablantes no nativos de español, cuya lengua materna es el finés, hemos observado que, si se dejan de lado matices fonéticos y ciertos aspectos gramaticales, lo que hace especialmente marcado el discurso del no nativo son precisamente los errores de competencia colocacional. Esta observación nos ha llevado a la elaboración de un test de colocaciones que hemos distribuido entre veinticinco de nuestros estudiantes más avanzados para tener una idea aproximada de los errores más frecuentes.

El test realizado consta de veinte oraciones. Cada una de ellas contiene un nombre en negrita (base de la colocación) y un espacio en blanco para el verbo colocativo con el que concurre. Para cada hueco en blanco hemos ofrecido tres opciones. En las diez primeras oraciones, sólo una de las opciones es correcta; en el resto, sólo una es incorrecta. Entre las respuestas que se ofrecen se incluye siempre una traducción literal del colocativo correspondiente en finés.

---

<sup>2</sup> No incluimos el diccionario REDES (Bosque 2004) en el grupo de los diccionarios de colocaciones. El propio Bosque, en la introducción al diccionario, dedica una sección a las coincidencias y divergencias entre REDES y los diccionarios de colocaciones y manifiesta explícitamente que el concepto de 'colocación' no se ajusta con propiedad a REDES (op. cit.: CLVII). En REDES, los lemas son colocativos y las bases de las colocaciones se agrupan en clases léxicas.

<sup>3</sup> En <http://dicesp.com> y <http://dicesp.cesga.es> se ofrece una demo de este diccionario.

Para el análisis de los resultados y su interpretación, hemos utilizado una clasificación, propuesta por Alonso Ramos (2004: 91-92), en la que se jerarquiza el carácter más o menos vacío de significado de los colocativos verbales. De acuerdo con la autora, existen cinco tipos:

1. Verbos puros: su significado es sólo general y taxonómico; por ejemplo *tener*, que concurre con propiedades y *estar*, con estados.
2. Verbos con componentes semánticos comunes a los de su contrapartida libre: así, *gozar* se combina con nombres que contienen el componente semántico 'agradable'. Se dice *gozar de la admiración de X*, pero no \**gozar de su odio*.
3. Verbos sin contrapartida libre: una parte del sentido de *cometer* (*un error, un crimen*), por ejemplo, es taxonómica ('hacer') y otra parte es específica, relacionada con los componentes semánticos genéricos de los nombres que concurren con él.
4. Verbos con una contrapartida libre homónima: constan de una parte taxonómica y otra específica como en el caso precedente, por ejemplo *librar* (*una batalla*).
5. Verbos semánticamente plenos: mantienen su significado básico, el mismo que tienen como verbos plenos, por ejemplo, *decir*.

En nuestro cuestionario, la mayoría de los verbos son del tipo 1 (ej.: *hacer, dar, tener*); hay, no obstante, representantes de los tipos 2 (ej.: *sufrir*), 4 (ej.: *librar*) y 5 (ej.: *decir*).

De acuerdo con los resultados obtenidos, podemos clasificar en tres grupos los errores. El mayor número de errores se sitúa en los verbos cuyo contenido semántico está comprendido en el tipo 4, es decir, verbos que presentan una contrapartida libre homónima (entre 15 y 20 errores). Registramos 20 errores (un 80%) en la colocación verbo + *jaleo*. Al estudiante se le ha pedido que indique el colocativo incorrecto dando como posibles opciones *armar* (tipo 4), *montar* (4) y *levantar* (2). De 25 encuestados, 18 señalaron *armar* como la combinación incorrecta, 2 apuntaron hacia *montar* y sólo 5 reconocieron que *levantar* no era la correcta calco del colocativo en finés (*nostaa meteli*). Otra colocación en la que se registró un alto porcentaje de errores es *librar una batalla*. La mayoría de los encuestados señaló como opción correcta \**hacer una batalla* y un escaso número indicó \**andar una batalla*. Aquí, tanto *librar*, opción correcta en español, como *andar*, calco del colocativo en finés (*käydä taistelu*) son del tipo 4. Por último, la tercera colocación con un índice elevado de errores ha sido *faltar a su palabra*. Si bien en español *faltar* no es un colocativo del tipo 4 sino 2 presenta componentes semánticos comunes a los de su contrapartida libre la traducción del colocativo en finés, *comer* (*syödä sanansa*), sí pertenece al tipo 4. De 25 estudiantes, 13 señalaron \**comer su palabra* como opción correcta, 5 \**saltar a su palabra* (colocativo inventado del tipo 4) y sólo 7 seleccionaron *faltar a su palabra*.

El segundo grupo de errores (entre 11 y 12) se da, sobre todo, con colocativos del tipo 1 cuando se incluye entre las opciones alguna variante más formal o más coloquial. Así por ejemplo les cuesta aceptar *celebrar una reunión* y *echar un discurso*.

El tercer grupo de errores (entre 5 y 1) se da, en su mayoría, cuando los colocativos son del tipo 1 en ambas lenguas. Dentro de este grupo, el mayor número de errores ocurre en tres colocaciones que en finés se construyen con el verbo *nähdä* ('ver'): *nähdä unta* ('tener un sueño'), *nähdä vaivaa* ('hacer un esfuerzo') y *nähdä nälkää* ('tener hambre').

Por último, destacaremos que sólo en la colocación verbo + *verdad* todos los encuestados señalaron la respuesta correcta. Las opciones ofrecidas eran *hablar*, *contar* y *decir*, colocativos del tipo 5, semánticamente plenos. Los estudiantes indicaron que la única opción incorrecta era *hablar*, precisamente el calco del colocativo en finés (*puhua totta*).

Tras los resultados, podemos confirmar que efectivamente los estudiantes de nivel avanzado cometen errores de colocaciones.<sup>4</sup> Comprobamos que los colocativos del tipo 1, que suelen ser los más difíciles de aprender en el nivel elemental, no ofrecen tanta dificultad en el nivel avanzado. Por el contrario, parece que en este nivel los colocativos que tienen una contrapartida libre son los que ofrecen mayor resistencia (cfr. Nesselhauf 2003a: 232-234).

### 3 Diccionarios de colocaciones e información semántica colocacional

La lexicografía en lengua inglesa ha desarrollado ya varios diccionarios de colocaciones: *The BBI Dictionary of English Word Combinations* (BBI) (1997, 2ª ed.), el *Dictionary of Selected Collocations* (LTP) (1997) y, últimamente (2002), el *Oxford Collocations Dictionary* (OCD). Estos diccionarios, si bien tienen el mérito de ser pioneros en la materia, presentan un problema fundamental: la ausencia de información semántica que ayude al usuario a discernir entre los diferentes grupos de colocativos que se listan.

En concreto, el OCD, que se introduce en el mercado como un diccionario para estudiantes de inglés información que aparece en el subtítulo, incluyendo unas páginas con ejercicios que supuestamente pueden realizarse a partir de las entradas del diccionario, no podemos considerarlo un verdadero diccionario para aprendices. La falta de información semántica no permite al aprendiz seleccionar el colocativo adecuado. Bajo la entrada de *idea*, por ejemplo, se nos presentan varios grupos de colocativos adjetivales; sin embargo, el hecho de no hacer explícita la información semántica no nos permite saber qué diferencia existe entre *great* y *big*, que se consignan en diferentes grupos. Al contrario, en la entrada de *admiration*, *have* y *be full of* se incluyen en el mismo grupo a pesar de que estos verbos presentan una diferencia en el grado de intensidad.<sup>5</sup> Asimismo, los ejercicios que se proponen no pueden realizarse sin la ayuda de un diccionario monolingüe de inglés, destinado a proporcionar la información semántica que aquí se omite (Alonso Ramos 2005).

Si a los estudiantes que han realizado el test descrito en el apartado anterior, se les hubiese ofrecido la posibilidad de utilizar un hipotético diccionario de colocaciones de español sin información semántica, los resultados habrían sido perfectos. Recordemos que este test sólo mide la capacidad del aprendiz para reconocer el colocativo adecuado o inadecuado en español y las interferencias que se producen con su lengua nativa, el finés. Para ilustrarlo, tomemos como ejemplo dos de las cuestiones del test en las que *hambre* aparece como base de una colocación y en las que el encuestado debe señalar la opción incorrecta:

<sup>4</sup> La misma conclusión ha sido ampliamente verificada en estudios realizados con aprendices de inglés. Véanse Banhs y Eldaw 1993, Biskup 1992, Granger 1998, Nesselhauf 2003a y 2003b, entre otros.

<sup>5</sup> También se consignan en el mismo grupo, bajo la entrada de *respect*, los verbos *win/ gain* y *inspire*.

- (1) Hubo épocas en las que \_\_\_\_\_ hambre.
- a) tuvimos
  - b) vimos
  - c) pasamos
- (2) Desayuné bien para que no me \_\_\_\_\_ hambre a media mañana.
- a) viniese
  - b) entrase
  - c) diese

El encuestado no tendría más que consultar la entrada de *hambre* para comprobar que *ver* y *venir* no aparecerían entre los colocativos de este nombre. En definitiva, para contestar correctamente a todas las preguntas, bastaría que el diccionario consistiese en una lista de colocativos que se combinan con el lema. Es más, ni siquiera necesitaríamos recurrir a un diccionario especializado en colocaciones, pues el *Diccionario de uso del español* de María Moliner podría proporcionarnos la misma información. Bajo la entrada de *hambre*, leemos: *tener; pasar; hacer; apaciguar, apagar, matar, saciar, satisfacer; distraer, engañar, entretener; aguantarse; de.*

En nuestra opinión, la especificidad de un diccionario de colocaciones está en la descripción semántica de la colocación, información que los diccionarios generales, por lo regular, no incluyen. Si, en lugar de las cuestiones mencionadas, hubiésemos realizado un test con preguntas como:

- (3) 'Si el hambre empieza a existir en una persona', entonces:
- a) esa persona *pasa* hambre;
  - b) a esa persona le *da* hambre;
  - c) esa persona *distrae* el hambre.
- (4) ¿Cuál es el significado de *faltar a* en la secuencia *El hombre faltó a su palabra*?
- a) no entendió
  - b) se olvidó de
  - c) no cumplió

Diccionarios del tipo del OCD, BBI o LTP no nos ayudarían a encontrar la respuesta adecuada porque en (3), por ejemplo, todas las opciones contienen un colocativo del nombre (*hambre*) que aparecería registrado en sus listas. Necesitaríamos un diccionario que nos indicase que *X pasa hambre* cuando 'X tiene hambre durante un periodo de tiempo', *a X le da hambre* cuando 'el hambre empieza a existir en X' y *X distrae el hambre* cuando 'X hace algo para que el hambre desaparezca temporalmente'. De igual modo, para (4), el diccionario debería explicarnos que *faltar a la palabra* es 'no hacer lo que se debe hacer con la palabra'.

El ejercicio (3) sería un ejemplo del tipo de información que un usuario busca cuando quiere producir un texto; el ejercicio (4), en cambio, mostraría qué necesita el usuario para interpretar un texto. Ambas tareas, la producción o codificación y la comprensión o descodi-

ficación son las necesidades reales a las que se enfrenta el aprendiz o usuario del diccionario. En ambos casos, la información semántica es prioritaria.

#### 4 La descripción lexicográfica de las colocaciones en el DiCE: las FFLL y sus glosas

El DiCE nace con el propósito de elaborar un primer diccionario de colocaciones del español. El proyecto, dirigido por la profesora Margarita Alonso Ramos, se centra, en su primera fase, en los nombres de sentimiento, que presentan una gran riqueza en el dominio de las colocaciones.

Todo artículo lexicográfico en el DiCE presenta tres zonas principales: la semántica, la sintáctica y la de combinatoria léxica. En la zona semántica se consigna una etiqueta semántica, que representa el significado central de la UL y su forma proposicional o estructura de actantes. Así, *cariño* aparece etiquetada como ‘sentimiento’ en una de sus acepciones y como ‘manifestación’ en otra. Su forma proposicional como ‘sentimiento’ es: ‘cariño de individuo X por individuo/ objeto Y’. En la zona sintáctica se indica cómo se realizan en sintaxis superficial los actantes semánticos. Por ejemplo, *el cariño del hijo, su cariño, el cariño filial*.

Finalmente, en la zona de combinatoria léxica se consignan todas las UULL que forman colocaciones con el lema. Las FFLL describen los colocativos aportando tanto información semántica como sintáctica, pero resultan extremadamente complejas para el usuario particular. Con la intención de hacer más accesible y transparente la información, hemos decidido incluir siempre una glosa que traduzca el sentido de la colocación. Por el momento se han registrado tres tipos de glosas (Alonso Ramos 2005): 1) una paráfrasis de la FL, esto es, se describe la colocación en términos más simples semánticamente, utilizando una metalengua o español básico. Ej.: *costrar (cariño)* se describiría con la FL IncepOper1 y se parafrasearía como ‘X empieza a sentir cariño por Y’; 2) un valor por defecto de la FL; por ejemplo, *aumentar* valdría para describir (*el odio*) *crece* (IncepPredPlus) y 3) una definición breve, en caso de no disponer de un valor por defecto. Ej.: (*fumador*) *empedernido* se definiría como ‘con el vicio muy arraigado’.

A medida que el DiCE avanza, también lo hace nuestro conocimiento teórico y práctico de las glosas; no obstante, como indica Alonso Ramos (op. cit.), formular una glosa no es una tarea fácil. Pensemos por ejemplo en la FL Oper1, que tiene como valores verbos vacíos de significado que toman el lema (base de la colocación) como su OD: *sentir admiración, cometer un crimen, correr peligro*, etc. Glosar esta FL presenta diferentes dificultades dependiendo del colocativo concreto y de la base con la que se combina. En general, los colocativos no podrán ni parafrasearse ni ser descritos en términos semánticamente más simples porque son verbos (cuasi) vacíos de significado, cuya función es situar en el tiempo el predicado expresado por el nombre. Algunas de estas colocaciones podrán parafrasearse por un verbo derivado del nombre: *admirar, peligrar*, pero en este caso, habremos sustituido la colocación completa y no sólo el colocativo. Si optamos por glosar el colocativo por un valor por defecto de la FL, nos enfrentamos a otro tipo de problemas: ¿qué colocativo es el más adecuado cuando el nombre puede combinarse con varios?<sup>6</sup>, ¿el que se combina con todos los nombres

<sup>6</sup> Vid. la sección “Verbos de apoyo de un mismo nombre: comodines y apropiados” en Alonso Ramos (2004:100-107).

de un campo semántico, por ejemplo *sentir* para los nombres de sentimiento, *sufrir* para las enfermedades, o *tener*, que concurre tanto con sentimientos como con enfermedades?, ¿la diferencia entre *profesar/ rendir y tener/ sentir (admiración)* es sólo estilística?, ¿qué preferiremos *cometer (un crimen)* y *correr (un peligro)* o bien *hacer (un crimen)* y *estar en (peligro)*? Estas y otras muchas preguntas son las que estamos tratando de responder para desarrollar un mecanismo eficaz que nos permita glosar las colocaciones del DiCE.

## 5 Conclusiones

En este trabajo hemos tratado de justificar la necesidad de dotar a los diccionarios de colocaciones de una descripción semántica de la propia colocación. Sin esta información, los diccionarios no son más que listas de colocativos que no ayudan al usuario, sea hablante nativo o no nativo, a discernir entre los diferentes colocativos. Hemos presentado el proyecto del DiCE que, aparte de describir las colocaciones por medio de FFLL, introduce una glosa o traducción de la FL. Además, este diccionario puede servir como base para desarrollar, por ejemplo, ejercicios destinados a aumentar la competencia colocacional.

## Referencias

### A. Diccionarios

- Benson, M., Benson, E., Ilson, R. (1997), *The BBI Dictionary of English Word Combinations*. (1ª edición 1986, 2ª edición 1997.) Amsterdam/ Philadelphia, John Benjamins.
- Bosque, I. (dir.) (2004), *REDES. Diccionario combinatorio del español contemporáneo*, Madrid, SM.
- Crowther, J., Dignen, Sh., Lea, D. (eds.) (2002), *Oxford Collocations. Dictionary for Students of English*, Oxford, Oxford University Press.
- Hill, J., Lewis, M. (1997), *LTP Dictionary of Selected Collocations*, London, LTP.
- Mel'čuk, I. et al. (1984-1999), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I-IV*. Montreal, Les Presses de L'Université de Montréal.
- Moliner, M. (1989), *Diccionario de uso del español*. (1ª edición.) Madrid, Gredos.

### B. Otras referencias

- Alonso Ramos, M. (2000), 'Presentación del Diccionario de colocaciones y marcadores del español: estructura y objetivos.' presentado en el *IV Congreso de Lingüística General*, Universidad de Cádiz.
- Alonso Ramos, M. (2002), 'Un vacío en la enseñanza del léxico del español como lengua extranjera', en Braasch, A., Povlsen, C. (eds.), *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, volume II*, Copenhagen, CST, pp. 551-561.
- Alonso Ramos, M. (2003), 'Hacia un Diccionario de colocaciones del español y su codificación', en Martí, M. A. Fernández, A. Vázquez, G. (eds.), *Lexicografía computacional y semántica*, Barcelona, Edicions de la Universitat de Barcelona, pp. 11-34.
- Alonso Ramos, M. (2004), *Las construcciones con verbo de apoyo*. Madrid, Visor Libros.
- Alonso Ramos, M. (2005), 'Semantic Description of Collocations in a Lexical Database', en Kiefer, F. et al. (eds.), *Papers in Computational Lexicography COMPLEX 2005*, Budapest, Linguistics Institute and Hungarian Academy of Sciences, pp. 17-27.
- Alonso Ramos, M., Sanromán, B. (2000), 'Construcción de una base de datos de colocaciones léxicas', *Revista de la Sociedad Española de Procesamiento del Lenguaje Natural* 24, pp. 97-98.
- Bahns, J., Eldaw, M. (1993), 'Should We Teach EFL Students Collocations?', *System* 21(1), pp. 101-114.
- Biskup, D. (1992), 'L1 Influence on Learners' Renderings of English Collocations: A Polish/German

- Empirical Study', en Arnaud, P. J., Béjoint, H. (eds.), *Vocabulary and Applied Linguistics*, London, Macmillan, pp. 85-93.
- Cowie, A. P. (ed.) (1998), *Phraseology. Theory, Analysis, and Applications*, Oxford, Oxford University Press.
- Granger, S. (1998), 'Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae', en Cowie, A. P. (ed.), pp. 145-160.
- Hill, J. (2000), 'Revising Priorities: From Grammatical Failure to collocational Success', en Lewis, M. (ed.), pp. 47-69.
- Lewis, M. (ed.) (2000), *Teaching Collocation: Further Developments in the Lexical Approach*. London, ETP.
- Mel'čuk, I. (1998) 'Collocations and Lexical Functions' en Cowie, A. P. (ed.), pp. 23-53.
- Mel'čuk, I., Clas, A., Polguère, A. (1995), *Introduction a la lexicologie explicative et combinatoire*. Louvain-la-Neuve, Duculot.
- Nesselhauf, N. (2003a), 'The Use of Collocations by Advanced Learners of English and Some Implications for Teaching', *Applied Linguistic* 24 (2), pp. 223-242.
- Nesselhauf, N. (2003b), 'Transfer at the Locutional Level: An Investigation of German-Speaking and French-Speaking learners of English', en Tschichold, C. (ed), *English Core Linguistics*, Bern, Peter Lang, pp. 269-286.
- Wanner, L. (ed.) (1996), *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam/ Philadelphia, John Benjamins.