

Diccionarios combinatorios

Margarita Alonso-Ramos

Universidade da Coruña

margarita.alonso@udc.es

Resumen

El objetivo de este trabajo es mostrar las diferentes concepciones existentes en la lexicografía española de lo que se entiende por diccionario combinatorio. Dado que no hay consenso sobre cuál debe ser el objeto de un diccionario combinatorio, nos centraremos en la concepción teórica de las colocaciones, tal y como se entiende en la Lexicología explicativa y combinatoria (Mel'čuk 2012). A continuación nos detendremos en la descripción de los tres diccionarios combinatorios del español actuales (*Redes*, *Práctico* y DiCE), para pasar a mostrar el diseño de una nueva herramienta en línea que sirve para ayudar a escoger colocaciones en el proceso de escritura.

Palabras claves: colocaciones, combinaciones léxicas libres, combinaciones léxicas restringidas, restricciones semánticas, concordancia semántica.

Abstract

The aim of this work is to show the different interpretations found in the Spanish lexicography about what a combinatorial dictionary should be. Since there is no consensus on the object of this kind of dictionary, I will focus on the interpretation of collocations such as they are understood in the Explanatory and Combinatorial Lexicology (Mel'čuk 2012). After that, I will describe in detail the three current Spanish combinatorial dictionaries (*Redes*, *Práctico* and DiCE). Finally, I will present the main features of a new online tool which is aimed to help to choose collocations in the writing process.

Keywords: collocations, free lexical combinations, restricted lexical combinations, semantic restrictions, semantic agreement.

1. Introducción

A pesar de que la naturaleza de un diccionario combinatorio está lejos de ser unánimemente aceptada por la comunidad científica, el objetivo de este trabajo es mostrar las diferentes concepciones existentes en la lexicografía española de qué es un diccionario combinatorio. El principal punto de discusión es la propia definición de su objeto: ¿qué entendemos por “combinaciones” en el contexto de un diccionario? Un usuario puede preguntarse si cabe encontrar en este tipo de diccionario combinaciones léxicas como *destapar*, *tramar* o *preparar un complot*, pero también *destapar*, *tramar* o *preparar una alianza*. Como veremos más tarde, las combinaciones con *complot* son restringidas, mientras que las combinaciones con *alianza* son libres. El término usual para referirse a las combinaciones léxicas restringidas es el de *colocaciones*, pero suscita un gran desacuerdo. De los diccionarios combinatorios del español, solo uno lleva el nombre “colocación” en su título y dos de ellos subrayan que no son

diccionarios de colocaciones. También empieza a usarse más recientemente el término *coaparición* (véase la interfaz de [CORPES XXI](#) de la RAE). Según cómo se defina el objeto de un diccionario combinatorio, los distintos enfoques teóricos dan respuestas a preguntas diferentes. Así, desde un enfoque se procura responder a preguntas como cuál es el verbo que se combina con *complot* para expresar el significado ‘hacer público’, mientras que desde otro enfoque, interesa proporcionar qué cosas pueden ser *destapadas*; es decir, cuáles son los argumentos del predicado ‘destapar’ en sentido figurado.

El panorama de la lexicografía española con respecto a los diccionarios combinatorios es poco halagüeño con respecto al de otras lenguas europeas. Frente al abundante número de diccionarios combinatorios del inglés, francés o italiano,¹ entre otras lenguas, el español dispone solo de tres, uno de los cuales es una versión adaptada de otro de ellos, y el tercero, restringido a un campo semántico.² En el mismo año vieron la luz *Redes* (Bosque 2004) y el *Diccionario de colocaciones del español* (DiCE, Alonso-Ramos 2004), de muy diferente cobertura y enfoque. Dos años más tarde, se publicó el *Diccionario combinatorio práctico* (*Práctico*, Bosque 2006), que es la versión de *Redes* dirigida al gran público. El modelo de *Redes* no existe para ninguna otra lengua, puesto que no es un diccionario combinatorio al uso y su publicación ha tenido impacto en la comunidad hispana en cuanto a la viabilidad del concepto de colocación.³ El término *colocación* no es tampoco habitual en los libros de texto ni entre los profesores de lengua, a pesar de que tanto los estudiantes de lengua española como L1 o de inglés como L2 hacen ejercicios sobre sustitución de verbos comodín (*pronunciar* en lugar de *dar una conferencia*, por ejemplo) o elección entre el verbo *to make* o *to do* en combinación con un nombre dado en inglés. A pesar de todo, desde los 90⁴ no han dejado de publicarse trabajos académicos sobre colocaciones que, sin embargo, se concentran casi exclusivamente en la delimitación del concepto de colocación más que en la descripción de las colocaciones propiamente dichas.⁵ El debate gira en torno a diferentes aspectos, algunos de los cuales trataremos en detalle más abajo, como la pertenencia o no de las colocaciones a la fraseología, puesto que son composicionales semánticamente. Si las colocaciones no son unidades fraseológicas (= combinaciones no libres), solo pueden ser combinaciones libres. Sin embargo, esta línea de razonamiento también niega la existencia de combinaciones libres, puesto que toda combinación léxica sigue ciertas reglas semánticas y sintácticas y, por lo tanto, los elementos de estas expresiones no se combinan libremente. Lo que se sigue es que no hay razón para distinguir entre combinaciones libres y combinaciones restringidas (o colocaciones) porque todas las combinaciones presentan restricciones que, sin embargo, pueden ser explicadas por reglas (Bosque 2004a: LXXXIII). También es objeto de debate la frecuencia como rasgo inherente a las colocaciones; por ejemplo, de un lado, *miedo cerval* es muy poco frecuente, pero la atracción mutua entre sus componentes es muy fuerte; de otro lado, *profundo miedo* / *gran amistad* es frecuente pero, precisamente por eso, parece que no tiene interés su registro lexicográfico.⁶

Todo esto puede llevar a la idea de que el concepto de colocación es vago e inherentemente inestable. No es casual que una reciente monografía se titule *Defining collocations for lexicographic purposes* (Orlandi y Giacomini 2016). Sin embargo, aquí defenderemos que no se trata de un concepto vago sino de que se incluyen bajo el mismo término fenómenos diferentes. Como ya señalé hace años, creo que es posible definir unos criterios para decidir si una expresión dada es o no una colocación. Si otro autor utiliza otros criterios cuya aplicación a esa expresión produce unos resultados

diferentes, no quiere decir que no existen fronteras entre las expresiones fraseológicas, que el carácter fraseológico sea gradual, sino que los criterios son diferentes (Alonso-Ramos 2010). Este estudio está enmarcado en la Lexicología explicativa y combinatoria (Mel'čuk 2012), por lo que necesariamente partimos de una concepción teórica de las colocaciones. En la siguiente sección, trataremos de perfilar las principales propiedades de lo que nosotros consideramos que debe ser objeto de un diccionario combinatorio, las colocaciones, a la vez que simultáneamente caracterizamos fenómenos colindantes. El objetivo es encontrar una intersección entre los diversos modelos teóricos y ver si hay algo común que pueda ser inequívocamente objeto de un diccionario combinatorio. En la sección 3 entramos de lleno en la descripción de los tres diccionarios mencionados para ir encaminándonos en la sección 4 hacia un nuevo modelo de herramienta electrónica que sirva de ayuda en el uso de las colocaciones. Finalmente, en la sección 5 intentaremos dar algunas respuestas aclaratorias dirigidas especialmente al usuario novel de un diccionario combinatorio.

2. Propiedades de las colocaciones

Necesitamos establecer una terminología operativa. Por el momento, basta decir que una colocación es una unidad fraseológica formada por dos elementos, la *base*, que es seleccionada por su significado y el *colocativo*, que es seleccionado bajo el control de la base.⁷ Así, por ejemplo, los nombres *ley*, *promesa* o *secreto* se combinan respectivamente con *violar*, *romper* o *desvelar* para expresar aproximadamente el mismo sentido: 'incumplir'. En este caso se trata de tres colocaciones verbales, en donde el nombre constituye la base y el verbo funciona como colocativo. También podemos tener colocaciones adjetivales (*firme promesa*), nominales (*sarta de mentiras*) y adverbiales (*prometer firmemente*). A continuación, nos detendremos en tres propiedades clave para delimitar adecuadamente la noción de colocación: 1) su composicionalidad, 2) el control léxico ejercido por la base sobre el colocativo, y 3) la concordancia semántica entre los componentes de la colocación.

2.1. Sobre la composicionalidad de las colocaciones

Existe cierta tendencia a pensar que toda unidad fraseológica es no composicional, es decir, que la suma del significado de sus constituyentes no equivale a la suma global. Este es el caso dentro de lo que se entiende comúnmente por *locuciones* como *estirar la pata*. Sin embargo, como ha explicado en detalle Mel'čuk (2012, 2015), las colocaciones son composicionales puesto que el significado global equivale a la suma del significado de sus componentes. No debe confundirse la composicionalidad con la transparencia semántica. Así, por ejemplo, *cosechar una derrota* es perfectamente composicional, a pesar de no ser completamente transparente, dado que el significado global 'ser objeto de una derrota' se puede asignar a cada uno de sus elementos. El hecho de que *cosechar* signifique 'ser objeto de' con *derrota*, *crítica* y pocos nombres más no impide decir que este es el significado de esta unidad léxica, puesto que *cosechar*, como la mayoría de las palabras de todas las lenguas, es polisémica y el significado que se realiza en esta colocación es uno de sus significados. Por lo tanto, la llamada "paradoja de las colocaciones" (Orlandi 2016: 35) no es tal. La restricción léxica no está vinculada a la opacidad de las colocaciones, puesto que no existe menor restricción en el caso de una colocación transparente. No hay tampoco contradicción entre restricción léxica y composicionalidad.

La transparencia semántica no es especialmente pertinente para la descripción de las colocaciones. Una expresión dada puede ser más o menos transparente para un sujeto, nativo o no, dependiendo de si conoce el sentido que tiene el colocativo en coocurrencia con esa base. Así, la colocación *cosechar una derrota* será perfectamente transparente para el que conozca el sentido de esta unidad léxica verbal. Y con respecto a los aprendices de español, la transparencia de una colocación depende, entre otras cosas, de la proximidad entre las lenguas en cuestión. Así, entre francés y español, las colocaciones *gros fumeur* y *gran fumador* son muy transparentes. Sin embargo, desde el punto de vista de la producción, el hablante no nativo no puede prever cuándo un colocativo coincide con su lengua materna y cuándo no: obsérvese el contraste entre *grosse fièvre* y *fiebre alta* (**gran fiebre*). Si se evita la inclusión en el diccionario de colocaciones transparentes, se puede conducir a error a los hablantes no nativos al pensar que la combinación está controlada por reglas semánticas. Ya Hausmann (1999: 131) aludió a la importancia de incluir combinaciones transparentes como *quitter l'hôtel* basándose en el enfoque onomasiológico.

2.2. Sobre el control léxico y el enfoque onomasiológico

Desde este enfoque, es donde se entiende mejor por qué las colocaciones, a pesar de ser composicionales, forman parte de las unidades fraseológicas o combinaciones no libres. Empezaremos por explicar que el adjetivo *libre* en este contexto refiere al modo de selección de las unidades léxicas que forman un sintagma. Así, cuando un sintagma es libre, cada uno de sus componentes léxicos se selecciona según su significado, independientemente de la identidad léxica de los otros componentes (Mel'čuk 2012: 33); por ejemplo, *destapar una alianza* es una combinación libre, lo que significa que uno puede reemplazar cualquier componente con sus sinónimos preservando al mismo tiempo el significado del sintagma: por ejemplo, *descubrir las conversaciones para llegar a un acuerdo*, *revelar el acuerdo pactado*, etc. El hecho de que un sintagma sea libre no significa que no tenga restricciones semánticas. Así, el predicado 'destapar' impone una restricción semántica a su segundo argumento: tiene que ser un hecho susceptible de ser ocultado o descubierto. El adjetivo *libre* debe ser estrictamente comprendido como que permite la selección de una unidad léxica independientemente de las otras unidades léxicas de la misma expresión (Mel'čuk 2012: 33).

Una colocación es una combinación no libre porque uno de sus componentes léxicos es seleccionado por el hablante dependiendo de la identidad léxica del otro componente; es decir, porque la base ejerce un control léxico sobre el colocativo. Por *control léxico* nos referimos a la capacidad de la base de una colocación para controlar la selección del colocativo. Así, para seleccionar el verbo *destapar* con el significado aproximado 'incumplir [un requisito incluido en el sentido del nombre]', la identidad léxica del nombre tiene que ser conocida previamente. Si se trata de *complot* o *conspiración*, por ejemplo, el verbo *destapar* es una elección correcta, mientras que en el caso de *ley* o *requisito*, por ejemplo, se deberían seleccionar *violar*, *desobedecer* o *incumplir*; y si se trata de *promesa*, habría que escoger el verbo *romper*. A pesar de que estos verbos tienen aproximadamente el mismo significado en estos contextos, no es posible intercambiarlos porque cada nombre impone un requisito diferente y en consecuencia, una selección de un conjunto diferente de verbos: **destapar una ley* o **violar un complot*.

Se puede argumentar que en los casos anteriores la elección de *destapar* no depende de la identidad léxica del nombre objeto sino de su significado, de modo que *complot* (y otros nombres con significados similares) se pueden combinar con *destapar* porque este verbo restringe semánticamente la elección de sus posibles complementos. Este argumento, sin embargo, solo se mantiene desde el punto de vista de la comprensión, no de la producción. Dentro de esta perspectiva, no es posible distinguir entre sintagmas libres tales como *destapar una alianza* y una colocación como *destapar un complot*. Ambas combinaciones son semánticamente composicionales y en ambas el verbo *destapar* expresa el mismo significado. La diferencia radica en cómo se produce cada combinación. En una combinación libre, el punto de partida puede ser el nombre o el verbo, dependiendo del contexto; es decir, el hablante puede querer hablar sobre una situación de descubrir algo o sobre una alianza que se descubre. En consecuencia, en este caso, la selección léxica es conducida semánticamente, lo que implica que cualquier verbo que exprese un significado similar: *destapar*, *descubrir* o *revelar* podrían ser usados, puesto que el nombre no impone ninguna restricción. Sin embargo, en el caso de una colocación, el componente léxico que nombra la situación es el nombre: cuando se utiliza el nombre *complot*, el hablante busca una unidad léxica que exprese el significado ‘no cumplir el requisito’, sin tener en cuenta el significado de *destapar*. Obsérvese que este significado ‘hacer público lo que se supone que debe permanecer secreto’ no está activo en el momento de la producción. El significado del nombre *complot* incluye el requisito de permanecer secreto⁸ y por lo tanto, activa la combinación con el significado ‘incumplir el requisito’, como veremos en la sección siguiente. Para poder elegir *destapar* el hablante necesita conocer la otra unidad léxica que ya ha sido elegida. Por lo tanto, en el caso de las colocaciones, la elección de la base es conducida semánticamente, mientras que la elección del colocativo es además conducida léxicamente (Mel’čuk 1996: 37).

El hecho de que una base ejerza el control sobre su colocativo no significa que la relación sea completamente arbitraria; puede estar motivada semánticamente, pero esta motivación no se percibe hasta que se produce la colocación, como veremos en la siguiente sección. Si una colocación es más o menos semánticamente motivada no es relevante para la producción, solo para la comprensión y, como se ha señalado con frecuencia, las colocaciones son un fenómeno de producción.⁹ Esto significa que mientras que las colocaciones son más o menos semánticamente transparentes, un aprendiz de lenguas no tiene forma de saber con certeza cuándo una colocación puede ser traducida literalmente de su lengua materna o no.

El punto de vista de la descodificación o la comprensión busca generalizaciones. Así, a partir de colocaciones ya producidas, es posible clasificar las bases en grupos semánticos, como veremos hace el diccionario *Redes*. Desde esta perspectiva, no hay combinaciones binarias, debido a que los colocativos pueden ser semánticamente compatibles con largos paradigmas de bases (Bosque 2011, 2017). Por el contrario, desde el punto de vista de la codificación o la producción, la agrupación de bases no es relevante, ya que en el proceso de producción de textos un hablante parte de una unidad léxica, no de un grupo de unidades léxicas. Por ejemplo, cuando se habla de un *complot*, el hablante busca una unidad léxica que exprese un significado determinado y se puede combinar con el nombre en cuestión, pero en este proceso el hecho de que *complot* pertenezca al mismo grupo semántico de *conspiración* no es relevante. Este enfoque contrasta, por tanto, con la tendencia actual que tiende a contrarrestar la potencial arbitrariedad de las colocaciones (Higueras 2011, entre otros).

2.3. Sobre la concordancia semántica

Hasta aquí he defendido la existencia de las colocaciones basándome especialmente en el control léxico que ejerce la base sobre la selección del colocativo. Ahora voy a presentar la noción de concordancia semántica para expresar la especial cohesión semántica entre los componentes de una colocación. Seguiré principalmente la concepción de *concordancia semántica* de Apresjan (Apresjan and Glovinskaja 2007; Apresjan 2009), que no difiere esencialmente de la visión de Bosque (2011).¹⁰ La diferencia reside en que mientras el primero apunta a que dentro de las colocaciones se pueden formular expectativas basándose en la semántica, el segundo utiliza la noción de redundancia o concordancia para negar la existencia de las colocaciones. Así, coinciden en que la concordancia semántica se da tanto en las combinaciones libres como en las colocaciones. La concordancia semántica requiere que los dos elementos de una combinación compartan al menos un componente semántico no trivial en sus definiciones. Cuanto mayor sea el número de componentes recurrentes, mayor será la cohesión semántica. Así, a pesar de la potencial polisemia de los elementos de la combinación libre *to cook fish*, la única interpretación semánticamente cohesiva es ‘to prepare the flesh of a water animal as food’, debido al número de sentidos recurrentes compartidos entre ‘cook’ y ‘fish’.

La coocurrencia semántica (es decir, coocurrencia que está basada en la definición de los componentes léxicos) también se da en las colocaciones. Así, por ejemplo, el hecho de que, en inglés, *blunder* se combine fácilmente con *perpetrate* y con *commit*, a diferencia de *mistake*, se debe a que el primer nombre es un error mayor con potenciales efectos dañinos, lo que coincide con esos verbos que implican una más seria infracción. Sin embargo, a pesar de la motivación semántica de la elección de colocativos, en muchos casos, no hay una seguridad absoluta de que basándose solo en la definición de los elementos podamos construir la colocación. La concordancia semántica permite formular expectativas, pero no crear reglas. Dado el significado de una base, podemos prever con qué tipo de colocativos se puede combinar. Así, desde lo más obvio, como que un nombre de estado no tendrá un verbo de apoyo de acción, hasta detalles semánticos más finos como que los nombres de emoción de causa externa momentánea como *susto* o *disgusto* pueden combinarse con *llevarse*, puesto que este verbo incluye también este componente aspectual acotado (Sanromán 2011). Con todo, es imprescindible incluir en el diccionario con qué verbo se combinan cada uno de estos nombres, principalmente por dos razones: en primer lugar, hay excepciones a las posibles generalizaciones, y en segundo lugar, el análisis semántico necesario para poder deducir de qué tipo es cada nombre no es siempre evidente. Por ejemplo, no es obvio contestar a qué tipo de clase semántica pertenece el nombre *cita*; no parece que sea concebible como una acción, sino como el resultado de una acción (‘una cita es algo que ha sido acordado’). Sin embargo, en inglés este nombre se combina con *to make*, verbo prototípico para las acciones, mientras que en español, diríamos *pedir* o *concertar una cita*.

Como se ha señalado muchas veces, el espacio de variación fraseológica en el que se mueven las colocaciones va desde las que se acercan mucho a las combinaciones libres hasta las que se funden con las locuciones. Apresjan (2009) señala que para expresar el significado ‘no bueno’, el abanico de expresiones posibles va desde el adjetivo *bad*, que es posible con un gran número de bases y por lo tanto, predecible, a adjetivos más

idiomáticos como *raw* (con *deal*), *baneful* (con *influence*) o *improper* (con *behaviour*). Sin embargo, tanto los predecibles como los idiomáticos deben estar descritos en el diccionario porque el usuario no puede saber con certeza si se trata o no de una combinación completamente libre o no. La motivación semántica posible entre los elementos de una colocación se percibe solo después de que la colocación se ha producido y no participa, por tanto, en el proceso de producción de la combinación. Lo mismo ocurre en el caso de las locuciones: unas son más transparentes semánticamente que otras, pero en el momento de la producción el hablante escoge la unidad entera sin combinar cada uno de sus elementos.

De hecho, puesto que la gran mayoría de las colocaciones están entre estos dos extremos de variación, cabe preguntarse si en lugar de examinar la motivación semántica entre los miembros de una colocación, no sería más útil valorar el grado de concordancia semántica entre sus componentes. Como pasaré a defender a continuación, la concordancia semántica es mucho más fuerte en las colocaciones que en las combinaciones libres.

El mayor grado de concordancia puede ser usado como criterio para decidir si una combinación léxica dada es una colocación o no. Examinemos ahora la concordancia semántica entre estas dos combinaciones verbales: *destapar una alianza* y *destapar un complot*. Pensamos que solo la segunda es una colocación y no porque sea especialmente idiomática ni frecuente, sino porque hay mayor cohesión semántica proporcionada por la mayor concordancia entre el verbo y el nombre. Así, como ya hemos dicho, el significado ‘complot’ incluye el componente semántico ‘secreto’ y por tanto, cuando se destapa un complot, se está haciendo público algo que era secreto.¹¹ El componente ‘secreto’ se repite en los sentidos del verbo y del nombre. Sin embargo, esta redundancia no se da en la combinación *destapar una alianza*, puesto que las alianzas no son necesariamente secretas. Por tanto, el nombre *alianza* acepta la predicación ‘destapar’ pero no la espera, a diferencia del nombre *complot*.

Analicemos otro ejemplo con el verbo *barajar*. Este verbo se combina con diferentes nombres: *posibilidad*, *opción*, *hipótesis*, *ideas*, *fechas*, *nombres*, entre otros. Pues bien, solo con los tres primeros nombres diríamos que se trata de una colocación, porque solo con esos nombres hay una fuerte concordancia semántica. Este verbo significa aproximadamente ‘considerar las posibilidades de algo’, así que cuando lo combinamos con *posibilidad*, *opción* u otro nombre que no designa una posibilidad pero que la incluye, como *hipótesis*, hay una fuerte concordancia semántica entre el verbo y el nombre. Por tanto, el verbo *barajar* junto con *posibilidad* es altamente redundante, pero desde la perspectiva de la producción lo que importa es que cuando hablamos de posibilidades, necesitamos un verbo que expresa lo que hacemos con las posibilidades: compararlas o considerarlas, y el español ofrece un verbo específico para expresar este significado. Sin embargo, en el resto de los ejemplos, no se da tal concordancia, puesto que la definición de ‘idea’, ‘fecha’ o ‘nombre’ no incluyen el componente semántico ‘posibilidad’. Este componente es añadido por el verbo *barajar* sin que se produzca ninguna concordancia de componentes semánticos del nombre. Por tanto, tenemos combinaciones libres dado que estos nombres aceptan la predicación de *barajar*, pero no la esperan. De hecho, este verbo puede formar combinaciones libres con todo nombre que admita la predicación de ‘posibilidad’ y por lo tanto, no parece haber límites.

A continuación veremos el diferente tratamiento que se da a estas combinaciones en los diccionarios combinatorios del español. Dado que sus objetivos son distintos, *Redes* y *Práctico* incluyen tanto combinaciones libres (*barajar fechas*) como combinaciones restringidas (*barajar hipótesis*), mientras que desde la perspectiva del DiCE solo se consideraría objeto del diccionario la segunda combinación.

3. Descripción de diccionarios combinatorios del español

Podemos empezar por el tipo de preguntas que un usuario puede hacer a un diccionario combinatorio. Si quiere saber cuál es el verbo que sirve para expresar qué se hace cuando se consideran varias hipótesis, la información la tiene que buscar en la entrada del nombre *hipótesis*. Esta es la disposición del *Práctico* y del DiCE y de todos los otros diccionarios combinatorios de otras lenguas. Si, en cambio, la pregunta que se hace el usuario es qué cosas pueden ser barajadas, la información solo puede ser proporcionada en la entrada del verbo *barajar*. En *Redes*, el usuario encontrará organizados en clases los diferentes nombres que pueden ser objeto de este verbo, mientras que en *Práctico* son meramente listados y agrupados semánticamente. Pasemos a explicar con más detalle las principales características de estos diccionarios.¹²

3.1. *Redes*. Diccionario combinatorio del español contemporáneo

Como ya hemos avanzado, *Redes* no es un diccionario combinatorio al uso, puesto que difiere tanto en la macroestructura como en la microestructura del resto de los diccionarios combinatorios. Se diferencia también de otros diccionarios combinatorios en que está más orientado a la descodificación que a la codificación.¹³

3.1.1. Macroestructura de *Redes*

Los lemas de las entradas principales, las llamadas *entradas analíticas*, son predicados. Como indica Bosque (2004a: CXVI) en la amplia introducción al diccionario, la expresión que introduce la información “se combina con” debe interpretarse como ‘se predica de’. La mayoría de los lemas son verbos, adjetivos y adverbios. Se podría decir que los lemas de *Redes* son los colocativos que aparecen en otros diccionarios de colocaciones dentro de las entradas de las bases. Sin embargo, esta afirmación es solo aproximada porque cuando *Redes* abre una entrada, por ejemplo, para *barajar* o para *destapar*, no lo hace en tanto que colocativos seleccionados léxicamente por sus bases, sino en tanto que predicados que seleccionan sus argumentos.¹⁴

Existen también las *entradas abreviadas*, que son generadas automáticamente a partir de las entradas analíticas. Los lemas de estas entradas se corresponden con los argumentos que aparecen descritos en las entradas analíticas, es decir con las palabras seleccionadas por los predicados lemas de las entradas analíticas. Así, por ejemplo, el nombre *hipótesis* es el lema de una entrada abreviada que se ha generado a partir de las entradas analíticas de lemas como *barajar*, *formular*, *plantear*, entre otros. La entrada abreviada es más similar a la de los diccionarios de colocaciones convencionales. En el caso de un nombre como lema, se mencionan adjetivos y verbos ordenados alfabéticamente que aproximadamente se corresponden a los colocativos en un diccionario de colocaciones (*vid.* Fig. 1). Las palabras con superíndices reenvían a las entradas analíticas y a su posición en ellas; por ejemplo, en el caso de *certero* el

superíndice indica que en la entrada analítica *certero*, el nombre *hipótesis* aparece listado en la posición 26.

hipótesis ♦ acertado, arriesgado, atinado, atrevido, aventurado, burdo, certero²⁶, convincente, desacertado, descabellado¹⁵, de trabajo, disparatado, endeble, erróneo, fundado¹², fundamentado, gratuito, implausible, inconsistente, inobjetable, insostenible, irrefutable, manido⁴⁵, mero, peregrino, plausible, rebuscado, remoto, retorcido¹³, sin pies ni cabeza, sin sentido, trivial ♦ a tenor (de)¹², a título (de)⁶ ♦ afianzar(se)⁶², airear²⁶, alimentar⁷⁵, analizar, apoyar, argumen-

Fig. 1. Extracto de la entrada de *hipótesis* en *Redes*

Por último, queda por mencionar las *entradas conceptuales*, que son de dos tipos. El primer tipo son las referencias cruzadas a los conceptos, como es el caso del concepto PERDÓN representado por el lema en versalitas. Aquí se listan los lemas de las entradas analíticas cuyos argumentos están en relación con el concepto ‘perdón’. Así, en este caso, se menciona *deshacerse en e implorar* en cuyas entradas aparecerán sustantivos relacionados con el concepto ‘perdón’ agrupados en la clase E y en la clase C. Si el usuario va a la entrada analítica *deshacer*, encontrará combinaciones como *deshacerse en excusas, explicaciones, disculpas*. El segundo tipo son también una referencia cruzada, pero, en lugar de apuntar al contenido de las entradas analíticas, organizan el propio leuario; por ejemplo, bajo el lema *PENSAMIENTO*, se agrupan diferentes palabras que figuran como lemas de las entradas analíticas como *deductivo, calcular, concebir*, etc. Por lo tanto, puede darse el caso de que un lema esté repetido varias veces en cada tipo de entrada. Es lo que ocurre con el nombre *pensamiento*, que tiene una entrada abreviada, la más similar a los diccionarios de colocaciones convencionales, y también una entrada conceptual donde se reenvía a sustantivos o a verbos de pensamiento.

3.1.2. Microestructura de *Redes*

Nos detendremos en particular en la información contenida en las entradas analíticas que son las principales de este diccionario. Si los lemas de las entradas analíticas son los predicados, el contenido de la entrada lo proporcionan los argumentos, organizados en *clases léxicas*. El concepto de clase léxica, aunque no recibe una definición explícita en *Redes*, refiere una descripción semántica de un conjunto de argumentos seleccionados por el predicado lema de la entrada. Así, en la entrada de *barajar* se mencionan las diferentes clases. *Vid.* la Fig. 2 con un extracto de la entrada de este verbo.

barajar *v.* ■ Con el sentido de ‘considerar o tener en cuenta’ se construye generalmente con complementos en plural (*La empresa baraja nombres de varios candidatos*). Se combina con...

A SUSTANTIVOS QUE DENOTAN POSIBILIDAD, ALTERNATIVA Y OTRAS NOCIONES DE SIGNIFICACIÓN PROSPECTIVA: **1** posibilidad ++: ...anunció que *baraja* la posibilidad de expulsar a los directores y maestros de escuelas públicas que se plegaron a la huelga. LRE120103 **2** hipótesis ++: Días después, cuando su cuerpo fue hallado sin vida (...) miembros de la Subregión del Callao *barajaron* una primera hipótesis. CAP030797 **3** opción ++: Otra de las opciones *barajadas* por Pebsa es el traslado de personal a Cerca Grande. LVG301091 **4** expectativa: Las expectativas que se *barajan* tienden a considerar que su resultado potenciará el alza... LVE240296

B SUSTANTIVOS QUE DESIGNAN DIVERSAS UNIDADES DEL PENSAMIENTO, MÁS FRECUENTEMENTE SI SE RELACIONAN CON LA INTENCIÓN ORGANIZADA DE ACTUAR: **5** propuesta +: ...lo somete ahora a la previsión de inflación de 1996 del 2,6, después de *barajar* propuestas mucho más flexibles. EME050896 **6** idea +: Se han escrito infinidad de informes que, con matices, *barajan* todas esas ideas. LVE030696 **7** teoría: Una teoría *barajada* con frecuencia como fuerza motivadora de este «salto» literal es la de un brusco cambio en las condiciones meteorológicas. ABC080494 **8** plan: ...en fuentes políticas vascas conocedoras de los planes que a tal efecto se *barajan* en el Ministerio de Justicia. LVE030294 **9** proyecto: Uno de los proyectos *barajados* por Sanz pasaría por dotar a Madrid y Barça del porcentaje más elevado. EME021096 **10** tesis -: La tesis que se *baraja* es que uno o varios mantuvieron contactos con mercenarios de origen francés... LVE010495

C SUSTANTIVOS QUE DENOTAN DATO O RESULTADO: **11** cifra ++: Es imperdonable además *barajar* cifras erróneas, la información hablaba de 200 millones de pesetas... EME151096 **12** dato +: La campaña de entrega de remolacha se prolongará hasta el mes de marzo, según los datos que *barajan* la industria Azucarera Ebro y las organizaciones agrarias. ENC120101 **13** resultado: Según los resultados preliminares *barajados* ayer, la coalición Re-

Fig. 2. Extracto de la entrada de *barajar* en *Redes*

La idea subyacente a la clase léxica es la defensa de que las colocaciones no son combinaciones de dos unidades léxicas sino que el predicado ‘barajar’ selecciona diferentes argumentos, agrupados por ciertos rasgos semánticos comunes que forman una clase. Como ya dijimos en la sección 2, las clases están concebidas desde la perspectiva de la comprensión o de la descodificación. En cambio, desde la perspectiva de la producción, las colocaciones son combinaciones binarias, puesto que cuando el hablante busca un verbo que exprese el significado ‘considerar’ referido a una hipótesis, el hecho de que este nombre forme parte de una clase no parece relevante. Sí puede serlo, en cambio, para la comprensión o descodificación: a un aprendiz de español que aprenda el significado de *barajar* con *hipótesis* puede serle útil saber que otros nombres similares semánticamente también se combinan con el mismo verbo para expresar ese significado, pero no puede tener la certeza de qué nombres entran en la clase y cuáles no, además de las potenciales excepciones. No es obvio poder listar qué nombres encajan bajo la clase A. Así, el nombre *suposición* no parece muy alejado de *hipótesis*

pero no parece combinarse con *barajar*, quizás porque no designa una alternativa, pero curiosamente el nombre *alternativa* no aparece mencionado en la clase A.

Las posibles ausencias como esta pueden deberse a cómo se ha obtenido la información. *Redes* está basado en un corpus periodístico de 250 millones de palabras, pero no se explicitan las técnicas que se utilizaron para obtener las combinaciones léxicas. Como se observa en la entrada de *barajar*, se incluye el símbolo “+” para indicar la frecuencia, pero no puede decirse estrictamente que *Redes* sea un diccionario guiado por el corpus (en inglés, *corpus-driven*), puesto que incluye combinaciones no documentadas, aunque “naturales”. En el diccionario se combina la *frecuencia*, concepto estadístico, con la *naturalidad*, concepto lingüístico (Bosque 2004a: CLVIII). De hecho, han primado el concepto de naturalidad cuando optan por marcar con el símbolo de “sumamente frecuente” a combinaciones encontradas pocas veces pero que el lexicógrafo las percibe como sumamente naturales. Por esta razón, este diccionario se reclama como perteneciente a un tipo de *lingüística con corpus* (frente a *lingüística de corpus*): “una lingüística en la que el corpus está al servicio del investigador, de forma que los datos encontrados se filtran por su introspección, se evalúan y se completan con otros que el corpus no proporciona, pero que la introspección considera naturales” (Bosque 2004a: CLIX).

3.2. *Diccionario combinatorio Práctico del español contemporáneo*

El *Práctico* es una versión simplificada de *Redes*¹⁵ y más similar a los diccionarios de colocaciones de otras lenguas. Con todo, Bosque (2006: XVII) insiste en no presentarlo como diccionario de colocaciones, puesto que no quiere asumir “el riesgo conceptual” de que las combinaciones contenidas sean exactamente colocaciones o no. *Práctico* duplica el número de entradas con respecto a *Redes*, pero aligera mucha información, dada su orientación más pedagógica.¹⁶ Una pérdida importante son los ejemplos de corpus que desaparecen en el *Práctico*, así como la falta de información de frecuencia.

3.2.1. *Macroestructura de Práctico*

Los lemas son tanto palabras seleccionadoras como seleccionadas; es decir, aumenta el número de nombres como lemas que en *Redes* solo figuraban como lemas de entradas abreviadas. En la Guía de Uso, se señala que “los lemas son palabras o locuciones elegidas porque su combinatoria léxica se puede restringir” (Bosque 2006: XXXIII). Sin embargo, como ya hemos avanzado antes, no es fácil ponerse de acuerdo en qué consiste esa restricción. Cambia la información de la entrada según la categoría del lema. Así, en las entradas con nombres como lemas, la información se corresponde con un diccionario de colocaciones convencional, pero en las entradas de verbos, adjetivos, preposiciones y adverbios se da una ambivalencia de información que comentaremos en la sección siguiente. A pesar de que este diccionario es más similar a los diccionarios convencionales, sigue siendo diferente en el hecho de abrir lemas para los colocativos. Coincide en esta opción con el McMillan (Rundell 2010) que también abre entradas para colocativos como *to commit* con la finalidad complementaria de facilitar la comprensión de la combinación. Así, un aprendiz que ha averiguado que el verbo *commit* se combina con *crime* puede querer saber más sobre este verbo aprendiendo con qué otros nombres se combina. Como señala McGee (2012), el autor del McMillan no está desafiando la distinción entre base y colocativo, puesto que no dice que un usuario no comienza la producción de una combinación verbal por el nombre, sino que la

entrada del verbo como colocativo sirve para dar una mejor comprensión de su significado.

Existen también *entradas genéricas* que corresponden *grosso modo* al hiperónimo de diferentes nombres. Así, en la entrada genérica de *deporte* se incluirá información útil para *bádminton*, *balompié*, *baloncesto*, etc. De este modo se evita tener que incluir repetidamente en la entrada de estos nombres que, por ejemplo, todos se combinan con el verbo *practicar*.

3.2.2. Microestructura de *Práctico*

La organización inicial viene dada por la clase de palabras. Así, en la entrada de un nombre se listan los adjetivos que se combinan con él, así como otros sustantivos, verbos y preposiciones. No abre distinto grupo para los verbos cuando el nombre funciona como sujeto y cuando funciona como complemento, lo que puede dar lugar a alguna confusión, especialmente para el usuario aprendiz de español. Dentro de cada clase, se agrupan semánticamente en lo que llaman *subgrupos combinatorios* listas de unidades léxicas separadas por el signo ||. Así, en la entrada de *hipótesis* (vid. Fig. 3), se agrupan semánticamente la siguiente lista de adjetivos: *acertada*, *atinada*, *certera*, *válida*, *cierta*, *fuerte* y *falsable*.

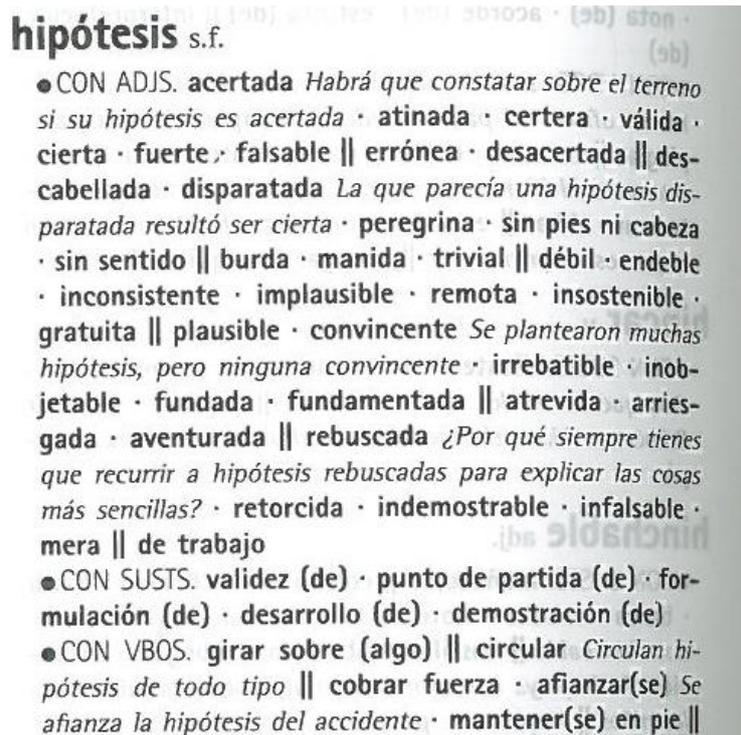


Fig. 3. Extracto de la entrada de *hipótesis* en *Práctico*

El hecho de que *Práctico* desdeñe el llamado “problema de la direccionalidad” (Bosque 2003) hace que la información de las entradas sea ambivalente en muchos casos, puesto que se da simultáneamente información como colocativo y como base. Así, por ejemplo, en la entrada de *coger* en combinación con sustantivos, el verbo está concebido como colocativo (*vuelo*, *cariño*, *infección*, etc.); o si no se quiere llamar “colocativo”, como predicado. Sin embargo, en la información que se da con adverbios, *coger* es concebido como la base (*con las manos en la masa*, *in fraganti*, etc.), o si se

quiere, como la palabra seleccionada por los predicados expresados por esas locuciones. También se da una información ambivalente en el caso de entradas nominales como *alud*, para la que se abre una subentrada como base que selecciona adjetivos (*arrollador*) o verbos (*sepultar*, *arrollar*) y otra subentrada para el colocativo seleccionado por otros sustantivos (*alud de nieve*, *tierra*, *críticas*, etc.).

3.3. *Diccionario de colocaciones del español (DiCE)*

El DiCE ha pasado por varias versiones desde que se puso en Internet en 2004 y está continuamente sometido a cambios tanto en contenido como en el modo de acceso a su información. A diferencia de otros diccionarios electrónicos que han partido de una versión en papel, el DiCE ya ha sido diseñado con una arquitectura que permite explotar al máximo las posibilidades de acceso a la información que da la versión electrónica. El usuario puede acceder de un modo tradicional, a través de la lista de lemas, cliqueando en [Consulta general](#), o puede hacer preguntas más específicas a través de Consultas avanzadas: 1) [Qué significa](#), orientada hacia la comprensión del significado de la colocación; 2) [Cómo se dice](#), dirigida hacia la producción de un colocativo correspondiente a un significado dado dependiendo de la base; 3) [Búsqueda por colocativo](#), que parte del colocativo para proporcionar las distintas bases con que se combina para expresar un sentido o diferentes sentidos; 4) [Búsqueda por FL](#), más dirigida a los especialistas ya que se puede preguntar por una función léxica (FL) en concreto aplicada a una base dada.¹⁷

3.3.1. Macroestructura del DiCE

Los lemas son nombres pertenecientes al campo semántico ‘sentimiento’. Es frecuente leer que el DiCE no está “acabado”, puesto que se “limita” al campo semántico de los nombres de sentimiento. Sin embargo, un diccionario, tal y como lo concibo, es necesariamente un proyecto en curso: los “acabados” son procesos más vinculados al mundo editorial y económico que el propiamente científico. La limitación a un campo semántico permite además mostrar las regularidades y las posibilidades de generalización que no serían posibles con una muestra de lemas más aleatoria.

Desde la perspectiva de la codificación, los nombres son productivos como base de colocaciones, pero no hay ninguna objeción teórica a tener como lemas verbos, adjetivos o adverbios. A diferencia de los diccionarios dirigidos por Bosque, en el DiCE un lema verbal, adjetival o adverbial sería siempre una base y, por lo tanto, en la entrada se proporcionarían sus colocativos, los adverbios que son las categorías que pueden modificar un verbo (*prohibir terminantemente*), un adjetivo (*perdidamente enamorado*) o un adverbio (*rematadamente mal*); con un lema adjetival incluiría además colocativos verbales (*quedar contento*). Esto no impide poder consultar la información tanto por la base como por el colocativo, de tal modo que se pueda examinar cuáles son las bases con las que, por ejemplo, [cumplir](#) es un colocativo.

3.3.2. Microestructura del DiCE

Siendo precisos, deberíamos decir que el DiCE no es exactamente un diccionario de colocaciones, sino de *correlatos léxicos*. La razón se debe a que la herramienta teórica utilizada son las FLL (‘frecuencias léxicas’) que permiten describir para una unidad léxica dada todas las elecciones paradigmáticas y sintagmáticas que dependen de elecciones léxicas hechas previamente. Así, por ejemplo, bajo la entrada de *alegría*,

incluimos el verbo *dar* (correlato sintagmático *alegría*, como en *dar una alegría* [a alguien]) y el verbo *alegrar* (su correlato paradigmático). Si queremos registrar en un diccionario las distintas maneras de expresar ‘causar una alegría’, debemos incluir tanto la colocación verbal como el verbo relacionado paradigmáticamente. Gracias a las FLL subyacentes, el DiCE ofrece una descripción semántica y sintáctica de las colocaciones, a diferencia de otros diccionarios 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.

Cada lema está asociado con una lista de unidades léxicas (UL). A partir de cada UL, el usuario puede consultar los dos tipos principales de información: 1) la información central, que consiste en información semántica y combinatoria; y 2) la información de uso que concierne la frecuencia y nivel del *Marco común europeo* (Consejo de Europa 2001). Vid. la siguiente figura con información de la entrada de *admiración* 1.

Unidades léxicas del lema: **admiración**

Desplegar todo Contraer todo

Etiqueta semántica | **Información combinatoria**

admiración 1 f. (Sentimiento) [ver ejemplos]
Juan siente admiración hacia su profesor por su capacidad de trabajo
Frecuencia notable
Nivel B2
Nivel MCER

Colocaciones
ver todas, atributo de los participantes, admiración + adjetivo, verbo + admiración, admiración + verbo, nombre de admiración

admiración 2a f. (Sentimiento) [ver ejemplos]
Ante la admiración del público, el mago hizo aparecer un conejo
Frecuencia moderada

Colocaciones
ver todas, atributo de los participantes, admiración + adjetivo, verbo + admiración, nombre de admiración

admiración 2b f. (Algo) [ver ejemplos]
El espectáculo fue la admiración de los asistentes
Frecuencia baja

Colocaciones
ver todas, admiración + adjetivo, verbo + admiración

Fig. 4. Información principal de la entrada de *admiración*

En cuanto a la información semántica, el artículo lexicográfico de cada UL proporciona: 1) la *etiqueta semántica* que representa el significado genérico; 2) la *forma proposicional*, en la que aparecen los participantes de la situación designada por el nombre; 3) ejemplos de corpus, la mayoría del *Corpus de Referencia del español actual*; y 4) los cuasi-sinónimos y cuasi-antónimos (con el número de la UL correspondiente).

La información combinatoria ofrecida es de dos tipos: sintáctica, por medio del *esquema de régimen*, y léxica, por medio de las funciones léxicas formateadas en glosas en lengua natural. Con el esquema de régimen se proporciona de forma sucinta la información sintáctica concerniente a la realización lingüística de los actantes. Así, el actante X de *admiración* 1 puede realizarse por medio de un sintagma preposicional (*la admiración de Juan*), por un adjetivo posesivo (*su admiración*) o por un adjetivo relacional (*la admiración nacional*). El actante Y se realiza siempre mediante un sintagma preposicional: *admiración a / hacia su profesora / por su trabajo*, etc. Como vemos, el DiCE es también un diccionario de régimen, pero distingue y separa los dos tipos de información combinatoria.¹⁸

» admiración : I

Esquemas de régimen

Actantes	Realizaciones
1 - X	de N [ver ejemplos] Apos [ver ejemplos] A [ver ejemplos]
2 - Y	a N [ver ejemplos] hacia N [ver ejemplos] por N [ver ejemplos] ante N [ver ejemplos]
3 - Z	por N [ver ejemplos] por Vinf [ver ejemplos]

1. su admiración

Fig. 5. Esquema de régimen de *admiración*

La combinatoria léxica se muestra en la sección “Colocaciones”. Véase la Fig. 6 del artículo lexicográfico de *admiración*:

ver todas, atributo de los participantes, admiración + adjetivo, verbo + admiración, admiración + verbo, nombre de admiración

Fig. 6. Información de combinatoria léxica de *admiración*

Como vemos, se pueden hacer cinco tipos de consulta. La primera necesita alguna explicación. Bajo *Atributo de los participantes* agrupamos atributos o nombres que refieren a los participantes de la situación designada; por ejemplo, *digno de admiración* o *admirable* refiere al participante Y, el objeto de la admiración. En el primer caso, hay una relación sintagmática con el nombre *admiración*, mientras que en el segundo, la relación es paradigmática. Obsérvese también que codificamos con diferente rúbrica cuando la base funciona como sujeto (UL+Verbo, *su admiración se dirige a [Y]*) o cuando funciona como objeto o primer complemento (Verbo+UL, *profesar admiración [hacia Y]*). Se trata obviamente de una convención para marcar esta diferencia sintáctica aunque, como sabemos, el sujeto en español pueda ir antes o después del verbo.¹⁹ Véase la siguiente figura con algunas colocaciones verbales de *admiración*:

Encontradas **40** colocaciones

admiración 1 (Sentimiento)

Colocaciones
 ver todas , atributo de los participantes, admiración + adjetivo, verbo + admiración, admiración + verbo, nombre de admiración

Desplegar todo
 Contraer todo
 Mostrar funciones léxicas
 Ordenar por frecuencia

<input type="checkbox"/> sentir ~ deber, dispensar, profesar, rendir, sentir, tener, tributar
<input type="checkbox"/> ser objeto de ~ gozar, tener
<input type="checkbox"/> continuar sintiendo ~ conservar
<input type="checkbox"/> dejar de ser objeto de la ~ de alguien perder
<input type="checkbox"/> causar que la ~ sea mayor aumentar 2
<input type="checkbox"/> causar ~ en alguien causar, despertar, infundir, inspirar, producir, provocar, suscitar
<input type="checkbox"/> causar ~ hacia sí atraer, atraerse, captar, concitar, conquistar, conseguir, cosechar, ganar , ganarse, granjearse, propiciarse
<input type="checkbox"/> causar que alguien sienta ~ llenar
<input type="checkbox"/> hacer que ~ no se note disimular, ocultar
<input type="checkbox"/> deber ser objeto de ~ merecer
<input type="checkbox"/> verbo cuasisinónimo admirar 1
<input type="checkbox"/> manifestar ~ confesar, demostrar, expresar, mostrar, testimoniar

Fig. 7. Colocaciones V+*admiración* 1

Pasemos ahora a la información de uso, empezando por la información de frecuencia. El contenido del DiCE fue derivado de corpus, principalmente del CREA. Sin embargo, a la hora de decidir qué combinaciones léxicas incluir, la frecuencia no fue tomada en cuenta. Como resultado, algunas combinaciones pueden ser consideradas representativas del español, mientras que otras lo son menos. Por esa razón, optamos por incluir información de frecuencia. La obtención de esa información fue una tarea ardua principalmente porque hemos asignado frecuencias a sentidos, no a palabras (*vid.* Vincze y Alonso-Ramos 2013a). Hemos optado por asignar un índice de frecuencia a cada UL que constituye la base de la colocación, mientras que asignamos un índice de frecuencia a la colocación entera ponderada por la frecuencia de la base. Cada UL recibe una etiqueta de frecuencia baja, moderada, notable o alta, siguiendo los parámetros de Almela *et al.* (2005). Los colocativos son presentados en el DiCE bajo la entrada léxica de la base según su tipo sintáctico y su significado, y ordenados por su frecuencia. De esta información, podemos derivar una lista de todas las colocaciones incluidas en el DiCE con información de frecuencia, de importancia vital tanto para propósitos didácticos como evaluativos. Véase la siguiente figura con la información de frecuencia en cada UL de *admiración*.

Desplegar todo Contraer todo

<input type="checkbox"/> admiración 1 f. (<i>Sentimiento</i>) [ver ejemplos]	Frecuencia notable
admiración de individuo X a individuo Y por hecho Z Ejemplos 1. Juan siente admiración hacia su profesor por su capacidad de trabajo Cuasisinónimos adoración 2, aprecio 1a, consideración III.1, devoción 2, entusiasmo 1, respeto I.1a, veneración 2 Cuasiantónimos aversión 1, desdén 1, desprecio 1 Ver esquema de régimen Colocaciones ver todas, atributo de los participantes, admiración + adjetivo, verbo + admiración, admiración + verbo, nombre de admiración	
<input checked="" type="checkbox"/> admiración 2a f. (<i>Sentimiento</i>) [ver ejemplos]	Frecuencia moderada
<input checked="" type="checkbox"/> admiración 2b f. (<i>Algo</i>) [ver ejemplos]	Frecuencia baja

Fig.8. Índice de frecuencia a cada UL

En la siguiente figura se observan algunos colocativos verbales de *admiración* ordenados por frecuencia, con la posibilidad de ordenarlos alfabéticamente.

Encontradas **40** colocaciones

admiración 1 (*Sentimiento*)

Colocaciones
 ver todas , atributo de los participantes, admiración + adjetivo, verbo + admiración, admiración + verbo, nombre de admiración

Desplegar todo Contraer todo Mostrar funciones léxicas **Ordenar alfabéticamente**

<input type="checkbox"/> sentir ~ tener, sentir, rendir, deber, tributar, profesar, dispensar
<input type="checkbox"/> ser objeto de ~ tener, gozar
<input type="checkbox"/> continuar sintiendo ~ conservar
<input type="checkbox"/> dejar de ser objeto de la ~ de alguien perder
<input type="checkbox"/> causar que la ~ sea mayor aumentar 2
<input type="checkbox"/> causar ~ en alguien despertar, causar, provocar, suscitar, producir, infundir, inspirar
<input type="checkbox"/> causar ~ hacia sí ganarse, conseguir, atraer, atraerse, granjearse, conquistar, cosechar, captar, concitar, ganar , propiciarse
<input type="checkbox"/> causar que alguien sienta ~ llenar
<input type="checkbox"/> hacer que ~ no se note ocultar, disimular
<input type="checkbox"/> deber ser objeto de ~ merecer
<input type="checkbox"/> verbo cuasisinónimo admirar 1
<input type="checkbox"/> manifestar ~ expresar, mostrar, demostrar, confesar, testimoniar

Fig. 9. Colocativos de *admiración* ordenados por frecuencia

En cuanto a la información del nivel del *Marco* (A1, A2, B1, etc.), hemos asignado un nivel tanto a las bases como a las colocaciones. Para ello, nos basamos principalmente en el vaciado de información en el *Plan curricular del Instituto Cervantes* (PCIC). Aunque no todas las UULL ni las colocaciones del DiCE aparecen en el PCIC, hemos elaborado un criterio de derivación basado en frecuencias. Establecemos una correlación entre el nivel asignado en el PCIC y su frecuencia en un corpus de referencia (esTenTen, Kilgarriff y Renau 2013). La metodología detallada puede encontrarse en García Salido y Alonso-Ramos (en prensa). En la siguiente figura, se muestra el nivel de diferentes colocativos del nombre *alegría*.

Encontradas 36 colocaciones

alegría 1a (Sentimiento)

Colocaciones
ver todas , atributo de los participantes, **alegría + adjetivo**, verbo + alegría, alegría + verbo, nombre de alegría

Desplegar todo Contraer todo **Mostrar funciones léxicas** Ordenar por frecuencia

intensa **a raudales, desbordante, enorme, extraordinaria, grande, impagable, impresionante, indecible, indescriptible, inefable, inmensa, intensa, loca, tremenda**

a raudales	Nivel C2
Glosa intensa	
Ejemplos 1. Alegría a raudales, que diría un cursi. 2. Y entonces nos pusimos a firmar papeles y papeles, mientras Matias, Paula y Gonso derrochaban alegría a raudales. (web)	
desbordante	Nivel C2
Glosa intensa	
Ejemplos 1. Opuesto al de la depresión, se caracteriza por la alegría desbordante 2. Observó con sorpresa que todos los detalles nimios de la vida diaria, le producían una alegría desbordante 3. A veces cuando muchos nos ven pasar por los caminos rocieros, con nuestra alegría desbordante, con nuestros cantos y rezos a la virgen	
enorme	Nivel B2
Glosa intensa	
Ejemplos 1. hubo música, canciones, fuegos artificiales, rayos láser y globos en un ambiente informal y de enorme alegría.	

Fig. 10. Nivel del *Marco* para las colocaciones

3.4. Síntesis comparativa de los diccionarios objeto de estudio

Intentaremos sintetizar en los siguientes aspectos las principales semejanzas y diferencias entre *Redes* y el *Práctico*, de un lado, y el DiCE, de otro.

1) Objetivo principal

El objetivo principal de *Redes* es listar las clases léxicas de argumentos seleccionados por un predicado. Por ejemplo, las clases de nombres que se combinan con el verbo *albergar*: i) sustantivos de significación intencional como *esperanza*, *deseo*, etc.; ii) sustantivos que designan estados y sensaciones de incertidumbre como *duda*, *sospecha*, etc.

El objetivo principal de *Práctico* es ofrecer combinaciones léxicas sin priorizar la dirección de la relación. De ese modo, en la entrada de *coger*, aparecen tanto sus propios colocativos como las bases cuando el verbo funciona como colocativo.

El objetivo principal del DiCE es proporcionar las UL que son seleccionadas de un modo restringido por otra UL previamente elegida. Así, tiene como tarea registrar cuáles son los verbos que, *grosso modo*, significan ‘tener’ cuando se habla de *esperanza*, de *duda* o de *sospecha*: entre otros verbos, estará el verbo *albergar*, pero también *abrigar* y *tener*.

2) Lemas

Los lemas de las entradas principales de *Redes* (las *entradas analíticas*) son verbos, adjetivos o preposiciones, y unos pocos nombres que funcionan también como colocativo (como *sarta*, *oleada*, etc.). Así, por ejemplo, *albergar* será un lema de *Redes*, pero *esperanza* solo será el lema de las *entradas abreviadas*, producidas automáticamente como rastreo de la base de datos de las entradas analíticas.

Los lemas de las entradas de *Práctico* (las *entradas simples*) son tanto nombres como verbos, adjetivos y preposiciones.

Los lemas del DiCE son las UL que seleccionan restringidamente a sus “acompañantes”, los colocativos. Por el momento, el DiCE solo incluye lemas nominales, pero no hay objeción teórica a abrir entradas para verbos, adjetivos o adverbios. Así, el verbo *cumplir* también podría ser el lema, pero en tanto que base de una colocación: *cumplir*_{Base} a *rajatabla*_{Colocativo} y no en tanto que colocativo (*cumplir*_{Colocativo} la *ley*_{Base}). Sí pueden recuperarse automáticamente, en cambio, todas las bases que seleccionan el colocativo *cumplir* (*ley*, *sueño*, *promesa*, *deseo*, etc.).

3) Fundamentación

La fundamentación en la que se basa *Redes* en la distinción entre predicados y argumentos es esencialmente semántica. Por esa razón, se puede considerar que *Redes* es un diccionario de restricciones semánticas, más que de colocaciones.

La fundamentación en la que se basa *Práctico* es también esencialmente semántica pero, dado que no se mencionan las clases léxicas, aparece menos visible.

La fundamentación en la que se basa el DiCE es de naturaleza esencialmente léxica. Por ejemplo, semánticamente el predicado ‘incumplir’ predica algo de ‘complot’, pero el hecho de que ese predicado se realice por *destapar* es algo léxico, no semántico: es el nombre *complot* el que selecciona léxicamente ese verbo; si se tratara del nombre *promesa*, la selección léxica sería otra, *romper*.

Para poder valorar la eficacia de estos diccionarios es necesario someterlos a test de usabilidad. No conocemos ningún test hecho a *Redes*, pero sí hicimos una comparación del *Práctico* y el DiCE con aprendices de español (Alonso-Ramos 2008) en donde el DiCE tuvo mejores resultados debido principalmente a la presencia de ejemplos. Además sometimos al DiCE a un test de usabilidad con traductores, profesores y estudiantes (Vincze y Alonso-Ramos 2013b). Los resultados de dicha prueba mostraron que la calidad global de la interacción entre usuarios y diccionario era buena, aunque también se identificaron áreas de mejora. En ese test ya se trasluce una tendencia general de los usuarios: quieren una herramienta subyacente al proceso de escritura, que no tengan que aprender a utilizar. Volveremos sobre esto en la siguiente sección.

Después de analizar los diccionarios combinatorios actuales, nos gustaría abordar las perspectivas futuras.

4. Hacia un nuevo concepto de diccionario combinatorio: herramienta de ayuda a la escritura

Es necesario preguntarse cómo debe ser el recurso ideal que sirva a un usuario a resolver sus necesidades colocacionales. El concepto de diccionario está cambiando hacia una herramienta más flexible y dinámica que se adapte mejor a las necesidades de los usuarios (Verlinde y Peeters 2012). La lexicografía guiada por el corpus (*corpus-driven lexicography*) ha dado nacimiento a lo que podríamos llamar *corpus guiados léxicamente*; es decir, recursos que en lugar de proporcionar información léxica en la forma de un diccionario lo hacen en forma de un programa de concordancias explotando un corpus. De este modo, los límites entre diccionarios y corpus empiezan a desdibujarse (*vid.* Alonso-Ramos 2009). En el mismo sentido, ya Cobb (2003) proponía una mezcla de diccionario y concordancias como el recurso ideal. Igualmente, Kilgarriff (2009) ofrecía sus *word sketches* como algo a medio camino entre corpus y diccionario: “output from the corpus, making it corpus-like, but it is a condensed summary of what was found there, making it dictionary-like.” Independientemente de cómo nos refiramos a estos nuevos recursos o herramientas, el hecho es que los diccionarios han dejado de ser productos independientes y tienden a integrarse con otros recursos como los corpus y otros diccionarios, como parte de herramientas de ayuda a la redacción (Wanner *et al.* 2013; Granger y Paquot 2015). También sirven para complementar y a su vez son complementados por aplicaciones de enseñanza de lenguas asistida por ordenador (Abel 2010).

Conforme a esta línea de pensamiento, estamos construyendo una herramienta de ayuda a la redacción, centrada en las colocaciones (HARenEs,²⁰ Alonso-Ramos 2016, Alonso-Ramos *et al.* 2014). Esta herramienta integra el DiCE, pero también un diccionario personalizado del usuario, así como la consulta directa al corpus sin pasar por el diccionario. Aunque todavía no está completamente operativa, su arquitectura está ya diseñada. Véase la siguiente figura con sus principales funcionalidades.



Fig. 11. Pantalla inicial de HARENES

El entorno consta de tres componentes principales: 1) el verificador de colocaciones, que propone sugerencias en caso de que la colocación consultada sea incorrecta; 2) la búsqueda de colocativos de una base dada, por medio principalmente de medidas estadísticas en el corpus; 3) el diccionario de colocaciones personalizado, en donde el usuario puede guardar las colocaciones que le interesen con sus ejemplos, así como el registro de los errores y sus soluciones. En la funcionalidad “comprobar”, el usuario verifica la corrección de una colocación específica y, en el caso de una combinación incorrecta, la herramienta propone sugerencias junto con ejemplos. Véase la figura siguiente en donde el usuario ha pedido comprobar *tomar un paseo*.



Fig. 12. Sugerencias de corrección proporcionadas por HARENES

El segundo componente, el buscador, es similar a otras herramientas léxicas que usan corpus.²¹ En estas herramientas, el usuario obtiene los colocativos de la palabra que

consulta según su estructura sintáctica (V+N, Adj+N, etc.). Sin embargo, nuestro objetivo es poder obtener los datos también a través de una tipología semántica. Por ejemplo, poder obtener los colocativos que expresan la fase inicial de *miedo*, independientemente de su estructura sintáctica: *coger*, *entrar*, *asaltar*, *invadir*, etc. (vid. Vincze 2015).

HARenEs ofrece también la posibilidad de administrar recursos personales para que el usuario pueda almacenar las colocaciones que aprende, los errores frecuentes que comete junto con sus correcciones, ejemplos, etc. Vid. la siguiente figura con un ejemplo de un diccionario personalizado.



Fig 13. Un ejemplo de diccionario personalizado

La herramienta todavía necesita ser mejorada. Por ejemplo, el orden de las sugerencias no es el mejor y la identificación de las colocaciones no es siempre correcta. Con todo, no dudamos que herramientas como HARenEs son el camino a seguir y lo esperado por los usuarios, que cada vez piden más diccionarios en línea integrados en su entorno de trabajo.

5. Conclusión

Nos gustaría dar algunas respuestas aclaratorias dirigidas especialmente al usuario novel de un diccionario combinatorio del español. Pensamos que los diccionarios combinatorios son poco conocidos popularmente a pesar de su gran potencial. Cada vez que alguien busca una palabra precisa para expresar un significado dado cuyo control depende de una elección léxica previa, está necesitando un diccionario combinatorio (aunque no lo sepa). Así, al escribir, por ejemplo, *el intenso debate fue sobre...*, podemos sentir que el estilo mejoraría si sustituimos *intenso* por *acalorado* y el verbo *ser* por *girar*. Estas respuestas las puede dar un diccionario combinatorio en la entrada de *debate*. La facilidad con que se encuentre esas respuestas en el diccionario es otra cuestión que debe medirse con test de usabilidad, como anunciamos arriba. Lo que aquí importa ahora es que para ese tipo de dudas es para lo que se suele consultar un diccionario combinatorio: para producir una combinación léxica restringida o una colocación. Es posible también consultar qué significa una colocación, aunque no es frecuente que los diccionarios de colocaciones den información semántica de la colocación.

En cambio, una pregunta de otro cariz, más metalingüística, es qué cosas pueden *girar* en sentido figurado. Para responder este tipo de preguntas se diseñó *Redes*. Obsérvese que la respuesta a esta pregunta no es una combinación, sino agrupaciones de posibles combinaciones, las diferentes clases léxicas que aparecen en la entrada de *girar*. No se trata, por tanto, de una consulta que sirva para satisfacer la necesidad de producir una combinación dada sino, más bien, una curiosidad intelectual. La respuesta a esta pregunta hará que sepamos más sobre la semántica del verbo *girar*, pero el usuario que quiera utilizar ese verbo debe empezar por elegir uno de los nombres mencionados en las diferentes clases, como *debate*, con lo que de nuevo volvemos a la entrada del nombre.

La idea de que se puedan agrupar en clases diferentes las bases ha hecho poner en cuestión la diferencia entre una combinación libre y una combinación restringida o colocación. Creemos haber mostrado que esa diferencia debe seguir manteniéndose, basándonos principalmente en la noción de control léxico ejercido por la base y en la mayor concordancia semántica entre los componentes de una colocación. Así, mientras que en *destapar una alianza*, la elección del verbo está guiada solo semánticamente, en *destapar un complot*, la elección está guiada también léxicamente. Igualmente, mientras que en la combinación libre no hay una fuerte intersección de componentes semánticos comunes, la concordancia semántica de la colocación es grande, puesto que se comparte el componente ‘secreto’.

Es probable que no *haya desvelado* ningún *secreto* a muchos lectores ni espero *haber violado* ninguna *ley*, pero sí me gustaría *hacer un llamamiento*: dejemos la discusión teórica y trabajemos más. Hacen falta más diccionarios combinatorios con distintos formatos y destinatarios, monolingües y bilingües, en papel y electrónicos, con mediciones de su usabilidad, con integración de actividades didácticas. Para ello, son necesarios equipos interdisciplinares y mucho tesón, como en cualquier obra (lexicográfica).

Agradecimientos

Este trabajo da muestra de las actividades realizadas dentro del proyecto de investigación FFI2016-78299-P financiado por el MINECO y por fondos FEDER. Agradezco la atenta lectura de Marcos García Salido y su ayuda en la edición de figuras, así como la revisión de los editores.

Referencias bibliográficas

Abel, Andrea. 2010. Towards a systematic classification framework for dictionaries and CALL. En S. Granger y M. Paquot, eds. *eLexicography in the 21st century: New Challenges, New Applications. Proceedings of eLex 2009*. Louvain-la-Neuve: Cahiers du Cental 7, Presses universitaires de Louvain, pp. 3-11.

Almela, Ramón; Cantos, Pascual; Sánchez, Aquilino; Sarmiento, Ramón; Almela, Moisés. 2005. *Frecuencias del español: Diccionario y estudios léxicos y morfológicos*. Madrid: Universitas.

Alonso-Ramos, Margarita 1993. *Las funciones léxicas en el modelo lexicográfico de I. Mel'čuk*, tesis doctoral, Madrid: UNED.

Alonso-Ramos, Margarita. 2004. *Diccionario de colocaciones del español* (DICE).

Disponible en: <<http://www.dicesp.com>>

Alonso-Ramos, Margarita. 2007. Actantes y colocaciones. *Nueva revista de filología hispánica* 55.2: 435-458.

Alonso-Ramos, Margarita. 2008. Papel de los diccionarios de colocaciones en la enseñanza de español como L2. En E. Bernal y J. De Cesaris, eds. *Proceedings of the Thirteenth EURALEX International Congress*. Barcelona: IULA, Universitat Pompeu Fabra, pp. 1215-1230.

Alonso-Ramos, Margarita. 2009. Hacia un nuevo recurso léxico: ¿fusión entre corpus y diccionario? En P. Cantos Gómez y A. Sánchez Pérez, eds. *A Survey of Corpus-based Research. Panorama de investigaciones basadas en corpus*. Murcia: AELINCO, pp. 1191-1207.

Alonso-Ramos, Margarita. 2010. No importa si la llamas o no colocación, descríbela. En C. Mellado, P. Buján, C. Herrero, N. Iglesias y A. Mansilla, eds. *La fraseografía del s. XXI. Nuevas perspectivas para el español y el alemán*. Berlin: Frank & Timme, pp. 55-80.

Alonso-Ramos, Margarita. 2015. Discovering hidden collocations in a bilingual Spanish-English dictionary. En I. Kosem, M. Jakubicek, J. Kallas y S. Krek, eds. *Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference*, Trojina-Institute for Applied Slovene Studies/Lexical Computing Ltd., Ljubljana/Brighton, pp. 170-185.

Alonso-Ramos, Margarita. 2016. Learning resources for Spanish collocations: From a dictionary towards a writing assistant. En B. Sanromán Vilas, ed. *Collocations Cross-Linguistically. Corpora, Dictionaries and Language Teaching*, volume C of Mémoires de la Société Néophilologique de Helsinki, [Société Néophilologique de Helsinki](http://www.societe-neophilologique.de/helsinki/), Helsinki, pp. 65-95.

Alonso-Ramos, Margarita. 2017. Can collocations be deduced? A lexically-driven analysis from the perspective of language production. En S. Torner y E. Bernal, eds. *Collocations and other lexical combinations in Spanish. Theoretical, lexicographical and applied approaches*. London/New York: Routledge, pp. 21-40.

Alonso-Ramos, Margarita; García-Salido, Marcos; Vincze, Orsolya. 2014. Towards a collocation writing assistant for learners of Spanish. En G. Faaß y J. Ruppenhofer, eds. *Workshop Proceedings of the 12th Edition of the Konvens Conference*. Hildesheim: Universitätsverlag Hildesheim, pp. 77-88.

Apresjan, J. 2009. The Theory of Lexical Functions: An Update. En *Proceedings of the Fourth International Conference on MTT*. Montreal: OLST, pp. 1-14. Disponible en: <<http://mtt.upf.edu/mtt2009/ProceedingsMTT09.pdf>>

Apresjan, Jurij D.; Glovinskaja, Marina Ja. 2007. Two projects: English ECD and Russian production dictionary. En *Proceedings of the Third International Conference on MTT*. München: Wiener Slawistischer Almanach. Disponible en: <http://meaningtext.net/mtt2007/proceedings/03Apresjan_GlovinskajaFinal.pdfv>

Barrios Rodríguez, M.^a Auxiliadora. 2007. Diccionarios combinatorios del español: diferencias y semejanzas entre *Redes* y *Práctico*. *redELE revista electrónica de didáctica / español lengua extranjera*, 11.

Beauchesne, Jacques. 2001. *Dictionnaire des cooccurrences*. Montréal: Guérin.

- Benson, Morton; Benson, Evelyn; Ilson, Robert. 1986. *The BBI Combinatory Dictionary of English. A Guide to Word Combinations*. Amsterdam/Philadelphia: John Benjamins.
- Boneu, Javier. 2001. *Diccionario euléxico. Para expresarse con estilo y rigor*. Barcelona: Editorial Juventud.
- Bosque, Ignacio, dir. 2004. *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: SM.
- Bosque, Ignacio, dir. 2006. *Diccionario combinatorio práctico del español contemporáneo*. Madrid: SM.
- Bosque, Ignacio. 2003. La direccionalidad en los diccionarios combinatorios y el problema de la selección léxica. En T. Cabré, ed. *Lingüística Teòrica: Anàlisi i perspectives I*, Catalan Journal of Linguistics Monographies, pp. 13-58. Disponible en: <<https://sites.google.com/site/ignaciobosquemunoz/publicaciones-y-presentaciones/2004>>
- Bosque, Ignacio. 2004a. Combinatoria y significación. Algunas reflexiones. En I. Bosque, dir. *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: SM, pp. LXXVII-CLXXIV.
- Bosque, Ignacio. 2011. Deducing Collocations. En L. Boguslavsky y L. Wanner, eds. *Proceedings of the 5th International Conference on Meaning-Text Theory*. Barcelona: UPF, pp. vi-xxiii. Disponible en: <<http://olst.ling.umontreal.ca/pdf/proceedingsMTT2011.pdf>>
- Bosque, Ignacio. 2017. On the conceptual bases of collocations: restricted adverbs and lexical selection. En S. Torner y E. Bernal, eds. *Collocations and other lexical combinations in Spanish. Theoretical, lexicographical and applied approaches*. London/New York: Routledge, pp. 21-40.
- Buendía, Miriam; Faber, Pamela. 2014. Collocation dictionaries: A comparative analysis. *MonTi: Monografías de Traducción e Interpretación* 6: 203-235.
- Capra, Daniela. 2016. What do we talk about when we talk about collocation in Spanish? En A. Orlandi y L. Giacomini, eds., pp. 151-173.
- Charest, S.; Fontaine, J.; Saint-Germain, J. 2012. *Le Grand Druide des cooccurrences*. Montréal : Editions Druide.
- Cobb, Tom. 2003. Do corpus-based electronic dictionaries replace concordancers? En B. Morrison, G. Green y G. Motteram, eds. *Directions in CALL: Experience, experiments, evaluation*. Hong Kong: Polytechnic University, pp. 179-206.
- Consejo de Europa. 2001. *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: MEC y Anaya. Disponible en: <http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/>
- Corpas, Gloria. 2016. Collocation dictionaries for English and Spanish: the state of the art. En A. Orlandi y L. Giacomini, eds., pp. 175-210.
- Corpas, Gloria. 2017. Collocations in e-bilingual dictionaries: from underlying theoretical assumptions to practical lexicography and translation issues. En S. Torner y E. Bernal, eds. *Collocations and other lexical combinations in Spanish. Theoretical and Applied approaches*. London: Routledge.

Crowther, Jonathan; Dignen, Sheila; Lea, Diana, eds. 2002. *Oxford Collocations Dictionary for Students of English* [OCD]. Oxford: Oxford University Press.

Ferrando, Verónica. 2013. El tratamiento de las colocaciones en la lexicografía española y alemana: estudio contrastivo. *Revista Internacional de Lenguas Extranjeras* 2: 31-53.

García Salido, Marcos; Alonso-Ramos, Margarita. (En prensa). Asignación de niveles de aprendizaje al *Diccionario de Colocaciones del Español*. *Signos*.

Granger, Sylviane; Paquot, Magali. 2015. Electronic Lexicography goes local: Design and structures of a needs-driven online academic writing aid. *Lexicographica* 31.1: 118-141.

Hausmann, Franz Josef. 1998. O diccionario de colocacións. Criterios de organización. En X. Ferro Ruibal, coord. *Actas do I Coloquio Galego de Fraseoloxía (1997)*. Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades, Xunta de Galicia, pp. 63-81.

Hausmann, Franz Josef. 1999. Le dictionnaire de collocations. Critères de son organisation. En N. Greiner, J. Kornelius y G. Rovere, eds. *Texte und Kontexte in Sprachen und Kulturen. Festschrift für Jörn Albrecht*. Trier: Wissenschaftlicher Verlag Trier, pp. 121-139.

Higueras García, Marta. 2006. El diccionario *Práctico* en la práctica docente del español como lengua extranjera. En *Actas del II Congreso Internacional de Lexicografía Hispánica*. Alicante, pp. 436-443.

Higueras García, Marta. 2011. Lexical collocations and the learning of Spanish as a foreign language. En J. L. Cifuentes Honrubia y S. Rodríguez Rosique, eds. *Spanish Word Formation and Lexical Creation*. Amsterdam/Filadelfia: John Benjamins, pp. 439-464.

Instituto Cervantes. 1997-2017. *Plan Curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid: Biblioteca nueva. Disponible en: <http://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/>

Kilgarriff, Adam. 2009. Corpora in the classroom without scaring the students. En *Proceedings from the 18th International Symposium on English Teaching*. Taipei. Disponible en: <www.kilgarriff.co.uk/Publications/2009-K-ETA-Taiwan-scaring.doc>

Kilgarriff, Adam; Renau, I. 2013. esTenTen, a Vast Web Corpus of Peninsular and American Spanish. *Procedia-Social and Behavioural Sciences* 95.25: 12-19.

Le Fur, Dominique. 2007. *Dictionnaire des combinaisons de mots*. Paris: Le Robert.

Lo Cascio, Vincenzo. 2012. *Dizionario combinatorio compatto italiano*. Amsterdam/Philadelphia: John Benjamins.

Makkai, Adam. 1972. *Idiom Structure in English*. The Hague: Mouton.

Mayor, Michael. 2013. *Longman Collocations Dictionary and Thesaurus*. Harlow: Pearson/Longman.

McGee, Iain. 2012. Collocation Dictionaries as Inductive Learning Resources in Data-driven Learning: An Analysis and Evaluation. *International Journal of Lexicography* 25.3: 319-361.

Mel'čuk, Igor. 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.
- Mel'čuk, Igor. 2012. Phraseology in the Language, in the Dictionary, and in the Computer. *Yearbook of Phraseology* 3.1: 31-56.
- Mel'čuk, Igor. 2015. Clichés, an understudied class of phrasemes. *Yearbook of Phraseology* 6.1: 35-54.
- Mel'čuk, Igor; Polguère, Alain. 2007. *Lexique actif du français*. Louvain-la-Neuve: De Boeck.
- Orlandi, Adriana. 2016. Monolingual collocation lexicography: State of art and new perspectives. En A. Orlandi y L. Giacomini, eds., pp. 19-70.
- Orlandi, Adriana; Giacomini, Laura, eds. 2016. *Defining Collocation for Lexicographic Purposes*. Bern: Peter Lang.
- Rundell, Michael, ed. 2010. *Macmillan Collocations Dictionary for Learners of English*. Oxford: Macmillan Publishers Ltd.
- Russo, Domenico. 2010. *Modi di dire. Lessico italiano delle collocazioni*. Roma: Aracne.
- Sanromán Vilas, Begoña. 2016. Searching and understanding collocations to improve dictionaries and language teaching: A tale in eight languages. En B. Sanromán Vilas, ed. *Collocations Cross-Linguistically. Corpora, Dictionaries and Language Teaching*, volume C of Mémoires de la Société Néophilologique de Helsinki, [Société Néophilologique de Helsinki](http://www.societe-neophilologique.de/helsinki/), Helsinki, pp. 9-64.
- Sanromán Vilas, Begoña. 2011. The Unbearable Lightness of Light Verbs. En I. Boguslavsky y L. Wanner, eds. *Proceedings of the 5th International Conference on Meaning-Text Theory*. Barcelona: Universitat Pompeu Fabra, pp. 253-263.
- Serra, Susana. 2009. Las restricciones de selección en los diccionarios generales de lengua española. *Boletín de Filología* XLIV.2: 187-213.
- Tiberii, Paola. 2012. *Dizionario delle collocazioni. Le combinazioni delle parole in italiano*. Bologna: Zanichelli.
- Urzi, Francesco. 2009. *Dizionario delle combinazioni lessicali*. Luxembourg: Convivium.
- Verlinde, Serge; Peeters, Geert. 2012. Data access revisited: The Interactive Language Toolbox. En S. Granger y M. Paquot, eds. *Electronic lexicography*. Oxford: Oxford University Press, pp. 147-162.
- Vincze, Orsolya. 2015. *Learning multiword expressions from corpora and dictionaries*, tesis doctoral, A Coruña: Universidade da Coruña.
- Vincze, Orsolya; Alonso-Ramos, Margarita. 2013a. Incorporating frequency information in a collocation dictionary: Establishing a methodology. *Procedia - Social and Behavioral Sciences* 96: 241-248.
- Vincze, Orsolya; Alonso-Ramos, Margarita. 2013b. Testing an electronic collocation dictionary interface: Diccionario de Colocaciones del Español. En I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets y M. Tuulik, eds. *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference*. Ljubljana/Tallinn, Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut, pp. 328-337.

Wanner, Leo; Verlinde, Serge; Alonso-Ramos, Margarita. 2013. Writing assistants and automatic lexical error correction: word combinatorics. En I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets y M. Tuulik, eds. *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference*, Ljubljana/Tallinn: Institute for Applied Slovene Studies/Eesti Keele Instituut, pp. 472-487.

Wu, Shaoqun; Witten, Ian H; Franken, Margaret. 2010. Utilizing lexical data from a web-derived corpus to expand productive collocation knowledge. *ReCALL* 22.1: 83-102.

Notas

- ¹ Sin ser exhaustivos, podemos listar los siguientes diccionarios combinatorios: 1) para el inglés, el BBI (Benson *et al.* 1906), el OCD (Crowther *et al.* 2002), el McMillan (Rundell 2010), el Longman (Mayor 2013); 2) para el francés, el Robert (Le Fur 2007), el Beauchesne (2001), el Grand Druide (Charest *et al.*, 2012), el *Lexique actif* (Mel'čuk y Polguère 2007); 3) para el italiano, Urzi (2009), Russo (2010), Lo Cascio (2012), Tiberii (2012). En Sanromán (2016), así como en Orlandi (2016) se puede encontrar una presentación del estado de la cuestión sobre diccionarios monolingües de colocaciones de las principales lenguas europeas. No hay, sin embargo, mucha riqueza en diccionarios bilingües de colocaciones (*vid.* Alonso-Ramos 2015 y Corpas 2017).
- ² Existe una obra previa, el *Diccionario euléxico* (Boneu 2001) que proporciona principalmente colocativos verbales y adjetivales de 1800 nombres, aunque en ningún momento se menciona el término *colocación*. A pesar de su innegable interés, hay que reprocharle la ausencia de ejemplos, así como la falta de distinción entre las distintas acepciones de los nombres.
- ³ Afirmaciones como estas dan muestra: “las ‘colocaciones’, si es que realmente existen como fenómeno particular en las lenguas [...]” (Serra 2009: 202) o “el concepto de colocación, entendido como ‘combinación habitualizada de *dos* unidades léxicas’, parece resultar inoficioso” (Serra 2009: 210).
- ⁴ Por lo que sabemos, la primera referencia en la bibliografía española donde se usa el término *colocación* es Alonso-Ramos (1993), en donde se repasa el tratamiento de nociones afines a las colocaciones como las solidaridades léxicas o el contorno de la definición. Para una revisión actual del tratamiento de todas estas nociones en la bibliografía española, *vid.* Capra (2016).
- ⁵ En la lista de referencias sobre colocaciones compilada por Kazumi Koike hay alrededor de 2000 publicaciones entre 2000 y 2012. Puede consultarse en: <<http://collocate.blogspot.com.es/p/bibliografia.html>>
- ⁶ Barrios (2007: 2) señala que el *Práctico* no es un diccionario de colocaciones, puesto que no incluye combinaciones frecuentes como *buen libro* o *libro genial*, a diferencia del OCD que sí incluye *great* y *good* en la entrada de *book*, “con el consiguiente peligro de que el concepto de colocación se vacíe de contenido”.
- ⁷ Aunque los términos *base* y *colocativo* fueron introducidos por Hausmann, no difieren de cómo se usan en la Lexicología explicativa y combinatoria. Para Hausmann (1998), las bases de las colocaciones son *autosemánticas*, es decir semánticamente autosuficientes, mientras que los colocativos son *sinsemánticos*, porque dependen semánticamente de la base.
- ⁸ Una definición aproximada de *complot* puede ser: ‘plan **secreto** hecho por un grupo de personas para hacer algo ilegal’.
- ⁹ Makkai (1972) los llamó *idioms of encoding*.

-
- ¹⁰ Apresjan y Glovinskaja (2007: 6) la definen como “recurrence of certain components in the meaning of two collocating items”.
- ¹¹ Una definición aproximada de este sentido de *destapar* podría ser la siguiente:
X destapa Y= X hace público un hecho secreto Y en el que los implicados en Y no quieren que sea público.
- ¹² Para una evaluación de los diccionarios combinatorios del español, en ocasiones en comparación con diccionarios de otras lenguas, puede consultarse: Ferrando (2013), Buendía y Faber (2014) o Corpas (2016).
- ¹³ No coincidimos por tanto con Serra (2009) ni con Higuera (2011) cuando se afirma que *Redes* es un diccionario onomasiológico. Solo cumple ese papel en las entradas abreviadas, que son marginales.
- ¹⁴ Se puede encontrar una reflexión sobre la distinción entre los diccionarios de colocaciones y *Redes* en Bosque (2004a: XXII) y la diferencia entre selección semántica y selección léxica en Alonso-Ramos (2007).
- ¹⁵ Vid. Barrios (2007) para una comparación de los dos diccionarios.
- ¹⁶ Vid. Higuera (2006) para una exploración de cómo se puede explotar este diccionario en la enseñanza de español como lengua extranjera. Asimismo, véase Alonso-Ramos (2008) para un estudio experimental con aprendices de español sobre la usabilidad de este diccionario y del DiCE.
- ¹⁷ Aunque la noción de FL es clave dentro de los diccionarios enmarcados en la Lexicología explicativa y combinatoria, baste decir aquí que se trata de una herramienta que además de describir semántica y sintácticamente las colocaciones, modeliza el control léxico ejercido por la base. En una FL, la base es su argumento y el colcativo, su valor. Así, la misma FL cuando se aplica a diferentes nombres produce diferentes valores. Vid. Mel'čuk 1996 para más detalle.
- ¹⁸ Algunos autores tratan la relación de régimen entre un núcleo léxico y una preposición regida como un tipo de colocación, las *colocaciones gramaticales* (entre otros, Benson *et al.* 1986). Pensamos que este término solo puede dar lugar a confusión.
- ¹⁹ En cambio, en el *Práctico* (Bosque 2006), no se distingue el papel sintáctico desempeñado por el nombre en una colocación verbal, lo que puede dar lugar a confusión, especialmente a aprendices de español.
- ²⁰ HARenES fue desarrollada conjuntamente por el equipo de la Universidad Pompeu Fabra, dirigido por L. Wanner, y por mi equipo, en el marco del proyecto de investigación del Ministerio FFI2011-30219-C02-01.
- ²¹ Vid. por ejemplo el componente *Learning collocations* component de FLAX (Wu *et al.* 2010), o la interfaz *For better English* (<<http://forbetterenglish.com>>) o la utilidad *Combinations* utility de *Just the word* (<<http://www.just-the-word.com>>).