

IMPLICACIONES DE LA FRECUENCIA DE USO DE LAS LOCUCIONES EN LA ELABORACIÓN DE UN DICCIONARIO¹

INMACULADA PENADÉS MARTÍNEZ
Universidad de Alcalá
inmaculada.penades@uah.es

Recibido: 26/05/2015

Aceptado: 03/06/2015

Resumen

En este artículo se reflexiona, en primer lugar, sobre la frecuencia de las locuciones en relación con los límites que los investigadores establecen para el periodo de tiempo calificado como *actual*. En segundo lugar, se muestra cómo, aun limitándose a un corpus determinado, la situación de los estudios en fraseología, a día de hoy, vuelve extremadamente compleja la tarea de indicar la frecuencia absoluta de una locución concreta respecto a un conjunto de locuciones o la relativa en función de la distribución de las locuciones en clases o en relación con las unidades léxicas simples con las que la locución se vincula. En tercer lugar, el análisis de un amplio conjunto de locuciones ya llevado a cabo ha mostrado que, en la frecuencia de uso de una locución, influye su pertenencia a un determinado estado de lengua, la existencia de variantes no marcadas de la locución en cuestión y la posibilidad de que se dé una combinación de palabras homónima de la locución, pero de naturaleza no fraseológica. Los tres puntos mencionados se examinan para establecer sus implicaciones en la redacción de un diccionario de locuciones idiomáticas del español.

PALABRAS CLAVE: fraseología, lexicografía, locuciones, frecuencia.

Abstract

In the first place, this article is a reflexion on the frequency of idioms in relation to the limits established by researchers for the period of time which can be defined as the present. Secondly, we have shown how, despite having restricted the study to a specific corpus, the situation of phraseology studies today makes it extremely difficult to indicate the absolute frequency of a certain idiom with regard to a set of idioms or its relative frequency according to the distribution of idioms in classes or in relation to the simple lexical units with which the idiom is linked. Thirdly, the previous analysis of a broad set of idioms has shown that the frequency of use of an idiom is influenced by its belonging to a certain period of the language and also by the existence of unmarked variants of the idiom in question and the possible existence of a group of words that is homonymous with the idiom but not phraseological. The three points mentioned have been examined in order to establish their implication in the compilation of a dictionary of Spanish idiomatic expressions.

KEYWORDS: phraseology, lexicography, idioms, frequency.

1. Introducción

La elaboración de un diccionario de locuciones² idiomáticas del español³ ha conducido a plantearse cuáles deben formar parte de su macroestructura y cuáles deben excluirse de ella, en

1 Una primera versión muy reducida de este trabajo se presentó, bajo el título «Factores que influyen en la frecuencia de uso de las locuciones», en el XI Congreso Internacional de Lingüística General, celebrado en Pamplona del 21 al 23 de mayo de 2014.

2 Por locución se entiende la «combinación fija de palabras que funciona como elemento de la oración y cuyo significado no se corresponde con la suma del significado de sus componentes» (Penadés Martínez, 2012b: 23).

3 El DILEA (Penadés Martínez, 2008; Penadés Martínez, 2010).

función de su carácter de actuales o no actuales. Por otra parte, se ha hecho evidente, asimismo, la necesidad de marcar aquellas que presentan tanto un alto grado de frecuencia como un bajo grado de esta misma particularidad en relación con el estado de lengua presente. Desde este punto de partida, la búsqueda y el registro de ejemplos de uso para varios miles de locuciones han puesto de manifiesto diversas cuestiones que constituyen el objeto de estudio de este artículo. En primer lugar, se reflexiona (apartado 2.) sobre la relación entre la frecuencia de las locuciones y la periodización que se establece para la lengua, pues los límites que los investigadores fijan para el estado de lengua actual no son homogéneos ni fruto de un acuerdo consensuado. Esta situación conlleva la posibilidad de extender la sincronía presente más allá de los últimos cincuenta años, y de ahí la justificación del registro de locuciones arcaicas o antiguas en un diccionario fraseológico del español correspondiente al periodo actual, pero, al mismo tiempo, surge la obligación de marcarlas en relación con su baja utilización.

En segundo lugar, se muestra (apartado 3.) cómo la marcación de la frecuencia de las locuciones en un diccionario se ve dificultada por cuestiones todavía pendientes de resolución en fraseología, como la pertenencia de determinadas unidades fraseológicas a una u otra clase, pues son muchos todavía los desacuerdos existentes en relación con unidades concretas, de manera que las de clasificación dudosa no pueden ponerse en relación de frecuencia con las de una clase específica ni con las unidades léxicas simples de la clase correspondiente. Por otra parte, la naturaleza polisémica u homónimica de las locuciones, *per se* o en relación con una combinación de palabras no fraseológica, impide apoyarse de manera absoluta en los instrumentos que facilitan las nuevas tecnologías de la información para alcanzar resultados completamente satisfactorios en el estudio de su frecuencia.

En tercer lugar, el examen de numerosísimos ejemplos de uso de locuciones ha puesto en evidencia (apartado 4.) que la baja frecuencia de muchas de ellas va acompañada de la presencia de variantes no marcadas y de la existencia de combinaciones de palabras no fraseológicas homónimas. De este modo, cuando una locución cuenta con dos variantes no marcadas, es muy habitual que una de ellas sea más frecuente que la otra; de manera semejante, ha podido comprobarse que la colisión homónimica entre una locución y una combinación de palabras no fraseológica, ambas con idéntico significado, repercute en la baja frecuencia del fraseologismo.

Las implicaciones lexicográficas que todas estas cuestiones suponen para la elaboración de un diccionario de locuciones son presentadas en el apartado 5., en el que, de manera específica, se ofrecen las claves para determinar el umbral a partir del cual se puede empezar a fijar la frecuencia de una locución y para diferenciar si aquella es baja o elevada.

La naturaleza de este estudio, el examen de la frecuencia de las locuciones a partir del rastreo y registro de sus ocurrencias en los corpus académicos y en la red, impide la presentación de una hipótesis de trabajo que deba ser probada. El planteamiento adoptado es otro: el análisis de múltiples ejemplos de uso, es decir, la asunción de una metodología inductiva, permite demostrar las tesis de que la frecuencia de las locuciones está condicionada por la concepción que se mantenga del concepto estado de lengua actual, por la decisión que se tome en relación con la clasificación de estas unidades y por las características que de las locuciones muestren los datos examinados: locuciones con variantes y locuciones homónimas de combinaciones de palabras no fraseológicas.

La adopción de una metodología inductiva en una investigación concreta no entra en contradicción, claro está, con su fundamentación teórica. En este trabajo, se parte del estructuralismo y funcionalismo europeos representados por los trabajos de Saussure, Hjem-

slev, Trubetzkoy y Coseriu citados en las referencias bibliográficas, más los ya clásicos de semántica de Ullmann, así como de las aportaciones que enmarcadas en una fraseología claramente descriptiva ofrecen las obras mencionadas de Zuluaga, Corpas Pastor, Ruiz Gurillo, Mellado Blanco o Penadés Martínez.

2. La frecuencia de las locuciones en relación con la periodización de la lengua⁴

Si se vincula el concepto de frecuencia con el de actualidad no se origina una relación biunívoca, pues si bien es cierto que podría considerarse que una locución frecuente es actual, no ocurre lo contrario, es decir, puede darse el caso de que una locución actual no sea frecuente, y ello por su escasa presencia en corpus lingüísticos que la documenten para este periodo. Así, por ejemplo, no puede afirmarse que *mandar a hacer gárgaras* sea muy frecuente si se tiene en cuenta que el CREA solo registra dos ejemplos, el primero de 1987 y el segundo de 1995:

- (1) *Sus diferencias databan de 1985, cuando Hussein y Arafat no pudieron coordinar sus puntos de vista sobre negociaciones de paz con Israel. Entonces, el Monarca hachemita **mandó a Yasser a hacer gárgaras.***
- (2) *Un día le pregunté a mi madre que Rosa de quién era hija, para ser prima nuestra. Mi madre estuvo a punto de **mandarme como siempre a hacer gárgaras**, pero luego, no sé por qué, se lo pensó mejor y me dijo lo que Rosa ya me había dicho.*

Además, las ocurrencias que figuran en las diez primeras páginas de la red⁵ no pueden considerarse usos reales, pues corresponden a ejemplos inventados para ilustrar repertorios lexicográficos, o, al menos, parecen serlo. Análoga es la situación de *llamar Dios a su santo seno*, el CREA solo documenta el primero de los siguientes ejemplos y la red, el segundo:

- (3) *a tu papá lo ahorcaron en Pitiquito pero murió con mucha dignidad y decencia, un hombre debe ser digno y decente hasta que **Dios le llame a su santo seno.***

4 El análisis de la cuestión enunciada en el título de este apartado va a ceñirse al periodo de tiempo relativo a la actualidad, pues el diccionario de locuciones mencionado, el DILEA, corresponde al estado de lengua ahora vigente.

5 La búsqueda de ejemplos de uso en la red para este artículo se ha llevado a cabo, en los meses de noviembre y diciembre de 2013, y en enero de 2014 mediante el buscador *Google* y limitándola siempre a las diez primeras páginas de España y en español ofrecidas por el buscador. Esto significa que los textos de donde se han obtenido los ejemplos pertenecen a la modalidad escrita y oral de la lengua, pues el buscador ofrece indistintamente páginas de medios de comunicación, de blogs, de chats, de foros, etc. Los ejemplos de los corpus académicos también han sido extraídos en el mismo periodo de tiempo; por otra parte, aunque el corpus oral académico constituye aproximadamente un 10% de los últimos periodos del CREA, es posible acceder a través de él a casi 9 millones de formas procedentes de transcripciones de la lengua hablada, con más de 1600 textos encuadrados en dos grandes grupos: textos procedentes de grabaciones de radio o de televisión que han sido transcritos y codificados siguiendo el sistema de codificación de la parte oral de CREA, y textos ya transcritos procedentes de otros corpus orales que, tras la revisión del texto y/o del audio, se han convertido al sistema de transcripción y codificación del CREA oral (Real Academia Española, <<http://www.rae.es/recursos/banco-de-datos/crea-oral>>. De ahí que, al hablar de la frecuencia de una locución concreta a partir de los datos obtenidos en la red y en los corpus académicos, se trate de su frecuencia absoluta, es decir, independientemente de si es mayor su aparición en textos orales que escritos, por ejemplo, pues no se busca la frecuencia en relación con el tipo de texto en que la locución pueda ser más habitual, dado que tal proceder no sería acorde con un diccionario que pretende recoger locuciones idiomáticas del español actual (lo que supone una limitación diacrónica), peninsular (lo que implica una limitación diatópica), pero pertenecientes a las distintas modalidades de lengua (oral, escrita), propias de diferentes situaciones de habla (formal, informal) o específicas de los varios registros (vulgar, neutro), sin limitaciones diafásicas, por tanto.

- (4) *Ahora nos enteramos que, según la forense, el enfermo tiene leves síntomas y que todavía le queda un tirón para que **Dios lo llame a su santo seno.***

Como se verá en el apartado 4., la baja frecuencia de una locución en un estado de lengua puede estar relacionada con factores relativos a la propia unidad: la existencia de variantes no marcadas o su colisión con homónimos no fraseológicos. Pero también se podría explicar por los desajustes que se dan en el establecimiento de los límites para los periodos o estados por los que ha pasado una lengua. Quiere decirse con ello que, en función de la amplitud o reducción con que se tome el término *actual*, una locución puede presentar distinta frecuencia.

En efecto, la situación con que nos encontramos en relación con la lengua española, al delimitar el estado de lengua presente, no es uniforme. El calificativo *actual*, usado sobre todo en diccionarios generales de lengua y específicos de fraseologismos, no es resultado de aplicar una periodización homogénea y consensuada. Así, los autores del DEA, para ceñirse al léxico contemporáneo, establecieron su propio catálogo léxico creando una base documental constituida por textos reales de lengua española no anteriores a la segunda mitad del siglo XX, de este modo se puede afirmar que el adjetivo *actual*, cuando se publicó el diccionario, hacía referencia a las palabras usadas por los españoles a lo largo de los cincuenta últimos años del siglo anterior (Seco, Andrés y Ramos, 1999: XIII).

No siempre los redactores de diccionarios son tan precisos en sus explicaciones sobre la vigencia temporal del léxico recopilado. Así, por ejemplo, en la «Introducción» al GDUEA, que salió a la luz en 2001, los límites del término *actual* son más difusos, pues se habla de la lengua «usada en nuestros días», «en el presente histórico en que nos toca vivir», «en los últimos años del siglo XX, uso que ciertamente se prolongará en los inicios del siglo XXI» (Sánchez, 2001: 7 y 9).

Los mismos compiladores del DEA, a principios de este siglo, concretamente en 2004, publicaron una nueva obra lexicográfica, esta vez limitada a unidades fraseológicas, en cuyo título, *Diccionario fraseológico documentado del español actual*, también aparece el término de la compilación anterior y casi con la misma extensión, aunque no exactamente, pues las unidades registradas pertenecen al español contemporáneo –hasta el 2004, no hasta el fin del XX– y su existencia real está acreditada con testimonios escritos de los últimos cincuenta años. El DFDEA es, en palabras de sus autores (Seco, Andrés y Ramos, 2004: XII y XV), un diccionario sincrónico cuya documentación está integrada por citas procedentes de textos publicados entre 1955 y 2004, es decir, el último medio siglo en relación con la fecha de aparición de la obra, periodo en el que viven los hablantes de español de ese momento, tanto los más viejos como los más jóvenes⁶.

6 No siempre el ámbito temporal de un diccionario queda claro desde el propio título, pues, en otros casos, aparecen denominaciones como *moderno*, en el DFEM, o *usual*, en el DFE, término este último que obliga a ampliar el análisis a los diccionarios calificados como *de uso*, ya sean de lengua (DUE) o fraseológicos (JAM). No obstante, este trabajo no va a abordar estas obras, baste con indicar que *moderno* queda sin precisar en el DFEM, pues sus autores solo indican que recogen las unidades fraseológicas que «proceden de la moderna lengua hablada del español de España» (sic.) (DFEM, 1994: X), mientras que los compiladores del DFE, sin explicar explícitamente que entienden por *usual*, afirman que el diccionario recoge numerosas locuciones y expresiones de uso normal y frecuente en español, expresiones de la lengua de todos los días (DFE, 2007: 8), lo que no ha sido obstáculo para incluir locuciones como *descubrirse la hilaza*, ‘Aparecer o ponerse de manifiesto un defecto que se procuraba ocultar o disimular’, no recogida en el DFEM ni en el DFDEA, aunque sí existe una ocurrencia de esta locución

Fuera del campo de la lexicografía, *actual* califica el que, posiblemente, es el mejor corpus del español, el CREA de la Real Academia Española, compuesto, como es de sobra conocido, de una amplia variedad de textos escritos y orales, producidos en todos los países de habla hispana desde 1975 hasta 2004, y estructurado en períodos de cinco años: 1975-1979; 1980-1984; 1985-1989; 1990-1994; 1995-1999; y 2000-2004. Por su parte, el CORPES XXI (*Corpus del Español del Siglo XXI*) se solapará parcialmente con el CREA, pues, cuando finalice la primera parte del proyecto en diciembre de 2014, los textos procederán del periodo comprendido entre 2001 y 2012⁷, además de originarse la paradoja de restringir la denominación *actual* al siglo pasado al no utilizarse para los textos de este.

Según lo que se acaba de indicar, los límites para incluir o excluir de un diccionario las locuciones que presenten el carácter de actuales o carezcan de él varían en función de la posición metodológica adoptada, dando lugar, a partir de los datos anteriores, a un periodo fluctuante que comprende desde el 1950 al 1974. En consecuencia, locuciones que se documenten solo en ese espacio temporal –no resulta ahora de interés que también se registren en otros anteriores a él– podrán ser caracterizadas como actuales o no, y, por tanto, recogidas o desechadas para un diccionario, en función de la perspectiva adoptada por el lexicógrafo.

Más todavía, y enlazando ahora el factor de actualidad con el de frecuencia, si una locución es frecuente en el tercer cuarto del siglo XX, el que va de 1950 a 1974, pero se documenta con una baja frecuencia o no se documenta del 1975 hasta el día de hoy, cabría plantearse la posibilidad de excluirla de un diccionario cuya macroestructura esté constituida por locuciones actuales, siempre y cuando el inicio del estado de lengua contemporáneo se fije en el 1975. Unos ejemplos servirán para concretar más la cuestión. La consulta del DFDEA indica que forman parte de su macroestructura locuciones con ocurrencias en textos cuya publicación se sitúa en los años comprendidos entre 1950-1974. Así, el uso de la locución *estar hecho con rabos de lagartija*, clasificada en el DFDEA (2004: 860) como locución verbal que ‘*Se emplea, referido a pers, para ponderar su rápido y continuo movimiento o su inquietud*’⁸, está documentado en una obra publicada en 1963; pero hay que tener en cuenta que de esta locución no existen ejemplos de uso en el CREA ni el CORDE, tampoco se encuentra ninguno en las diez primeras páginas de la red; cabría cuestionarse, pues, su carácter actual. Algo semejante ocurre con *echar lumbre*, ‘Dar muestras de mucho enfado’ (DILEA), pues el ejemplo del DFDEA es también de 1963, el CREA no registra ninguno y los más modernos del CORDE son de 1909, documentados en dos obras de Benito Pérez Galdós:

- (5) *No te quiero decir cómo se puso el hombre al ver la profanía... Su cara echaba lumbre, redió; le tembló la papada, apretó los puños...*
- (6) *En Calatañazor había gran escándalo, y los Gaitines de Soria echaban lumbre, abrasados de ira y furor de venganza.*

Por supuesto, no habría que percibir como actual una locución documentada solo con anterioridad al 1950, si ahí se fija el límite de inicio de este periodo de la lengua. Aunque no tengamos la certeza de que no se ha respetado este límite inferior, podría pensarse que no ha

en el CREA, en un texto de Fernando Arrabal: *Tarsis considera la red que le tiende su adversario como un arte que descubre la hilaza, le da carrete y toma el peón 19... Db2xa2.*

⁷ <<http://rae.es/recursos/banco-de-datos/crea>>; <<http://rae.es/recursos/banco-de-datos/corpes-xxi>>.

⁸ Cuando se proporciona el significado de una locución, se indica el diccionario del que se toma.

sido una práctica inusual en el DFDEA por los ejemplos que en este diccionario figuran con un asterisco que les precede, indicando que la falta de testimonio escrito se ha suplido por uno oral (DFDEA, 2004: XXVI). Por ejemplo, *dar el estallido*, con el significado ‘Morir por exceso de presión o de fatiga’ (DFDEA), viene documentada de ese modo, con un enunciado que parece construido *ad hoc*:

- (7) *Un día **darás el estallido** si no dejas de trabajar a ese ritmo.

Ni el CORDE ni el CREA registran usos de esta locución y en la red, buscándola con el verbo conjugado en distintas formas, solo se encuentra la siguiente ocurrencia de *Fortunata y Jacinta*, novela de Pérez Galdós con fecha de publicación 1886-87:

- (8) *Para que lo sepas; tu mujer te ha faltado con aquel señor de Moreno, que se murió de repente, una noche. La suerte tuya fue que **dio el estallido**; y es que los corazones revientan, de la fuerza del querer.*

Aunque la locución forma parte de la macroestructura del DFDEA, bien podría pensarse que no es actual, en este caso por no estar fehacientemente documentada en textos de la segunda parte del siglo XX.

Como se habrá ido intuyendo esta situación problemática tiene que ver con el concepto de estado de lengua o sincronía, y con varias cuestiones relacionadas con él. Situándose en la perspectiva saussureana, la lingüística sincrónica se ocupa, como se sabe, «de las relaciones lógicas y psicológicas que unen términos coexistentes y que forman sistema, tal como aparecen a la conciencia colectiva» (Saussure, 1945: 174). No obstante, el establecimiento de los principios generales de un sistema o de los factores constitutivos de un estado de lengua no deja de suponer una complejidad reconocida por el propio maestro ginebrino, por cuanto que un estado de lengua no es un punto, sino una extensión más o menos larga, de diez años, de una generación, de un siglo o incluso mayor. De este modo, se comprende su reconocimiento explícito de que la limitación en el tiempo es una de las dificultades que se encuentran en la definición de un estado de lengua (Saussure, 1945: 175-177). Por eso el adjetivo *actual* para referirse al estado presente de la lengua española puede abarcar desde el 1950 hasta el último año del siglo XX, según la percepción del DEA, desde el 1955 hasta el 2004, desde la concepción del DFDEA, desde el 1975 hasta el 2004, según el punto de vista de la Real Academia Española al confeccionar el CREA, o incluso todo el siglo XX y lo que llevamos del XXI, si aceptamos que un hablante relativamente culto del español actual es capaz de comprender las obras literarias escritas en torno al 1900, porque tiene, si bien parcialmente, una conciencia, con otro término: una *competencia*, de la lengua de aquella época, aunque sea pasiva.

Más todavía, aunque Coseriu (1978) se mostró fuertemente crítico con la concepción de Saussure sobre sincronía / diacronía, acabó reconociendo el relativismo en la noción de estado de lengua, pues no de otro modo cabe interpretar sus afirmaciones acerca de que 1) en un estado de lengua pueden comprobarse arcaísmos, que, en cuanto existen y funcionan, son elementos actuales; 2) los hablantes tienen la conciencia de que ciertos elementos son más viejos o más nuevos; 3) la lengua sincrónica es un sistema actual de tradiciones lingüísticas antiguas y recientes; y 4) en la lengua conviven durante largo tiempo lo viejo y lo nuevo (Coseriu, 1978: 18, 17, 61 y 128).

En consecuencia, la imposibilidad de establecer claramente los límites de un estado de lengua y la existencia en él de unidades –en el caso que aquí interesa fraseológicas– que para la conciencia de los hablantes, y como fruto de su reflexión metalingüística, pueden calificarse de arcaicas, antiguas o viejas conduce a afirmar que locuciones como *echar lumbre* o *dar el estallido* pueden verse, sí, como actuales, si se amplía esta sincronía hasta principios del siglo XX, y de ahí la justificación de su registro en un diccionario fraseológico del español correspondiente al periodo actual, pero, al mismo tiempo, como no frecuentes, y su escasa frecuencia de uso obliga a plantearse también la posibilidad de su exclusión en un diccionario calificado de *actual*, o, al menos, a marcarlas en relación con su baja utilización.

Por otra parte, visto que actualidad no implica frecuencia, surge la necesidad de examinar esta última característica de las locuciones *per se*, sin establecer relación alguna con aquella, tan solo acotando el número de veces que se presenta una locución en conjuntos de textos más o menos extensos. Ahora bien, la determinación de la frecuencia de las locuciones tampoco es una cuestión libre de dificultades, como se muestra en el siguiente apartado.

3. Dificultades en el análisis de la frecuencia de las locuciones

Al observar la situación de otras lenguas, se comprueba que para el inglés se cuenta con datos sobre la frecuencia de las locuciones en corpus de la magnitud del COBUILD, analizado por Moon (1997)⁹, y, como explica Corpas Pastor (2013), se pueden obtener, igualmente, en el corpus BYC-BNC; asimismo, para el alemán existen proyectos de bases de datos que ofrecen información acerca del nivel de la frecuencia de uso de los fraseologismos (Mellado Blanco, 2012: 150). Sin embargo, en el caso del español, los estudios que arrojen luz sobre esta cuestión en corpus de extensión similar son prácticamente inexistentes exceptuando la investigación de Boers y Stengers (2008), quienes han analizado la frecuencia de aparición de 500 locuciones del inglés y del español en el *Collins Cobuild Online Wordbanks* de ambas lenguas; no obstante, es necesario indicar que las unidades contabilizadas fueron previamente extraídas, para el caso del español, de dos diccionarios: el de Buitrago Jiménez, *Diccionario Espasa de dichos y frases hechas*, y el de Giménez, *Del dicho al hecho*. Por otra parte, las locuciones estaban agrupadas según distintos dominios fuente tales como juegos y deportes, fauna y flora, guerra y agresión, vehículos y transportes o religión y superstición; en este sentido, los tantos por ciento que se ofrecen corresponden a los grupos de locuciones, no a unidades particulares.

Disponemos, en cambio, de análisis sobre conjuntos de textos relativamente reducidos (Penadés Martínez, 2004; Penadés Martínez, 2012a), concretamente sobre dieciocho conversaciones y dieciséis entrevistas, ambas calificables como conversaciones coloquiales periféricas, que ocupan unas veintisiete horas de grabación y contienen una suma aproximada de 322 000 palabras, y, por otra, sobre treinta y seis entrevistas que constituyen el corpus PRESEEA-Madrid (Barrio de Salamanca), que comprenden unas treinta y seis horas de grabación¹⁰.

9 Esta investigadora comprobó (Moon, 1997: 52) respecto a 4000 locuciones que un 30% aparece menos de una vez por diez millones de palabras; el 35% aparece entre una y tres veces por diez millones de palabras, y el 20% tiene una frecuencia de, al menos, un ejemplo por dos millones de palabras.

10 Contamos, además, con las investigaciones de Corpas Pastor (1998) sobre la frecuencia de las paremias en el corpus *Vox-Biblograf*. Junto a ello Ruiz Gurillo (1998) ofrece porcentajes de los sintagmas prepositivos fraseo-

Los distintos objetivos de los estudios citados en el párrafo anterior, tanto en el cuerpo del texto como en la nota a pie de página, hacen difícil establecer vinculaciones entre ellos para obtener datos concretos sobre la frecuencia de las locuciones. Aun así, limitándose a Penadés Martínez (2004), se desprende, en primer lugar, que las 10 locuciones verbales más frecuentes del total de 515 emitidas son, con sus respectivos porcentajes: *darse cuenta* (12,2%), *hacer falta* (10,8%), *llevarse bien* (6,9%), *dar igual* (5,6%), *tener en cuenta* (3,3%), *tener que ver* (3,3%), *dar lo mismo* (2,7%), *echar de menos* (2,3%), *querer decir* (1,9%) y *darse bien* (1,7%)¹¹; y, en segundo lugar, que son muy numerosas las que tienen la menor frecuencia relativa hallada, 0,1%, concretamente 103 locuciones verbales de un total de 158 distintas registradas.

Independientemente de los datos reales obtenidos en los estudios que se acaban de citar, lo que encontramos sobre esta cuestión son referencias bibliográficas en las que se comenta la posibilidad de estudiar la frecuencia de las locuciones 1) en bases de datos con determinadas características (Mogorrón Huerta, 2012; Ghezzi, 2012), 2) apoyándose en la lingüística de corpus (Mellado Blanco, 2012) y 3) incluso mediante la utilización del buscador *Google* (Sevilla Muñoz, 2012). Es previsible que esta deficiencia de investigaciones sobre la frecuencia de uso de las locuciones en español se subsane con el tiempo. Con todo, hay una serie de cuestiones que afectan al estudio de las locuciones en esta lengua desde la situación actual de la fraseología y que vuelven extremadamente compleja la tarea de indicar la frecuencia absoluta y relativa de una locución concreta aun limitándose a un corpus determinado.

lógicos, de las unidades fraseológicas verbales y de los esquemas fraseológicos documentados en un corpus de español oral. Castillo Carballo (2000), por su parte, apunta, como resultado del estudio de su corpus, que la base de datos recoge un total de mil seiscientos treinta y cinco registros de unidades pluriverbales, de las que cuatrocientas siete se catalogan como locuciones verbales. Por último, el trabajo de Sosiński (2010) presenta los resultados de un análisis cuantitativo de las unidades fraseológicas empleadas por cuatro inmigrantes marroquíes residentes en Granada y por un hablante nativo de español entrevistado como parte del corpus PRESEEA-Granada.

11 Obsérvese que la mayor parte de ellas no presenta la característica de la idiomatidad. En efecto, la locución *hacer falta* significa ‘Ser necesario’ (DFDEA, 2004: 452), y el verbo *faltar*, que tiene el mismo significado léxico que el elemento *falta* de la locución, en su acepción 3, es ‘Hacer falta o ser necesario’ (DEA, 1999: 2117); por su parte, *llevarse bien* significa ‘Tener buena relación’ (DFDEA, 2004: 590) y el verbo *llevarse*, en la acepción 22, significa, según el DSLE (1996: 953), ‘Mantener < una persona > [buenas o malas relaciones con una o varias personas]’; la locución *dar igual* significa, según el DFDEA (2004: 545), ‘Ser indiferente’, y el elemento *igual* de la locución, en su acepción 1: ‘[Pers. o cosa] de las mismas características externas o internas [que otra [...]], en todo o en el aspecto que se considera’ (DEA, 1994: 2550); *tener en cuenta* significa ‘Tener presente o en consideración’ (DFDEA, 2004: 345) y *contar*, verbo con el mismo significado léxico que el elemento *cuenta* de la locución, en su acepción 10, quiere decir ‘Tener presente [...] para algo que se proyecta’ (DEA, 1999: 1217); *dar lo mismo* tiene como significado, según el DFDEA (2004: 653), ‘Ser indiferente’, y el elemento *misimo* de la locución, en su acepción 5, se define así: ‘Precediendo o siguiendo al n, o siguiendo a un pron, indica que se elige una entre varias posibilidades sin que haya motivo de preferencia’ (DEA, 1999: 3090); por último, *darse bien* tiene como significado ‘Tener buenas condiciones para actuar’ (DFDEA, 2004: 362), mientras que el verbo *darse*, según el DUE (1966: 859), en su acepción 5, significa ‘Resultar fácil para alguien una cosa por tener habilidad o aptitud para hacerla o aprenderla’. Por otra parte, casi todas ellas muestran una fijación formal relativa: *darse (perfecta) cuenta*, *hacer (mucho / poca) falta*, *llevarse (bastante / muy) bien*; *dar (exactamente) igual*, *tener (muy) en cuenta*, *tener (mucho / poco) que ver*, *dar (exactamente) lo mismo*, *echar (mucho / tanto) de menos* y *darse (bastante / muy) bien*. Significa esto que prácticamente todas las locuciones verbales que aparecen con mayor frecuencia están situadas en la periferia, no en el centro, de la clase de las locuciones.

En primer lugar, no son pocas las locuciones sobre las que recaen dudas acerca de su ubicación en una u otra clase. Como se ha indicado en Penadés Martínez (2006), *el fin del mundo, moco de pavo, plato de gusto, trigo limpio, agua pasada, el espíritu de la golosina, el pan nuestro de cada día* y muchas más son nominales, tal como las recoge el DFDEA, no verbales, y, consecuentemente, pese a su inclusión en esta clase por parte del DSLE, su forma no corresponde a *no ser el fin del mundo, no ser moco de pavo, no ser plato de gusto, no ser trigo limpio, ser agua pasada, ser el espíritu de la golosina, ser el pan nuestro de cada día*. También para las clases de las verbales y las adverbiales hay unidades que, según la obra lexicográfica tomada como referencia, quedan encuadradas entre las de la primera clase o las de la segunda. Así, las siguientes, por ejemplo: *andar/estar a dos velas, andar/estar como el perro y el gato, andar de boca en boca, andar/estar de uñas, andar/estar ojo avizor, anunciar/gritar/pregonar/publicar a los cuatro vientos*, son verbales para el DFEM, si nos atenemos a la paráfrasis definitoria, ya que en este diccionario no se marcan gramaticalmente de manera explícita; sin embargo, la consulta del CREA muestra que está más acertado el DFDEA al considerarlas adverbiales, pues, *a dos velas* se combina con *dejar, estar y quedar*; *como el perro y el gato*, con *andar, estar y llevarse*; *de boca en boca*, con una amplia variedad de verbos: *andar, circular, correr, extenderse, ir, transmitirse*, etc.; igual que *de uñas*, que modifica a *esperar, estar, poner, recibir*, etc.; *ojo avizor*, por su parte, lo hace sobre *andar, estar, mantenerse, pasear*; y *a los cuatro vientos* aparece también con una extensa serie de verbos, como, además, paradójicamente, muestra el propio lema del DFEM: *airear, anunciar, clamar, contar, pregonar, proclamar, propagar, repetir, vocear*, etc.

Los ejemplos están puestos para hacer ver que, en tanto no tengamos establecida sin discusión la clasificación de una locución dudosa, no tendrá mucho sentido analizar su frecuencia, pues no podrá ponerse en relación con la de las otras locuciones de su clase, ni con la de las locuciones de clases distintas ni con la de las unidades léxicas simples de la clase correspondiente. Y, si se hace, los datos estarán sujetos a una provisionalidad tal que se convertirán en poco fiables para marcar las locuciones en un diccionario en relación con su frecuencia de uso.

En segundo lugar, de manera análoga a como ocurre con las unidades léxicas simples, existen locuciones que cuentan con la particularidad de la polisemia u homonimia¹², es decir, que presentan más de un significado. No existe acuerdo entre los fraseólogos a la hora de considerar si es un fenómeno que se da en las locuciones de manera usual o no¹³. En cualquier caso, su existencia es innegable, y también lo es la dificultad añadida que supone para contabilizar la frecuencia de uso de estas unidades fraseológicas, pues no basta con su simple registro en un conjunto de textos, es necesario, asimismo, que el propio investigador discrimine si una ocurrencia concreta se refiere a uno u otro significado de una misma cadena formal, de un único significante¹⁴. Piénsese, por ejemplo, en locuciones como *meter caña*, para la que pueden establecerse seis significados distintos a tenor de sus ejemplos

12 Para las razones que conducen a no diferenciar entre los fenómenos de la polisemia y la homonimia véase Penadés Martínez (2012b: 267-277).

13 Así, para Corpas Pastor (1996: 115) y Mogorrón Huerta (2012: 71) la polisemia en las locuciones es un fenómeno frecuente, mientras que Larreta Zulategui (2001: 199) y Mellado Blanco (2004: 75) ponen en duda esta particularidad.

14 No proceden así, sin embargo, Boers y Stengers (2008), que hacen caso omiso de los distintos significados de una locución sumando simplemente las distintas ocurrencias de ambos; igual actúan con aquellas locuciones que son parcialmente homófonas: *por los pelos* y *traer por los pelos*, no se computan de manera separada.

de uso, exactamente ‘Golpear’, ‘Atacar criticando’, ‘Tratar con dureza y con exigencias’, ‘Provocar’, ‘Hacer que un vehículo aumente de velocidad’ y ‘Actuar con fuerza o con agresividad’ (DILEA):

- (9) *Los Domecq, Alfonso Domecq senior y Alfonso Domecq junior, armaron la trapatiesta en la corrida goyesca de Ronda, el año 75, y parece que le metieron caña a un modesto representante de la autoridad.*
- (10) *Ten en cuenta que a Guardiola le metieron caña después de dos partidos empatados. Si lo hicieron con él, que es considerado un icono allí, lo hacen con Vilanova seguro.*
- (11) *Iñaki les mete caña. El técnico sigue aumentando la intensidad de los entrenamientos, en el segundo día, con doble sesión de trabajo.*
- (12) *El líder del PP utilizó con «los mayores» un lenguaje diferente al que suele utilizar en sus mítines normales, intentando dar a los jubilados mensajes positivos y de confianza, de forma que, aunque habló de escándalos y corrupción, sin citar para nada a Roldán u otros casos, fueron los propios asistentes los que «metieron caña» y en cuanto hablaba de austeridad, los jubilados pedían «que devuelvan el dinero que han robado», «Felipe a la cárcel».*
- (13) *he metido caña a la moto para tratar de ganar una etapa. Finalmente no ha sido posible, pero estoy contento porque ayer me encontraba mal físicamente mientras que hoy me siento muchísimo mejor.*
- (14) *El disco lo hicimos a lo grande y ahí está la gracia del directo. El disco suena con gracia, muy pulido, pero en directo, como somos diez y no treinta, metemos más caña.*

En el CREA se recogen solo siete ocurrencias de *meter caña*, por lo que su análisis no resulta complejo, pero contabilizarlas simplemente y ponerlas en relación, de manera conjunta, con las de otra locución monosémica carece de sentido, ya que uno de los significados anteriores no está registrado en el CREA (‘Golpear’) y otros lo están aunque con una mínima presencia; de ‘Atacar criticando’ solo hay dos ejemplos, igual número tiene ‘Actuar con fuerza o con agresividad’, mientras que se encuentra un registro de cada uno de estos significados ‘Tratar con dureza y con exigencias’, ‘Provocar’ y ‘Hacer que un vehículo aumente de velocidad’. Los datos de *meter caña* en conjunto contrastan con los de la variante *dar caña*, de la que el CREA contiene cincuenta y seis ocurrencias, que deberían ser distinguidas de acuerdo con su significado para, posteriormente, estudiar la frecuencia de las muestras de lengua de cada significado en relación con las correspondientes a la variante *meter caña* y en relación, asimismo, con las locuciones de la clase verbo, sean monosémicas o tengan más de un significado.

Lo que se quiere indicar con estos datos es la necesidad de una revisión individual, locución por locución, por parte del investigador, pues, a partir de ellos, no existe una locución *meter caña*, sino seis. Para determinar las locuciones más frecuentes de un determinado corpus, no basta entonces con el simple recuento de una cadena de significantes en una base de datos –efectuado, si así se desea, mediante programas informáticos procesadores de textos–, lo impide el fenómeno de la polisemia u homonimia de las locuciones, y también la homonimia entre una locución y una combinación de palabras que no es fraseológica, tercera cuestión que afecta al establecimiento de la frecuencia de las locuciones dificultándola.

En efecto, junto al caso de *meter caña* o *dar caña*, locuciones polisémicas u homónimicas, existen también situaciones como las ilustradas en los siguientes ejemplos, uno de la locución *abrir la puerta*, con el significado de ‘Hacer posible una cosa o hacer que ocurra’

(DILEA) bajo la forma *abre la puerta*, y el otro de la combinación de palabras *abre la puerta*, literal, no fraseológica a partir de la acepción 2. de *abrir* en el DRAE¹⁵:

- (15) *La ciencia nos abre la puerta al pensamiento crítico, y este nos da libertad, autonomía y control sobre el propio destino.*
- (16) *Mientras nos abre la puerta de su casa, nos habla de otra de sus aficiones: la recogida de fósiles por los distintos andurriales.*

El hecho de que existan locuciones que se relacionan con combinaciones no fijas de palabras homónimas ha sido explicado por algún autor considerando que los fraseologismos se caracterizan por poseer, junto a su significado literal, deducible de los significados de cada uno de sus componentes, un significado global en su conjunto, el significado fraseológico (López Roig, 2002: 33-34). Ahora bien, considerar que una unidad fraseológica cuenta con un significado literal, «resultante de la suma de los significados de los componentes que forman la UF tomando ésta como combinación libre de palabras» (López Roig, 2002: 114), conduce a eliminar su condición intrínseca de unidad fraseológica, para asignarle la naturaleza de combinación no fija de palabras, con lo que deja de ser un fraseologismo. No es posible, por lo tanto, hablar de significado literal y fraseológico de locuciones como *abrir la puerta*, *ir arreglado* ('Estar equivocado'), *llamar a la puerta* ('Pedir ayuda'), *marcar un gol* ('Sorprender') (DILEA), etc., etc.¹⁶. Lo que la realidad de la lengua muestra es la existencia de combinaciones de palabras homónimas: una no fijada y otra fijada en su forma y en su significado, esta última, pues, fraseológica.

Los casos en que la lengua dispone de una locución homónima de una combinación no fija de palabras no son infrecuentes, al contrario, y su existencia tiene el mismo efecto que la polisemia u homonimia en una misma unidad fraseológica: la necesidad de que el investigador examine ocurrencia por ocurrencia para deslindar la que corresponde a la locución y la que constituye una combinación de palabras no fraseológica, de modo que los recuentos estadísticos de las locuciones no queden falseados¹⁷, pues, como afirma Corpas Pastor

15 'Separar del marco la hoja o las hojas de una puerta o ventana, haciéndolas girar sobre sus goznes, o quitar o separar cualquier otra cosa con que esté cerrada una abertura, para que deje de estarlo'.

16 Véanse los siguientes pares de ejemplos, el primero de cada par es de la locución y el segundo, de la combinación no fraseológica de palabras:

- a) *No hay reforma que valga, y Rajoy si cree que ya en la Moncloa va a poder hacer algo por lo suavín **va arreglado**.*
- b) *¿Con vaqueros no se **va arreglado** acaso? ¿Qué puede pasar si no **vas arreglado**?*
- a) *Su lema es que «sin riesgo no hay éxito» y que lo justo es «dar oportunidades a la gente que **llama a la puerta**».*
- b) *Alguien dice: «Podríamos pedir unas pizzas y unas cervezas». Aprobación inmediata. Al cabo de veinte minutos, un chico con un casco, que ha dejado la moto en la acera, **llama a la puerta** y reparte la cena.*
- a) *¿Que interpretáis de la jugada de Mou? ¿Es un estrechamiento de manos entre entrenador-capitán? ¿Han sido todas las movidas una estrategia para despistar? ¿Qué opináis? Yo reconozco que Mou me **ha marcado un gol** en esa acción.*
- b) *Leandro – Delantero (fichado por una pasta sólo me **ha marcado un gol** en 5 partidos, poca titularidad así...).*

17 Así, en Mogorrón Huerta (2012: 78-79), se muestra un ejemplo de la aplicación del programa Unitex a un conjunto de textos con la finalidad de obtener las expresiones que contienen el verbo *abrir*. Una de ellas está registrada en el siguiente fragmento: «[...] **Antes de que nadie intentara abrir la puerta**, un juez ha suspendido el desalojo y ha [...]», fragmento que, a su vez, forma parte de este otro:

«Todo eso hizo que hoy activistas y vecinos se acantonaran en el lugar preparándose para una defensa numantina. Una joven ha llegado a encadenarse por el cuello a una de las puertas de entrada al recinto. Un cartel

(2013: 351), actualmente los sistemas de identificación de las expresiones multiverbales (y las unidades fraseológicas lo son) no disponen de las técnicas necesarias para poder procesarlas ni interpretarlas contextualmente de manera adecuada.

El análisis de la frecuencia de las locuciones a partir de los instrumentos que facilitan las nuevas tecnologías de la información es una realidad que cada vez se va haciendo más presente en el quehacer de la investigación fraseológica¹⁸, pero las posibilidades de ahí derivadas no deben ocultar las dificultades que giran en torno a él por estar todavía pendientes de resolución en fraseología cuestiones que afectan a la esencia de las unidades fraseológicas, como es el hecho de su pertenencia a una u otra clase y su naturaleza polisémica u homónima *per se* o en relación con una combinación de palabras no fraseológica. Conviene, pues, que desde la fraseología queden bien delimitadas estas cuestiones antes de abordar el estudio de la frecuencia de unidades fraseológicas concretas.

4. La frecuencia en locuciones con variantes no marcadas y en locuciones homónimas de una combinación no fija de palabras

La redacción del DILEA, iniciada por las locuciones pertenecientes a la clase de las verbales, ha evidenciado que, en no pocos casos, la frecuencia de uso de una locución determinada es distinta para sus diferentes variantes no marcadas y se reduce al mínimo cuando existe una combinación no fija de palabras homónima de la locución. Estos dos casos, que son explicados en los siguientes apartados, obligan a tomar decisiones acerca de la inclusión o no, en el diccionario, de una variante y de una locución homónima de una combinación no fija de palabras, o, al menos, acerca de la necesidad de marcar de algún modo su baja frecuencia.

4.1. La distinta frecuencia de variantes no marcadas

Antes de abordar esta cuestión, conviene indicar lo que se entiende por variante y por variación no marcada y variación marcada en las unidades fraseológicas. Siguiendo a Zuluaga (1980: 109): «Las variantes fraseológicas son parcialmente idénticas en su estructura y en sus componentes», o a Mellado Blanco (2004: 91): «las variantes fraseológicas se reconocen porque el número de elementos que no varían en ellas supera al de elementos que varían», se considera que *llevarse el demonio*, por ejemplo, debe analizarse como una variante léxica en relación con *llevarse el diablo*. Por su parte, *no marcado* se utiliza para indicar un uso neutro de una variante de una locución; con un ejemplo, las variantes *esconder la cabeza bajo el ala ~ meter la cabeza bajo el ala* pueden aparecer indistintamente en

advertía: «Peligro de muerte. Detrás de esta puerta una mujer está encadenada por el cuello. Si intentan abrir esta puerta la matarán a ella». Antes de que nadie intentara abrir la puerta, un juez ha suspendido el desalojo y ha ordenado al ayuntamiento que informara a los vecinos de sus planes de desalojo parcela a parcela».

La contextualización de la supuesta locución *abrir la puerta* indica a las claras que se trata de una combinación no fraseológica que no debe contabilizarse para obtener la frecuencia de la locución homónima. Téngase en cuenta que el contexto constituye una de las salvaguardas contra cualquier posibilidad de confusión provocada por dos homónimos (Ullmann, 1983: 181-182; Ullmann, 1986: 297-298).

18 Ya indica Corpas Pastor (2013: 368): «Quizá en un futuro no muy lejano podamos disponer de los corpus necesarios y las herramientas adecuadas para detectar, extraer, interpretar, representar y traducir todo tipo de unidades fraseológicas».

diferentes contextos, hecho que queda representado por la ausencia de una marca que las caracterice en los diccionarios, de ahí su carácter de variantes no marcadas de la invariante de contenido ‘No querer ver la realidad’ (DILEA)¹⁹. Por el contrario, *marcado* señala que una variante fraseológica es propia de alguna sincronía o estado de lengua, de una determinada diatopía, específica de hablantes caracterizados por alguna variante social o particular de alguna situación comunicativa o de una modalidad de lengua; en este sentido, *mandar a paseo* y *mandar a la mierda*, la una informal por la imagen mental de acción cotidiana que provoca y la otra vulgar por el elemento *mierda* que la forma (Penadés Martínez, 2013), constituyen dos variantes marcadas de la invariante de contenido ‘Echar o rechazar a una persona o una cosa’ (DILEA).

Aunque el estudio de la frecuencia de las variantes de las locuciones puede llevarse a cabo tanto para las que son marcadas como para las no marcadas, aquí va ceñirse el análisis a estas últimas, pues la ausencia de un valor diatópico, diastrático o diafásico²⁰ en el uso de una u otra posible variante permite que su comparación se lleve a cabo en relación con la propia locución. Si contaran para el análisis los factores lugar, característica social y situación comunicativa (formal, informal), registro (neutro, vulgar) o modalidad (oral, escrita), las distintas frecuencias de las variantes deberían vincularse a estos factores y no a las distintas formas de una misma locución²¹.

Pues bien, no resulta difícil encontrar ejemplos de locuciones con variantes no marcadas de distinta frecuencia para cada una de estas. Por ejemplo, examinando el caso de *cerrarse a la banda* / *cerrarse en banda*, se comprueba que la primera variante está documentada solo en el CORDE en cuatro textos que, para el siglo XX, van de 1916 (uno de Unamuno) hasta el 1951 (uno de Guerra Navarro y dos de Cela); por el contrario *cerrarse en banda* tiene cuarenta y cuatro ocurrencias en el CREA, la distinta frecuencia de una y otra variante es, pues, notable. En cuanto a *pagar con la misma moneda* / *pagar en la misma moneda*, la situación no es la misma, pero el resultado sí: una variante es más frecuente en la actualidad que la otra; la prueba nos la ofrecen, de nuevo, los corpus académicos: en el CREA hay diecisiete ocurrencias de *pagar con la misma moneda*, frente a cinco de la otra variante; el CORDE, por su parte, recoge este mismo número de ejemplos para esta variante y cuatro para *pagar con la misma moneda*, lo que significa que el uso de esta última se va incrementando frente a la obsolescencia de la otra.

19 Como se habrá deducido, esta concepción de invariante y variante está, asimismo, vinculada al estructuralismo europeo, sobre todo al praguense y al danés, más en concreto a las reglas de Trubetzkoy (1987: 41-46) para diferenciar sonidos de fonemas y al interés de Hjelmslev (1974: 90-108) por establecer las invariantes de un sistema lingüístico.

20 Como viene siendo habitual en los estudios sobre la variación (Casas Gómez y Escoriza Morera, 2009; Kabatek, 2002), no hay obstáculo alguno para que la diafásica incluya la llamada *variación diamésica*, es decir, la determinada por el medio de realización, gráfico o fónico, de un texto.

21 Obsérvese que no se menciona el valor diacrónico, relativo al factor tiempo, porque, justamente, la mayor o menor frecuencia de una variante respecto de otra puede llevar a la conclusión de que la menos frecuente pertenece a un estado de lengua anterior, con lo que entraría en juego la variación marcada diacrónicamente. Aunque, como esa interpretación no tiene por qué ser la única posible, puesto que una variante no frecuente no ha de serlo necesariamente por ser de una sincronía precedente, el factor tiempo no va a tenerse en cuenta en el análisis de la frecuencia que se presenta a continuación. Dicho de otra manera, cuando se indique de los ejemplos que se van a estudiar que una de las variantes no es frecuente, no se da por hecho que sea más usada en el estado de lengua inmediatamente anterior, puede que sea así o puede que no; la certeza solo podría obtenerse tras un análisis específico con el punto de mira puesto en la variación diacrónica.

En ocasiones, el DFDEA se hace eco del distinto comportamiento de dos variantes, aunque no siempre; así, de *cerrarse a la banda* indica que es *más raro*, pero en la entrada correspondiente a *pagar con/en la misma moneda* no aparece ninguna marcación, al igual que en otras locuciones ninguna de cuyas dos variantes es frecuente, como *levantar de cascos / levantar los cascos*; en efecto, en el CREA no hay ningún ejemplo de *levantar de cascos* que corresponda a la acepción definida por Seco: ‘Alborotar[le] o inquietar[le] por la atracción sexual’, mientras que en el CORDE se encuentran tres, el último de 1927; por lo que respecta a *levantar los cascos*, no hay ningún registro en ninguno de los dos corpus de la Academia; los datos son suficientes para considerarla poco frecuente, aunque no presenta la marca *raro* en el DFDEA.

Los ejemplos sobre esta cuestión podrían multiplicarse, de ahí la necesidad de estar alerta acerca de la posibilidad de que dos variantes no marcadas de una locución tengan distinta frecuencia y, por supuesto, de emprender análisis que la certifiquen, pues es muy probable que una de ellas, al no tener valor distintivo por ser ambas variantes no marcadas, quede abandonada. La explicación puede encontrarse en las causas del cambio lingüístico tal como las entiende Coseriu (1978: 203-204): la libertad lingüística de los hablantes abandona aquello que resulta prácticamente inútil desde el punto de vista funcional. De este modo, un significante de una locución (una variante no marcada) que no sirve para distinguir, que resulta inútil, se abandona.

4.2. La colisión homonímica entre una locución y una combinación no fraseológica de palabras

El interés por estudiar la homonimia y, especialmente, las colisiones homonímicas entre unidades léxicas simples se hace patente ya en la semántica tradicional, que se ocupó, de manera hasta cierto punto extensa, de esta cuestión (Ullmann, 1983: 176-192; Ullmann, 1986: 295-317). Pero la vinculación de la ciencia que estudia el significado al nivel de la palabra, con exclusión de los otros niveles lingüísticos y de las unidades fraseológicas equivalentes a las palabras, también, claro está, el retraso en la constitución de la fraseología como disciplina lingüística independiente, impidieron que las unidades fraseológicas, al menos las locuciones, fueran examinadas en relación con las posibles colisiones homonímicas que presentan, no ya entre dos unidades fraseológicas que coinciden en la forma y difieren en el significado, sino entre una combinación de palabras no fraseológica y otra que lo es. Sin embargo, el fenómeno se da, y con repercusiones en la frecuencia de uso de las locuciones, como se mostrará a lo largo de este apartado. Pero antes va a presentarse el fenómeno de la homonimia entre dos combinaciones de palabras, una fraseológica y la otra no, en los términos en que la semántica tradicional lo ha hecho para las palabras.

Reconocida la existencia de la homonimia como una forma de ambigüedad, se hace necesario indicar que los homónimos relativos a combinaciones de palabras, al igual que los homónimos léxicos, no causan problemas en el uso de la lengua al menos por la existencia de dos salvaguardas que limitan las posibilidades de confusión: el contexto y la pertenencia a clases distintas, entre las distintas salvaguardas o soluciones terapéuticas delimitadas por Ullmann (Ullmann, 1983: 181-185; Ullmann, 1986: 296-302) para diferenciar las palabras homónimas. En el caso que aquí se está analizando, puede entenderse por contexto la distinta valencia que una locución y una combinación de palabras no fraseológica presentan, es decir, el distinto número de actantes y su diferente valor sintáctico-semántico. Esta

salvaguada tiene especial incidencia sobre las locuciones verbales, pues con excepción de las avalentes, el resto de clases, como los verbos, presentan la particularidad de la valencia. Así, la locución *abrir la puerta*, citada anteriormente, exige los actantes *alguien / algo abre la puerta a algo*, frente a la combinación homónima no fraseológica, que, aunque puede construirse con un actante sujeto correspondiente a un *alguien* o a un *algo*, no exige el complemento *a algo*: *alguien / algo abre la puerta*:

- (17) a **El PSOE**, junto a un plan completo, con nuevos impuestos y una optimista previsión de lucha contra el fraude, **abre la puerta a una bajada de impuestos a las clases más desfavorecidas y a los parados**.
- (17) b **Letona** recoge sus cosas, se levanta, **abre la puerta** y se va.

Igual ocurre con otras locuciones como *cubrir la espalda*, *darse patadas* o *echar gasolina*²² en relación con su correspondiente combinación de palabras homónima no fraseológica. Los actantes de unas y otras son: *alguien cubre la espalda a alguien / algo cubre la espalda de alguien* o *algo cubre la espalda a alguien*, *dos cosas se dan patadas* o *una cosa se da patadas con otra / dos personas se dan patadas*, *alguien echa gasolina / alguien echa gasolina a algo*:

- (18) a **Nadie le cubre la espalda** cuando sale a marcar y ya hay un jugador, el checo Poborsky, que puede decir que lo superó ampliamente.
- (18) b *El impresionante tatuaje que cubre la espalda de Rajon Rondo. El base de los Boston Celtics Rajon Rondo ha decidido cubrir su espalda con este llamativo tatuaje con su logo: una erre montada sobre otra.*
- (18) c **Una larga capa sostenida por un collar de plumas y cadenas plateadas le cubre la espalda**.
- (19) a **Espejos granadinos y sillones Luis XV talaveranos convivían dándose patadas**²³.
- (19) b *Recuerdo que, en la hora del patio, había niños que estaban todo el rato dándose patadas.*
- (20) a *cuando lleguen a la Campana aprovecharán para salir a «echar gasolina», beber un trago en cualquier bar*²⁴.
- (20) b *¿Es bueno echarle gasolina o aceite al gasoil cuando hace frío?*

Lo que no significa que en todos los casos la valencia actúe como elemento que desambigua una locución frente a una combinación de palabras homónima no fraseológica, pues también existen ejemplos en que el número y la naturaleza de los actantes son los mismos en ambas: *alguien da en la cabeza a alguien*, *alguien da una charla a alguien*, *alguien mete en la cama a alguien*²⁵:

- (21) a *hay que seguir y seguiremos dándoles en la cabeza; y hay que combatirlos hasta acabar con ellos.*

22 Los significados de las locuciones son, respectivamente, ‘Proteger a una persona contra un riesgo o una eventualidad’ (DILEA), ‘Combinar muy mal’ (DFDEA) y ‘Echar un trago (tomar bebidas alcohólicas)’ (DFDEA).

23 Ejemplo tomado del DFDEA (2004: 753).

24 El ejemplo procede del DFDEA (2004: 490).

25 Los significados de las locuciones en el DILEA son, respectivamente: ‘Contrariar las pretensiones de una persona, procurar que no se cumplan’, ‘Reprender a una persona’ y ‘Hacer que realice el acto sexual’.

- (21) b *uno de los hombres lo sujetó por el cuello, y otro, con la tranca, le dio en la cabeza, y lo golpeó de tal modo, que lo dejó muerto.*
- (22) a «*La segunda operación coincidió con un Everton-Chelsea*», recuerda, «*al despertarme tenía al lado en la sala de recuperación a un aficionado del Everton. Él tenía la radio puesta y me iba dando el resultado. Luego otro me dio una charla por los pasillos: ‘El otro día aposté por vosotros y perdí dinero’. ¡Y yo allí medio mareado!*».
- (22) b *Me pidió que si iba hacia Cibeles le dejara en el Gijón. No pude negarme. Por el camino me dio una charla interminable sobre la libertad.*
- (23) a *El marido dice: «yo, cuando la veo triste, lo que le digo es «ven aquí, cariño» y la meto en la cama».*
- (23) b *A esta edad tu pequeño será más reactivo a irse a la cama, y la clave aquí radica en ser firme, coherente y seguir siempre el horario y el ritual de cada día [...] Mételo en la cama cuando esté despierto, evita que se duerma en el salón o en otra cama.*

En cuanto a la pertenencia a clases de palabras distintas como salvaguarda contra la homonimia según Ullmann (Ullmann, 1983: 182-183; Ullmann, 1986: 303-304), si, en principio, evita el riesgo de confusión entre homónimos léxicos, pues resulta muy improbable que un verbo (*bajo*) ocupe la misma posición que una preposición (*bajo*) o un adjetivo (*bajo*), en ocasiones puede dejar de funcionar, como ejemplifica este autor para el sustantivo francés *part* y la preposición *par* de la misma lengua. También una locución verbal y una combinación de palabras no fraseológica homónima de aquella, cuyo elemento nuclear sea un verbo, son, de hecho, unidades de clases distintas: sintagma fijo / sintagma no fijo, pero su equivalencia funcional en una oración: ser el predicado, contrarresta esta potencial salvaguarda, de manera que solo queda la interpretación del enunciado para dilucidar si se está ante un significado idiomático, el de la locución, o uno literal, el de la combinación de palabras no fraseológica. En consecuencia, cabría añadir el tipo de significado: fraseológico / literal, a la relación de salvaguardas enunciadas por el semantista citado.

Si aun a pesar de las salvaguardas, se produce una colisión homonímica, la consecuencia en los homónimos léxicos es la desaparición de uno de ellos y su sustitución por otro elemento, efectuada de múltiples maneras. Conocer hasta qué punto los modos de llenar el vacío creado por la desaparición de una palabra homónima son aplicables a las locuciones exige unos análisis diacrónicos de estas unidades de los que, por el momento, carecemos. Por otra parte, aquí no interesa averiguar si una locución verbal, homónima de una combinación no fraseológica de palabras, ha desaparecido del uso lingüístico en un estado de lengua, sino constatar su baja frecuencia, que puede estar determinada, justamente, por la existencia de la combinación de palabras no fraseológica.

En efecto, en los siguientes pares de ejemplos, están documentadas una combinación de palabras que no corresponde a una locución y una unidad fraseológica de esta clase, en concreto *irse al agua*, ‘Fracasar’ (DILEA), *llevar la cesta*, ‘Acompañar a una pareja de enamorados, esp. cuando se muestran afecto’ (DFDEA), *colocar el disco*, ‘Exponer un tema conocido y que resulta pesado, esp. por su reiteración’ (DFDEA), y *meter las manos*, ‘Intervenir’ (DILEA):

- (24) a *Krieger la cogió en brazos y se fueron al agua.*
- (24) b *Sentía que éramos [...] bastante felices, pero nuestro amor se fue al agua.*
- (25) a *De pronto, mi hermano se volvió para ofrecerse a llevar la cesta de la merienda, como si hasta aquel momento no hubiera reparado en ella.*

- (25) b ¡**Llevar la cesta!** *Eso ya no se estila. Era de nuestros tiempos, cuando un beso era pecado.*
- (26) a **Coloque el disco** *en su caja protectora, si la tiene, e insértelo en la unidad.*
- (26) b –*De este permiso no pasa el que la **coloque el disco** –decía un combatiente enamorado mientras trazaba sus planes respecto a la señora de sus pensamientos²⁶.*
- (27) a *Cuando volví a **meter las manos** en el agua, el extraño ser dio un salto hacia atrás y estuvo a punto de perder el equilibrio.*
- (27) b *No puedo dejar de responder a todo, a toda cosa en la que se pueda **meter las manos**. Y no es deseo de mangonear, no, es que esa emoción, ese arrebató que me produce la perfección –la vista o la imaginada– lo sufro por las cosas más increíbles.*

De ninguno de los casos anteriores puede afirmarse, con absoluta seguridad, que se ha producido la desaparición del homónimo fraseológico ni que esta se haya compensado de las maneras que se ha visto para las palabras (Ullmann, 1983: 186-188; Ullmann, 1986: 309-313), pero sí es posible considerar que la colisión homonímica repercute en la baja frecuencia de uso de las locuciones, y existen muchos más casos, de manera que no parece cumplirse la afirmación de Martinet (1955: 183) acerca de que los conflictos homonímicos reales son mucho más raros que los teóricos, aunque este lingüista se esté refiriendo al francés. Así, la consulta del CREA muestra que de la combinación literal *irse al agua* o *ir al agua* hay cinco ocurrencias, frente a las tres de la locución homónima. Lo cual, sin embargo, no supone ninguna merma en las posibilidades de expresión para los hablantes, dado el amplio número de locuciones sinónimas y de variantes con las que se puede expresar el significado de ‘Fracasar’²⁷. No se ha llegado, pues, a producir la desaparición de la locución homónima *irse al agua*, pero su lugar está siendo ocupado por un amplio conjunto de sinónimos y variantes, lo que repercute en su baja frecuencia, tres ocurrencias en un corpus de 160 millones de formas frente a cinco de la combinación de palabras no fraseológica.

En cuanto a la locución *llevar la cesta*, en el CREA no se encuentra ningún ejemplo que la documente, las seis ocurrencias que aparecen son de la combinación de palabras homónima no fraseológica, igual que en la red, donde, no obstante, buscando con el verbo en infinitivo, se encuentra el ejemplo anotado anteriormente. El DFDEA, que sí registra la locución, la marca indicando *hoy raro*; por otra parte, está recogida en el DRAE, pero no en el DUE ni en el GDUEA, y tampoco en diccionarios fraseológicos como el DFEM o el DDEE²⁸. Posiblemente, la baja frecuencia de uso de esta locución encuentre su razón de ser, no tanto en la colisión homonímica con la combinación de palabras no fraseológica, cuanto en los cambios producidos en la sociedad en relación con los usos y comportamientos en las relaciones amorosas, que ya no exigen la presencia de una carabina que lleve la cesta. En cualquier caso, resulta probada su baja frecuencia de uso.

26 El ejemplo está tomado del DFDEA (2004: 403).

27 Véase este listado: *dar el gatillazo, dar gatillazo, hacer agua, hacer aguas, irse a freír espárragos, irse a freír monas, irse a hacer gárgaras, irse a hacer puñetas, irse a la eme, irse a la m, irse a la mierda, irse a la porra, irse a paseo, irse a tomar por culo, irse a tomar por el culo, irse a tomar viento, irse al carajo, irse al cuerno, irse al demonio, irse al diablo, irse al garete.*

28 El autor de este último es, a la vez, coautor del DFE, que, en cambio, sí incluye la locución *llevar la cesta*.

Peor es la situación de la locución *colocar el disco*. El CREA tampoco registra de ella ninguna ocurrencia, y las cuatro que existen corresponden a la combinación de palabras homónima de la locución, combinación que no es fraseológica, pues tiene un significado literal a partir del verbo *colocar* y del sustantivo *disco*. Por otra parte, la búsqueda a través de *Google*, con distintas formas verbales de *colocar*, tampoco ofrece resultados positivos. Además, la lengua no dispone de sinónimos de la locución que puedan compensar la colisión homonímica con la combinación no fraseológica. Todo esto explicaría que distintos diccionarios del español, generales y fraseológicos, no la recojan²⁹. Sin tener ningún afán adivinatorio, se podría prever la desaparición de esta unidad fraseológica.

El caso de *meter las manos* presenta particularidades distintas a las de las locuciones anteriores, pues junto a ella está la variante no marcada *meter mano*. La búsqueda en el CREA de la cadena *meter las manos*, en infinitivo, para facilitar la localización de ocurrencias, ofrece solo una de esta locución, la que figura en (27), frente a ocho de la combinación homónima no fraseológica. Por su parte, de la otra variante, *meter mano*, consultándola también en infinitivo, se encuentran nueve ejemplos. Debe pensarse, pues, que esta última variante le está ganando terreno a *meter las manos*, y la explicación puede estar en la homonimia de esta última con la combinación de palabras no fraseológica.

De lo expuesto en este apartado no cabe concluir, necesariamente, una relación de causa-efecto entre la colisión homonímica de locuciones y combinaciones no fraseológicas de palabras, y la baja frecuencia de las locuciones; en otras palabras, no se está afirmando aquí que la colisión homonímica sea la causa que provoca el efecto del escaso uso de determinadas locuciones, aunque desde un punto de vista teórico nada obsta para que así sea, pues, como apunta Coseriu en el análisis de las causas del cambio lingüístico, si un significante no se distingue de otro con significado distinto (aquí una unidad fraseológica homónima de una combinación no fija de palabras) se modifica o se sustituye. La existencia de variantes de las locuciones *irse al agua* y *meter las manos*, que también son poco frecuentes, son prueba de que estas locuciones se están modificando; por su parte, la baja frecuencia de *colocar el disco* y *llevar la cesta* indican que pueden encontrarse en un proceso de sustitución que, claro está, no es puntual sino gradual.

En cualquier caso, lo que se ha intentado mostrar, simplemente, es que la baja frecuencia en la utilización de ciertas locuciones va acompañada, para muchas de ellas, de la existencia de combinaciones de palabras no fraseológicas, homónimas de las locuciones. Que la relación entre colisión homonímica y baja frecuencia sea casual o causal no tiene interés desde el punto de vista aquí adoptado, el lexicográfico. Sí tiene, sin embargo, implicaciones para la labor lexicográfica, pues la cuestión examinada en este apartado y las presentadas en los anteriores repercuten en la confección de un diccionario de locuciones.

5. Implicaciones lexicográficas

En efecto, del apartado 2. se deduce que existen locuciones actuales, es decir, pertenecientes a un estado de lengua de límites difusos para la teoría lingüística, desde la

²⁹ Es el caso, por ejemplo, del DRAE, del DUE, del GDUEA, del DFEM o del DFE.

perspectiva de los investigadores y desde la percepción de los hablantes, las cuales, no obstante, deben marcarse como poco frecuentes en un diccionario que las recoja por su escasa representación en corpus de textos del periodo considerado. La marcación sería inexcusable realizarla si se opta por la inclusión de estas locuciones en la macroestructura del diccionario, pero también cabe su exclusión. En uno y otro caso se plantea el problema de qué criterio o criterios seguir para marcar como poco frecuentes o excluir determinadas locuciones. ¿Es suficiente su no documentación en el CREA? ¿Debe acompañarse este criterio del examen del CORDE, al menos desde 1900 hacia delante? ¿Cuántos ejemplos de una locución en el CREA y en el CORDE determinan su baja frecuencia o su exclusión? ¿Qué peso debería darse a las ocurrencias de una locución en relación con la cronología de los documentos que la registran? ¿Qué número de registros de la locución debería documentarse en la red para considerarla de baja frecuencia, en caso de no estar en los corpus de la Academia?

Es fácil imaginar la complejidad de dar respuesta exacta a estas interrogantes y la amplia casuística que se puede derivar de su aplicación potencial al conjunto de locuciones de una lengua. La tarea de su resolución parece titánica y el esfuerzo que conlleva puede ser descorazonador. De ahí la necesidad de tomar unas decisiones que, aunque provisionales y discutibles, faciliten la labor de redacción de un diccionario, la cual nunca se llevaría a cabo si hubiera que esperar a tener todas las locuciones completamente analizadas.

Partiendo de que el CREA es un corpus de 160 millones de formas, el CORDE cuenta con 250 millones de registros y la red, aunque limitada al español de España, es un universo, un mundo lingüístico, cabe pensar que encontrar menos de tres ocurrencias de una locución tras la consulta de los dos corpus académicos y de la red³⁰ es indicio de una baja frecuencia. Aunque no disponemos de datos para fijar el umbral a partir del cual se puede empezar a considerar la frecuencia de una locución. Piénsese que el GDUEA, uno de los pocos diccionarios de la lengua española que ofrecen información relativa a la frecuencia de uso de las voces registradas, ha fijado el valor para acceder al nivel más bajo de frecuencia en cinco ocurrencias sobre el total del corpus *Cumbre* utilizado para elaborar el diccionario: 20 millones de palabras (GDUEA, 2001: 10). Pero debe quedar claro que se trata de la frecuencia de uso de las unidades léxicas simples, no de las locuciones registradas en él.

En consecuencia, la locución para la que no se hayan obtenido tres ejemplos podría o bien ser excluida de un diccionario de locuciones del español actual o bien ser marcada como poco frecuente. Las que se corresponden con esta situación, a tenor de lo observado tras el vaciado del DFDEA, son numerosas, pero baste este pequeño conjunto para probarlo; todas van en la forma y con el significado que proporciona el DFDEA: *no caber la menor* ('No caber la menor duda'), *no descubrir la pólvora* ('Se usa para ponderar la falta de inteligencia o perspicacia'), *no echarse nada en el bolsillo* ('No obtener ningún beneficio económico'), *no haber cosa con cosa* ('No existir relación entre una cosa y otra de las mencionadas'), *no haber (ni) para un diente* ('Ser muy escasa o insuficiente'), *no oírse más allá del cuello de la (o su) camisa* ('Hablar en voz muy baja'), *no poner los pies en el suelo* ('Correr o caminar muy rápidamente'), *no quedar ni el apuntador* ('Producirse

30 Recuérdese que la consulta de la red se limita a las diez primeras páginas que ofrece de la cadena objeto de búsqueda.

muchas muertes’) y *no tener más que pellejo* (‘Estar sumamente delgado o consumido’)³¹. No deja de ser curioso, sin embargo, que solo una: *no haber cosa con cosa*, se marca como *raro* en el DFDEA.

Por el contrario, tener ampliamente documentada una locución en relación con otras, después del vaciado de un conjunto de textos orales y escritos tomados como fuentes lingüísticas, puede ser criterio para marcarla como muy frecuente. En este sentido, aunque las fuentes utilizadas para el DILEA no están cerradas, pues se van incrementando paulatinamente³², es posible afirmar que, frente a otras locuciones todavía no documentadas a partir de tales fuentes, se cuenta ya con un conjunto de ejemplos relativamente amplio para locuciones nominales como *callejón sin salida*, *canto de sirena*, *el más allá*, *idas y venidas* o *ir y venir*; adjetivas como *de carne y hueso*, *de derechas*, *de mala muerte* o *de medio pelo*; verbales como *darse cuenta*, *echar de menos*, *echar una mano*, *sacar pecho* o *salirse con la suya*; y adverbiales como *a bocajarro*, *a cal y canto*, *a estas alturas*, *contra viento y marea*, *de buenas a primeras* o *de vez en cuando*.

Lo expuesto incide sobre la necesidad de marcar las locuciones en un diccionario en relación con su frecuencia de uso, si no la elevada —que puede tener interés en relación con cuestiones más específicas como la enseñanza de la lengua—, sí la baja, para que el usuario del diccionario sea consciente de esta particularidad que condiciona, al menos, la comprensión de muchas locuciones. En general, los diccionarios específicos de esta clase de unidades fraseológicas no establecen ni utilizan marcas para la frecuencia. Una excepción la

31 Como se ha indicado, todas ellas están registradas en el DFDEA, aunque ilustran distintas situaciones. En cuanto a *no haber la menor*, solo se documenta en el CORDE en dos ejemplos de 1927; en la red, mediante la búsqueda de la cadena «no cabe la menor», no se encuentra ningún ejemplo en que no aparezca el sustantivo *duda*; téngase en cuenta que esta locución parece haberse formado por la elipsis de *duda* en *no haber la menor duda*. En los corpus académicos solo se registra un ejemplo, en el CREA, de *no descubrir la pólvora*; en la red no se encuentra ninguno; por otra parte, no la recoge el DUE, el GDUEA ni el DFEM. Los corpus de la Academia no documentan ningún ejemplo que se corresponda con el lema *no echarse nada en el bolsillo*; la búsqueda en la red solo ofrece una ocurrencia de esta locución. De *no haber cosa con cosa* se documentan diez ejemplos en el CORDE, aunque todos de Delibes y de la misma obra, *Diario de un emigrante* (1958); en la red no se encuentra ningún ejemplo adecuado; no la registran el DUE, el GDUEA ni el DFEM. La Academia no documenta ningún ejemplo de *no haber (ni) para un diente*; en la red se obtiene una ocurrencia buscando la cadena «ni para un diente»; el ejemplo del DFDEA va precedido de un asterisco, lo que significa que la falta de testimonio escrito se ha suplido por uno oral. Los corpus académicos no proporcionan ningún ejemplo de *no oírse más allá del cuello de la (o su) camisa* y la red no ofrece ningún ejemplo adecuado. La locución *no poner los pies en el suelo* no está documentada ni en el CREA ni en el CORDE; en la red aparece una ocurrencia y, además, ejemplos de esta locución en *El Quijote*, más otros correspondientes a la combinación de palabras homónima no fraseológica; por otra parte, el ejemplo del DFDEA no corresponde a ningún documento escrito al ir precedido de un asterisco. Tampoco existen ocurrencias de *no quedar ni el apuntador* en la Academia; la búsqueda en la red solo proporciona dos ejemplos. Por último, en el CREA y el CORDE no se encuentran ocurrencias de *no tener más que pellejo*; los ejemplos que se obtienen usando *Google* corresponden a la combinación de palabras homónima no fraseológica; y, como en otros casos, el ejemplo de uso del DFDEA no ha sido extraído de ningún documento escrito, parece, más bien, creado *ad hoc*: «*Esta niña no tiene más que pellejo».

32 En estos momentos están constituidas por cuarenta y ocho textos escritos de distinta extensión, correspondientes a novelas y ensayos, que han sido vaciados al completo en relación con las locuciones nominales, adjetivas, pronominales, verbales y adverbiales que contienen; los periódicos EL PAÍS y EL MUNDO, pues su lectura ocasional pero frecuente es aprovechada para vaciar todas las locuciones de las clases anteriores de cada ejemplar leído; y seis corpus de textos orales (los recopilados para las investigaciones expuestas en Cestero Mancera (2000), Blanco Canales (2004) y Ruiz Martínez (2003); el de Briz y Grupo Val.Es.Co (2002); y los del PRESEEA publicados en Cestero Mancera, Molina Martos y Paredes García (2012), y Moreno Fernández, Cestero Mancera, Molina Martos y Paredes García (2002)), asimismo vaciados.

constituye el DFDEA, que usa *frec* (frecuente o frecuentemente) para señalar construcciones frecuentes en las que aparece la unidad fraseológica de la que se aporta esta información, no para referirse a la propia frecuencia de esta, y *raro* u *hoy raro* para indicar la cuestión que se analiza en este artículo, pues se entiende que la unidad fraseológica que se marca así es una «Expresión infrecuente; en especial, expresión anticuada empleada hoy solo de manera ocasional» (DFDEA, 2004: XXXI), pero ya se ha visto a lo largo de todo el trabajo la necesidad no tanto de revisar las así marcadas, cuanto de asignar la marca a muchas que no la presentan.

También REDES, aunque no es un diccionario de unidades fraseológicas, sino de combinaciones, proporciona marcas de frecuencia no tanto de las locuciones que registra cuanto de la combinación de una locución con determinadas palabras, pero, vaciando las locuciones contenidas en este diccionario, se obtendría un conjunto de locuciones y de combinaciones frecuentes de ellas que, de manera indirecta, mostraría que aquellas no tienen una baja frecuencia de uso. Por ejemplo, *caer en saco roto* es una locución verbal registrada en REDES que presenta una combinación sumamente frecuente con los sustantivos *consejo*, *advertencia* y *palabra*, y bastante frecuente con *propuesta*, *iniciativa*, *petición*, *demanda*, *denuncia*, *promesa*, *idea*, *reflexión* y *esfuerzo*. De estos datos cabría inferir que la locución verbal *caer en saco roto* es, al menos, frecuente.

Por último, el DELE ha hecho un uso completamente sistemático de marcas como pf (poco frecuente), f (frecuente) y mf (muy frecuente), hasta el punto de que todas las locuciones recogidas aparecen con alguna de las tres marcas indicadas, pero desconocemos los criterios que han llevado a sus autores a asignarlas a locuciones como *dar/meter caña*, que tiene la marca (mf); *hacer sus primeras armas*, con la marca (pf); *tragárselas dobladas*, marcada como (f); *(un) animal de bellota(s)*, cuya marca es también (mf)³³, además de la falta de información acerca de los datos que les han conducido a considerar una locución frecuente en vez de muy frecuente, o al revés, y otra frecuente, pero no poco frecuente, o lo contrario³⁴.

6. Conclusiones

De este trabajo se desprenden una serie de conclusiones que se desglosan a continuación. 1.^a) El término *actual* no se restringe necesariamente a los últimos cincuenta o veinticinco

33 Téngase en cuenta, sin embargo, que, como ya se ha indicado, las ocurrencias de *dar/meter caña* difieren mucho en el CREA, de la primera variante de esta locución, que tiene seis acepciones distintas, hay un total de cincuenta y seis, mientras que de la segunda solo hay siete, no se puede, en consecuencia, asignar la misma marca de frecuencia a las dos variantes; de *hacer sus primeras armas* existen doce registros en el CORDE y diez en el CREA, lo que incide en la necesidad de tener establecido un criterio para determinar qué locuciones son poco frecuentes; *tragárselas dobladas*, a pesar de su marca de frecuente en el DELE, solo tiene un registro en el CORDE y otro en el CREA; *(un) animal de bellota(s)* está documentada en el CORDE, ya sea bajo la forma *animal de bellota* o *un animal de bellota*, con siete ocurrencias y en el CREA, bajo esas mismas formas, con cinco ejemplos, pero en ninguno de los dos corpus se registra *animal de bellotas* o *un animal de bellotas*, lo que prueba lo indicado en este trabajo: que no todas las variantes de una locución tienen la misma frecuencia, de manera que no se puede marcar *(un) animal de bellotas* como muy frecuente sin más. Puesto que es relativamente fácil obtener ejemplos del DELE de locuciones cuya marca de frecuencia contrasta con su documentación en los corpus académicos, baste, pues, con los presentados.

34 Añádase a todo esto que tampoco los diccionarios generales de lengua dan un tratamiento sistemático y coherente a este aspecto de las locuciones.

años, pues los límites de un estado de lengua no son fáciles de fijar, además de estar sometidos a modificación continua, por el paso del tiempo, cuando el estado de lengua es el actual. La imposibilidad de establecer claramente los límites de un estado de lengua y la existencia en él de unidades —en el caso que aquí interesa fraseológicas— que para la conciencia de los hablantes, y como fruto de su reflexión metalingüística, pueden calificarse de arcaicas, antiguas o viejas conduce a considerar que determinadas locuciones pueden verse como actuales, si se amplía esta sincronía hasta principios del siglo XX.

2.^a) El análisis de la frecuencia de las locuciones se ve dificultado por cuestiones todavía pendientes de resolución en fraseología, tales como la pertenencia de las unidades fraseológicas a una u otra clase y su naturaleza polisémica u homónima *per se* o en relación con una combinación de palabras no fraseológica. Sería necesario, pues, que desde la fraseología quedaran bien delimitados estos aspectos de los fraseologismos para alcanzar resultados completamente satisfactorios en el estudio de su frecuencia, el cual, además, todavía no puede apoyarse de manera absoluta en los instrumentos que facilitan las nuevas tecnologías de la información.

3.^a) Al analizar la frecuencia de las locuciones, el investigador debe prestar especial atención a aquellas que cuentan con variantes, pues existe la posibilidad de que dos variantes no marcadas de una locución tengan distinta frecuencia. Es altamente probable que una de ellas, al no tener valor distintivo por ser ambas variantes no marcadas, quede abandonada, sea muy poco frecuente. La libertad lingüística de los hablantes deja de lado aquello que resulta prácticamente inútil desde el punto de vista funcional. De este modo, un significante de una locución (una variante no marcada) que no sirve para distinguir, que resulta inútil, acaba por no ser usado.

4.^a) La baja frecuencia en la utilización de ciertas locuciones va acompañada, para muchas de ellas, de la existencia de combinaciones de palabras no fraseológicas, homónimas de las locuciones. Este hecho no implica una relación de causa-efecto entre la colisión homónima de locuciones y combinaciones no fraseológicas de palabras, y la baja frecuencia de las locuciones. Dicho de otro modo, no se debe concluir que la colisión homónima sea la causa que provoca el efecto del escaso uso de determinadas locuciones, aunque desde un punto de vista teórico nada obsta para que así sea, pues en el análisis de las causas del cambio lingüístico está comprobado que, si un significante no se distingue de otro con significado distinto (aquí una unidad fraseológica homónima de una combinación no fija de palabras), se modifica o se sustituye. Lo que sí cabe deducir, simplemente, es que el estudio de numerosos ejemplos muestra una relación entre la baja frecuencia de una locución y la existencia de una combinación de palabras no fraseológica homónima de la unidad fraseológica.

5.^a) Desde la perspectiva lexicográfica, hay una serie de puntos que deben tenerse en cuenta en la redacción de diccionarios de locuciones y en el tratamiento lexicográfico de las locuciones en diccionarios generales de lengua, todos ellos relativos a la frecuencia de uso de estas unidades fraseológicas: a) un diccionario de locuciones calificado de *actual* obliga a plantearse la posibilidad de excluir aquellas que presenten una baja frecuencia de uso; b) no obstante, una obra lexicográfica de tal naturaleza puede incluir locuciones anticuadas siempre que sean convenientemente marcadas en relación con su escaso uso; c) es necesaria, pues, una marca (*rara*, *poco frecuente*) para señalar esta particularidad de las locuciones que están escasamente documentadas; y c) la marcación de frecuencia tendrá que llevarse a cabo a partir de unos criterios previamente fijados y sistemáticamente aplicados tanto para establecer el umbral a partir del cual se puede empezar a determinar la frecuencia de una locución, como para diferenciar si aquella es baja o elevada.

Referencias bibliográficas

- Blanco Canales, A. (2004): *Estudio sociolingüístico de Alcalá de Henares*. Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.
- Boers, F. y Stengers, H. (2008): «A Quantitative Comparison of the English and Spanish Repertoires of Figurative Idioms». En F. Boers y S. Lindstromberg (eds.): *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*. Berlin / New York, Mouton de Gruyter, págs. 355-374.
- Briz, A. y Grupo Val.Es.Co (2002): *Corpus de conversaciones coloquiales*. Madrid, Arco/Libros.
- Casas Gómez, M. y Escoriza Morera, L. (2009): «Los conceptos de diastratía y diafasia desde la Teoría Lingüística y la Sociolingüística Variacionista». En M. V. Camacho Taboada, J. J. Rodríguez Toro y J. Santana Marrero (eds.): *Estudios de lengua española: descripción, variación y uso. Homenaje a Humberto López Morales*. Madrid / Frankfurt am Main, Iberoamericana / Vervuert, págs. 151-178.
- Castillo Carballo, M.^a A. (2000): «Lingüística de corpus y fraseología: algunas consideraciones sobre las locuciones verbales». En G. Corpas Pastor (ed.): *Las lenguas de Europa: estudios de fraseología, fraseografía y traducción*. Granada, Comares, págs. 95-126.
- Cestero Mancera, A. M.^a (2000): *El intercambio de turno de habla en la conversación (análisis sociolingüístico)*. Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.
- Cestero Mancera, A. M.^a, Molina Martos, I. y Paredes García, F. (2012): *La lengua hablada en Madrid. Corpus PRESEEA-Madrid (distrito de Salamanca), Volumen I. Hablantes de instrucción superior*. Alcalá de Henares, Universidad de Alcalá.
- Corpas Pastor, G. (1996): *Manual de fraseología española*. Madrid, Gredos.
- Corpas Pastor, G. (1998): «El uso de paremias en un corpus del español peninsular actual». En G. Wotjak (ed.): *Estudios de fraseología y fraseografía del español actual*. Frankfurt am Main / Madrid, Vervuert / Iberoamericana, págs. 365-390.
- Corpas Pastor, G. (2013): «Detección, descripción y contraste de las unidades fraseológicas mediante tecnologías lingüísticas». En I. Olza y E. Manero Richard (eds.): *Fraseopragmática*. Berlin, Frank & Timme, págs. 335-373.
- Coseriu, E. (1978, 3.^a ed.): *Sincronía, diacronía e historia. El problema del cambio lingüístico*. Madrid, Gredos.
- Ghezzi, M. (2012): «Creación de una base de datos para el estudio de las comparaciones estereotipadas y su explotación en la enseñanza de ELE». En M.^a I. González Rey (ed.): *Unidades fraseológicas y TIC*. Madrid, Instituto Cervantes, págs. 203-216.
- Hjemslev, L. (1974): *Prolegómenos a una teoría del lenguaje*. Madrid, Gredos.
- Kabatek, J. (2002): «Oralidad, proceso y estructura», *Pandora*, 2, 2, págs. 37-54.
- Larreta Zulategui, J. P. (2001): *Fraseología contrastiva del alemán y el español*. Frankfurt am Main, Peter Lang.
- López Roig, C. (2002): *Aspectos de fraseología contrastiva (alemán-español) en el sistema y en el texto*. Frankfurt am Main, Peter Lang.
- Martinet, A. (1955): *Économie de changements phonétiques. Traité de phonologie diachronique*. Berne, A. Francke.
- Mellado Blanco, C. (2004): *Fraseologismos somáticos del alemán*. Frankfurt am Main, Peter Lang.

- Mellado Blanco, C. (2012): «Optimización de los recursos TIC en la fraseografía del par de lenguas alemán-español». En M.^a I. González Rey (ed.): *Unidades fraseológicas y TIC*. Madrid, Instituto Cervantes, págs. 147-166.
- Mogorrón Huerta, P. (2012): «Explotación informática de una base de datos multilingüe de unidades fraseológicas». En M.^a I. González Rey (ed.): *Unidades fraseológicas y TIC*. Madrid, Instituto Cervantes, págs. 63-81.
- Moreno Fernández, F., Cestero Mancera, A. M.^a, Molina Martos, I. y Paredes García, F. (2002): *La lengua hablada en Alcalá de Henares. Corpus PRESEEA – ALCALÁ. I. Hablantes de instrucción superior*. Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.
- Moon, R. (1997): «Vocabulary Connections: Multi-word Items in English». En N. Schmitt y M. McCarthy (eds.): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge, Cambridge University Press, págs. 40-63.
- Penadés Martínez, I. (2004): «¿Caracterizan las locuciones el registro coloquial?». En M. Villayandre Llamazares (ed.): *Actas del V Congreso de Lingüística General, León, 5-8 de marzo de 2002*, III. Madrid, Arco/Libros, págs. 2225-2235.
- Penadés Martínez, I. (2006): «La información gramatical sobre la clasificación de las locuciones en los diccionarios». En M. Alonso Ramos (ed.): *Diccionarios y fraseología*. A Coruña, Universidade da Coruña, Servizo de Publicacións, págs 249-259.
- Penadés Martínez, I. (2008): «Proyecto para la redacción de un diccionario de locuciones del español». En E. Bernal y J. DeCesaris (eds.): *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)*. Barcelona, IULA, Universitat Pompeu Fabra, págs. 1379-1384.
- Penadés Martínez, I. (2010): «El *Diccionario de locuciones del español actual (DILEA)*». En C. Mellado, P. Buján, C. Herrero, N. Iglesias y A. Mansilla (eds.): *La fraseografía del S. XXI. Nuevas propuestas para el español y el alemán*. Berlin, Frank & Timme, págs. 191-210.
- Penadés Martínez, I. (2012a): «La variación en las locuciones a partir de materiales del PRESEEA (Barrio de Salamanca, Madrid)». En A. M. Cestero, I. Molina y F. Paredes (eds.): *La lengua, lugar de encuentro. Actas del XVI Congreso Internacional de la ALFAL (Alcalá de Henares, 6-9 de junio de 2011)*. Alcalá de Henares, Universidad de Alcalá, págs. 2081-2091.
- Penadés Martínez, I. (2012b): *Gramática y semántica de las locuciones*. Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.
- Penadés Martínez, I. (2013): «La imagen subyacente a las locuciones como criterio de marcación». En I. Olza y E. Manero Richard (eds.): *Fraseopragmática*. Berlin, Frank & Timme, págs. 23-47.
- Ruiz Gurillo, L. (1998): *La fraseología del español coloquial*. Barcelona, Ariel.
- Ruiz Martínez, A. M. (2003): *Estudio fonético del nordeste de la Comunidad de Madrid*. Alcalá de Henares, Servicio de Publicaciones de la Universidad de Alcalá.
- Saussure, F. de (1945): *Curso de lingüística general*. Publicado por Ch. Bally y A. Sechehaye con la colaboración de A. Riedlinger. Traducción, prólogo y notas de A. Alonso. Buenos Aires, Losada.
- Sevilla Muñoz, M. (2012): «Utilización de recursos en línea en la enseñanza/aprendizaje de traducción de unidades fraseológicas». En M.^a I. González Rey (ed.): *Unidades fraseológicas y TIC*. Madrid, Instituto Cervantes, págs. 283-298.

- Sosiński, M. (2010): «Aproximación a la fraseología de los inmigrantes (comparación de los materiales del corpus ISPIE y del corpus PRESEEA-Granada)», *Lengua y Migración / Language and Migration*, 2, 2, págs. 113-125.
- Trubetzkoy, N. S. (1987): *Principios de fonología*. Madrid, Cincel.
- Ullmann, S. (1983, reimp.): *Semantics. An Introduction of the Science of Meaning*. Oxford, Basil Blackwell
- Ullmann, S. (1986, 2.^a reimp.): *Introducción a la semántica francesa*. Traducción y anotación por Eugenio de Bustos Tovar. Madrid, C.S.I.C.
- Zuluaga, A. (1980): *Introducción al estudio de las expresiones fijas*. Frankfurt am Main, Peter D. Lang.

Diccionarios

- DDEE: Cantera Ortiz de Urbina, J. (2011): *Diccionario de dichos y expresiones del español. Su interpretación al alcance de todos*. Madrid, Abada Editores.
- DEA: Seco, M., Andrés, O. y Ramos, G. (1999): *Diccionario del español actual*. Madrid, Aguilar.
- DELE: Martínez López, J. A. y Jørgensen, A. M. (2009): *Diccionario de expresiones y locuciones del español*. Madrid, Ediciones de la Torre.
- DFDEA: Seco, M., Andrés, O. y Ramos, G. (2004): *Diccionario fraseológico documentado del español actual. Locuciones y modismos españoles*. Madrid, Aguilar.
- DFE: Cantera Ortiz de Urbina, J. y Gomis Blanco, P. (2007): *Diccionario de fraseología española: locuciones, idiotismos, modismos y frases hechas usuales en español [su interpretación]*. Madrid, Abada Editores.
- DFEM: Varela, F. y Kubarth, H. (1994): *Diccionario fraseológico del español moderno*. Madrid, Gredos.
- DILEA: Penadés Martínez, I. (en preparación): *Diccionario de locuciones idiomáticas del español actual*.
- DRAE: Real Academia Española (2001, 22.^a ed.): *Diccionario de la lengua española*. Madrid, Espasa Calpe.
- DSLE: Gutiérrez Cuadrado, J. (dir.) (1996): *Diccionario Salamanca de la lengua española*. Barcelona, Círculo de Lectores.
- DUE: Moliner, M. (1984, reimp.): *Diccionario de uso del español*. Madrid, Gredos.
- GDUEA: Sánchez, A. (dir.) (2001): *Gran diccionario de uso del español actual*. Alcobendas, Sociedad General Española de Librería.
- JAM: Calles Vales, J. y Bermejo Meléndez, B. (2001): *Jergas, argot y modismos. Recopilación de palabras y expresiones de uso cotidiano que no suelen aparecer en los diccionarios*. Alcobendas, Libsa.
- REDES: Bosque, I. (dir.) (2004): *Las palabras en su contexto. REDES. Diccionario combinatorio del español contemporáneo*. Madrid, Ediciones SM.