

Léxico gramaticalizado y lengua formulaica: algunas precisiones al enfoque léxico

Anna Sánchez Rufat - Universidad de Córdoba
asrufat@uco.es

Rebut / Received: 18-5-11

Acceptat / Accepted: 14-7-11 (provisional); 31-7-11 (final)

Resum. Lèxic gramaticalitzat i llengua formulaica: algunes precisions a l'enfocament lèxic. Per tal d'establir maneres eficaces d'abordar el lèxic a l'aula de llengües estrangeres i aconseguir la producció d'enunciats normativament correctes —sense caure en les sobregeneralitzacions pròpies de les regles gramaticals tradicionals— cal definir, en primer lloc, els diferents tipus de combinacions lèxiques que conformen l'àmbit de la interfície lèxico-sintàctica, els quals han de designar conceptes pròpiament lingüístics. En segon lloc, partint dels resultats dels treballs sobre llengua formulaica que subratllen l'omnipresència de blocs lingüístics que es recuperen directament de la memòria —que no són resultat de l'aplicació de regles gramaticals—, cal distingir les combinacions fruit de les restriccions sistemàtiques de les que no ho són, i explotar a l'aula el seu potencial combinatori.

Paraules clau: combinacions lèxiques, col·locació, restriccions lèxico-sintàctiques, ensenyament i aprenentatge de vocabulari, espanyol llengua estrangera.

Abstract. Grammaticalised lexis and formulaic language: some clarifications about a lexical approach. With the aim of establishing a more efficient way of dealing with vocabulary in the classroom (L2) which contributes to the production of correct lexical sequences —without falling into the overgeneralizations of the traditional grammar rules—, certain concepts concerning the lexical-syntax interface have been clarified and modified. Secondly, taking account of the finds in formulaic language which highlight the omnipresence of semi preconstructed phrases or lexical chunks in both, native language acquisition and production, the differences between the combinations based on systematic restrictions and those on lexical preferences will be established in order to make the most of their combinatory potential in the L2 classroom.

Keywords: lexical combinations, collocation, lexical-syntactic restrictions, teaching and learning vocabulary, Spanish as foreign language.

1. Introducción

Hasta no hace mucho tiempo, teníamos la visión en las aulas de lenguas extranjeras de que la gramática constituía el elemento vertebrador del conocimiento lingüístico; era esencial ensamblar el esqueleto primero, de modo que las palabras que hubieran de ser aprendidas pudieran ser encajadas en el lugar adecuado. Esta manera de proceder en las aulas refleja la tradición chomskyana, que fundamentaba el dominio de la lengua en una dicotomía: el control de reglas gramaticales por un lado, y el dominio de unidades léxicas por el otro.

Pese a que esta concepción del lenguaje estaba muy extendida en los años setenta, persistían en esta época los trabajos de tendencia estructuralista que prestaban atención a las propiedades imitadoras del lenguaje natural y a la lengua formulaica¹ (Stengers 2009) —donde la anterior dicotomía no tiene cabida—. Ya en 1933, Palmer defendía que las coapariciones frecuentes de palabras debían ser enseñadas como “an integral whole or independent entity, rather than by the process of piecing together their component parts” (Stengers 2009, p. 6, *apud*). Fillmore (1976, p. 9), uno de los fundadores de la gramática constructivista, afirmaba que “an enormously large amount of language is formulaic, automatic and rehearsed, rather than propositional, creative or freely generated”. Bolinger (1976, p. 2) manifestaba que “speakers do at least as much remembering as they do putting together”. En este nuevo paradigma de investigación léxica, Brown (1974), siguiendo a Firth (1957), introdujo la noción de *colocación* en el área de la didáctica de la lengua para destacar la importancia del contexto a la hora de determinar el significado de una palabra. Por su parte, Nattinger (1980) propuso que la enseñanza del vocabulario se centrara en las expresiones léxicas *multipalabra*, también denominadas *pluriverbales*, ignoradas hasta entonces debido a la prioridad concedida a las estructuras gramaticales.

Gracias a las contribuciones de estos y otros trabajos que constituyen los primeros estudios de tendencia lexicalista, unidas a las aportaciones de la lingüística de corpus, la lingüística del texto, la semántica léxica y el análisis del discurso, el léxico —y, en especial, las expresiones multipalabra de coaparición frecuente— acaba adquiriendo un papel fundamental en las teorías lingüísticas y de adquisición². Tras el reconocimiento,

1. Según la mayoría de los estudiosos, muchas y variadas son las unidades integradas en el término *lengua formulaica* —también denominado *expresiones multipalabra*, *colocación* en sentido amplio, o *bloques semi-construidos*— como marcadores del discurso (*por un lado*), locuciones (*tomar el pelo*), símiles estandarizados (*fresco como una lechuga*), proverbios (*a caballo regalado...*), exclamaciones (*¡hay que ver!*), fórmulas sociales (*¡que te mejores!*), citas (*ser o no ser*), esquemas fijos abiertos (*cuanto más* [presente subjuntivo], *más* [futuro]), combinaciones libres frecuentes (*reservar habitación*), colocación en sentido estricto (*cometer* + sustantivos que denotan error); etcétera. Todas estas combinaciones tienen en común la coaparición frecuente y la recurrencia automática y natural por parte de una comunidad de hablantes.

2. Los modelos de descripción de la lengua que se desarrollan en estos últimos años son sintagmáticos o léxicos, como Pattern Grammar (Hunston y Francis 2000) y Linear Unit Grammar (Sinclair y Mauranen 2006). Una visión sintagmática de la lengua toma en consideración la relación entre sentido y sintaxis.

por un lado, de la importancia del léxico como componente esencial de la estructura lingüística, y, por otro, de la omnipresencia de la lengua formulaica, esta última ha sido propuesta como el elemento clave para lograr una producción nativa (Pawley y Syder 1983, Wray 2002). En consecuencia, es coherente que hayan surgido métodos que promuevan el aprendizaje de la lengua formulaica, entre los que sobresale el enfoque léxico de Lewis (1993, 1997, 2000).

Este estudio nace de los trabajos que destacan la presencia de la lengua formulaica a la hora de adquirir, procesar y usar la lengua³. No obstante, se pretende poner de manifiesto la dificultad para trabajar en el aula con este concepto de lengua formulaica —caracterizado por criterios difícilmente interpretables— que, a su vez, integra combinaciones léxicas de naturaleza muy diversa cuyo tratamiento en el aula y en los materiales didácticos es homogéneo, lo que resulta confuso e improductivo. A continuación, veremos en qué consisten estos conceptos y, de acuerdo con ello, todas nuestras reflexiones, incluso las más puramente teóricas, estarán determinadas por intereses didácticos.

2. Contribuciones de la lingüística del corpus: lengua formulaica y léxico gramaticalizado

A partir de los importantes hallazgos de los nuevos trabajos de análisis de corpus informático se ha podido llevar a cabo una mejor descripción de la lengua —de su naturaleza y de la forma en la que la procesamos⁴. Sinclair (1991, p. 110) explica el mecanismo del uso del lenguaje como la interacción de dos principios diferentes: el principio de idiomatidad y el de selección libre⁵. El primero describe que “a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments”, mientras que el principio de selección libre —el único que existía hasta ese momento— basa el proceso de producción lingüística en la elaboración de estructuras sintácticas —determinadas por unas reglas gramaticales— en las que se insertan las piezas léxicas. En cuanto a este

3. La gran cantidad de evidencia lingüística obtenida de los estudios de corpus en las dos últimas décadas (Stubbs 1995, Moon 1998) demuestra que la actuación comunicativa se fundamenta en gran parte en el almacenamiento y recuperación de secuencias formulaicas.

4. La primera publicación fruto de las pioneras investigaciones basadas en corpus y que abrió camino en cuanto a la cobertura de este tipo de trabajo, data de 1982; el trabajo titulado *Word Frequencies in British and American English* fue llevado a cabo por los lingüistas escandinavos K. Hoffland y S. Johansson, quienes compararon el trabajo de Brown —*The Brown Corpus*— con el que ellos acababan de concluir en el *LOB corpora*. Pese a su innovación, estas técnicas ya habían sido empleadas en un trabajo más revolucionario que implicaba un cambio en la concepción del lenguaje, sobre todo en la relación entre gramática y léxico; fue dirigido por Sinclair y concluido en 1970 —*The OSTI Report*—, pero no llegó a ver la luz hasta el año 2004. Sobre los primeros gérmenes, avances y progresos en la descripción del lenguaje como fruto de la aparición de la lingüística de corpus, *vid.* Sinclair *et al.* 2005.

5. El germen de estas ideas ya se encontraban en Fillmore y Bolinger, como se ha mencionado anteriormente.

segundo principio, nosotros, al igual que reconocen muchos sintactistas, no admitimos que en un acto de predicación pueda darse una ausencia de restricción semántica entre un predicado y sus argumentos⁶, pues lo que varía es la naturaleza de la restricción⁷. Con el paso del tiempo, Sinclair (Sinclair *et al.* 2005, p. XVIII) volverá a referirse al principio de selección libre y a aquellos primeros tiempos de los programas de concordancias con las siguientes palabras: “lexis was entirely concerned with word meaning and not with pattern”. Ahora, la dicotomía entre léxico y gramática se encuentra difuminada y se ha dado paso al nuevo concepto de léxico gramaticalizado.

Lewis (2000, p. 147) parte de estas consideraciones y explica en qué se basa esta nueva concepción de la lengua caracterizada por el léxico gramaticalizado —concepto fundamental de su enfoque: “language consists of many smaller patterns, which exhibit varying degrees of fixedness or generalisability, each based on a word; in a sense, each word has its own grammar”. Las posibilidades combinatorias de una palabra pueden ser de carácter más restringido —por tener un contenido intensional alto— o más generativo en aquellas palabras cuyo significado posee menos carga intensional. Así, por ejemplo, el verbo *entartimar* es muy restrictivo, ya que solo existe en tanto que vaya acompañando a su argumento *suelo*, que es el único complemento directo que admite el predicado y, por tanto, forma parte de la definición de dicho verbo. Por otro lado, el adjetivo *marrón* es mucho menos restrictivo que *castaño*, por lo que el campo extensional del primero es mucho mayor que el del segundo; así, mientras que no es correcto decir *su bonito pelo y ojos castaños apenas se lucen con *el traje castaño que se ha comprado*, sí es posible *su bonito pelo y ojos marrones apenas se lucen con el traje marrón que se ha comprado*. Y aun cuando en español se pueden *lavar y fregar los platos*, solo pueden ser lavados —y no fregados— *la conciencia, el alma, el honor, la honra, el nombre, la reputación y la imagen*. En este último grupo de palabras menos restrictivas se encuentran, lógicamente, muchas de las usadas con más frecuencia, como los llamados verbos soporte del tipo *hacer* o *dar*. Estas palabras frecuentes siempre se han estudiado dentro del vocabulario, pero la dificultad de muchas de ellas no se encuentra en su significado, como hemos visto en los anteriores ejemplos *marrón* y *castaño*, y *fregar* y *lavar*, sino en su combinatoria, en su gramática; de ahí que en el aula se deba evitar definir estas palabras aportando únicamente el sinónimo conocido por el alumno, dado que es raro que dos palabras sean sinónimos absolutos y, por tanto, intercambiables en todos los contextos. Esto nos devuelve a la conocida afirmación de Firth (1957, pag. 11): “you shall know a word by the company it keeps”; esto es, a la hora de aprender una lengua hay que tenerla presente siempre tal y como es empleada. El enfoque léxico promueve que las palabras no sean definidas ni enseñadas

6. Sobre la polémica distinción entre combinatoria libre y combinatoria restringida, *vid.* Bosque 2005, p. lxxxiii-lxxxiv.

7. En nuestro trabajo (Sánchez Rufat 2010) proponemos un modelo de clasificación de las combinaciones léxicas en función de la naturaleza de la restricción, que resumimos en el apartado 4 “Sobre la rentabilidad de las restricciones sistemáticas”.

de forma aislada y que no se marginen las expresiones frecuentes ni las combinaciones posibles en las que estas formas ocurren, ya que el alumno, en todo acto comunicativo, ha de juntar las palabras que previamente aparecían unidas en contexto pero que han sido separadas para ser definidas en el aula, en lo que es un esfuerzo contraproducente e inútil.

Esta concepción didáctica del léxico contrasta, en parte, con la del paradigma comunicativo vigente (PPP) en el que, a juzgar por un análisis de los libros de texto⁸, se espera que el aprendizaje de este componente lingüístico se lleve a cabo de forma autónoma e inconsciente. No obstante, las tareas de producción abierta —actividad fundamental del método— basadas en instrucciones estrictamente comunicativas no conllevan un desarrollo del léxico productivo. Dado que la prioridad del método reside en el éxito comunicativo, en detrimento de la precisión, a menudo se concede escasa importancia en el aula a los errores de producción que se corresponden con usos lingüísticos que, o bien no reflejan las restricciones del sistema, o no reproducen las sanciones normativas:

- (1) *Es necesario *hacer una decisión* ya
- (2) *Todas mis *esperanzas se han caído*
- (3) *Mis *ojos se pusieron poco claros* de repente, y me desmayé
- (4) *Es la tercera vez que *haces* el mismo *error*
- (5) *Transportaba al cadáver el *coche mortuorio*

Estos errores no se solucionan con una explicación gramatical tradicional. En el paradigma comunicativo, la corrección —si la hay— se limita a sustituir *tomar* por *hacer*, *derrumbado* por *caído*, *se nublaron* por *se pusieron poco claros*, *cometes* por *haces* y *fúnebre* por *mortuorio*. En cambio, el enfoque léxico va más allá y persigue que el alumno, de manera autónoma, aprenda a identificar del *input* las estructuras prefabricadas —a través de su actividad clave denominada *chunking* o segmentación— en lugar de ir analizando palabra por palabra; así, tendrá más facilidad a la hora de almacenarlas como una sola unidad en el lexicón y recuperarlas de igual manera en la producción. No obstante, como veremos, todavía se puede dar un paso más en el desarrollo de la competencia léxica, y más concretamente, en el de la subcompetencia combinatoria⁹, si se quiere explotar las posibilidades que ofrece profundizar en el análisis de la palabra, en su comportamiento léxico-sintáctico.

8. En los libros de texto comunicativos consultados no se trata el léxico de forma sistemática, la programación destina escasa carga horaria al léxico, la estrategia de aprendizaje más extendida es la inferencia de significados en contexto y las actividades, cuando las hay, requieren muy poco esfuerzo cognitivo.

9. Entendemos por esta subcompetencia la capacidad de usar las combinaciones sistemáticas y la de saber distinguir las combinaciones que podrían haber sido sancionadas por la comunidad de hablantes de las que realmente han acabado alcanzado ese rango. Precisamente, este conocimiento es el que caracteriza el dominio lingüístico.

3. Hacia una enseñanza —explícita¹⁰— de la lengua formulaica

Si se tiene en cuenta que se estima que alrededor del 50% del discurso nativo está basado en estructuras prefabricadas (Erman y Warren 2000) y que la importancia del dominio de estas secuencias por parte del aprendiz está reconocida (Pawley y Syder 1983, Skehan 1998, Wray 2002), es lógico, como ya se ha mencionado, que hayan surgido enfoques didácticos (como el de Lewis) que se fundamenten en estas nuevas teorías lingüísticas —fruto, en gran medida, de los análisis de corpus— y de adquisición, y que, por lo tanto, promueven la enseñanza de las expresiones multipalabra. La necesidad y las ventajas de su enseñanza son evidentes si partimos de que son imprescindibles para conseguir un dominio nativo. Por un lado, al ser almacenadas y recuperadas como una sola unidad —tanto en L1 como en L2—, su procesamiento es holístico, esto es, contribuyen al conocimiento basado en la memoria, que se aplica para codificar y descodificar elementos formulaicos predecibles o esperados; así, reducen el esfuerzo al tiempo que elevan la fluidez y liberan el procesamiento cognitivo de ir palabra por palabra. Por otro lado, contribuyen a la precisión pues bloquean las interferencias de la L1 o de una L3, como en **me toma una hora llegar a casa*, **el tiempo contará quién tiene razón*, **he tomado un examen hoy*, o —de los ejemplos del apartado 2— **hacer una decisión* y **hacer un error*. Al mismo tiempo, evitan las estructuras poco naturales resultado de la gramática abstracta, de los circunloquios y parafraseos, como en **si yo lo sé correctamente, hoy no hay clase, es festivo*; en lugar de *que yo sepa, hoy no hay clase*, o en el ejemplo que ya hemos visto en el apartado 2: **mis ojos se pusieron poco claros de repente, y me desmayé*, por *mis ojos se nublaron de repente*.

La enseñanza de estas expresiones ha de ser obligatoria si tenemos en cuenta que todo nativo adquiere la mayoría de ellas sin esfuerzo, por estar expuesto al *input* en la L1; sin embargo, la exposición al *input* en la L2 es infinitamente inferior al de la L1, por lo que en un contexto formal no podrán repetirse las expresiones con tanta frecuencia como para que se adquieran de forma inconsciente. Asimismo, han de ser trabajadas en el aula no solo porque gran cantidad de ellas no se genera activando los conocimientos gramaticales, sino porque tampoco podrán construirse ni descodificarse muchas de ellas a partir de la información que encontramos en los diccionarios normales ni en los de expresiones idiomáticas¹¹. Tomemos como ejemplo el adjetivo *encendido* y su combinatoria. De la definición¹² se deduce automáticamente que de su restricción léxico-sintáctica se obtiene

10. En Sánchez Rufat (en prensa), hacemos referencia no solo a una enseñanza expresa de las combinaciones —frente a un aprendizaje casual—, sino a un papel del profesor en el aula más explícito de lo que propone Lewis. Dicho aspecto, sin embargo, escapa a las pretensiones de este estudio.

11. Hasta ahora no hay acuerdo sobre el contenido y los límites exactos de la disciplina de la fraseología.

12. Definición tomada del *DRAE*: “de color rojo muy subido”. A partir de ahora, todas las definiciones que incluyamos en el presente trabajo proceden de la vigesimosegunda edición del *DRAE*.

un sustantivo que designa color (*rojo*), junto a los adjetivos *color* y *cromatismo*. Asimismo podemos relacionar este adjetivo con otros sustantivos que designan ciertas partes del cuerpo humano susceptibles de adquirir ese color (o tonos semejantes) como *mejillas* o *cara*. Pero en español también *los discursos, las polémicas, los debates, las controversias, las protestas, las críticas y los elogios* pueden ser *encendidos*¹³.

De esta manera, de la información que se puede obtener de los diferentes tipos de obras lexicográficas, el hablante nativo no puede deducir directamente que no solo *la botella, la olla o el cubo se destapan*, sino también *el fraude, la corrupción, el delito o el crimen*, entre otros sustantivos¹⁴. Al mismo tiempo, cuando en el diccionario este verbo es definido como *descubrir una cosa oculta o cubierta*, da lugar a un paradigma extensional mayor del que el sistema permite, pues, normalmente, este es más restrictivo de lo que, en un principio, parece; no podemos **destapar el arma del crimen* por muy escondida que la hayan dejado o **destapar los zapatos que estuve buscando el mes pasado* por muy ocultos o cubiertos que estuvieran por otros objetos del armario¹⁵.

Asimismo, y según se ha dicho ya, es importante tener cuidado con los sinónimos que se aportan en el aula o con los que puedan aparecer en las definiciones de los diccionarios. Si buscamos *disolver*, encontramos *deshacer, destruir, aniquilar y separar lo que estaba unido*. No obstante, *un edificio se destruye, el camino y la cama se deshacen, y los siameses pueden llegar a separarse* pero ninguno puede disolverse. Por otro lado, *el azúcar se disuelve o se deshace* en el café, pero no se destruye. Al mismo tiempo, estas definiciones no nos permiten codificar automáticamente que una *manifestación, concentración, disturbio, pacto o rumor se disuelven*. Este problema lo encontramos con gran cantidad de términos: *derrumbarse* se define como *caer (una cosa, persona) de golpe. Se derrumba o se cae un edificio*, pero también *la bolsa, los precios, el mercado, el poder, el sistema*; sin embargo, *se derrumban*, pero no *se caen, los sueños, la esperanza o la ilusión*.

Un último argumento a favor de la enseñanza explícita de las secuencias formulaicas es que muchas de ellas no se perciben, ya sea porque están formadas por palabras conocidas o porque constituyen una expresión muy similar a la combinación en la L1, por lo que se acaba produciendo una interferencia negativa: “picking up target language forms from input when they do not carry information crucial to the task, appears unlikely for adults. Paying attention to language form is hypothesized to be facilitative in all cases” (Schmidt 1990, p. 49). La siguiente transcripción refleja este proceso:

13. Ejemplos extraídos del diccionario combinatorio *Redes*, p. 904 y 905.

14. Aun así, dado que los sentidos figurados son extensiones naturales de los literales, es posible que el hablante nativo pueda reconocer el sentido de las combinaciones, y que el estudiante asienta ante un mayor o menor número de ellas en función de su conocimiento del idioma (Bosque 2005); esto es, poder comprender o descodificar no implica poder codificar una combinación; requieren un conocimiento y un esfuerzo de procesamiento diferentes.

15. Entiéndase que no nos referimos en estos dos ejemplos a *destapar* en el sentido literal de *abrir, descubrir o quitar la tapa*.

—Charles, en esta situación... ¿tú qué decisión tomarías?
 —Yo haría la decisión de Kate, es la más inteligente que [sic] todas

Queda, por lo tanto, expuesta la importancia de enseñar estas expresiones de forma explícita en el aula. A continuación, veremos de qué manera se puede trabajar con estas expresiones.

4. Sobre la rentabilidad de las restricciones sistemáticas

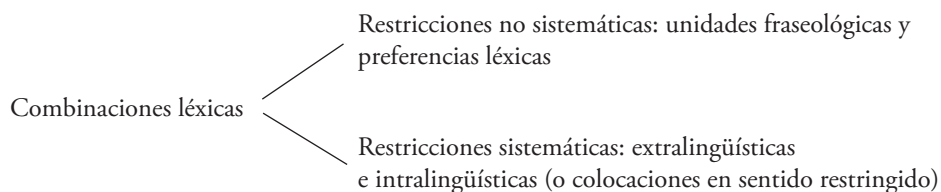
Lewis, a la hora de seleccionar las expresiones multipalabra o colocaciones —en su sentido más laxo— que se deben enseñar, parte de la definición aportada por la corriente estadística, que es la que se toma como referencia en la enseñanza del inglés y la que se está extendiendo en los estudios sobre aprendizaje y enseñanza del léxico en ELE, así como en los materiales didácticos de esta lengua. El autor del enfoque léxico (Lewis, 2000, p. 132) ofrece la siguiente definición de colocación: “collocation is the way in which words co-occur in natural text in statistically significant ways”. Colocaciones en sentido restringido, fórmulas sociales/rutinarias, marcadores discursivos, compuestos, locuciones, símiles estandarizados, proverbios, clichés, exclamaciones, combinaciones “libres” frecuentes, etcétera, campan a sus anchas en el territorio *abarcado* por *lengua formulaica*, con el único denominador común de la coaparición frecuente de sus integrantes y la recurrencia automática y habitual por parte de los usuarios de la lengua. Precisamente, los principales intentos que ha habido de caracterización de las expresiones prefabricadas han sido realizados tanto desde una perspectiva basada en un corpus como desde una perspectiva psicolingüística, de ahí que los criterios de determinación de estas combinaciones sean el de la frecuencia, en el primer caso, y el de la naturalidad, en el segundo, la cual ayuda a estimar —a través de la discusión con nativos y aprendices de modelos de adquisición y procesamiento de estas expresiones— si las secuencias se almacenan como una sola unidad en el lexicón mental (Stengers 2009). Como exponemos en anteriores trabajos (Sánchez Rufat 2010), la frecuencia no es un criterio que aporte información sobre la naturaleza lingüística de una combinación (Bosque 2001); frecuentes pueden ser estados, acciones o eventos que sean repetidos o habituales sin tener por qué serlo desde el punto de vista del sistema lingüístico, como sucede con las acciones diarias del tipo *lavar los platos*, *sacar la basura* o *ver la televisión*. Es, además, un concepto relativo, pues una combinación puede ser frecuente dependiendo de los tipos de textos que constituyen el corpus. Por estos motivos, el alumno no puede trabajar con este criterio de forma autónoma¹⁶, esto es, no es capaz de descri-

16. Recordemos que el objetivo del enfoque léxico (Lewis 1993, 1997, 2000) es lograr la autonomía del aprendiz equipándolo con estrategias para identificar en un *input* auténtico las expresiones multipalabras de la L2 dentro y fuera de clase.

minar lo que es una secuencia formulaica de lo que no lo es¹⁷. Lo mismo sucede con la naturalidad, pues incluso los nativos no tienen por qué ponerse de acuerdo en si una secuencia de palabras es natural, y se almacena, por tanto, en nuestra memoria como una sola unidad.

Por todo esto, si el término *lengua* o *secuencia formulaica*, (o *colocación*, *expresión multipalabra* o *bloque semiprefabricado*) no restringe —en términos léxicos y sintácticos— los dominios que necesitan ser establecidos en estudios del lenguaje, consideramos que el empleo de este término y el concepto que designa no es útil como unidad funcional (diferenciada) de descripción en la didáctica de lenguas¹⁸. Para poder enseñar a identificar las combinaciones léxicas propias de la norma y del sistema, independientemente de su frecuencia, y evitar así los errores combinatorios, es necesario sistematizar estas combinaciones atendiendo a criterios puramente lingüísticos, dado que es evidente que no es posible dar un contenido homogéneo —al menos, no comprensible conceptualmente en términos teóricos— al variado grupo de fenómenos al que se aplica el término *secuencia formulaica*.

Con el fin de facilitar la labor de profesores y aprendices de lenguas extranjeras, hemos realizado una propuesta de organización conceptual y terminológica partiendo del concepto de selección léxica de Bosque (2001 y 2005), basado en la motivación de la restricción de las combinaciones. A continuación, incluimos el esquema y unas breves anotaciones de las diferentes combinaciones léxicas¹⁹:



17. Stengers (2009) hace referencia a experimentos que ponen de manifiesto el fracaso del alumno a la hora de tener que identificar estas secuencias partiendo de este criterio.

18. Del mismo modo que Bosque (2001) y Sinclair *et al.* (2005) señalan la falta de utilidad de que ciertas unidades se encuentren recogidas en un diccionario combinatorio por el mero hecho de ser frecuentes, nosotros cuestionamos su validez en el ámbito de la enseñanza de lenguas.

19. Tratamos el asunto con más profundidad en Sánchez Rufat (2010).

En las *restricciones no sistemáticas* incluimos las unidades fraseológicas²⁰ y las preferencias léxicas, en las que la restricción no viene impuesta por el sistema lingüístico, sino por factores en ocasiones difíciles de sistematizar, ya sean culturales, sociales o históricos. En muchos casos de restricciones no sistemáticas desconocemos de qué son producto estas agrupaciones, cuál es su motivación; y si las consideramos ajenas a la descripción del sistema lingüístico, hablamos de norma. Un ejemplo de preferencia léxica lo encontramos en el ejemplo (5) del apartado 2: *coche fúnebre*. El alumno debe aprender que *fúnebre* selecciona a *coche* y a *honras*, mientras que su sinónimo *mortuorio* se combina con *cámara* o *esquela*. Así, no conocemos explicación semántica que justifique esta diferente preferencia léxica, que, por otro lado, muestra el problema que supone definir un término en el aula por medio de un sinónimo al tiempo que se ignora su comportamiento léxico-sintáctico.

El segundo grupo de nuestra clasificación lo constituyen las *restricciones sistemáticas*²¹, en las que entran en juego las propiedades combinatorias de las palabras; es decir, estas combinaciones son producto del carácter restrictivo que forma parte del sistema lingüístico. Los predicados seleccionan sus argumentos restringiendo el conjunto de piezas léxicas que pueden denotar de acuerdo con sus rasgos semánticos, por lo que las unidades seleccionadas no son piezas aisladas —como en el caso de la preferencia léxica—, sino clases léxicas. La teoría de los predicados de Bosque, a la que nos hemos sumado, explica la capacidad de generación de estas restricciones sistemáticas: atender a estas restricciones no como combinaciones en las que la relación entre los elementos es biunívoca, sino como casos de selección léxica, implica adquirir —de entrada— un conocimiento de la lengua generativo (basado en reglas)²², pues una misma relación léxico-sintáctica permite generar diferentes combinaciones²³.

Tanto en las preferencias léxicas como en las restricciones sistemáticas estamos ante la combinación de dos piezas léxicas que mantienen su significado estable; no obstante,

20. Incluimos bajo el término *unidad fraseológica* un primer grupo formado por las unidades que son conmutables por un enunciado, como las paremias (refranes, proverbios, citas y enunciados de valor específico), así como las fórmulas rutinarias; y un segundo grupo constituido por unidades fraseológicas que no conforman un enunciado completo, esto es, las locuciones. A diferencia de las demás combinaciones léxicas, las unidades fraseológicas no son composicionales desde un punto de vista semántico, y no se forman aplicando las reglas de la gramática con libertad.

21. En español hay dos diccionarios combinatorios: *Redes*, ya mencionado, se centra en las restricciones sistemáticas, principalmente en las intralingüísticas, que son las más complejas; *Práctico* muestra las combinaciones más habituales en las que suele aparecer una palabra, ya sea preferencia léxica o combinación sistemática.

22. Frente al conocimiento memorístico o formulaico, basado en la memoria, que es el único que se puede desarrollar en los casos de restricciones no sistemáticas.

23. Obsérvese que esta productividad se pierde si la direccionalidad de la selección de naturaleza semántica parte de los argumentos para llegar a los predicados, como sucede con los trabajos derivados de la Teoría Sentido-Texto.

el hecho de que estas combinaciones sean composicionales no quiere decir que —como acabamos de comprobar con el ejemplo de los sinónimos *fiébre* y *mortuorio*— su formación esté exenta de problemas.

Las restricciones que los predicados imponen a sus argumentos pueden ser de dos tipos, de ahí que haya dos tipos de restricciones sistemáticas: si están fundamentadas únicamente en nociones de naturaleza física (Bosque 2005), podemos construir el paradigma extensional de acuerdo con el conocimiento que tenemos del mundo. Por el contrario, si la restricción no está basada únicamente en nociones físicas, tendremos que recurrir a nuestro conocimiento del idioma para obtener las piezas léxicas que conforman la extensión, de ahí la denominación de restricciones sistemáticas intralingüísticas o colocaciones —en sentido restringido. Para ilustrar este concepto, volvamos al ejemplo (2) del apartado 2, *Mis esperanzas se han derrumbado*, y analicemos brevemente las combinaciones de base semántica resultado del carácter restrictivo del verbo *derrumbarse*. Si atendemos a la restricción de naturaleza extralingüística de esta predicación, resulta una cantidad de argumentos que podemos reunir en una clase léxica de gran extensión: los sustantivos que designan edificaciones y construcciones. Una vez que conocemos el significado del predicado, seamos hablantes nativos o no (en otras lenguas existe un término equivalente con igual extensión), podemos añadir piezas léxicas a esta clase aplicando nuestro sentido común. Por otra parte, como acabamos de señalar, las restricciones sistemáticas intralingüísticas continúan siendo casos de selección léxica, pero, a diferencia del grupo anterior, la restricción viene dada en función de las propiedades del predicado, ya sea este un adverbio, un adjetivo, un verbo, una proposición o un sustantivo. No nos basta con nuestro conocimiento del mundo, necesitamos un conocimiento profundo del predicado para obtener el paradigma extensional. Este no tiene por qué coincidir en otra lengua, de ahí que sea una fuente de errores para el estudiante. De esta manera, y continuamos con el mismo ejemplo, en su sentido figurado *derrumbarse* se combina —entre otras clases léxicas— con la clase formada por sustantivos que designan lo que se anhela o se pretende conseguir; también con otros que expresan el sentimiento de confianza que acompaña a los deseos que se pretenden alcanzar, como *sueño*, *esperanza*, *ilusión*, *aspiración*, *expectativa* y *confianza*²⁴. Pese a que este verbo se combina con sustantivos pertenecientes a diferentes clases léxicas que, *a priori*, partiendo del significado del verbo, no podemos señalar, el hecho de que el sentido figurado no sea más que una extensión natural del literal —de ahí que sean combinaciones composicionales— hace que la deducción de aquel nos permita descodificar, en muchos casos, el significado de la combinación, pero no codificarla.

Así, si volvemos a los ejemplos del apartado 2, todavía se puede dar un paso más con respecto a la propuesta de Lewis y ahondar en el concepto de léxico gramaticalizado, que

24. Vid. *Redes*, p. 720-722.

él mismo plantea, pero cuyo tratamiento es, a nuestro juicio, algo superficial²⁵. En *tomar una decisión*, en lugar de tratar de que el alumno fije y almacene en la memoria esta expresión como relación biunívoca, se puede aprovechar la oportunidad para considerar esta expresión como la restricción intralingüística (o colocación) que propiamente es. De esta manera, se obtienen varias combinaciones léxicas de una misma relación léxico-sintáctica —lo que resultaría muy rentable para el alumno; así, junto a *decisión*, también constituyen argumentos de *tomar* los sustantivos que denotan elección o resolución, como *resolución*, *determinación*, *elección*, *acuerdo*, *iniciativa* y *opción*. Cuando, como en el ejemplo (3), *nublarse* adopta el sentido figurado de *oscurecer*, *enturbiar*, *empañar* o *turbar*, también se combina con otros sustantivos que denotan facultad u órgano de la visión, como *vista*, *visión* y *mirada*²⁶. Con respecto a *cometer* (4), no solo se combina con el sustantivo *error*, sino con todos los que denotan error —entre otras clases léxicas—, a saber: *fallo*, *equivocación*, *desliz*, *torpeza*, *lapsus*, *patinazo* o *desacierto*²⁷.

5. Conclusiones

Conocer una pieza léxica no se limita a saber el significado, sino que también atañe a cómo se combina en el eje sintagmático —lo que supone una fuente de errores para los alumnos. Poder usar la palabra en las distintas clases de combinaciones léxicas en las que puede aparecer implica tener desarrollada la (sub)competencia combinatoria, que dota al alumno de precisión y fluidez. Con el objetivo de facilitar el proceso de enseñanza y aprendizaje de los conceptos que conforman esta interfaz léxico-sintáctica, hemos propuesto un modelo de clasificación de estas combinaciones basado en la naturaleza de su restricción. De acuerdo con lo expuesto, proponemos centrar el trabajo en el aula en desarrollar las relaciones léxico-sintácticas más rentables y más productivas para el alumno desde un punto de vista lingüístico, esto es, las restricciones sistemáticas. Su aprendizaje —a diferencia de las restricciones no sistemáticas— contribuye al conocimiento generativo, por lo que se incrementa la precisión, pero también al memorístico —que se corresponde con la fluidez— en aquellos casos en los que se almacene en el lexicón una pareja concreta frente al resto de las combinaciones que pueden resultar de una misma relación léxico-sintáctica —pese a que no estamos ante relaciones biunívocas²⁸. Explotar

25. No es el objetivo de este estudio centrarse en las estrategias didácticas más efectivas para enseñar las combinaciones léxicas. Una introducción al respecto puede verse en Sánchez Rufat (*en prensa*).

26. Ejemplos extraídos de *Redes*, p. 1351.

27. Ejemplos extraídos de *Redes*, p. 500.

28. No se debe desatender el esfuerzo dedicado en el aula al conocimiento memorístico de las restricciones sistemáticas más útiles y frecuentes para el alumno que tanto aportan a la fluidez de la codificación y descodificación de mensajes, especialmente en las situaciones de oralidad en tiempo real, donde no suele haber espacio para la creatividad y el esfuerzo cognitivo. Esto permite que las combinaciones pasen a formar parte de la lengua formulaica, cuyas ventajas ya han sido expuestas en el apartado 3.

el procesamiento lingüístico a través del conocimiento basado en reglas, además, supone beneficiarse de la tendencia de todo aprendiz adulto a utilizar fundamentalmente el procesamiento de adquisición del léxico de modo generativo o analítico, en detrimento del memorístico —la competencia nativa se caracteriza por el equilibrio entre ambos.

De las anteriores restricciones sistemáticas, las extralingüísticas no tienen por qué causar problemas a un estudiante, pues el significado total de la combinación es la suma del sentido literal de cada una de las piezas léxicas que la conforman. Por eso, pese a su rentabilidad, no es necesario detenerse demasiado en estas combinaciones, y bastará con que el alumno tome conciencia del significado estable de la palabra y se preste atención a posibles interferencias negativas de la L1 o una posible L3²⁹. En cambio, exige un esfuerzo más explícito del profesor el que el alumno logre desarrollar la capacidad para usar las colocaciones (en sentido restringido), esto es, (sub)competencia colocacional.

Referencias

- Bolinger, D. (1976). "Meaning and memory", *Forum Linguisticum*, 1, 1-14.
- Bosque, I. (2001). "Sobre el concepto de colocación y sus límites", *Lingüística Española actual* xxiii/1, 9-41.
- Bosque, I. (coor.) (2005). *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: SM.
- Bosque, I. (coor.) (2006). *Práctico. Diccionario combinatorio práctico del español*. Madrid: SM.
- Brown, D. F. (1974). "Advanced vocabulary teaching: the problem of collocation", *RELC Journal*, 5: 2, 1-11.
- Erman, B. y B. Warren (2000). "The idiom principle and the open-choice principle", *Text*, 20, 87-120.
- Fillmore, C. J. (1976). "The Need for a Frame Semantics within Linguistics". *Statistical Methods in Linguistics*, 5-29. Estocolmo: Skriptor.
- Firth, J. R. (1957). *Papers in Linguistics 1934-1951*. Londres: Oxford University Press.
- Hunston, S. y G. Francis (2000). *Pattern Grammar: Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.
- Lewis, M. (1993). *The Lexical Approach*. Londres: Language Teaching Publications.
- Lewis, M. (1997). *Implementing the Lexical Approach*. Londres: Language Teaching Publications.
- Lewis, M. (2000). *Teaching Collocation. Further Developments in the Lexical Approach*. Londres: Language Teaching Publications.
- Moon, R. (1998). *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press.

29. Compárese *tener* + sustantivos que designan una sensación física (*frío, calor, hambre*) con su expresión equivalente inglesa *to be (cold, hot, hungry)*.

- Nattinger, J. R. (1980). "A lexical phrase grammar for ESL", *TESOL Quarterly*, 14, 337-344.
- Pawley, A. y F. Syder (1983). "Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency". En J. Richards y R. Schmidt (eds.), *Language and Communication*, 191-226. Londres, Longman.
- Real Academia Española. *Diccionario de la lengua española*, 22ª ed. Disponible en: <http://www.rae.es>. Acceso: 25.08.2010.
- Sánchez Rufat, A. (2010). "Apuntes sobre las unidades fraseológicas y el concepto de colocación", *Anuario de estudios filológicos*, XXXIII, 291-306.
- Sánchez Rufat, A. (en prensa). "El aprendizaje de las unidades fraseológicas a partir de planteamientos lingüísticos cognitivos". En M.ª L. Calero Vaquera y M.ª A. Hermosilla Álvarez (coords.), *Lenguaje, literatura y cognición*. Córdoba: Servicio de publicaciones de la Universidad de Córdoba.
- Schmidt, R. W. (1990). "The role of consciousness in second language learning", *Applied Linguistics*, 11: 2, 129-158.
- Sinclair, J. M. y A. Mauranen (2006). *Linear Unit Grammar: Integrating Speech and Writing*. Amsterdam: John Benjamins.
- Sinclair, J. M. (1991). *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.
- Sinclair, J. M. , S. Jones y R. Daley (2005). *English Collocation Studies: The OSTI Report*. Londres: Continuum.
- Skehan, P. (1998). *A Cognitive Approach to Language Teaching*. Oxford: Oxford University Press.
- Stengers, H. (2009). *The Idiom Principle Put to the Test: An Exercise in Applied Comparative Linguistics*. Tesis doctoral inédita. Bruselas: Universidad Vrije de Bruselas.
- Stubbs, M. (1995). "Collocations and Semantic Profiles: On the Cause of the Trouble with Quantitative Studies", *Functions of Language*, 2: 1, 23-55.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.