

PÉREZ SERRANO, MERCEDES

UNIVERSIDAD DE COLUMBIA, ESTADOS UNIDOS

ANÁLISIS DE ERRORES COLOCACIONALES EN UN CORPUS DE APRENDIENTES DE ELE

BIODATA

Mercedes Pérez Serrano es lectora de español en la Universidad de Columbia. Es doctoranda en el programa *Lengua española: investigación y enseñanza* en la Universidad de Salamanca. Obtuvo su máster en *Enseñanza de Español como Lengua Extranjera* en Instituto Cervantes y la Universidad Internacional Menéndez Pelayo en 2009. Es autora de materiales y sus trabajos han sido publicados en revistas y presentados en conferencias. Su interés investigador se centra en la adquisición, enseñanza y aprendizaje del léxico.

RESUMEN

Las colocaciones y la rentabilidad de su enseñanza-aprendizaje en el proceso de adquisición del español como L2 está ya reconocida por la comunidad investigadora, de profesores de creadores de materiales didácticos. En esta línea de reconocimiento de la importancia de la dimensión combinatoria del léxico en la enseñanza-aprendizaje de una L2, en este trabajo realizamos un análisis de errores que cometen los aprendientes de nivel inicial al producir colocaciones léxicas a partir de un corpus de textos producidos por los mismos aprendientes y comparamos sus producciones con las de hablantes nativos de español.

Este estudio fue realizado bajo la tutela del profesor Javier de Santiago Guervós. Constituye su *Trabajo de doctorado* y fue entregado en junio de 2012 en la Universidad de Salamanca con una calificación de sobresaliente (9).

1. INTRODUCCIÓN

Con varios los autores que han abogado por el final del debate sobre la naturaleza sintáctica, léxica y semántica de las colocaciones y han llamado la atención a la comunidad investigadora sobre el peligro de caer en definir como continuum todo aquello que nos cuesta definir o delimitar (Bosque, 2001: 1, Alonso Ramos, 2010).

Ello nos hace pensar que la literatura que recoge disquisiciones teóricas sobre la naturaleza de las colocaciones y sobre qué disciplina debería acogerlas –sintaxis, lexicología, fraseología-, es más que abundante y ya han surgido voces abogando por poner fin al debate terminológico para *ceder el paso a una fase de desarrollo* (Alonso Ramos, 2010, 1).

Las colocaciones y la rentabilidad de su enseñanza-aprendizaje en el proceso de adquisición del español como L2 está ya reconocida por la comunidad de profesores y por los creadores de materiales didácticos¹. No obstante, en la comparación con las fuentes bibliográficas y los materiales didácticos de nuestra lengua con los del inglés, observaremos que la *diferencia es abismal* (Fernando Aramo, 2009). Asimismo, como han reconocido otros estudiosos del tema, *se echan en falta estudios sobre las colocaciones más frecuentes y rentables del español* (Álvarez Cavanillas, 2008:123).

¹ De hecho, un reciente manual E. Martín Peris, N. Sans, y A. Garmendia (2011) se cede protagonismo al vocabulario en todas las fases del aprendizaje. En la sección de “Palabras en compañía” se presenta el funcionamiento de las agrupaciones más frecuentes y útiles del vocabulario de cada unidad.

En este contexto queremos enmarcar nuestro trabajo de investigación, que pretende, siguiendo la metodología empleada por el grupo COLOCATE², realizar un análisis de errores colocacionales cometidos por los aprendientes de nivel inicial a partir de un corpus de textos producidos por los mismos aprendientes. El análisis de errores proporciona una información inestimable a la hora de guiar al aprendiente en el proceso de enseñanza- aprendizaje de las colocaciones, tal y como han reconocido Alonso Ramos, Wanner y otras autoras:

To adequately support students in learning collocations, it is crucial to identify and classify the collocation errors made by them and then offer targeted exercises and adequate illustrative material. This presupposes the availability of collocation tagged learner and general corpora: a learner corpus allows us to identify the most common collocation errors: a general corpus is needed as a source of illustration and training material. (Alonso Ramos, Leo Wanner et al. 2010: 375)

2. PUNTO DE PARTIDA Y OBJETIVOS DEL TRABAJO

Como veremos, el término *colocación* es polisémico y se ha abordado desde diversas disciplinas. En este trabajo, partiremos de la concepción de colocación como un caso de restricción léxica orientada desde la base. De esta manera, a partir de una base como *cariño*, si el hablante quiere expresar la función ‘empezar a sentir’, seleccionará el colocativo *coger*, para formar la colocación *coger cariño*. Esta selección no tiene por qué estar motivada semánticamente, es frecuente que varíe de unas lenguas a otras y

² El grupo COLOCATE es un grupo de investigación de la Universidad de La Coruña que centra su interés en el estudio y la descripción de las colocaciones. El lector interesado puede consultar su blog: <http://collocate.blogspot.com.es/>

por ello resulta un punto particularmente difícil para el estudiante de lenguas extranjeras.

Así, en este trabajo pretendemos alcanzar los siguientes objetivos:

- Ofrecer un breve repaso sobre las principales corrientes que han abordado las colocaciones y la perspectiva desde la que lo han hecho.
- Ampliar el análisis de errores colocacionales a los aprendices de nivel inicial.
- Encontrar las colocaciones empleadas por los aprendices de español de nivel inicial y clasificar las erróneas de acuerdo con una tipología de errores establecida.
- Comparar las colocaciones empleadas por los aprendices de español con aquellas utilizadas por los hablantes nativos.
- Trazar líneas futuras de investigación a la luz del análisis y la interpretación llevados a cabo.

3. LA NOCIÓN DE COLOCACIÓN

Las colocaciones se han estudiado desde diversas perspectivas y parcelas de la lingüística. A continuación, ofrecemos un breve repaso de las corrientes más significativas que han abordado el fenómeno en el ámbito internacional, haciendo hincapié en los estudios más representativos de la lingüística hispánica.

3.1. LAS COLOCACIONES DESDE LA LINGÜÍSTICA DE CORPUS: CORRIENTE PROBABILÍSTICA

Aunque se ha señalado la restricción combinatoria ha interesado a los estudiosos desde hace más de cinco siglos (Corpas Pastor: 2001: 89), se coincide en apuntar que el primero en utilizar el término colocación fue Firth en los años 50, que inaugura la corriente estadística representada por autores del contextualismo británico como Halliday (1966) y Sinclair (1991).

De acuerdo con los autores de esta corriente de base estadística o probabilística para que dos palabras formen una colocación, solo tienen que tener una alta probabilidad de coexistencia en un corpus. De esta concepción del término se ha criticado la falta de especificidad y la vaguedad que implica. De acuerdo con sus detractores, si adoptamos tal punto de vista, entonces cualquier combinación frecuente de palabras. Pensemos en un verbo como *leer*. El hecho de que la combinación *leer un libro* no solo sea probable, sino frecuente, no hace de ella una colocación léxica porque *leer* acepta como complemento todo aquello que sea susceptible de ser leído.

3.2. LAS COLOCACIONES COMO CASOS DE RESTRICCIÓN LÉXICA

Otro punto de vista desde el que se ha abordado el fenómeno de las colocaciones es el de la restricción léxica. Como apunta Higuera (2011: 441), es Hausman (1989) quien adjudica un estatus diferente a los componentes de una colocación: para este autor, uno determina la elección del otro. A uno lo denomina la *base*, elegido libremente, y al otro *colocativo*, cuya elección viene condicionada por la base. Así pues, en la colocación *concertar una cita* la base *cita* determinaría la elección del colocativo *concertar*.

La corriente de la restricción léxica está íntimamente unida a la elaboración de diccionarios, por lo que haremos alusión a las obras lexicográficas bajo las que subyacen tales visiones.

En el panorama de la lingüística española actual, cabe destacar los trabajos de Margarita Alonso Ramos y el equipo de Investigación de la Universidad de La Coruña (Colocate), quienes defienden un enfoque desde el punto de vista del lexicógrafo que elabora un diccionario de colocaciones. El *Diccionario de Colocaciones del Español (DICE en adelante)* engarza en la línea de investigación sobre diccionarios de colocaciones (Mel'čuk y Polguère 2007) y tal y como se declara en su introducción, *pretende cubrir una laguna ya que, en español, no existe ningún diccionario especializado en el fenómeno léxico de las colocaciones.*

A grandes rasgos, se puede afirmar que estos trabajos se encuadran en el marco teórico de la Teoría Sentido- Texto (Mel'čuk y Polguère, 1987) y la lexicología explicativa y combinatoria (Mel'čuk et al. 1995).

Esta teoría, como apunta Barrios (2010), nació al hilo de elaboración de diccionarios, con lo que está estrechamente vinculada a la lexicografía. Para la elaboración del *Dictionnaire Explicatif et combinatoire du français contemporain* (DEC en adelante) que cuenta hasta ahora con cuatro volúmenes (1984, 1988, 1992, 1999). Mel'čuk y su equipo se sirven del concepto de función léxica para dar cuenta de las colocaciones:

Grosso modo, una FL es una función f que asocia a una unidad léxica $L1$ un conjunto de unidades léxicas cuasi sinónimas que son escogidas en función de $L1$ para expresar el sentido correspondiente a la FL f . En términos de FFLL, una colocación $L1$ (base) y $L2$ (colocativo) se presenta como $f(L1)=L2$, en donde $L1$ es la PALABRA LLAVE de la correspondiente FL y $L2$ su VALOR. (Alonso Ramos y Sanromán Vilas, 2000)

Algunos ejemplos de colocaciones léxicas expresadas en términos de funciones léxicas tomados de M. A. Barrios:

Magn 'intenso, muy' Magn (lluvia)= lluvia torrencial, fuerte
Oper 'dar' (hacer) Oper (orden)= dar una orden
Sing 'elemento único de' Sing (arena)= grano de arena

Lo más interesante bajo nuestro punto de vista es que es un diccionario elaborado desde la perspectiva de la codificación más que desde la descodificación: cuando un hablante quiere expresar la noción de intensidad, no utilizará la misma unidad léxica si esta noción de intensidad se aplica a la base odio *odio –odio mortal-, a crimen –crimen atroz- o a paciencia –paciencia infinita-;* de la misma manera que si un hablante quiere aplicar la noción de 'hacer', cuando lo haga a la base paseo utilizará *dar*, como en dar un paseo; pero utilizará *hacer* cuando la aplique a la base favor, como en *hacer un favor*. Así pues, mientras que las bases se seleccionan libremente, la elección del colocativo está condicionada por la base. A esto se le llama restricción léxica.

La idea clave en la teoría de Alonso Ramos es la de restricción selectiva orientada desde la base. Para esta autora, los criterios de transparencia semántica, impredecibilidad o la restricción semántica no son criterios operativos y eficaces, pero sí lo es la orientación de la restricción, ya sea semántica o léxica. Así, en colocaciones como *el barco atraca, enhebrar una aguja o balar una oveja*, lo que lleva a considerarlas como colocaciones es que la relación está orientada desde la base, es decir, la base se elige libremente (barco, aguja, oveja) y los colocativos son seleccionados por su significado en función de la misma. Se selecciona *balar*, porque estamos hablando de una oveja: la selección del colocativo está restringida por la base. Es, así pues, como declara la autora, una perspectiva desde la

producción. Por eso en su proyecto lexicográfico los lemas o entradas son las bases y no los colocativos.

Sin embargo, el análisis de las colocaciones puede ser abordado estableciendo una dirección distinta. El proyecto de lexicografía combinatoria dirigido por Ignacio Bosque, está compuesto por dos diccionarios, *Redes* (2004) y *Práctico* (2006). Como advierte el mismo autor, no se trata de diccionarios de colocaciones exclusivamente, pero ambos las contienen. En *Redes* a la relación base- colocativo se le da una orientación distinta, aunque quizá complementaria. En este diccionario, las entradas están compuestas principalmente por colocativos y nos ofrece su combinatoria organizada, en el caso de las entradas analíticas, por clases léxicas.

Tal y como explica Bosque en su artículo “El concepto de colocación y sus límites” (2001) y siguiendo con el mismo ejemplo de relación colocativa de tipo intensivo o elativo, en el caso del DICE la entrada sería para el sustantivo *odio* y entre las muchas especificaciones, aparecería el colocativo *mortal*, seleccionado cuando se quiere aplicar la función *Magn* a la base *odio*. En el caso de *Redes*, la entrada es el colocativo *mortal*, con el objetivo de mostrar qué uso hace el idioma de la palabra *mortal*. Como apunta Bosque, el objetivo en este caso no es encontrar el mejor adjetivo intensificador para el sustantivo *odio*, sino *entender el concepto mismo de intensidad como noción lingüística, para lo que resulta enormemente útil poder disponer de un estudio detallado de su ámbito de aplicación* (2001:33).

3.3. LAS COLOCACIONES DESDE LA FRASEOLOGÍA: LAS COLOCACIONES COMO UNIDADES FRASEOLÓGICAS

Corpas Pastor reclama el estudio de las colocaciones desde la

fraseología, por lo que entiende las colocaciones como

unidades fraseológicas formadas por dos unidades léxicas en relación sintáctica que no constituyen por sí mismas, actos de habla ni enunciados; y que, debido a su fijación en la norma, presentan restricciones de combinación establecidas por el uso, generalmente de base semántica: el colocado autónomo semánticamente (la base) no sólo determina la elección del colocativo, sino que, además, selecciona en éste una acepción especial, frecuentemente de carácter abstracto o figurativo (1996: 66).

En su artículo “En torno al concepto de colocación” (2001), Corpas señala, como características definitorias de esta unidad léxica el hecho de que es una unidad pluriverbal, la frecuencia de coaparición, la institucionalización –reconocimiento de la unidad como familiar por parte de los hablantes, que la perciben como un segmento prefabricado- y la variación lingüística y discursiva a la que están sujetas.

Koike, autor de una de las monografías más extensivas dedicadas a las colocaciones en español, señala que una colocación es un vínculo léxico-semántico condicionado por la sintaxis (2001: 76). Señala una serie de características tanto formales como semánticas que resumimos a continuación:

- a) Coocurrencia frecuente, característica importante, aunque no un rasgo exclusivo.
- b) Restricciones combinatorias impuestas por el uso tradicional.
- c) Composicionalidad formal, que permite flexibilidad formal.
- d) Vinculación de lexemas, que explica el cambio de categoría gramatical de algunas colocaciones.

- e) Relación típica entre los componentes.
- f) Precisión semántica de la combinación.

3.4. LAS COLOCACIONES DESDE LA DIDÁCTICA DE LENGUAS

Michael Lewis, considerado el padre del enfoque léxico, un desarrollo del enfoque comunicativo que concede especial relevancia al contenido léxico, en su monografía *The lexical approach* (2000), habla de la necesidad de segmentar el léxico en unidades complejas (*pedagogical chunking*) y la consiguiente atención a unidades léxicas superiores a la palabra, entre las que se encuentran *colocaciones* y *expresiones idiomáticas*, concediendo a las primeras una especial relevancia.

Así, Lewis, tal y como recoge Fernando Aramo (2009: 18), describe las colocaciones como unidades léxicas pluriverbales caracterizadas por la coaparición de sus elementos integrantes, su arbitrariedad y su variación tanto en el eje sintagmático como en el paradigmático, adoptando así una visión más amplia que las otras dos descritas con anterioridad.

En esta línea, Higuera (2006) trata de conciliar la definición lingüística y la práctica de la enseñanza de lenguas extranjeras, abogando por una definición más amplia pero igualmente precisa. Esta autora distingue entre las colocaciones prototípicas y no prototípicas. Afirma que las colocaciones no prototípicas, aquellas que quizá no serían consideradas como tales por otros lingüistas, son las más numerosas y, por consiguiente, las que más peso tendrían en el proceso de enseñanza-aprendizaje de una segunda lengua.

En la siguiente figura, Higuera expone las características de ambos conceptos y añade los siguientes ejemplos: una colocación prototípica como *reservar una mesa*, presentaría todas las características enumeradas; *hacer la compra*, considerada una colocación no prototípica, sería muestra solo de algunas de ellas; y por último, *comprar un libro*, como combinación libre que es, solo presentaría regularidad sintáctica y transparencia semántica.

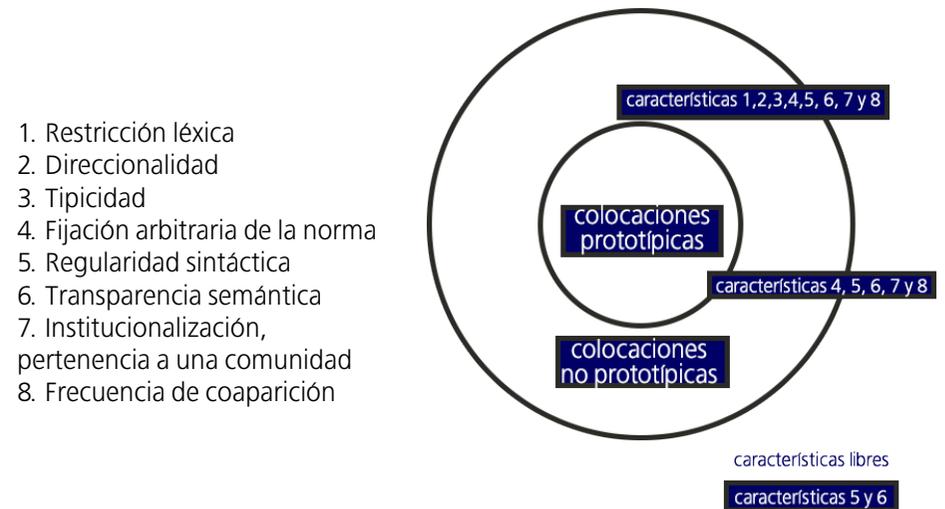


FIGURA 1. Clasificación de las colocaciones de Higuera (2006: 27)

4. METODOLOGÍA

4.1. LA MUESTRA

La muestra se ha extraído principalmente del corpus de aprendices de español L2 CEDEL2, un proyecto que surge ante el creciente interés por el estudio de la adquisición de español L2.

Se trata de un corpus formado por redacciones escritas por hablantes de español como lengua segunda y cuya L1 es el inglés. Está acompañado por un corpus de hablantes nativos de español de características similares.

CEDELE2 surge en el seno del proyecto *Word Order in Second Language Acquisition Corpora* (WOSLAC), dirigido por Amaya Mendikoetxea, de la Universidad Autónoma de Madrid.

El contenido del corpus no está diseñado de acuerdo con criterios lingüísticos, -si así fuera, predominaría una estructura determinada o el léxico de un campo semántico en concreto-, sino de acuerdo con criterios comunicativos, que Lozano llama criterios externos. Así pues, los aprendices que han participado en CEDELE2, han podido elegir entre 12 temas diferentes³, con lo que pretenden ofrecer muestra de una amplia gama de estructuras y de léxico, pero de ninguna en particular.

En cuanto a los participantes en el corpus, se trata de aprendices de todos los niveles (principiante, intermedio, avanzado) según un test de nivel estandarizado (Universidad de Wisconsin, 1998) habilitado para que los aprendices pudiesen completarlo en línea.

Los datos de CEDELE2 son recogidos en línea a través de formularios electrónicos que se dividen en tres partes:

³ 1. ¿Cómo es la región donde vives? 2. Habla de una persona famosa; 3. Resume una película que has visto recientemente; 4. ¿Qué hiciste el año pasado durante las vacaciones?; 5. ¿Cuáles son tus planes para el futuro?; 6. Describe un viaje que has hecho recientemente; 7. Cuenta una experiencia que hayas vivido; 8. Habla del problema del terrorismo en el mundo; 9. ¿Qué opinas de la nueva ley anti-tabaco?; 10. ¿Crees que las parejas gay tienen el derecho de casarse y adoptar niños?; 11. ¿Crees que la marihuana se debería legalizar?; 12. Analiza los principales aspectos de la inmigración.

1. *Learning background* (datos personales, datos lingüísticos, autovaloración del nivel en español y en otras lenguas.)

2. *Placement test*: se trata de un test de nivel estandarizado (Universidad de Wisconsin, 1998).

3. *Essay* (redacción). Además de la redacción escrita se proporciona información sobre el lugar de redacción, si ha habido investigación previa sobre el tema y las herramientas lingüísticas empleadas.

Para este trabajo, se han seleccionado tres de los 12 temas disponibles de composiciones que ofrece el corpus, a saber, ¿cuáles son tus planes para el futuro?, ¿cómo es la región donde vives? y Habla de una persona famosa. La razón que nos ha llevado a escoger estos tres temas es que representan funciones y abarcan nociones correspondientes a los niveles del MCER A1 y A2 y, por lo tanto, aparecen glosadas en el *Plan curricular del Instituto Cervantes* en dichos niveles:

- Expresar planes e intenciones (2006: 230)

- Geografía y naturaleza; Espacios urbanos o rústicos: Clima y tiempo atmosférico.

- Individuo: dimensión física y dimensión perceptiva y anímica. (2006: 336).

En cuanto a la selección dentro de la muestra del Corpus, se han seleccionado aquellas redacciones que, en el test de nivel estandarizado mencionado con anterioridad, han obtenido una calificación inferior al 60%.

Tal y como se comprobará en el apartado de *Análisis y comentario de los datos*, una de las referencias para determinar la corrección de una colocación determinada ha sido el corpus de nativos. El corpus CEDEL2 ofrece, además de las redacciones escritas por aprendices de español, otra muestra de redacciones que responden a los mismos títulos esta vez escritas por hablantes nativos. Hemos encontrado muestras suficientes para los temas “¿Cuáles son tus planes para el futuro?” y para “¿Cómo es la región donde vives?”. Para el tema de “Habla de una persona famosa”, al no contar con producciones escritas de CEDEL2 hemos acudido a participantes voluntarios con la característica común de tener el español como primera lengua que han accedido a escribir una redacción sobre este tema.

4.2. BÚSQUEDA Y ETIQUETADO DE COLOCACIONES: EL PROCEDIMIENTO

4.2.1. ¿Es una colocación o no?

El primer problema que hemos encontrado en el proceso de anotación y etiquetado de los errores reside en la noción misma de *colocación*. En esta práctica, tal y como reconocen Vincze, Ramos y otras autoras (2011: 280), resulta problemático trazar los límites entre las combinaciones libres de palabras, las colocaciones y las expresiones idiomáticas.

Por ejemplo, para la función de intensificación (*Magn*, según las funciones léxicas de Mel'čuk), el español cuenta con diversos recursos léxicos. No hemos considerado como colocación las secuencias como *muy creativo* o *mucho dinero* porque estos intensificadores no vienen exigidos léxicamente por el adjetivo *creativo* ni por el sustantivo *dinero* y porque el error no radica en cuestiones léxicas sino en no conocer el funcionamiento gramatical

de *muy mucho*. Sin embargo, para el sustantivo *población*, hemos considerado como colocación (errónea en este caso), la secuencia *grande población*. En este caso, la base *población* para ser intensificada exige un adjetivo como *alta* o *elevada*.

En el proceso de anotación, hubo dudas también a propósito de si secuencias como *hijo único* deberíamos considerarla como colocación o como combinación libre. La conclusión es que sí, porque el sustantivo *hijo* selecciona al adjetivo *único* con la significación de ‘sin hermanos’.

En otras ocasiones, he encontrado problemas a la hora de dilucidar si se trataba de una colocación o de una expresión idiomática. Así, por ejemplo, en una primera lectura *perros calientes* fue considerado como una colocación mejorable (sustituible por *perritos calientes*), pero en una segunda reflexión se ha incluido dentro del grupo de expresiones idiomáticas, dado que no tiene un significado composicional.

Se han considerado colocaciones las llamadas construcciones con verbo de apoyo (CVA en adelante). Estas construcciones del tipo *hacer un viaje*, *tomar una decisión*, *tener éxito*, son casos de selección léxica de tal forma que la base (*viaje*, *decisión*, *miedo*) selecciona uno u otro verbo de apoyo o ligero que, desde un punto de vista semántico cede todo el peso semántico al sustantivo, quedando casi vacío de significado y adquiriendo la función de portador de marcas gramaticales. De acuerdo con Alonso Ramos (2004), el verbo de apoyo no es elegido de acuerdo con una base semántica sino arbitrariamente, en función del nombre. Dichos nombres se consideran predicados en el sentido que apunta la *Nueva Gramática de la Real Academia Española categoría que designa estados, acciones, propiedades o procesos en los que intervienen uno o varios participantes* (2009: 64). Estos participantes

son los argumentos, que vienen exigidos por su significado léxico. Así, lo que hace el verbo de apoyo es proporcionar posiciones sintácticas para que los actantes del nombre puedan aparecer en un contexto oracional. El verbo *hacer* en *Juan hizo un viaje por Latinoamérica*, permite la aparición del argumento locativo del sustantivo *viaje*: *por Latinoamérica*.

No siempre es fácil distinguir cuándo un verbo funciona como verbo de apoyo o es portador de significado y cuando el sustantivo que hemos denominado predicado no es un hecho o una situación. Esta tarea ha resultado particularmente ardua en el caso de las construcciones con *tener* en construcciones como *tener hijos*, *tener los ojos negros*, etc, donde decidir si el verbo *tener* significa posesión es problemático, y donde los sustantivos no denotan hechos o situaciones. Para resolverlo, hemos acudido al concepto de *cuasipredicado* de Mel'cuk y Polguère (2008). Para estos autores, un cuasi predicado es una entidad pero que está relacionada con una situación o hecho que implica la intervención de otros argumentos.

est particulière, en ce sens qu'elle est définie plutôt par son implication dans un fait que par ses propriétés intrinsèques ; à cause de cela, un quasi-prédictat contrôle des positions actanciels. Par exemple, 'professeur', qui dénote une entité (un individu), dénote avant tout une entité impliquée dans une situation particulière : un individu X qui enseigne Y à des individus Z (dans une institution W)⁴.

Los nombres relacionales como *madre*, *hijo*, *ojos*, *pelo*, etc. de los

⁴ Un cuasipredicado es particular, en el sentido de que es definido más por su implicación en un hecho que por sus propiedades intrínsecas; a causa de esto, un cuasipredicado controla posiciones actanciales. Por ejemplo, "profesor", que denota una entidad (un individuo), denota ante todo una entidad implicada en una situación particular: un individuo X que enseña Y a unos individuos Z (en una institución W). (*t.p.*)

que hemos encontrado muchos ejemplos en el corpus, son cuasipredicados que seleccionan el verbo *tener* para vincularse con su argumento. Así pues, las construcciones *tener hijos*, *tener los ojos negros*, *tener el pelo largo*: los nombres relacionales de partes del cuerpo y los de parentesco participan en estos casos de CVAs y se sirven del verbo *tener* para vincularse a su argumento. Se han considerado, por consiguiente, casos colocaciones léxicas.

4.2.2. ¿Es un error? ¿De qué tipo?

Muy ligada a la consideración de si una secuencia representa una colocación o no, está la consideración de qué tipo de error es. En un principio, se dudaba de si considerar una secuencia como **soy contento* como un error de elección del colocativo, en cuyo caso asumiríamos que el adjetivo *contento* selecciona léxicamente al verbo *estar*. No obstante, no nos parece que sea este el caso porque representa el uso regular del verbo *estar*: se ha optado por no considerar *estar contento* como un caso de selección léxica y por tanto obviar este error para los fines de esta investigación, considerándolo como un error no de tipo léxico sino gramatical. Lo mismo ocurre para los casos de elección de *muy* o *mucho* puesto que ante secuencias como **mucho guapa*, el aprendiz desconoce la regla gramatical de que para intensificar adjetivos y adverbios se utiliza *muy*, mientras que para intensificar verbos o sustantivos se utiliza *mucho*. Serían, por tanto, errores de tipo gramatical.

La decisión de considerar una colocación como correcta, mejorable o incorrecta ha descansado sobre tres pilares. En primer lugar se ha acudido al corpus de nativos para constatar sus preferencias combinatorias frente a las de no nativos. Además, se ha utilizado el Corpus del Español Actual (CREA) para comprobar si una secuencia efectivamente se usa –hemos exigido la aparición de al menos cinco casos– y al *Diccionario Combinatorio Práctico del Español* (2006).

Otra cuestión es la de la permisividad con respecto a la creatividad del que escribe. I. Bosque, en su introducción a *Práctico*, alude a ello como *el problema de la creatividad*, donde se defiende de las críticas que se hacen a la lexicografía combinatoria y consideran a los diccionarios combinatorios como un *listado de ramplonerías* o un *enorme cajón de rutinas* (2006: XV). Bosque reconoce la existencia de clichés idiomáticos pero cree necesario trazar la frontera entre esas expresiones manidas y *otras agrupaciones de palabras, igualmente acuñadas, que se sustentan en bases semánticas restringidas y están integradas en el sistema lingüístico* (2006: XV). Véase por ejemplo la secuencia encontrada **el sol es presente* que si sustituimos por *el sol está presente* sería una secuencia gramaticalmente correcta. ¿Debemos considerarla una combinación incorrecta que debiera ser sustituida por *hace sol*? Si bien somos conscientes de que a la hora de determinar si algo es correcto o incorrecto somos más permisivos con el nativo que con el aprendiente, al igual que Cowie, creemos que la creatividad se asienta sobre el conocimiento *a priori* de cómo se combinan las palabras: *the demands of creative expression in the foreign language rests, as it does for native speakers and writers, on prior knowledge of a repertoire of such expressions* (Cowie, 1992).

Otro problema que hemos tenido que enfrentar en el proceso de etiquetado de errores es el tratamiento de aquellas expresiones que designan realidades que no existen en nuestro entorno. Como consecuencia de ello, no contamos con expresiones que las designen. Así ocurre con la secuencia *agricultura fondo*, inserta en la secuencia "Mi hermana, Susan, terminaba en ciencias de animales, y yo estudiaba las letras con una agricultura fondo" en la que el aprendiz parece referirse a una realidad del sistema educativo americano.

Encontramos también problemas concernientes a la variación lingüística. Así, por ejemplo en un primer momento las secuencias *jugar algunos deportes, jugar voleibol* se habían considerado erróneas por la falta de preposición. Sin embargo, aparecen casos en el CREA que confirman que sí es adecuada.

Por otra parte, no hemos tenido en cuenta aquellos errores que se han catalogado como *lapsus calami* como *tomando muchas fojos* que directamente se ha sustituido por *tomando muchas fotos* o *tocaré química orgánica dos*, sustituido por *tomaré química orgánica dos*.

Hemos obviado, asimismo, lo que hemos denominado secuencias indescifrables del tipo: ** rindió también discutiblemente ella muy exitoso solo fechar*.

5. ANÁLISIS Y COMENTARIO DE LOS DATOS

Se han analizado noventa y cinco redacciones de estudiantes de español clasificados por debajo del porcentaje 60% por los administradores del corpus. En estas composiciones, se ha encontrado un total de 328 colocaciones, de las cuales 144 se han considerado correctas, 31 mejorables y 153 incorrectas. En la figura siguiente se pueden observar los porcentajes correspondientes:

Porcentaje de colocaciones empleadas por los aprendices

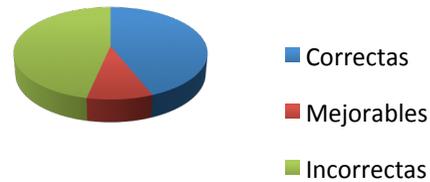


FIGURA 2: Porcentaje de colocaciones empleadas por los aprendices

Dado que un análisis cuantitativo de los datos no proporciona información significativa, tras este primer recuento hemos procedido al análisis cualitativo, que presentaremos en dos partes. En primer lugar, haremos una clasificación de los errores proporcionando los ejemplos encontrados para cada tipo; seguidamente, presentaremos un comentario de los errores con referencia a las producciones de los nativos, a corpus disponibles en línea –fundamentalmente al CREA– y al *Diccionario combinatorio Práctico del español*.

5.1. CLASIFICACIÓN DE LOS ERRORES

A continuación expondremos la tipología de errores colocacionales elaborada por Alonso Ramos, Wanner, Vincze y otros autores y que exponen en el artículo “Towards a motivated annotation schema of collocation errors in learner corpora” (2010).

A grandes rasgos, esta tipología de errores distingue tres dimensiones: la localización del error, es decir, si el error afecta a la base, al colocativo o a la colocación entera; la naturaleza del error –léxica, gramatical o de registro– y, por último, la explicación del error –si este es de carácter interlingual o intralingual–.

5.1.1. La localización del error

Los errores que presentan las colocaciones, como decíamos, pueden afectar a la base, como en *tengo muchos memorias*, por *tengo muchos recuerdos*; pueden afectar al colocativo, como en *ganar un grado*, en lugar de *obtener un grado* y a la colocación entera, como en *tenemos mucha diversión*, en lugar de *nos divertimos mucho* o *tener un precio* en lugar de *ganar un premio*.

5.1.2. La descripción del error

A nivel descriptivo, una colocación puede ser errónea desde el punto de vista léxico, gramatical, o de registro.

Los errores de tipo léxico que pueden afectar tanto a la base como al colocativo son los de *sustitución* y los de *creación*. En los de *sustitución* se emplea erróneamente una base o un colocativo que ya existe en español. Ocurre así en los ejemplos de *gasto tiempo*, en lugar de *paso tiempo* o de *ganar un grado* en lugar de *obtener un grado*. En los errores de creación se usa una palabra inexistente en español como base o como colocativo en una colocación. Ocurre así en *el corre de cuatro temporadas*, en lugar de *el pase de cuatro temporadas*.

El resto de subclases de errores de tipo léxico conciernen a la colocación entera. Pueden ser:

- Síntesis: el aprendiz crea una nueva unidad léxica en lugar de utilizar una colocación. Así ocurre en los casos de *forénsicas* en lugar de *medicina forense*,
- Análisis: el aprendiz crea una colocación en lugar de utilizar una unidad léxica simple, como ocurre en *tener mucha diversión*, en lugar de *divertirse mucho*.

c) Sentido diferente: el aprendiz utiliza una colocación correcta pero con un sentido diferente. Para ilustrar este tipo de error, puede servir el empleo de la colocación *hacer escuela* con el sentido de *ir a la escuela*. *Hacer* o *crear escuela* significa crear una tendencia que es seguida por otros.

d) Construcción diferente⁵: el aprendiz utiliza una colocación inexistente en español para expresar un sentido que en español se expresa con una construcción distinta, no necesariamente una colocación. Es el caso de *les dan un tiempo difícil* una secuencia de palabras que existen en español pero que en ningún caso forman colocación y cuyo sentido se expresa de forma totalmente diferente: *se lo hacen pasar mal*.

Los errores de tipo gramatical pueden estar también relacionados o bien con la base, bien con el colocativo o bien con la colocación entera. Entre los que afectan a la base, están los errores de determinante, de género, de régimen, o de especificación. La presencia o ausencia errónea de determinante, la vemos en ejemplos como *tiene el novio* por *tiene novio*, o *tiene pelo negro* por *tiene el pelo negro*; el género erróneo en *tiene muy bonita dientes*; errores de régimen preposicional, en *conseguir a un Maestro en...* en lugar de *conseguir una maestría/ un máster en...* y errores de especificación, donde un modificador obligatorio no aparece. Los ejemplos de esta categoría encontrados en el corpus van vinculados a una utilización errónea del artículo, como en *tiene una novia*, donde, dada la aparición del artículo indeterminado, se requiere la aparición de un modificador: *tiene una novia muy guapa/ muy simpática*.

⁵ Esta categoría no aparece en la tipología de Alonso Ramos, ha sido añadida por nosotros.

Los errores gramaticales de colocativo pueden estar también relacionados con el régimen, como en *tengo vacaciones A Florida*. Alonso Ramos et al. (2010) registran también errores de pronombre, en aquellos ejemplos donde se usa erróneamente o se omite un pronombre reflexivo. No se han encontrado ejemplos en este corpus.

Por último, hay errores gramaticales que afectan a la colocación entera como errores de orden de palabras, como por ejemplo: *grados buenos* por *buenas notas*; *libre tiempo* por *tiempo libre*.

En cuanto a los errores de registro, hay que apuntar que no se han encontrado en el corpus, quizá porque la variación diafásica no afecta tanto a la interlengua de los aprendientes con un bajo grado de dominio.

5.1.3. La explicación del error

A nivel descriptivo los errores se han clasificado en interlinguales, es decir, aquellos errores que están causados por la transferencia con la L1 del estudiante; y los intralinguales, cuyas causas hay que buscarlas únicamente en la lengua meta.

Todos los errores descritos con anterioridad, ya sean gramaticales, de registro o léxicos, pueden ser analizados como errores interlinguales o intralinguales. Por ejemplo, un error gramatical como *tiene pelo moreno* sería etiquetado como error interlingual gramatical, porque asumimos que la no presencia del artículo se explica por influencia del inglés: *she has dark hair*.

Los errores léxicos de carácter interlingual están divididos en dos subclases:

a) de importación. Este error consiste en la creación en español de una unidad léxica que proviene bien de su lengua materna o de otra. Los casos encontrados son aquellos en los que el

aprendiente no realiza adaptación alguna de la unidad léxica: la toma directamente de su lengua, como en *recibir un degree* por obtener un graduado.

b) de extensión, en los que el aprendiente extiende el significado de una unidad léxica existente en español, como en *liberar un álbum*, donde se elige *liberar* porque es una traducción posible de *release*.

Los errores léxicos de carácter intralingual son clasificados por Alonso Ramos et. al. (2010) en tres subtipos:

a) Derivación errónea: errores en los que el aprendiz produce una forma como resultado de una derivación errónea. Según Alonso Ramos y otros autores, se trata de una forma inexistente en la L2. Sin embargo, hemos incluido en esta categoría la colocación errónea *hacer una apariencia*, en lugar de *hacer una aparición*, donde *apariencia* existe en español pero no es la base adecuada en este contexto. En el caso de *persona influyente* en lugar de *persona influyente*, *influyente* sí es inexistente en español.

b) Generalización: el aprendiente selecciona una unidad léxica de significado más vago o más genérica. Sucede muy a menudo con las construcciones con verbo de apoyo *hacer*, como en *hago notas buenas*, en lugar de *saco buenas notas*; *hacer una vida* en lugar de *llevar una vida*.

c) Selección léxica errónea, error en el que se da una selección errónea de una unidad léxica sin aparente influencia de su lengua materna, como en *ganar medicina para la gente*.

5.2. EMPLEO DE COLOCACIONES POR PARTE DE NATIVOS Y NO NATIVOS

Se han analizado 24 redacciones que responden al título *¿Cuáles son tus planes para el futuro?* de muy distinta extensión, que van desde las 28 palabras hasta alrededor de 700 palabras.

Como ya hemos expuesto en el apartado de metodología, abordaremos un análisis cualitativo de las colocaciones empleadas por los aprendices y valoraremos su corrección a través de la comparación con el uso de los nativos, los corpus de la RAE y el de Mike Davies y el *Práctico*. Para simplificar la exposición, organizaremos el análisis de acuerdo con el campo semántico de la base de la colocación.

Con las bases *universidad*, *carrera* (o el nombre de una carrera concreta), *estudios* los nativos utilizan los colocativos *terminar*, *finalizar* mientras que los informantes no nativos eligen más frecuentemente el colocativo *completar*. (*completar escuela*, *completar universidad*, *completar mis clases*) resultante de la traducción de la combinación, muy frecuente en inglés *to complete a degree*. Con las bases que denotan títulos del sistema educativo, tales como *máster/ maestría*, *doctorado/ PHD*, *grado*, *licencia* para expresar la acción de obtener dicho título, los nativos han elegido los colocativos *terminar X*, *finalizar X* o las unidades sintéticas tales como *recibirse* o *doctorarse*. En las producciones de los no nativos no aparecen esas formas sintéticas –excepto *graduarse*, de la que sí se registra alguna aparición–, lo que podría deberse a que no existen formas equivalentes en su lengua materna. Sin embargo, las formas analíticas como *terminar mi doctorado*, *finalizar estudios*, son igualmente empleadas por los nativos, lo que nos lleva a considerarlas totalmente válidas. No podemos considerar acertada la elección del colocativo en ejemplos encontrados como **ganar un grado*, **ganar a un maestro y un PHD*. Tales casos, de acuerdo con

la tipología establecida, son errores léxicos de carácter interlingual de importación. El aprendiente extiende el significado de *earn* una unidad existente en su lengua materna y elige *ganar* porque es una posible traducción de *earn*, pero no en la expresión *earn a degree*. La elección de *obtener* y *conseguir*, que encontramos en algunos ejemplos del corpus, vienen recogidas por *Práctico* para la entrada de –doctorado. Para el sentido de ‘asistir’ se ha recogido **hacer escuela*, que (bajo la forma del infinitivo) solo aparece en cinco ocasiones en el CREA y con un significado idiomático, el de ‘crear una nueva tendencia que va a ser seguida o imitada’. Se recogen también *ir al la Emory para la escuela de medicina* con el sentido de *ir a la Emory*, así como la expresión errónea **attender* Morehose para medicina. Para expresar el significado de asistir, *Práctico* recoge *asistir (a)*, *ir (a)*, *acudir (a)*. En cuanto a las unidades que hacen referencia a los resultados académicos, se ha encontrado una aparición de *sacar buenas notas* y otra de *hacer notas buenas*, con elección errónea del colocativo en este último caso. No hay apariciones para tales expresiones en las redacciones de nativos analizadas, pero *Práctico* registra *sacar* y *obtener*. En cuanto a la elección de la base, es frecuente también un error de transferencia ya registrado por Vincze (et.al.) (2001: 283). Al igual que en este trabajo, hemos notado que la unidad *colegio* es empleada no con el sentido que en general se le da en español de ‘escuela primaria’ sino con el sentido de *universidad* y que muy probablemente se deba al parecido únicamente formal que hay entre las palabras *college* y *colegio*, fenómeno conocido como el de *cognados o falsos amigos*. En cuanto a las expresiones relativas a materias y asignaturas, se ha encontrado la colocación *idioma nativo*, además de otras que se han considerado erróneas, tales como *escuela de ley*, que nos parece una traducción literal de *School of Law* y que genera un error de creación por dar a la unidad *ley* un sentido que no tiene. Asimismo, se registra un error de síntesis por creación de una unidad léxica simple inexistente en español *forensicas* (adaptación de *forensics*

que en español se expresa como *medicina forense o medicina legal*) y el caso contrario de análisis, con la expresión *ciencias de animales*, que no aparece en los corpus consultados, en vez de *zoología*; error gramatical de número, como en *ciencia política* por *ciencias políticas* y un último error en la elección del colocativo en el ejemplo *tomaré química orgánica dos*. Para expresar este sentido *estudiar* y *cursar* son las formas que aparecen en las muestras de los informantes nativos. *Práctico* sugiere *elegir*, *cursar*, *estudiar*, *matricularse en* que acompañarían frecuentemente a la entrada *asignatura*.

Se han analizado, por otra parte, 45 redacciones que responden al título *¿Cómo es la región donde vives?* que van también desde la breve extensión de 45 palabras hasta redacciones de cerca de 1000 palabras. Como referencia, hemos extraído 41 redacciones del corpus escritas por hablantes nativos del español.

En este caso, nos hemos fijado en primer lugar en las colocaciones relativas al campo de la meteorología. Existen numerosas colocaciones relativas a este campo semántico que se expresan con el colocativo *hacer*: *hacer calor*, *hacer frío*, *hacer sol*, *hacer viento*, etc. Se han registrado varios ejemplos de empleo adecuado de estas colocaciones: *hace viento*, "*hace mucho calor [...] hace fresco*". Ejemplos también de acertada elección de aparición de este tipo de colocaciones pero insertadas en una estructura sintáctica errónea, como en *el otoño de Atlanta hace bueno [...] el verano de Atlanta hace mucho calor y sol* donde la estructura se interpreta como personal y el complemento locativo como sujeto gramatical del verbo *hacer*. En el corpus hemos encontrado ejemplos de elección errónea de este colocativo. En lugar de este, aparece muy frecuentemente el verbo *ser*: "*en la primavera es mucho sol [...] en el verano es mucho calor [...] en invierno es mucho frío*", "*cuando es frío*", "*es mucho frío en invierno*". Asimismo, se han encontrado

otros casos donde prima la elección del colocativo *tener*, como en "*tiene mucho calor también*", "*veces tiene muy calor y veces tiene muy frío*".

Por otra parte, sucede también que se extiende el uso del colocativo *hacer* a otros contextos donde resulta erróneo, como en "*hace llueva*", o "*hace hurricanes*" (34), que se construiría con el verbo *haber*. A nivel explicativo, se trataría de un error intralingual de generalización de acuerdo con la clasificación que manejamos. Aparece también la expresión *hay lluvia*: aunque no es incorrecta, vemos que los nativos prefieren la expresión sintética *llueve*.

Otra serie de colocaciones que interesa reseñar concernientes al campo de la meteorología son aquellas que se construyen con la base *temperatura*. Tenemos el ejemplo "*en invierno la temperatura es no caliente. La temperatura es mucho frío*". De acuerdo con *Práctico*, el adjetivo *frío* sí se puede aplicar a temperatura, mientras que su antónimo para el sentido contrario se elegirían los colocativos *cálida, alta, elevada*. Respecto al adjetivo *caliente* en *Práctico* no aparece junto a *clima, tiempo* o *temperatura*. Es *cálido* el que lo hace.

Por otra parte, en estas redacciones hemos seleccionado también aquellas expresiones relativas al campo del ocio y del tiempo libre. Aquellas colocaciones formadas con la base *deporte* o cualquiera de sus hipónimos tales como *baloncesto, fútbol, tenis*, se combinan con el verbo *practicar* entre los nativos y, según *Práctico*, también con *hacer* o *jugar a*. Entre las redacciones de no nativos analizadas, hemos encontrado muy frecuentemente la combinación con el verbo *jugar* y sin preposición. No hemos considerado errónea esta colocación porque aunque no aparecen casos en el CREA para *jugar deporte*, aparecen siete casos para *jugar baloncesto* y veintiún casos para *jugar tenis* (solo dos de ellos en documentos procedentes de

España). Encontramos también error de régimen preposicional en la colocación que aparece **andar el bicicleta* en lugar de *andar en bicicleta*. Asimismo, junto a la expresión sintética *pescar*, registramos la creación de la colocación *agarramos pez*. Respecto a las expresiones para designar el lugar donde se practican estos deportes, *Práctico* recoge *cancha de*, mientras que en las redacciones hemos encontrado la palabra en inglés directamente sin adaptación alguna, como en **fútbol de americano field* o en *basketball court*. Por último, hemos registrado la expresión **agua actividades*, donde se observa un error de colocativo que resulta ser la traducción literal de una palabra de su L1 (*water activities*) en lugar de lo que sería la correcta derivación morfológica, *actividades acuáticas*.

En cuanto a las colocaciones con la base *vacaciones*, en las redacciones escritas por nativos aparece siempre en plural y se combina para formar colocaciones V+S como *pasar las vacaciones, empezar las vacaciones, necesitar unas vacaciones*. En *Práctico* aparece la entrada en plural y se combina con verbos como *tomarse unas vacaciones* o *irse de vacaciones*. En el caso de los extranjeros encontramos errores en el número de la base, como en **tienes un vacación*; la construcción incorrecta **tengo vacaciones a Florida*, donde se cruza con *voy de vacaciones a Florida*. En cuanto a la base *viaje*, los nativos la combinan con el verbo *hacer* o *ir de* y entre los extranjeros encontramos la expresión *vamos viajes*, con omisión de la preposición requerida. Con esta misma estructura de *ir de* + sustantivo, encontramos entre los extranjeros *ir de compras* pero **ir en picnic*.

En las veintiséis redacciones analizadas con el título *Habla de una persona famosa*, abundan las colocaciones que son construcciones con verbo de apoyo con el verbo *tener*, tales como *tener el pelo rubio, tener los ojos negros*. La elección tanto de base como de

colocativo en estos casos suele ser acertada, quizá porque en inglés se elige el mismo verbo de apoyo, *to have*. Los errores en estos casos suelen ser por ausencia de determinante, como en *tiene pelo rubio*, *tiene pelo y ojos negros*. Sí se han encontrado más errores, como es de esperar, por la utilización del verbo *ser* en lugar de *tener* para la expresión de la edad, como en *mi hija es trece años*. Son construcciones correctas, asimismo, otras que se han encontrado también con el verbo tener como *tener corazón*, *tener compasión*, *tener éxito*, *tener potencial*. Conviene reseñar la creación por análisis de la colocación *tener un buen tiempo* en lugar de *divertirse*.

Dentro del campo de la música y el éxito, se han encontrado colocaciones erróneas como *golpear la radio*, traducción literal de *hit the radio*. En español, esta expresión no activa ningún tipo de significado idiomático, con lo que significa únicamente 'dar golpes a la radio'. Asimismo, se han encontrado junto a la base *álbum*, colocativos como *registrar* o *liberar*, casos para los que *Práctico* registra *presentar*, *lanzar* o *sacar*. *Liberar un álbum* es un error léxico de colocativo, de sustitución, interlingual de extensión del significado de *liberar*: una posible traducción de *release*, además de *lanzar* cuando se combina con *disco*, *película*, *álbum*, es *liberar*, cuando se combina con rehén o preso, como en *to release a prisoner*, *to release a hostage*.

6. CONCLUSIONES Y LÍNEAS DE INVESTIGACIÓN FUTURA

El análisis de la competencia colocacional de los aprendientes de ELE en los primeros estadios del proceso de aprendizaje nos permite reconocer la necesidad de atender a este tipo de unidades léxicas desde los niveles iniciales, dada la elevada frecuencia de aparición de las colocaciones y los errores colocacionales encontrados.

A pesar de que en los últimos años se ha producido un aumento de la bibliografía especializada en las colocaciones léxicas y de su aplicación al aula (Higueras, Ruiz Gurillo, Alonso Ramos), y se han publicado importantísimas obras lexicográficas en esta línea (Redes, Práctico, DICE) el salto al aula se está produciendo con lentitud, y los materiales disponibles para el profesor son escasos y *poco refinados* (COLOCATE, 2012).

El análisis de errores realizados nos permitirá, en investigaciones futuras, recoger y nivelar las colocaciones más frecuentes y rentables en español, completando así las propuestas del *PCIC*. Este documento curricular, si bien, tal y como reza su introducción *se sitúa en una línea de investigación que parte de la base de que un hablante cuenta, además de con unidades léxicas simples o palabras, con un número amplio de bloques semiconstruidos que puede combinar al hablar*, la combinatoria ofrecida está lejos de ser exhaustiva y necesita ser completada para ofrecer así una referencia tanto a profesores como a alumnos sobre las colocaciones que deberían formar parte de cada nivel.

A pesar de que, como ya se ha mencionado, ha aumentado considerablemente la bibliografía de colocaciones en español, se echan en falta también estudios que prueben empíricamente la rentabilidad su enseñanza- aprendizaje. Para poder defender y promover la inclusión de las colocaciones léxicas en currículos, *syllabus* y métodos, hacen falta estudios que den pruebas fehacientes de que su aprendizaje en un determinado contexto educativo, así como de que el entrenamiento en estrategias para segmentar el léxico tiene una incidencia positiva en el aprendizaje.

Por otra parte, a la luz de los resultados obtenidos, nos preguntamos si es realmente rentable a la hora de enseñar y aprender una segunda lengua, transmitir al aprendiz la delimitación entre casos de

colocaciones léxicas y de combinaciones libres, expresiones idiomáticas o institucionalizadas del tipo de *Buenos días, ¡que aproveche!* Más que esto, lo que nos parece necesario es entrenar al aprendiente para segmentar el léxico en unidades complejas, y quizá obviar las distinciones terminológicas manejadas por los expertos. *Pelo corto* no será analizado como un caso de coocurrencia léxica restringida si no una combinación libre de palabras, pues parece que su combinación correcta se deriva del conocimiento del significado de la palabra *corto* y *pelo* no selecciona un significado determinado de este adjetivo cuando se combina con él. Parece que si el aprendiente sabe la diferencia de significado que existe entre *corto* y *bajo*, entonces sabrá que a la hora de describir a alguien si quiere hablar de su estatura utilizará el adjetivo *bajo* y cuando quiere hablar de su pelo, utilizará *corto*. Sin embargo, hay otros pares binarios cuya diferencia en el significado no se aprecia si no es contemplándolos en el eje sintagmático, es decir, viendo con qué palabras se combinan. Así ocurre con el caso de *cálido* –que da calor- y *caliente* –que tiene o produce calor⁶-. Ante estas definiciones, un hablante de español como L2 –o incluso un hablante nativo- se puede preguntar ¿Hay verdaderamente una diferencia? ¿Cuál utilizaré al lado de *región*? ¿Y con *café*? Estas respuestas solo se pueden responder teniendo en cuenta la dimensión combinatoria del léxico. ¿Podemos decir entonces, que las palabras tienen un significado y en función de este se combinan? ¿O más bien, como apunta Bosque (2004, XVIII) en gran medida *se combinan de cierta manera porque expresan precisamente esos significados*? Independientemente del hecho de que *caliente* o *corto* en los casos de *pelo corto* y *café caliente* no sean casos de selección léxica como lo son *nariz respingona* o *calor sofocante*, el caso es que su combinatoria nos proporciona datos sobre su significado. Por ello de cara a la rentabilidad pedagógica de su aprendizaje en bloque

⁶ Definiciones tomadas del DRAE.

pueden ser tratados de igual manera en el aula que las colocaciones léxicas.

No obstante, algunos de los errores encontrados del tipo *tengo vacación a Florida, nos advierten que la enseñanza-aprendizaje de los bloques como *tengo vacaciones*, no está exenta de algunos problemas. El peligro reside en que aprender dichas fórmulas sin analizarlas –a parte del esfuerzo memorístico que supone- puede llevar a cometer este tipo de errores, de inserción de la construcción en una secuencia errónea. Por ello, no somos partidarios de llevar el enfoque léxico hasta el extremo y coincidimos con Gragner (1998) en que *it would be a foolhardy gamble to believe that it is enough to expose L2 learners to prefabs and the grammar will take care of itself*.

En líneas generales, el análisis y estudio de este corpus y de los errores colocacionales nos lleva a afirmar la importancia de enseñar y aprender a segmentar la lengua en lo que Lewis llamó los *pedagogical chunks*, también en nuestra lengua materna. Producir una secuencia como *ganar un grado*, revela una segmentación errónea del *chunk* en inglés *to earn a degree*. Al no contemplarlo como una unidad léxica, el aprendiente lo traduce palabra por palabra de forma aislada, lo que le lleva a producir *ganar un grado*, pasando por alto que los grados no se ganan, se obtienen. De hecho, se ha observado que la mayoría de colocaciones acertadas usadas por los aprendientes se dicen igual en inglés. Tal es el caso de *hacer dinero* (*make money*), *tener hijos* (*to have kids*) o *tomar fotos* (*to take pictures*).

7. REFERENCIAS BIBLIOGRÁFICAS

Alonso Ramos, M. (1994). "Hacia una definición del concepto de colocación", *Revista de Lexicografía*, 1, pp. 9-28. Disponible en http://ruc.udc.es/dspace/bitstream/2183/5383/1/RL_1-1.pdf

Alonso Ramos, M. y Sanromán Vilas, B. (2000)- «Construcción de una base de datos de colocaciones léxicas», *Revista de la Sociedad Española de Procesamiento del Lenguaje Natural* 24,97-98.

Alonso Ramos, M. (2004). *Las construcciones con verbos de apoyo*. Madrid: Visor Libros.

Alonso Ramos, M. (2006a). "Glosas para las colocaciones en el Diccionario de Colocaciones del Español". Alonso Ramos, M.(ed.), *Diccionario y Fraseología*, Coruña: Servicio de Publicaciones de la Universidad de La Coruña, Anejo de la Revista de Lexicografía, pp. 59-88.

Alonso Ramos, M. (2006b). "Towards a dynamic way to learn collocations in a second language". Corino, E., C. Marelló y C. Onesti (eds.). *Proceedings of the Twelfth EURALEX International Congress*. Turín: Accademia della Crusca, Università di Torino, Edizioni dell'Orso Alessandria, 2, pp. 909-923.

Alonso Ramos, M. Wanner, Leo et. al. (2009). «Tagging collocations for learners», en Granger, S y Paquot, M. (eds.), *Lexicography in the 21st century: New challenges, new applications, proceedings of Elex2009*. Louvain-la-Neuve, Presses universitaires de Louvain-UCL.

Alonso Ramos, M. (2010). "No importa si la llamas o no colocación, descríbela". Mellado, C. et al. (eds.), *La fraseografía del S. XXI*:

Nuevas propuestas para el español y el alemán, Frank & Timme, Berlin, pp. 55-80.

Alonso, M., Wanner, L., Vincze, O., Nazar, R., Ferraro, G., Mosqueira, E. y Prieto, S. (2011). "Annotation of collocations in a learner corpus for building a learning environment", *Learner Corpus Research*, Louvain-la-Neuve, Belgium.

Álvarez Cavanillas, J. L. (2008). *Algunas aplicaciones del enfoque léxico al aula de ELE*. Biblioteca virtual REDELE.

Barrios, M.A. (2007). "Diccionarios combinatorios del español: diferencias y semejanzas entre *Redes* y *Práctico*". *redELE*, núm. 11. Octubre. Disponible en: <http://www.mepsyd.es/redele/revista11/index.shtml>

Barrios, M.A. (2010). "El dominio de las funciones léxicas en el marco de la teoría sentido-texto", en *Estudios de lingüística del español*, vol. 30. Disponible en línea en: <http://elies.rediris.es/elies30/index30.html>

Bosque, I. (2001). "Sobre el concepto de colocación y sus límites". *Lingüística española actual* XXIII/1: 9-40.

Bosque, I. (dir.) 2004. *Redes. Diccionario combinatorio del español contemporáneo*, Madrid, ed. SM.

Bosque, I. (dir.) 2006. *Diccionario práctico combinatorio del español contemporáneo*, Madrid: ed. SM.

Corpas Pastor, G. (1996). *Manual de fraseología española*. Madrid: Gredos.

Corpas Pastor, G. (2001). «En torno al concepto de colocación», en *Euskera*, 46, 89-108, Bilbao, Real Academia de la Lengua Vasca.

Cowie, A. (1992). "Multiword lexical units and communicative language teaching". In *Vocabulary and applied linguistics*, Arnaud, P. And Béjoint, H. (eds.). London, MacMillan.

Fernando Aramo, V. (2010). *Materiales didácticos para la enseñanza-aprendizaje de las colocaciones. Análisis y propuestas*. En *redELE*, núm. 11. Disponible en línea en: <http://www.educacion.gob.es/redele/Biblioteca-Virtual/2010/memoriaMaster/1-Trimestre/VeronicaAramo.html>

Galisson, R. (1991). *De la langue à la culture par le mots*. Paris, CLE International.

Granger, S. (1998). «Prefabricated patterns in advanced EFL writing: Collocations and formulae», en en A.P. Cowie(ed.), *Phraseology. Theory, Analysis, and Applications*. Oxford, Clarendon Press: 145-160.

Hallyday, M.A.K. (1966). "Lexis as a linguistic level". En C. E. Bazell *et. al. : In Memory of John Firth*. Londres: Longmans.

Higueras García, M. (2011). "Lexical collocations and the learning of Spanish as a foreign language". En Cifuentes Honrubia, J.L. y Rodríguez Rosique, S. (eds.). *Spanish Word Formation and Lexical Creation*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 439-464.

Higueras, M. (2006). *Las colocaciones en la clase de ELE*. Madrid, Arco Libros.

Howarth, P. (1998). «The phraseology of learners' academic writing», en A.P. Cowie (ed.), *Phraseology. Theory, Analysis, and Applications*. Oxford, Clarendon Press: 161-186.

Instituto Cervantes (2006). *Plan curricular del Instituto Cervantes*, tomo 1, Madrid, Instituto Cervantes.

Koike, K. (2001). *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Madrid: Universidad de Alcalá y Universidad de Takushoku.

Lewis, M. (1993). *The Lexical Approach*. Londres, Language Teaching Publications.

Lewis, M. (2000). *Teaching Collocation. Further Developments in the Lexical Approach*. Londres, Language Teaching Publications.

LOZANO, C. (2009). *CEDEL2: Corpus Escrito del Español L2*. En: Bretones Callejas, Carmen M. et al. (eds) *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Hoy: Comprendiendo el Lenguaje y la Mente*. Almería: Universidad de Almería, pp. 197-212.

Mel'čuk, I. et al. (1984-1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques I-IV*, Montreal: Les Presses de l'Université de Montréal, Montréal.

Mel'čuk, I. (1984). "Le Dictionnaire Explicatif et Combinatoire du Français Contemporain", prolog. al *Dictionnaire Explicatif et Combinatoire du Français Contemporain. Recherches lexicosémantiques I*, Montreal: Les Presses de l'Université de Montréal, 3-15.

Mel'čuk, I. y Polguère, A. (1987) "A Formal Lexicon in Meaning-Text Theory", *Computational Linguistics*, vol. 13, pp. 261-275.

Mel'čuk, I., Clas, A. y Polguère, A. (1995): *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve: Duculot.

Mel'čuk, I., Polguère, A. (2007). *Lexique actif du français*. Bruxelles: Champs linguistiques.

Mel'čuk, I., Polguère, A. (2008). "Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique", en *Lidil*, 37. Disponible en línea <http://lidil.revues.org/index2691.html>

Nation, I.S.P. (2001). *Learning Vocabulary in another Language*. Cambridge: Cambridge University Press.

Real Academia Española (2009). *Nueva Gramática de la lengua española*. Madrid, Espasa Calpe.

Prieto, S., Mosqueira, E, Vázquez, N. (2009). "Córpora y enseñanza de lenguas: se buscan colocaciones", en Cantos Gómez, P., Sánchez Pérez, A. (eds.): *A Survey of Corpus-based Research. Panorama de investigaciones basadas en corpus*, AELINCO, Murcia, 366-373. (Grupos: UDC LYS).

Sinclair, J. (1991). *Corpus, concordance, collocation: Describing English language*. Oxford: Oxford University Press.

Vincze, O., Alonso Ramos, M., Mosqueira, E., Prieto, S. (2011). "Exploiting a learner corpus for the development of a CALL environment for learning Spanish collocations", en Kosem, I. y K. Kosem, eds. *Electronic lexicography in the 21st century: New*

applications for new users. Proceedings of eLex 2011. Ljubljana: Trojina, Institute for Applied Slovene Studies, 280-285.

FECHA DE ENVÍO: 12 DE NOVIEMBRE DE 2014