

LOS CORPUS LINGÜÍSTICOS EN EL PROCESO DE ENSEÑANZA-APRENDIZAJE DE LOCUCIONES VERBALES EN ELE

ANA MARTÍN RÍDER
Universidad de Córdoba

1. INTRODUCCIÓN

GRACIAS AL DESARROLLO CIENTÍFICO Y TECNOLÓGICO, en especial de la informática, han surgido nuevas herramientas digitales que facilitan la investigación en numerosos campos. Algunas de ellas pueden ser muy útiles dentro del ámbito de la enseñanza del español como lengua extranjera (ELE), como ocurre con los corpus lingüísticos.

Desde la década de los 50, la tecnología nos ha permitido disponer de manera muy fácil, rápida y fiable de numerosos textos digitalizados que pueden ser manipulados y usados para realizar búsquedas lingüísticas detalladas. Así, la aparición de estas bases de datos ha revolucionado los estudios lingüísticos, pues ha permitido el desarrollo de investigaciones a partir de grandes volúmenes de datos que posibilitan el estudio de la lengua contextualizada a través de ejemplos reales, ofreciendo una alternativa a los trabajos basados en la introspección y la intuición.

En los últimos años, se ha implementado el uso de corpus lingüísticos en el proceso de enseñanza-aprendizaje de lenguas extranjeras (LE). La combinación de estas nuevas herramientas con la enseñanza de ELE y, especialmente, de unidades fraseológicas (UF) abre un nuevo ámbito de investigación lingüístico. Estas unidades y, en concreto, las locuciones verbales, se caracterizan por su fijación e idiomatismo, rasgos que provocan que su adquisición sea más compleja respecto a la de otras unidades léxicas. Por tanto, la aplicación de una metodología adecuada basada en el aprendizaje autónomo que busca la constante motivación y atención a las necesidades del alumnado puede solventar la falta de atención que reciben las locuciones verbales (LV) dentro de las aulas de ELE.

El presente trabajo persigue esencialmente demostrar que el empleo de corpus lingüísticos dentro de la enseñanza de ELE mejora y facilita el aprendizaje del lé-

xico, particularmente de las LV. Considerando esto, nuestros objetivos específicos son los siguientes: (1) fomentar el aprendizaje autónomo del alumnado; (2) favorecer la inclusión de materiales en los que se presenten muestras de lengua reales; y (3) ofrecer a los docentes recursos de los que extraer ejemplos para enriquecer sus explicaciones.

2. LAS LOCUCIONES VERBALES

Numerosos investigadores dedicados al estudio de la fraseología española (Casares, 1950/1992; Zuluaga, 1980; Higuera, 1996; Corpas Pastor, 1997; Solano Rodríguez, 2012) incluyen las locuciones como parte de sus taxonomías.

En el presente trabajo hemos decidido partir de la definición de Casares (1950/1992: 170), quien caracteriza la locución como una «combinación estable de dos o más términos que funciona como elemento oracional y cuyo sentido unitario consabido no se justifica, sin más, como una suma del significado normal de los constituyentes». A pesar de la antigüedad de la obra del lexicógrafo, es innegable el reconocimiento y la aceptación que esta definición obtuvo entre las investigaciones fraseológicas. Además, sigue siendo válida, ya que recoge las características básicas de este tipo de UF. Como bien señala De Bobes Soler (2016, p. 23), en esta definición se recogen conceptos básicos para el trabajo fraseológico y, por supuesto, todas las características propias de las locuciones: «fijación («combinación estable»), pluriverbalidad («de dos o más términos»), unidad léxica que se integra en la oración («que funciona como elemento oracional») e idiomática («cuyo sentido unitario consabido no se justifica, sin más, como una suma del significado normal de los constituyentes»)». Consecuentemente, estas son las características asignadas a las locuciones por Casares. A partir de dicha definición, el autor distingue dos tipos: las conectivas (que incluyen las conjuntivas y las prepositivas) y las significantes o conceptuales (divididas entre las nominales, las adjetivas, las verbales, las participiales, las adverbiales, las pronominales y las interjectivas). Aunque presenta un análisis detallado de cada tipo, nos vamos a centrar en las locuciones verbales, sobre las que versa el presente estudio.

Según Casares (1950/1992: 171), las locuciones verbales son aquellas «que se componen de un verbo que, asimilando su complemento directo o preposicional, forma un predicado complejo», es decir, la estructura *verbo + complemento* es precisa para su formación. Por tanto, las locuciones verbales son concebidas como predicados, que pueden ser transitivos («*beber los vientos* por una cosa»), intransitivos («*ir a gusto en el machito*») o predicativos («*eso es harina de otro costal*») (p. 177).

Todos estos rasgos los sintetiza De Bobes Soler (2016: 95) con claridad cuando apunta que «una locución verbal se distingue por ser una unidad fraseológica que

incluye en su constitución un núcleo verbal que, por lo general, puede flexionar¹ en todas sus formas; desde el punto de vista de la sintaxis, la locución verbal es un predicado porque tiene argumentos con los que establece una relación de interdependencia». Estamos de acuerdo con esta afirmación, ya que el elemento central de toda locución verbal, como su propio nombre indica, es el verbo. Así, el núcleo de *dar la cara* es *dar* (que puede flexionar según su conjugación), mientras que *la cara* es el complemento que se mantiene fijo en todo momento, como se puede observar en (1).

1. a. *Siempre doy la cara y no me escondo de nadie.*
- b. *Gustavo Guzmán declaró que Márquez daría la cara y respondería ante la prensa en el lugar de el dirigente...*
- c. *Si diesen la cara no lo harían tan mal.*²

En definitiva, observamos que las locuciones verbales son predicados compuestos por un núcleo verbal que requiere de unos argumentos³. Estas se caracterizan esencialmente por su pluriverbalidad, fijación e idiomática.

Como ya hemos mencionado, la idiomática es una de las características básicas de las locuciones verbales. En palabras de Zuluaga (1980: 122),

Idiomática es el rasgo semántico propio de ciertas construcciones lingüísticas fijas, cuyo sentido no puede establecerse a partir de los significados de sus elementos componentes ni del de su combinación o, parafraseando la formulación de Bally («oublí du sens des éléments»), idiomática es la ausencia de contenido semántico en los elementos componentes.

Según esta perspectiva, Zuluaga (1980: 124) señala que son tres los pasos para generar una construcción idiomática: (1) los signos mantienen su apariencia de significantes autónomos; (2) la unidad carece o pierde identidad semántica; y (3) la estructura constituye con sus componentes una unidad de sentido. Es decir, generalmente, el significado de las expresiones idiomáticas no equivale a la suma de significados de sus componentes, sino que es la construcción como unidad la

¹ Sin embargo, dependiendo de la estructura sintáctica y de los elementos léxicos que compongan la locución verbal, la flexión del núcleo se encontrará más o menos restringida. Por ejemplo, raramente nos encontraremos la expresión *ser coser y cantar* en primera o segunda persona, ya que suele emplearse en tercera persona, concretamente, del singular.

² Ejemplos extraídos del *Corpus del español* de Davies (sección «NOW»).

³ Para conocer más acerca de la construcción de predicados, véase Subirats (2001), quien hace una aplicación de la teoría de la sintaxis léxica al español, o Lezcano (1994-1995), quien defiende la teoría de las valencias.

que posee un significado en bloque. Por tanto, en una expresión no idiomática, la identidad semántica de los componentes se mantiene, mientras que, en una idiomática, dicha identidad se pierde y todos los componentes se convierten en una única unidad significativa. De esta manera, el significado «molestar» que extraemos de la expresión *dar la lata* no puede deducirse de la unión de los significados del verbo *dar* y el sustantivo *lata*. Consecuentemente, el aprendizaje de las expresiones idiomáticas (sean o no locuciones) es mucho más complejo y debe estar adecuadamente planificado.

No obstante, la idiomaticidad no se presenta siempre en el mismo grado, sino que existen expresiones con una mayor o menor transparencia de su significado. Es una tarea compleja el establecimiento del grado de idiomaticidad que posee una construcción, ya que, generalmente, la percepción de un significado más o menos literal se debe a cuestiones puramente subjetivas y no estrictamente lingüísticas (García-Page, 2008: 27-28). A pesar de ello, se han realizado numerosas propuestas para ilustrar los diferentes grados de idiomaticidad. Un claro ejemplo lo encontramos en Zuluaga (1980: 129-134), quien distingue tres niveles: (a) expresiones que carecen de sentido literal-regular (*dar en el busilis, a pie juntillas, hacer vaca, en torno a*); (b) expresiones que mantienen cierto sentido literal o lo mantienen de manera inminente (*tomar el pelo, pagar los platos rotos, meter gato por liebre*); y (c) estructuras semiidiomáticas (*recibir con los brazos abiertos, contarse con los dedos de la mano*). Sin embargo, desde nuestro punto de vista, las diferencias que se establecen entre los grupos (b) y (c) no se encuentran claramente definidas, de forma que ciertas construcciones podrían incluirse en ambos niveles.

Por tanto, vemos más acertada la distinción entre dos niveles de idiomaticidad: uno total y otro parcial, como proponen, por ejemplo, Mogorrón Huerta (2002) y Corpas Pastor (2003). Así, establecemos la siguiente distinción en función del grado de idiomaticidad que presenten, en este caso, las locuciones verbales:

1. Locuciones verbales con significado semitransparente o semiidiomático: su significado puede deducirse, por ejemplo, porque se encuentra parcialmente motivado por la realidad a la que remite dentro de un contexto (*ir para largo, morderse la lengua*) o porque solo algunos de los componentes son idiomáticos (*llover a cántaros, saltársele las lágrimas*).
2. Locuciones verbales con significado opaco o idiomático: su significado no es fácilmente deducible (*cantar las cuarenta, salir por patas*).

A pesar de ello, creemos que no se puede establecer una clasificación de las locuciones precisa según su grado de idiomaticidad, dado que, por ejemplo, quizás algún aprendiente pueda ser capaz de deducir por contexto el significado de la locución verbal *sudar la camiseta* («esforzarse mucho en algo»), pero es un hecho que

no podemos asegurar. Asimismo, es determinante la inferencia de su lengua nativa y otras lenguas extranjeras, pues puede conocer en dichos idiomas UF equivalentes o similares a las que encontramos en español.

Este mismo fenómeno ocurre con otra de las características esenciales de las LV: la fijación. Es decir, existe una gradación en función de si la locución permite al hablante una mayor o menor modificación de sus elementos. Como ya indicó Zuluaga (1975), la fijación afecta a numerosos aspectos locucionales, como al orden de los componentes, la categoría gramatical o la sustituibilidad de sus elementos, entre otros. Así, la locución *ser uña y carne* posee una fijación total. No obstante, determinadas locuciones aceptan ciertas modificaciones, provocando que se pueda establecer una gradación de esta característica. Encontramos, por ejemplo, la locución *llegar sano y salvo*, que admite cambios en la persona y el número (*llegó sana y salva*, *llegaron sanos y salvos*), o *quedar en familia*, que permite la inserción de otros elementos entre sus componentes (*todo queda*, *como quien dice*, *en familia*). Por tanto, si establecemos los rasgos que determinan la fijación, se podrían diferenciar grupos de locuciones que, según cumplan con un mayor o menor número de rasgos se aproximarán más o menos a la fijación total. Estos rasgos que determinan la fijación deben conducirnos a establecer diferencias claras entre una locución conjuntiva como *aun así* y una nominal como *pez gordo*.

3. LA ENSEÑANZA-APRENDIZAJE DE LAS LOCUCIONES VERBALES EN ELE

Actualmente, la finalidad del proceso de enseñanza-aprendizaje de cualquier lengua, tanto materna (LM) como extranjera (LE), es el desarrollo de la competencia comunicativa. Para lograrlo, el conocimiento léxico es imprescindible, pero, como señala Lewis (1993: 117), la adquisición de vocabulario requiere mucho más que la simple memorización de unidades léxicas, y ampliarlo no solo significa aprender más de ellas. Es esencial, para aumentar nuestro lexicón, la adquisición de nueva información sobre las unidades que ya conocemos (como nuevas acepciones, los registros en los que se pueden emplear o sus posibilidades combinatorias).

Asimismo, Olimpio de Oliveira (2004: 3) afirmó que «el conocimiento fraseológico contribuye al desarrollo de la competencia comunicativa del aprendiz, a que este tenga una *performance* lo más parecida posible a la de un nativo y a su interacción social». Por tanto, una de las grandes ventajas del dominio de la combinatoria y, por tanto, de las UF es que el discente logre la máxima competencia comunicativa en la LE, ya que contribuye notablemente a la mejora de la fluidez y al alcance de un dominio comunicativo similar al de un hablante nativo. A esto sumamos que, cuando un estudiante adquiere nuevas UF y, por tanto, nuevas locuciones, no solo está ampliando su conocimiento léxico, sino también el cultural, ya que los

fraseologismos recogen numerosos aspectos de la cultura del lugar donde se habla una lengua (Castillo Carballo, 2002).

A pesar de que el aprendizaje de la fraseología supone un gran beneficio para los aprendientes, estos no se encuentran suficientemente presentes en las aulas de ELE, dado que, por sus características lingüísticas, la dificultad de su enseñanza aumenta. Sin embargo, ¿por qué se consideran unidades tan complejas?

Como hemos señalado anteriormente, las locuciones verbales se caracterizan por tres factores: la pluriverbalidad, la fijación y la idiomática. Los dos últimos aspectos hacen que la dificultad de su enseñanza se vea incrementada. Olimpio de Oliveira (2004: 7), refiriéndose a la fijación, resalta que el hecho de que estas unidades no puedan ser modificables ni morfosintáctica ni semánticamente obstaculiza el aprendizaje. Un hablante no nativo de español podría modificar de manera no intencional una UF y, así, cambiar su significado, lo que podría provocar la interrupción del intercambio comunicativo e, incluso, un malentendido. A raíz de esta idea, Szyndler (2015: 205) pone como ejemplo la UF *abrir alguien el ojo* («estar advertido para que no le engañen») frente a su modificación en plural *abrir alguien los ojos* («conocer las cosas como son, para sacar provecho y evitar las que pueden causar perjuicio o ruina»).

Por otro lado, la idiomática es una característica que aumenta indudablemente la complejidad de la enseñanza de locuciones. Así, Fernández Prieto (2004: 352) afirma que «el aprendizaje de las UF es un camino dificultoso por la gran variedad de funciones que poseen, las formas que presentan y en muchos casos por su opacidad y sentido figurado o metafórico, lo que hace que descifrar su significado no sea nada fácil para el estudiante extranjero». Además, desde nuestra perspectiva, puede resultar complejo no solo el conocimiento del significado de la locución, o de cualquier otra UF, en bloque, sino también la determinación de si se está utilizando con su significado idiomático o literal. Por ejemplo, en (2a), la expresión *dar la vuelta a la tortilla* tiene un significado literal, mientras que en (2b) aparece con su significado idiomático.

2. a. *Se le da la vuelta a la tortilla y se cocina durante otros cinco minutos.*
- b. [...] *el PSOE le da la vuelta a la tortilla y se sitúa al frente del grupo que se sienta en los escaños de la oposición [...]*

Por último, existe una serie de dificultades relativas al papel del docente: la selección de las locuciones y la creación de recursos didácticos. Ambos aspectos resultan de la falta de aplicación de las metodologías propuestas por los investigadores dedicados al estudio de la fraseodidáctica. Actualmente, existen diversos modelos de enseñanza de locuciones y de su proceso de selección dentro del ámbito de las lenguas extranjeras. Sin embargo, pocos de ellos se han llevado a las aulas de ma-

nera asidua, donde realmente se comprueba si cumplen su función. Esto, unido a la desinformación por parte de los docentes, causa que las metodologías empleadas en la enseñanza-aprendizaje de ELE no evolucionen, o lo hagan muy lentamente. Como consecuencia, los docentes que desean que los estudiantes aprendan las locuciones u otros fraseologismos no disponen de los conocimientos necesarios para elegir las unidades adecuadas a las necesidades y el nivel de sus discentes, al mismo tiempo que carecen de recursos para trabajarlas correctamente.

4. EL CORPUS LINGÜÍSTICO: DEFINICIÓN Y TIPOLOGÍA

Un corpus lingüístico es una base de datos electrónica en la que se recopilan numerosos textos bajo una serie de criterios y que se toma como una muestra representativa de una lengua en uso. Muchos de ellos están bien clasificados y adecuadamente anotados, con el fin de que los usuarios puedan realizar rápidamente consultas de carácter lingüístico en numerosos textos al mismo tiempo. De hecho, si consultamos el *Diccionario de la lengua española* (RAE, 23.^a ed.), encontramos que se define «corpus» como el «conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etc., que pueden servir de base a una investigación». Sin embargo, esta caracterización es algo deficiente, ya que no se consideran numerosos aspectos característicos de los corpus lingüísticos.

Teniendo en cuenta los estudios de Parodi (2008) y Villayandre Llamazares (2008; 2010), los corpus lingüísticos se caracterizan esencialmente por su carácter electrónico, ya que los textos se encuentran disponibles en entornos computacionales. Esto permite la automatización de ciertas tareas como la búsqueda rápida de información (palabras, secuencias de palabras, categoría gramatical...), la recuperación de todos los casos de una unidad en su contexto inmediato, el cómputo de la frecuencia de aparición de una unidad y la clasificación de los datos según diferentes criterios (frecuencia de aparición, procedencia geográfica, fecha, medio de publicación...). Esos textos se suelen almacenar, generalmente, en formato de tipo digital plano, lo que permite el etiquetaje computacional semiautomático, ya sea a nivel morfosintáctico, léxico, textual... Esto consigue que las búsquedas de información sean más detalladas. Por otro lado, el registro de los datos cuantitativos posibilita la comparación y generalización de cifras.

Es esencial la autenticidad de los datos, pues los textos deben ser muestras reales de uso de la lengua para obtener conclusiones válidas y fiables de su comportamiento. Además, a la hora de elaborar un corpus, se deben tener en cuenta ciertos principios estadísticos, ya que debe reflejar, a pequeña escala, el funcionamiento real de la lengua. Es decir, el corpus debe ser representativo de la lengua o el aspecto de la lengua seleccionado. Aunque la representatividad debe primar sobre el tamaño, los corpus suelen ser preferentemente extensos, por lo que se miden en

millones de palabras. Dicha cantidad de palabras debe estar fijada previamente a la recogida de las muestras, excepto en aquellos que aceptan la inclusión de textos *a posteriori*. Estos últimos destacan por su fácil actualización.

Teniendo en cuenta todas las características anteriores, encontramos algunas definiciones mucho más completas que la propuesta por la RAE anteriormente expuesta. Por ejemplo, Parodi (2008: 106) afirma que un corpus es

un conjunto amplio de textos digitales de naturaleza específica y que cuenta con una organización predeterminada en torno a categorías identificables para la descripción y análisis de una variedad de lengua. Este conjunto de textos debe mostrar, de preferencia, accesibilidad desde entornos computacionales y visibilidad de modo que se posibilite su uso en diversas investigaciones con el fin de asegurar acumulación de conocimientos e integración de la investigación de una lengua particular o en comparación con otra. También debe cumplir con aportar detalles relevantes acerca de su recolección y procedencia.

Por su parte, Torruella y Llisterri (1999: 52) también dan una definición que creemos bastante acertada:

Un corpus es un conjunto homogéneo de muestras de lengua de cualquier tipo (orales, escritos, literarios, coloquiales, etc.) los cuales se toman como modelo de un estado o nivel de lengua predeterminado. El conjunto de enunciados incluidos en un corpus, una vez analizados, debe permitir mejorar el conocimiento de las estructuras lingüísticas de la lengua que representan.

A partir de esta última cita, se deduce la relevancia de que dichas bases de datos se configuren de acuerdo con unas características determinadas preestablecidas para cumplir unos objetivos concretos. Así, de acuerdo con Torruella y Llisterri (1999: 9 y ss.) y Villayandre Llamazares (2010: 349 y ss.), los principales parámetros que se siguen para elaborar y, por tanto, clasificar los corpus son:

- La modalidad de la lengua, ya que pueden incluir textos orales, escritos o ambos.
- El número de lenguas a las que pertenecen los textos (monolingüe, bilingüe o multilingüe).
- El porcentaje y la distribución de los diferentes tipos de texto: depende del volumen de textos que recoja dentro de cada variedad textual.
- Los límites del corpus: dependiendo de si consta de un número finito de palabras preestablecido a su constitución o si se pueden añadir textos *a posteriori*.
- La especificidad de los textos: los textos recogidos en cada corpus variarán en función de si pretende reflejar la lengua general, un tipo particular de sublenguaje, un género textual concreto...

- El periodo temporal que abarcan los textos: si comprenden una época concreta para el estudio sincrónico de la lengua o años sucesivos para el estudio diacrónico.
- El tratamiento aplicado al corpus: con una anotación más o menos específica. Así, encontramos corpus que no disponen de información adicional, mientras que en otros se señalan aspectos relativos a la codificación textual (autor, año...) e, incluso, lingüísticos (categoría gramatical, estructuras sintácticas...).

En definitiva, a la hora de diseñar un corpus, es esencial que se determinen con exactitud la finalidad del corpus, sus límites temporales, geográficos y lingüísticos, su tamaño y los tipos de textos que lo compondrán.

5. VENTAJAS Y DESVENTAJAS DEL EMPLEO DE CORPUS EN EL AULA DE ELE

Los corpus pueden ser de gran utilidad no solo a la hora de realizar ciertos estudios lingüísticos, sino también para la didáctica de lenguas. Numerosos estudios (Pitkowski y Vásquez Gamarra, 2009; Bernardini, 2016; Vyatkina y Boulton, 2017; Abad Castelló, 2019) ponen de relieve las ventajas de su empleo dentro de las aulas de enseñanza de lenguas (maternas o extranjeras), aunque también señalan algunas desventajas. A pesar de todo, estos investigadores, al igual que nosotros, han llegado a la conclusión de que su empleo proporciona muchos más beneficios que inconvenientes.

El aprendizaje del uso de corpus lingüísticos para personas que no tienen experiencia puede ser algo arduo, ya que suelen ofrecer numerosas opciones a las que se accede, generalmente, mediante combinaciones de símbolos gráficos, que en ciertas ocasiones no se aclaran en la sección de «ayuda». Además, al encontrarse en un entorno computacional, no podremos utilizar los corpus en aquellas aulas que no dispongan de ordenadores o acceso a internet. A esto se añade que la conexión a internet podría fallar e, incluso, la página que alberga el corpus puede no estar disponible en el momento en que se desea usar. No obstante, todas estas desventajas son comunes en el uso de cualquier herramienta tecnológica.

De forma más específica, la desventaja que puede presentar el uso de corpus lingüísticos y que puede afectar directamente al aprendizaje de una lengua está relacionada con los textos que el alumnado puede encontrar dentro de la base de datos. El docente debe tener controladas en todo momento las búsquedas realizadas por los estudiantes, de manera que no se frustren o desmotiven por intentar trabajar con textos excesivamente complejos para su nivel lingüístico. Asimismo, el docente

debe ser consciente de que pueden aparecer incorrecciones gramaticales, a causa de que las muestras proceden de un uso real de la lengua:

Procede directamente del uso del lenguaje, y no de la norma académica o lo gramaticalmente correcto, que puede falsear el modelo de lengua que se aprende. Frente a las afirmaciones de gramaticalidad con criterios heterogéneos y a veces contradictorios, o con consideraciones extralingüísticas, se toma como medida objetiva, relativa y graduada el uso actual de la lengua en un corpus (Campillos Llanos, 2005-2006: 10).

A pesar de esto, los beneficios que proporciona el uso de corpus son mayores. En primer lugar, el hecho de que los corpus lingüísticos estén compuestos por textos reales lo convierte en una fuente ideal de ejemplos no artificiales de uso de distintos aspectos lingüísticos. Además, a través de ellos se puede observar la variedad de las lenguas por sus distintas tipologías textuales, modalidades de la lengua, características de los informantes, etc. Otro factor de gran importancia es que, en el corpus, generalmente, se incluye el texto completo donde aparece el elemento que hemos buscado, por lo que el alumnado va a poder observar un hecho lingüístico en distintos contextos junto con todas sus características formales y pragmáticas. Supone también una gran ventaja a la hora de trabajar con textos orales, pues se mostrarán de forma veraz aspectos como las elipsis, los elementos de cohesión, las vacilaciones, las reformulaciones, las marcas de cortesía... propios de esta modalidad.

Asimismo, el docente puede emplear dichas bases de datos como herramienta de consulta de cuestiones lingüísticas que no aparezcan en los libros de gramática o en los diccionarios, herramientas mucho más limitadas. Además, los corpus permiten al profesorado crear actividades novedosas no solo de carácter deductivo, sino también inductivo, consiguiendo que los aprendientes estén activos continuamente y puedan realizar un trabajo autónomo. Esto último facilita, en gran medida, la motivación y la participación en el aula.

Gracias a estos factores, los trabajos lingüísticos basados en corpus se han posicionado frente a los estudios tradicionales fundamentados sobre la introspección y la intuición. El uso de corpus para el desarrollo de investigaciones lingüísticas no solo aporta mayor veracidad y credibilidad a los resultados, sino que pueden servir para corroborar o desechar dichas creencias intuitivas.

A todo esto, añadimos que numerosos corpus lingüísticos se ofrecen de forma gratuita y son fácilmente accesibles, de manera que los estudiantes podrán trabajar con ellos incluso desde sus casas. Así, en poco tiempo, dispondrán de una gran cantidad de información donde pueden realizar consultas lingüísticas muy diversas y consultar numerosos textos de diversos géneros.

En definitiva, el uso de corpus lingüísticos en el proceso de enseñanza-aprendizaje de LE supone la inclusión de una nueva herramienta que ayuda tanto a los docentes como a los discentes. Por tanto, pensamos que el tiempo que se dedica a enseñar el funcionamiento de dichas bases de datos al alumnado no supone un gran problema en comparación con el elevado número de ventajas que aporta su empleo.

6. PROPUESTA DIDÁCTICA: APRENDIZAJE DE LOCUCIONES VERBALES MEDIANTE EL USO DE CORPUS LINGÜÍSTICOS

6.1. CORPUS DISPONIBLES PARA LA ENSEÑANZA DE ELE

Hoy día, existen numerosos corpus lingüísticos que recogen textos del español y que se encuentran disponibles en línea de forma gratuita. En primer lugar, la propia Real Academia Española ha trabajado en la creación de corpus lingüísticos desde 1995, por lo que actualmente podemos consultar las siguientes tres bases de datos:

- *Corpus Diacrónico del Español (CORDE)*: donde se recogen textos de todos los países de habla hispana desde los inicios del idioma hasta 1974. Incluye 250 millones de palabras pertenecientes a textos de distintos géneros. Por tanto, es un corpus enfocado al estudio diacrónico de la lengua.
- *Corpus de Referencia del Español Actual (CREA)*: está compuesto por algo más de 160 millones de palabras de textos tanto orales como escritos de diversos géneros. Estos han sido producidos en todos los países de habla hispana entre 1975 y 2004.
- *Corpus del Español del Siglo XXI (CORPES XXI)*: recoge miles de textos de distintos géneros de todos los países hispanohablantes. Estos fueron producidos entre 2004 y la actualidad. Por tanto, con este corpus se pueden consultar los usos actuales de la lengua.

A pesar de que este último corpus lingüístico es, probablemente, uno de los que más se emplearía en una clase de ELE, todos pueden ser útiles en función de la actividad que propongamos a nuestro alumnado y su nivel de competencia lingüística.

Otros corpus destacables de la lengua española son:

- *Corpus Oral y Sonoro del Español Rural (COSER)*: registra grabaciones de más de 2800 informantes caracterizados por haber nacido y vivir en un entorno rural de escasa escolarización. La edad media de dichos hablantes es de 73 años y pertenecen a 1852 enclaves rurales diferentes. Las muestras fueron

recogidas entre 1990 y 2020, por lo que permite trabajar la comprensión oral con hablantes actuales.

- *Corpus del español (CdE)*, creado por Mark Davies: recoge textos orales y escritos en español desde el siglo XVIII hasta el siglo XXI. Se distinguen, fundamentalmente, tres secciones: (1) para buscar cambios históricos y variación de géneros; (2) «webs/dialectos», donde se incluyen textos de no más de cuatro años de antigüedad extraídos de páginas webs de 21 países hispanohablantes; y (3) «NOW», que recoge 7,2 billones de palabras extraídas de periódicos y revistas digitales de distintos países hispanohablantes entre 2012 y 2019.

Para la aplicación de la propuesta didáctica que se presenta en este estudio, hemos empleado el *CdE*, en concreto la sección «NOW», ya que es la que recoge un mayor número de palabras y la más actual. Además, aunque este corpus ofrezca numerosas opciones de búsqueda avanzadas, la búsqueda básica es bastante intuitiva y, por tanto, de fácil empleo. De esta manera, nuestros estudiantes no sentirán frustración a la hora de aproximarse al funcionamiento del corpus lingüístico.

6.2. PAUTAS GENERALES QUE RIGEN NUESTRA PROPUESTA DIDÁCTICA

La propuesta didáctica que exponemos en este apartado está basada en las siguientes premisas:

1. El enfoque comunicativo, recomendado por las dos obras de referencia para la planificación del proceso de enseñanza-aprendizaje de ELE en nuestro país: el *Marco Común Europeo de Referencia para las Lenguas (MCERL)* (2002) y el *Plan Curricular del Instituto Cervantes (PCIC)* (2007). La adopción de este enfoque permite un mejor desarrollo de la competencia lingüística, ya que está enfocado al trabajo equitativo de las cuatro destrezas: la comprensión y la expresión oral y escrita. Actualmente, se habla de una quinta destreza, a la que se ha denominado «de interacción», que consiste en el intercambio comunicativo (oral o escrito) entre dos o más participantes. Teniendo en cuenta este último reajuste, a través de nuestra propuesta didáctica buscamos siempre el equilibrio de estas cinco destrezas que rigen el enfoque comunicativo. Así, incluimos ejercicios en los que se combinen, por ejemplo, la interacción y la comprensión y la expresión escrita. Además, fomentamos el intercambio comunicativo con el fin de que los aprendientes compartan las ideas y suposiciones lingüísticas que hayan podido extraer de sus investigaciones.

2. La constante atención a las necesidades del alumnado, esencial para el desarrollo de la enseñanza y, por tanto, para la didáctica de lenguas. En el caso de que desconozcamos el bagaje fraseológico de nuestros estudiantes, debemos elaborar una prueba inicial que nos ayude a caracterizar a nuestro alumnado, a determinar sus necesidades y a establecer su base lingüística relativa a este tipo de UF. De esta manera, u obviando este paso si hemos trabajado de forma constante con dicho grupo, podremos fijar las locuciones que deseamos enseñar. Estas pueden clasificarse según diversos factores (campos semánticos, funciones comunicativas, contextos...), que pueden ser adaptados tanto a las necesidades del alumnado como a los requerimientos de la planificación docente.
3. Actividades basadas en la autonomía del alumnado, pues, al trabajar con UF idiomáticas, se suele cometer el error de proporcionar a los estudiantes únicamente el significado de dicha unidad en bloque, sin ofrecerles la oportunidad de descubrirlo de manera autónoma. Esta equivocación proviene del papel que se ha asignado tradicionalmente al docente como principal agente del proceso de enseñanza-aprendizaje de lenguas que debe ofrecer todos sus conocimientos a los aprendientes. Por ello, las actividades que vamos a proponer para la enseñanza de LV a través de corpus lingüísticos van a estar basadas en la autonomía del alumnado. Es decir, el docente actuará como guía de este proceso, pero serán los estudiantes quienes irán descubriendo los significados de estas nuevas expresiones, sus usos y sus características lingüísticas y culturales. Esto no quiere decir que deban enfrentarse solos a la adquisición de las nuevas unidades léxicas, sino que podrán requerir la ayuda del docente en cualquier momento y, aún más importante, de sus compañeros. Esto es, adquirirán, además, la competencia de aprender a aprender.
4. Aprendizaje por descubrimiento, que se ha probado como más eficaz que la metodología expositiva (Herrera y Conejo, 2009; Eleizalde et al., 2010). Así, el estudiante adquirirá sus conocimientos de forma significativa, por lo que podrá retenerlos en la memoria a largo plazo. Asimismo, la motivación aumentará al presentarles este tipo de actividades.
5. El trabajo grupal, que es fundamental para distintos aspectos: (1) la consecución de un aprendizaje más profundo de los contenidos resultante de la experiencia personal y la interacción; (2) la transmisión de valores como la responsabilidad individual, la interdependencia, el respeto, la solidaridad, el apoyo...; (3) la mejora de las relaciones entre los estudiantes y el buen funcionamiento de la clase; y (4) el fomento del aprendizaje cooperativo y colaborativo. No obstante, aunque gran parte de nuestra propuesta se base en el trabajo grupal y por parejas, también proponemos alguna actividad

de carácter individual, dado que este tipo de agrupamiento es también esencial para el día a día: el proponer constantemente actividades en equipo puede provocar que el estudiante sea incapaz de enfrentarse a ciertos problemas por sí mismo. Por tanto, se debe establecer un equilibrio entre los diferentes tipos de agrupamientos para que el alumnado sepa adaptarse a las distintas situaciones de trabajo.

6. La promoción de la interculturalidad a través del intercambio lingüístico y cultural entre los estudiantes. Por ello, promovemos que comenten entre ellos expresiones en su lengua equivalentes a las que estamos enseñando. Este intercambio será mucho más enriquecedor, por supuesto, si el aula está compuesta por aprendientes procedentes de distintos países, aunque también se puede llevar a cabo si tienen un mismo origen, pues el intercambio sería con la propia cultura española y con otras LE.
7. El aprendizaje del léxico de forma inductiva, ya que será el alumnado quien, a través de la búsqueda y observación de distintos ejemplos contextualizados de las LV que se estudian, descubra (o intente averiguar) el significado de dichas unidades. Posteriormente, podrán consultar con el resto de compañeros y con el propio docente su significado y profundizarán en las características lingüísticas, extrayendo sus conclusiones a partir de los textos del corpus.
8. La inclusión de la gamificación para aumentar la motivación de los aprendientes, quienes verán el aprendizaje como una competición. De esta manera, quien más significados acierte y más características lingüísticas extraiga de su consulta de textos, obtendrá una mayor puntuación. De esta manera, fomentamos, además, una observación más atenta a los ejemplos que se consulten y una búsqueda más profunda.
9. El uso de las TIC en las aulas de LE son, pues, una herramienta que se encuentra en el día a día de nuestra sociedad y, por tanto, de nuestros estudiantes, por lo que deben saber emplearlas de forma adecuada. Al mostrarles cómo utilizar un corpus lingüístico, el estudiante aprenderá a solucionar sus dudas lingüísticas sin depender del docente y, por tanto, podrá consultar cualquier cuestión en el momento y el lugar que desee.

6.3. APLICACIÓN Y RESULTADOS

A partir de las pautas generales que hemos presentado en el anterior apartado, hemos elaborado una propuesta didáctica con la que aplicar todos nuestros supuestos y comprobar si realmente los corpus lingüísticos son adecuados para la adquisición de las LV. Para ello, hemos contado con la ayuda de la Asociación

Pro-Inmigrantes de Córdoba (APIC), una organización formada por voluntarios que busca la integración intercultural de las personas inmigrantes. Entre las numerosas actividades que realizan, destacamos sus clases de ELE, que van desde los niveles iniciales (incluyendo un nivel de alfabetización en el que enseñan a los estudiantes a leer y escribir en español) hasta un B1-B2 según el *MCERL*.

Para la aplicación de nuestra propuesta didáctica, hemos escogido trabajar con el nivel B1-B2. Debemos tener en cuenta la heterogeneidad de los grupos de esta asociación, por lo que el nivel que aquí indicamos es orientativo, debido a que determinados estudiantes se encontrarán en un nivel B2 más avanzado, mientras que otros se aproximarán más a un B1. El grupo con el que hemos trabajado está conformado por cuatro personas, ya que, a causa de la situación provocada por la Covid-19, el número de estudiantes por aula fue reducido a un máximo de cinco. Por tanto, somos conscientes de que los resultados que vamos a presentar a continuación no son concluyentes, sino orientativos, dado que ese número de alumnos no es representativo.

Por otro lado, como hemos indicado anteriormente, para el desarrollo de esta propuesta, hemos decidido emplear el *CdE* creado por Mark Davies, concretamente la sección «NOW» (*News on the Web*), ya que la búsqueda básica y la consulta de textos es muy intuitiva. Además, el uso lingüístico que se refleja en este corpus es muy actual.

Mediante la siguiente tabla sintetizamos la planificación de las tres sesiones (de una hora cada una) que ocuparemos para la aplicación de nuestra propuesta:

TABLA 1. Planificación de actividades a lo largo de las tres sesiones

	Actividades	Duración
Sesión 1	Prueba inicial: caracterización del alumnado, necesidades y conocimientos de LV	20 min.
	Presentación de los corpus lingüísticos y explicación sobre cómo usar NOW	20 min.
	Actividad práctica para comprobar si han comprendido su funcionamiento	20 min.
Sesión 2	Creación de grupos de trabajo y asignación de locuciones verbales	5 min.
	Búsqueda y deducción del significado	30 min.
	Puesta en común: explicación de las locuciones a los compañeros y ejemplificación en contexto	15 min.
	Establecimiento de rasgos gramaticales y de relaciones semánticas y búsqueda de equivalencias con su L1	10 min.
Sesión 3	Actividad de interacción oral en parejas	30 min.
	Actividad de expresión escrita individual	30 min.

Como se puede observar, comenzamos haciendo una prueba para (a) conocer a nuestro alumnado (país de origen, lengua materna, edad...); (b) verificar sus ne-

cesidades lingüísticas (dónde usan el español, si lo emplean con frecuencia, en qué contextos, cuánto tiempo llevan en el país...); y (c) comprobar sus conocimientos sobre LV y, así, asegurarnos de que no están familiarizados con las que trabajaremos en las próximas sesiones.

Consecuentemente, la prueba inicial se divide en dos partes: datos personales y prueba lingüística. Esta última incluye tres actividades sobre LV (para relacionar la locución con su significado, para completar oraciones con una serie de locuciones dadas y un ejercicio de interacción donde se emplean LV). Gracias a esto, hemos comprobado que la media de edad de los aprendientes es de 27 años, mientras que la media del tiempo que han pasado en España es de tres años. Su procedencia es diversa (Marruecos, Moldavia y Argelia), por lo que la lengua materna del alumnado será el rumano (para quien procede de Moldavia) y el árabe (para quienes proceden de Marruecos o Argelia). Su nivel de estudios es también muy distinto, lo que se refleja en las razones por las que aprenden la lengua: mientras tres de los estudiantes buscan poder comunicarse en el país para poder encontrar trabajo, la cuarta persona tiene un fin más académico (mejorar su expresión oral y escrita, buscando la corrección lingüística). Asimismo, esta última persona declara hacer un uso mucho más frecuente del español que el resto, que no lo emplean en su círculo cercano. Respecto al conocimiento de las LV, cada estudiante se enfrentó a un total de quince (seis en la primera actividad, seis en la segunda y tres en la última). El porcentaje de aciertos en cada ejercicio fue del 45,8%, 16,6% y 16,6% respectivamente, es decir, un 26,3% en total. No obstante, sabemos, a partir de la observación durante la realización de esta prueba, que parte de los aciertos de la primera actividad fueron fortuitos. Por tanto, podemos asegurar que nuestro alumnado no está familiarizado con las LV que trabajaremos en la siguiente sesión.

A continuación, como somos conscientes de que nuestro alumnado nunca ha empleado un corpus lingüístico, dedicamos 20 minutos de la sesión a explicar en qué consiste un corpus, qué información pueden encontrar en él y cómo pueden acceder a ella. Posteriormente, para cerciorarnos de que han comprendido el funcionamiento del corpus, les presentamos una breve actividad práctica de búsqueda de unidades léxicas en el corpus con el que estamos trabajando. Tras la corrección de las actividades, comprobamos que han entendido cómo realizar búsquedas de LV correctamente, consultar la frecuencia de uso de cada unidad, distinguir los usos literales de los idiomáticos y deducir dicho significado a partir de la observación de ejemplos de uso contextualizados.

Tras esta sesión introductoria, en el segundo día de clase, nos centramos ya en aplicar la propuesta didáctica cuyas bases hemos presentado anteriormente. De esta manera, comenzamos creando dos grupos de trabajo y asignándoles a cada uno cuatro locuciones verbales pertenecientes a un mismo campo semántico. El número de grupos de trabajo puede variar en función del número de estudiantes

que compongan la clase. Sin embargo, creemos excesivo trabajar con más de ocho LV, ya que tendrían que asimilar demasiada información. En nuestro caso, las escogidas se recogen en la siguiente tabla:

TABLA 2. Selección de LV para cada grupo según dos campos semánticos relacionados con los sentimientos: el enfado y el amor

Enfado (Grupo 1)	Amor (Grupo 2)
Sacar de * casillas	Tirar la caña
Perder los papeles	Hacer buenas migas
Cantar las cuarenta	Ser uña y carne
Dar la tabarra	Estar colado

Ambos grupos disponen de 20 minutos para buscar en el *CdE* esas cuatro unidades. A partir del contexto, intentarán averiguar su significado, así como extraer uno de los contextos que ejemplifique el uso de la locución. A continuación, cada grupo deberá exponer ante sus compañeros sus conclusiones y les mostrará el ejemplo escogido para ilustrar su uso. El otro grupo determinará si está o no de acuerdo con las suposiciones expuestas. Si fuera necesario, el docente intervendrá para corregir o precisar el significado que han atribuido a cada LV, intentando, previamente, que el resto de compañeros aporte sus ideas. Una vez queden claros todos los conceptos, podemos pasar a establecer sus rasgos gramaticales y relacionarlos con otras unidades léxicas: establecemos campos semánticos, buscamos sinónimos y antónimos, compartimos si conocemos otras expresiones similares, etc. Además, se pide a los aprendientes que comenten si han encontrado alguna expresión equivalente en su LM.

Como resultado de esta actividad, comprobamos que, tras la realización de la anterior prueba de uso del corpus, los estudiantes han adquirido mucha facilidad para realizar búsquedas en el corpus y acceder a sus textos⁴. Además, han deducido con cierta facilidad y rapidez los significados de sus cuatro LV, aunque fue más complejo que detectaran otras cuestiones gramaticales y semánticas. Sí fueron capaces de determinar en qué registros se emplea más la expresión, así como la modalidad lingüística más frecuente (oral o escrita). Además, para añadir más ejemplos, ellos mismos tomaron la iniciativa de elaborar sus propias oraciones y de buscar, a través de internet, construcciones sinónimas. Fue el docente quien tuvo

⁴ Debemos puntualizar que gran parte de nuestro alumnado, al carecer de recursos económicos, no dispone de un ordenador o una tableta que utilice con frecuencia. Estos son los dos dispositivos que emplearon en el aula para acceder al corpus. A pesar de ello, el aprendizaje del empleo del corpus fue muy rápido.

que señalar algunos rasgos gramaticales (como la preposición requerida tras la LV, el tipo de complemento con el que se puede combinar, etc.). También fue costoso que presentaran expresiones con un significado similar en su lengua. A pesar de todo, el trabajo conjunto y la ayuda del docente facilitaron mucho la comprensión de todos los aspectos que rodean a cada LV con la que trabajaron.

Gracias al planteamiento de dicha actividad como una competición en la que se conseguían puntos cada vez que se acertaba algún aspecto relativo a las LV, la motivación de los aprendientes aumentó enormemente. Todos participaron constantemente y deseaban extraer la máxima cantidad de información posible de sus búsquedas.

Para concluir, en la última clase se busca comprobar que los estudiantes han comprendido todo lo explicado en la sesión anterior y que lo han retenido en su memoria. De esta manera, realizamos dos actividades: una de interacción oral y otra de expresión escrita, con el fin de observar que emplean las locuciones de forma adecuada en las dos modalidades lingüísticas. Asimismo, para verificar que no solo han comprendido las locuciones con las que trabajaron en la sesión anterior, deberán poner en práctica las unidades que explicaron sus compañeros, de modo que las locuciones con las que trabaja cada grupo en cada actividad se distribuyen de la siguiente manera:

TABLA 3. LV con las que trabaja cada grupo en función de la actividad que se desarrolle

	Búsqueda en corpus	Interacción oral	Expresión escrita
Grupo 1	Enfado	Amor	Enfado
Grupo 2	Amor	Enfado	Amor

En primer lugar, se realiza la interacción oral en la que se mantienen los mismos grupos de la clase anterior. En este momento, a cada grupo se le otorgará un contexto concreto a partir del cual deben crear un diálogo empleando al menos dos LV. Los contextos presentados son los siguientes:

- Grupo 1 (amor): un amigo le cuenta a otro cómo conoció a su pareja.
- Grupo 2 (enfado): sois compañeros de piso y uno de vosotros siempre ve la televisión a las dos de la mañana con el volumen muy alto. El otro compañero se levanta temprano todos los días para ir a trabajar y le ha repetido varias veces que le molesta el volumen de la televisión. Está muy enfadado porque su compañero no cambia su comportamiento.

Tras unos diez minutos de preparación del diálogo, cada grupo lo teatralizará ante el resto de los estudiantes. A continuación, se realizará la actividad de expresión escrita de forma individual. Los discentes dispondrán de 30 minutos para elaborar una redacción en la que se incluyan al menos dos locuciones del campo semántico que se le haya asignado. Los contextos presentados son los siguientes:

- Grupo 1 (enfado): escribe un correo electrónico a un amigo para quejarte del comportamiento de uno de tus compañeros de trabajo.
- Grupo 2 (amor): escribe una carta de amor a tu pareja.

Tras la corrección de ambas actividades y la observación de su realización, los estudiantes demostraron un uso adecuado de las LV de ambos campos semánticos. No obstante, su aplicación en la expresión escrita les resultó más complicada debido a que el nivel general del alumnado en esa destreza es bastante más bajo (pues no la practican tanto como la interacción oral en su vida diaria). Por otro lado, las temáticas de ambas actividades consiguieron que los aprendientes se divirtieran y, por tanto, estuvieran más motivados e implicados durante toda la sesión.

7. CONCLUSIONES

Tras determinar la complejidad que presenta la inclusión de LV dentro del proceso de enseñanza-aprendizaje de ELE, así como las ventajas que nos aporta el trabajo de la lengua a través de los corpus lingüísticos, decidimos elaborar una propuesta didáctica aunando ambos aspectos. De esta manera, defendemos la enseñanza de LV mediante el empleo de estas bases de datos que muestran un uso de la lengua real a nuestro alumnado, permitiéndoles extraer información lingüística veraz.

Así, nuestra propuesta versa sobre ocho pautas generales: (1) partimos de un enfoque comunicativo, como recomiendan el *MCERL* y el *PCIC*, para que el alumnado desarrolle adecuadamente su competencia lingüística; (2) buscamos el equilibrio entre las distintas destrezas (comprensión y producción oral y escrita e interacción); (3) atendemos a las necesidades de los estudiantes, por lo que tanto la selección de LV como la metodología de trabajo son adaptables; (4) prima el aprendizaje autónomo del alumnado a través de la búsqueda de información, por lo que el docente actúa simplemente como guía; (5) la motivación de los estudiantes (por ejemplo, a través de la gamificación) es esencial para el buen funcionamiento de las dinámicas dentro del aula; (6) proponemos la variación de las agrupaciones a lo largo de las sesiones para que los estudiantes se adapten a diferentes entornos de trabajo; (7) promovemos el intercambio cultural; y (8) suscitamos el aprendizaje del uso educativo de las TIC.

Esta propuesta didáctica fue puesta en práctica, demostrándonos su eficacia, pues los estudiantes adquirieron adecuadamente las LV y se adaptaron con rapidez al uso de los corpus lingüísticos. No obstante, esta debe ser aplicada en grupos más numerosos para comprobar realmente su efectividad, pues su funcionamiento en un grupo tan reducido de estudiantes no es totalmente fiable.

A pesar de ello, a partir de esta experiencia, hemos extraído algunos aspectos que podríamos mejorar: (a) el tiempo de reflexión lingüística debería aumentar, ya que el alumnado podría haber extraído mucha más información si hubieran tenido algunos minutos más (esto es necesario para las actividades de corte inductivo); (b) avisar a los estudiantes del tipo de metodología que se va a emplear ayudaría a la aplicación de la propuesta, dado que están acostumbrados a la exposición tradicional donde el docente es el centro del aprendizaje, por lo que les es costoso mantener su papel de agentes; (c) sería útil la realización de una encuesta final anónima donde el alumnado pudiera expresar con total sinceridad su experiencia de aprendizaje mediante la aplicación de esta propuesta.

REFERENCIAS BIBLIOGRÁFICAS

- ABAD CASTELLÓ, M. (2019). Uso de corpus lingüísticos por y para profesores de español como lengua extranjera. *RedELE: Revista electrónica de didáctica del español lengua extranjera*, 31. <https://www.educacionyfp.gob.es/dam/jcr:39778ef0-dddf-422f-b3c2-a8f-08730225b/07-uso-de-corpus-linguisticos.pdf>
- BERNARDINI, S. (2016). Discovery learning in the language-for-translation classroom: corpora as learning aids. *Cadernos de Tradução*, 36, 14-35.
- CAMPILLOS LLANOS, L. (2005-2006). *Adaptación de corpus C-ORAL-ROM a la enseñanza de español para extranjeros* [tesis doctoral, Universidad Autónoma de Madrid]. Repositorio institucional: <https://www.educacionyfp.gob.es/dam/jcr:2564a546-773d-486a-9636-903a7196fa03/2007-bv-08-05campillos-pdf.pdf>
- CASARES, J. (1950/1992). *Introducción a la lexicografía moderna*. CSIC.
- CASTILLO CARBALLO, M. A. (2002). Conocimiento cultural en la adquisición de la L2: La fraseología. En M. PÉREZ GUTIÉRREZ y J. COLOMA MAESTRE (Eds.), *El español, lengua del mestizaje y la interculturalidad. Actas del XIII Congreso Internacional de la ASELE* (pp. 209-216). Universidad de Murcia.
- CORPAS PASTOR, G. (1997). *Manual de fraseología española*. Gredos.
- CORPAS PASTOR, G. (2003). *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. Vervuert.
- DE BOBES SOLER, E. (2016). *Información gramatical asociada a las locuciones verbales del español* [tesis doctoral, Universidad de Salamanca]. Repositorio institucional: https://gredos.usal.es/bitstream/handle/10366/128557/DLE_BobesSolerE_Locucionesverbales.pdf?sequence=1&isAllowed=y

- ELEIZALDE, M., PARRA, N., PALOMINO, C., REYNA, A. y TRUJILLO, I. (2010). Aprendizaje por descubrimiento y su eficacia en la enseñanza de la Biotecnología. *Revista de Investigación*, 71, 271-290.
- FERNÁNDEZ PRIETO, M. J. (2004). La enseñanza de la fraseología. Evaluación de recursos y propuestas didácticas. En M. A. CASTILLO CARBALLO, O. CRUZ MOYA, J. M. GARCÍA PLATERO y J. P. MORA GUTIÉRREZ (Coords.), *Las gramáticas y los diccionarios del español como segunda lengua: deseo y realidad* Actas del XV Congreso Internacional de la ASELE (pp. 349-356). Universidad de Sevilla-Secretariado de Publicaciones.
- FERNÁNDEZ-ORDÓÑEZ, I. (Dir.). (2004-2009). *Corpus Oral y Sonoro del Español Rural (COSER)*. Universidad Autónoma de Madrid.
- GARCÍA-PAGE SÁNCHEZ, M. (2008). *Introducción a la fraseología española: Estudio de las locuciones*. Anthropos.
- HERRERA, F. y CONEJO, E. (2009). Tareas 2.0: La dimensión digital en el aula de español lengua extranjera. *MarcoELE*, 9. https://marcoele.com/descargas/9/herrera_conejo_tareas2.0.pdf
- HIGUERAS GARCÍA, M. (1996). Aprender y enseñar léxico. En L. MIQUEL y N. SANS (Coords.), *Didáctica del Español como Lengua Extranjera 3* (pp. 111-126). Fundación Actilibre.
- INSTITUTO CERVANTES (2007). *Plan Curricular del Instituto Cervantes. Niveles de referencia para el español*. Biblioteca Nueva.
- LEZCANO, E. (1994-1995). Una aproximación a la gramática de las valencias. *Lenguaje y textos*, 6-7, 161-174.
- LEWIS, M. (1993). *The Lexical Approach: The State of ELT and Way Forward*. Teacher Training.
- DAVIES, M. 2002-2019. *Corpus del español*. Brigham Young University. <https://www.corpusdelespanol.org/>
- MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE-INSTITUTO CERVANTES (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Anaya.
- MOGORRÓN HUERTA, P. (2002). *La expresividad en las locuciones verbales españolas y francesas*. Universidad de Alicante.
- OLIMPIO DE OLIVEIRA, M. E. (2004). *Fraseología y enseñanza de español como lengua extranjera* [Tesis doctoral, Universidad de Alcalá]. Repositorio institucional: <http://www.educacionyfp.gob.es/educacion/mc/redele/biblioteca-virtual/numerosanteriores/2006/memoriainmaster/1- semestre/olimpio-d.html>
- PARODI, G. (2008). Lingüística de corpus: una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada*, 46(1), 93-119.
- PITKOWSKI, E. F. y VÁSQUEZ GAMARRA, J. (2009). El uso de los corpus lingüísticos como herramienta pedagógica para la enseñanza y aprendizaje de ELE. *Tinkuy*, 11, 31-51.
- REAL ACADEMIA ESPAÑOLA [en línea]. *Corpus del Español del Siglo XXI (CORPES XXI)*. <https://www.rae.es/banco-de-datos/corpes-xxi>
- REAL ACADEMIA ESPAÑOLA [en línea]. *Corpus de Referencia del Español Actual (CREA)*. <https://www.rae.es/banco-de-datos/crea>
- REAL ACADEMIA ESPAÑOLA [en línea]. *Corpus Diacrónico del Español (CORDE)*. <https://www.rae.es/banco-de-datos/corde>

- SOLANO RODRÍGUEZ, M. A. (2012). Las unidades fraseológicas del francés y del español: tipología y clasificación. *Paremia*, 21, 117-128.
- SUBIRATS, C. (2001). *Introducción a la sintaxis léxica del español*. Vervuert.
- SZYNDLER, A. (2015). La fraseología en el aula de E/LE: ¿un reto difícil de alcanzar? Una aproximación a la fraseodidáctica. *Didáctica, Lengua y Literatura*, 27, 197-216.
- TORRUELLA, J. y LLISTERRI, J. (1999). Diseño de corpus textuales y orales. En J. M. BLECUA, G. CLAVERÍA, C. SÁNCHEZ y J. TORRUELLA, (Eds.), *Filología e informática. Nuevas tecnologías en los estudios filológicos* (pp. 45-77). Editorial Milenio.
- VILLAYANDRE LLAMAZARES, M. (2008). Lingüística con corpus (I). *Estudios Humanísticos Filología*, 30, 329-349.
- VILLAYANDRE LLAMAZARES, M. (2010). *Aproximación a la lingüística computacional* [Tesis doctoral, Universidad de León]. Repositorio institucional: http://www.revistacontextos.es/econtextos/tesis_milka_contextos.pdf
- VYATKINA, N. y BOULTON, A. (2017). Corpora in language teaching and learning. *Language learning and technology*, 21(3), 66-89.
- ZULUAGA, A. (1975). *La fijación fraseológica*. Instituto Caro y Cuervo.
- ZULUAGA, A. (1980). *Introducción al estudio de las expresiones fijas*. Peter Lang.