

## **Uso de corpus propios especializados para la enseñanza de colocaciones en clase de español con fines específicos**

*Use of do-it-yourself specialized corpora for teaching collocations in Spanish for specific purposes*

GUADALUPE RUIZ YEPES  
UNIVERSIDAD DE HEILBRONN  
Guadalupe.Ruiz-Yepes@hs-heilbronn.de

### **Resumen:**

El personal docente frecuentemente tiene dificultades a la hora de encontrar material didáctico adecuado para los cursos de español con fines específicos (EFE), especialmente si se trata de niveles avanzados. Este estudio defiende que los profesores compilen su propio corpus hecho a medida utilizando textos especializados en su área de enseñanza. Como ejemplo se presenta un corpus multilingüe hecho a medida para la enseñanza del uso de colocaciones en clase de EFE en el ámbito del *marketing*. Además, se dan a conocer diferentes métodos de extracción de colocaciones de corpus electrónicos y se presenta una unidad didáctica basada en ejercicios tanto deductivos como inductivos.

**Palabras clave:** lingüística de corpus, colocaciones, español con fines específicos

### **Abstract:**

Teaching staff often have difficulties finding teaching materials for Spanish for specific purposes courses at advanced levels. This study proposes that teachers compile their own tailored corpus (do-it-yourself corpus) using specialized texts in their area of teaching. To illustrate the process, I present a multilingual corpus tailored for teaching the use of collocations in SSP class in the field of marketing. In addition, different methods for extracting collocations from electronic corpora are discussed and a teaching unit based on data-based and data-driven exercises is presented.

**Keywords:** corpus linguistics, collocations, Spanish for specific purposes

## 1. Introducción

En las últimas tres décadas se ha podido observar un crecimiento exponencial de los artículos de investigación publicados en revistas científicas en el ámbito de la didáctica de lenguas extranjeras que defienden el uso del lenguaje auténtico, ya sea de manera directa o indirecta, en la enseñanza de estas materias. En general, la argumentación gira en torno a la idoneidad del uso de corpus lingüísticos digitalizados en clase o para la creación de material didáctico. Los corpus, como herramienta lingüística, han ayudado a revelar patrones de uso auténtico del lenguaje y han contribuido a descubrir discrepancias entre el lenguaje ilustrado en los libros de texto y el lenguaje utilizado en la vida real (Ruiz Yepes y Krishnamurthy, 2010).

Hoy en día es una práctica habitual de las editoriales crear material didáctico con ejemplos auténticos del uso de la lengua. La mayoría de los manuales para la didáctica de la lengua española usan corpus (CREA, CORDE, ARTHUS, CUMBRE, etc) ya sea para ilustrar la gramática con ejemplos auténticos de nuestra lengua, o para diseñar ejercicios y crear unidades didácticas.

Esta tendencia, sin embargo, se reduce al ámbito del lenguaje general. De momento no se ha trasladado a la didáctica de lenguas con fines específicos, por lo que el personal docente sigue teniendo dificultades a la hora de encontrar material didáctico adecuado para los cursos de español con fines específicos (EFE), especialmente si se trata de niveles avanzados.

Este estudio defiende que los profesores compilen su propio corpus hecho a medida utilizando textos especializados en su área de enseñanza. Tanto el personal docente como los aprendientes de EFE se benefician de este tipo de corpus debido a su especificidad y a su idoneidad como recurso de referencia léxico-gramatical. Con un corpus compilado a medida el/la docente puede dar rienda suelta a su creatividad haciendo ejercicios en clase en los que los estudiantes usen el corpus directamente para observar patrones de uso del lenguaje con ayuda de la herramienta KWIC (Keyword In Context). Pero también se puede usar el corpus de una manera indirecta diseñando material didáctico para su uso en clase.

En este artículo se presenta un corpus comparable multilingüe hecho a medida para la enseñanza del uso de colocaciones en clase de EFE en el ámbito del *marketing*. Una vez presentado el corpus propio<sup>1</sup>, se hace una pequeña descripción de las colocaciones comparándolas con otras unidades fraseológicas y diferenciándolas de las combinaciones libres de palabras. El tercer paso en este estudio consiste en explicar los diferentes métodos de extracción automática de colocaciones de corpus electrónicos. Dado que las colocaciones son unidades fraseológicas, este artículo incluye

---

<sup>1</sup> El corpus que presentamos en este artículo se llama CorMaC: *Corpus del lenguaje del Marketing Científico*. De ahora en adelante nos referimos a este corpus propio hecho a medida con este nombre.

un apartado dedicado a la didáctica de la fraseología, o fraseodidáctica, como parte constituyente de la didáctica de lenguas extranjeras. Por último, se presenta una unidad didáctica basada en ejercicios tanto deductivos como inductivos para aprender el uso de las colocaciones que aparecen en CorMaC.

## 2. Corpus propio: CorMaC

Afortunadamente hoy en día existe una gran variedad de corpus electrónicos disponibles gratuitamente. Por este motivo, no siempre es necesario compilar un corpus propio para dar respuesta a una pregunta de investigación concreta. Sin embargo, es cierto que a menudo los corpus lingüísticos que encontramos online están compuestos principalmente de textos periodísticos. Por lo tanto, para satisfacer las necesidades específicas de un proyecto de investigación, a menudo tiene sentido crear uno o varios corpus con las características pertinentes para el objetivo de estudio. Lo mismo ocurre con los corpus que se usan en la didáctica de lenguas extranjeras. Es difícil encontrar un corpus que sea tan versátil como para cubrir las necesidades didácticas de cada uno de los niveles del *Marco Común Europeo de Referencia para las Lenguas (Council of Europe, 2020)*. Más aún si se trata de clase de lengua para fines específicos.

Según Hunston: “It is a truism that there is no such thing as a *good* or a *bad* corpus, because how a corpus is designed depends on what kind of corpus it is and how it is going to be used” (Hunston, 2009: 155). Esta afirmación es una de las más citadas en lingüística de corpus porque deja claro que la compilación de un corpus depende del uso que se le vaya a dar y, por tanto, otorga a investigadores y docentes absoluta libertad a la hora de diseñarlo y compilarlo.

Este estudio trata la enseñanza de colocaciones en clase de español con fines específicos. Las colocaciones se encuentran a caballo entre las locuciones y las combinaciones libres de palabras (Reder, 2006; Rothkegel, 1994) y suponen un gran desafío para los aprendientes de una lengua extranjera. Numerosas publicaciones reconocen que el uso correcto de las colocaciones es lo que diferencia a los hablantes nativos de una lengua de los que no lo son (Durrant, 2009). Dado que es difícil encontrar material didáctico que trate las colocaciones de los lenguajes con fines específicos, la creación de un corpus propio por el docente es la mejor herramienta para tratar este tipo de unidades fraseológicas en clase.

En este trabajo se utiliza un corpus comparable compuesto por tres corpus en tres idiomas (alemán, español e inglés). Son comparables porque están compuestos por textos del mismo género textual (artículos publicados en revistas de investigación en el campo del *marketing*), publicados en el mismo periodo de tiempo (entre los años 2000 y 2014) y con un tamaño parecido (entre las 900.000 y 1.300.000 palabras).

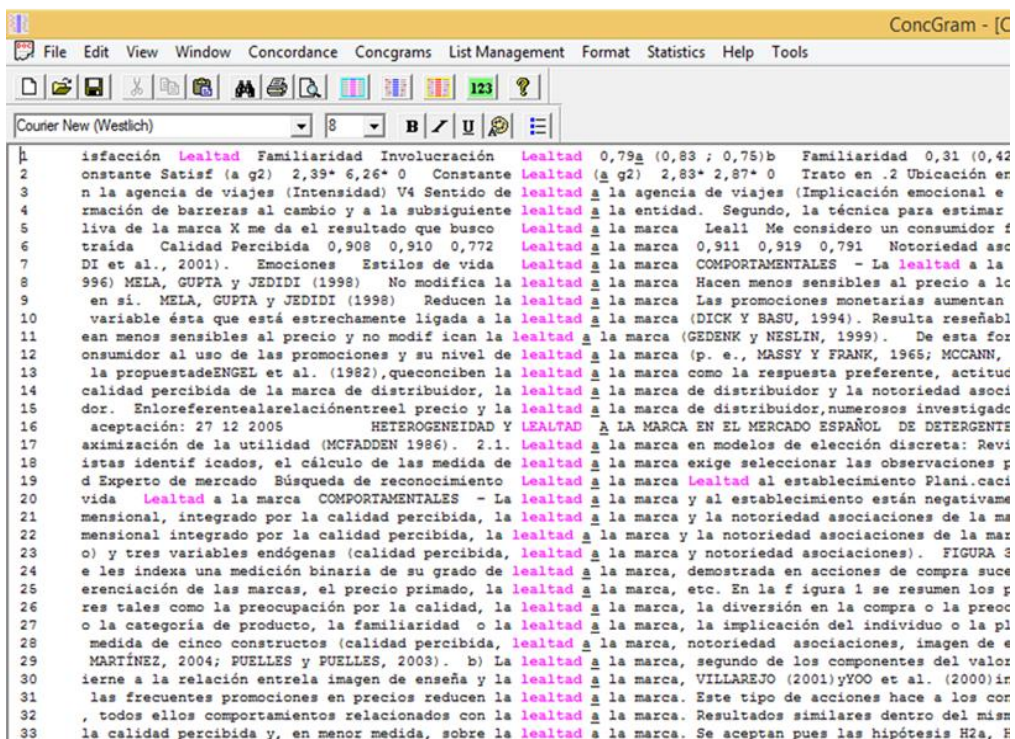
En el caso del corpus inglés se trata de todos los artículos publicados en la revista *The Journal of Marketing Management* entre 2000

y 2013 y consta de 1.131.744 palabras. El corpus alemán contiene todos los artículos aparecidos en la revista especializada *Der Markt – International Journal of Marketing* entre 2000 y 2012 y consta de 903.430 palabras. En el caso del corpus español, se compilaron todos los artículos publicados en la revista especializada *Revista española de Investigación de Marketing* entre 2004 y 2014, y contiene 1.279.954 palabras.

Una práctica muy común es enriquecer los corpus con información lingüística, lo que se denomina *anotación*. Las formas de anotación más conocidas y utilizadas son *Part-of-Speech-Tagging* (etiquetado de las partes de la oración), en el que se determina si las palabras son sustantivos, verbos, etc., la *lematización*, en la que se determinan las formas básicas o *lemas* de las palabras, y el *parsing* o análisis sintáctico automático. Dado que en este trabajo se emplea un software para la extracción de colocaciones (FipsCo) que tiene el *parser* incorporado, nuestro corpus multilingüe comparable está sin anotar. Para poder analizar las colocaciones en su contexto se usa la función *keyword in context* (KWIC) del software Concgram 1.0 (un KWIC muestra todas las apariciones de una palabra o conjunto de palabras de un corpus junto con un número determinado de palabras anteriores y posteriores de su contexto).

En la imagen 1 se muestra un ejemplo de KWIC realizado con el software Concgram 1.0 del corpus español de CorMaC, concretamente de la palabra *lealtad*, que da lugar a la colocación *lealtad a la marca*.

Imagen 1. KWIC de la palabra “lealtad” en el corpus español de CorMaC.



### 3. ¿Qué son colocaciones? Características y clasificaciones

Existen numerosas definiciones para el fenómeno “colocación”. Desde que Firth introdujo este término en 1957, investigadores de una amplia variedad de áreas se han ocupado de definir y clasificar este tipo de unidad fraseológica. Como consecuencia, las definiciones de “colocación” se pueden encontrar en la fraseología, en la lingüística computacional, en la lingüística de corpus, etc. De ello se deduce que el término es muy amplio y no tiene límites precisos. A pesar de esta ambigüedad, hay dos tradiciones a la hora de definir y clasificar colocaciones; una sigue los postulados empíricos de Firth dentro del marco del contextualismo británico, la otra tiene sus orígenes en la lexicografía franco-alemana (Hausmann, 1984). A diferencia del contextualismo británico, que se centra en la frecuencia de coexistencia de dos palabras, Hausmann pone el foco en la relación semántica y sintáctica entre estas palabras.

La tipología de colocación de Hausmann (cf. 1989: 1010) distingue seis patrones sintácticos diferentes de colocaciones:

- Verbo + sustantivo (objeto directo).  
Ejemplos<sup>2</sup>: *dar lugar, lanzar producto, prestar servicio, dar importancia, formar parte, ejercer influencia, guardar relación, etc.*
- Adjetivo + sustantivo (o sustantivo + adjetivo).  
Ejemplos: *mayor medida, igual modo, siguiente hipótesis, nueva tecnología, campaña publicitaria, evidencia empírica, etc.*
- Sustantivo (sujeto) + verbo.  
Ejemplos: *trabajo analizar, resultado sugerir, análisis permitir, etc.*
- Sustantivo + (preposición) + sustantivo.  
Ejemplos: *agencia de viaje, cuota de mercado, línea de investigación, punto de vista, cambio de actitud, etc.*
- Adverbio + adjetivo  
Ejemplos: *altamente novedoso, especialmente interesante, completamente independiente, particularmente relevante, etc.*
- Verbo + adverbio  
Ejemplos: *analizar detalladamente, demostrar empíricamente, influir positivamente, expresar libremente, suponer necesariamente, etc.*

A los efectos de este trabajo se tiene en cuenta en mayor medida el enfoque fraseológico de Hausmann.

Como ya se ha mencionado, las colocaciones se encuentran a caballo entre las combinaciones libres de palabras y las locuciones (Reder, 2006). Hay autores que no las consideran unidades fraseológicas por no ser lo suficientemente idiomáticas, por ejemplo, Burger (2003). Sin embargo, en los modelos con una concepción más amplia de la fraseología, solo la fijación se considera un rasgo constitutivo. En consecuencia, “las

<sup>2</sup> Todos los ejemplos han sido extraídos de CorMaC.

colocaciones como combinaciones de palabras no idiomáticas, pero fijas, forman una subcategoría de fraseologismos en estos modelos” (Reder, 2006: 44). En este contexto también se les llama *fraseologismos débiles* o *no idiomáticos*.

Para diferenciar a las colocaciones de las combinaciones libres de palabras se puede recurrir a la traducción a diferentes lenguas. Es decir, con mucha probabilidad, si la traducción literal de una combinación de palabras es correcta, se trata de una combinación libre de palabras. Un ejemplo de combinación libre de palabras en español con sus traducciones al inglés y al alemán sería: *ver la mesa, to see the table, den Tisch sehen*. Sin embargo, cuando se trata de una colocación, la traducción literal suele ser incorrecta. Por ejemplo, la traducción literal al inglés y al alemán de la colocación española *poner la mesa, \*to put the table, \*den Tisch stellen*, son ambas incorrectas. Las traducciones correctas serían: *to lay/to set the table* (inglés), *den Tisch decken* (alemán). Lo mismo ocurre si traducimos literalmente una colocación alemana al español y al inglés. Por ejemplo, la traducción literal al español y al inglés de la colocación alemana *Geld abheben, \*levantar dinero, \*to lift up money*, también son ambas incorrectas. Las traducciones correctas serían: *sacar dinero*<sup>3</sup> (español), *to withdraw money* (inglés).

Para diferenciar a las colocaciones de las locuciones se puede recurrir al concepto de la idiomaticidad: “aquella propiedad semántica que presentan ciertas unidades fraseológicas, por la cual el significado global de dicha unidad no es deducible del significado aislado de cada uno de sus elementos constitutivos” (Corpas 1996: 26). Por ejemplo, para el verbo *morir* existen las siguientes locuciones: *den Löffel abgeben* (alemán: *\*entregar la cuchara*), *to kick the bucket* (inglés: *\*patear el cubo*), *estirar la pata* (español).

Mientras que las locuciones presentan un grado alto de idiomaticidad, las colocaciones presentan un grado muy bajo. Incluso podría decirse que, en el caso de las colocaciones, el significado global de dicha unidad **sí** es deducible del significado aislado de cada uno de sus elementos constitutivos como, por ejemplo, en el caso de *lluvia torrencial* (*lluvia paracedia a un torrente*), *fumador empedernido* (*fumador con el hábito de fumar muy arraigado*), etc.

#### 4. Extracción de las colocaciones de CorMaC

Así como existen diferentes definiciones de colocación, también existen diferentes métodos para extraer colocaciones de corpus. Hasta ahora se han utilizado básicamente tres métodos:

- Método basado en un cálculo de coocurrencia.
- Método basado en los patrones sintácticos de las colocaciones.
- Métodos híbridos.

---

<sup>3</sup> *Sacar dinero del cajero automático.*

El primer método toma el concepto de colocación de Firth como punto de partida y se centra en la frecuencia de coexistencia de la base y el colocador. El segundo método toma la visión binaria de colocación de Hausmann como punto de partida y se centra en las partes del discurso de base y colocador. Por último, los métodos híbridos son una combinación de los dos anteriores (Ruiz Yepes, 2017).

Primer método: Para la extracción de colocaciones de corpus tomando como base la frecuencia de coocurrencia se han desarrollado diferentes medidas de asociación<sup>4</sup>. Para el lingüista de corpus o el investigador esto significa que debe decidir qué medidas de asociación aplicar. Según Evert (2009), no existe una medida “perfecta” de asociación. Por este motivo, conviene utilizar varias para poder conseguir resultados diferentes y poder compararlos entre sí. Lo que está claro es que el investigador debe preguntarse: “¿En qué medida la frecuencia de coocurrencia observada excede la frecuencia esperada?” (Evert, 2009: 1228). Las siguientes medidas han demostrado ser útiles: *Mutual Information (MI)*, *z-score*, *t-score*, *simple II*, *chi-squared*, *log-likelihood*. Kilgariff y Rundell (cf. 2002: 810) recomiendan filtrar las palabras funcionales de un corpus y lematizar el corpus antes de utilizar una medida de asociación, porque de lo contrario el número de colocaciones “incorrectas” (ruido) en el resultado final sería demasiado alto.

Segundo método: Como punto de partida se toma la definición de colocación de Bartsch: “coexistencias recurrentes restringidas léxicamente y/o pragmáticamente de al menos dos elementos léxicos que tienen una relación sintáctica directa entre sí” (2004: 76). Los patrones sintácticos que se usan en este método para la extracción de colocaciones suelen ser los propuestos por Hausmann (1989) y descritos en el apartado anterior, o versiones parecidas de otros autores. Para extraer colocaciones de los corpus grandes usando este método, se deben anotar al menos con etiquetas *PoS* (Part of Speech). Sin embargo, se puede lograr un mejor resultado si los corpus también se analizan sintácticamente (parsing).

Tercer método: Pamies y Pazo (2005) comparan las distintas medidas de asociación entre sí y llegan a la conclusión de que los métodos matemáticos por sí solos no son suficientes y que es preferible el uso de métodos “híbridos” (2005: 327). Estos autores defienden los métodos que utilizan una combinación de cálculo de coocurrencia y criterios lingüísticos en forma de patrones de colocación.

En este trabajo usamos el segundo método, concretamente el software *FipsCo*, un sistema desarrollado para extraer colocaciones de grandes corpus de texto, utilizando un potente analizador sintáctico (parser). *FipsCo* es capaz, incluso, de reconocer la presencia de colocaciones en oraciones subordinadas. Por ejemplo, en el caso de las oraciones de relativo, *FipsCo* es capaz de identificar el sustantivo al que se refiere el pronombre relativo, y establecer un vínculo entre este sustantivo y el verbo del que depende sintácticamente el pronombre relativo

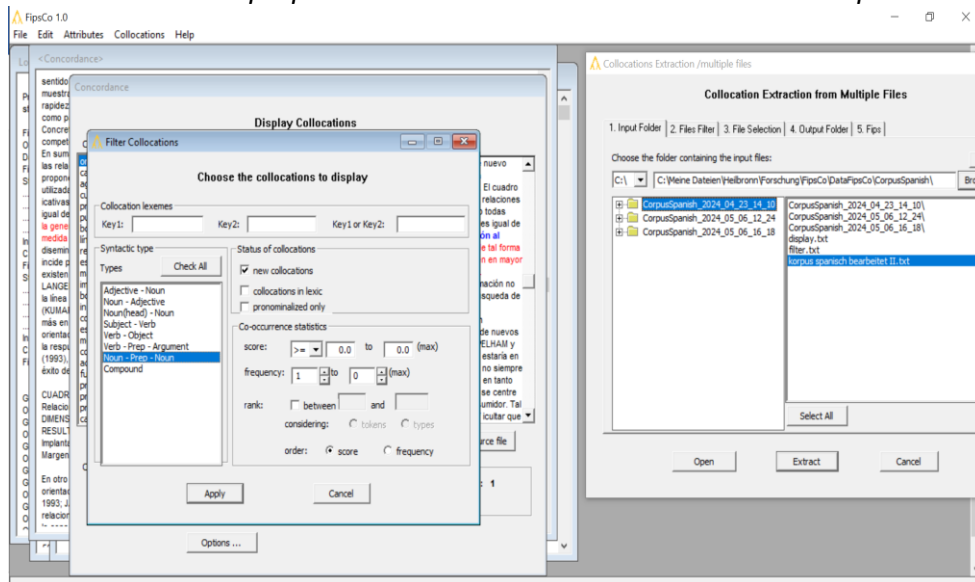
---

<sup>4</sup> Lo que equivale a desarrollar diferentes modelos estadísticos.

(Goldman y Wherli, 2001). De esta manera, no importa cuánta distancia haya entre la base y el colocador de la colocación; lo que importa es su relación sintáctica.

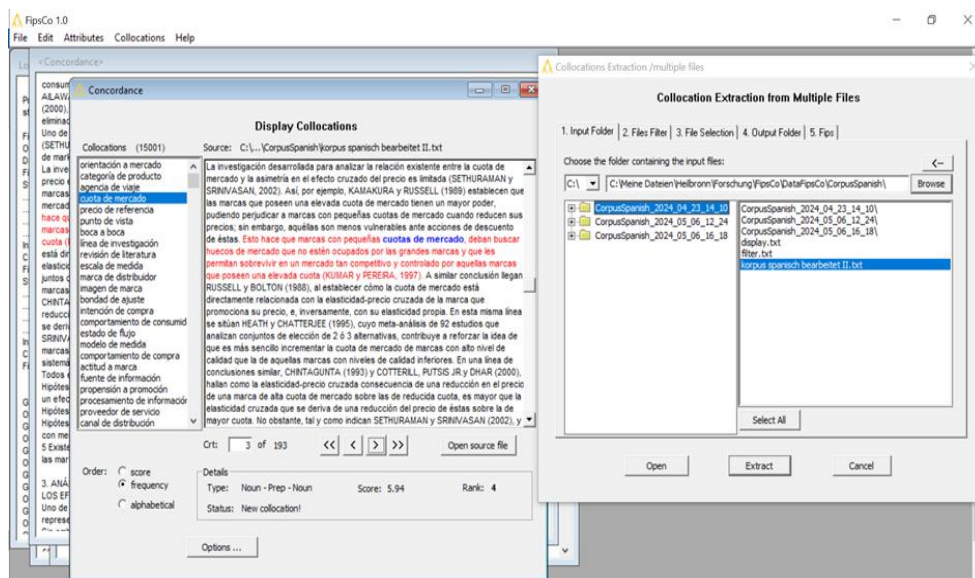
En la imagen 2 se muestra el procedimiento para extraer colocaciones del tipo *sustantivo + preposición + sustantivo* de CorMaC usando FipsCo.

Imagen 2: Extracción de colocaciones del tipo “sustantivo + preposición + sustantivo” de CorMaC usando FipsCo.



En la ventana *Filter collocations/Choose the collocations to display* bajo el rótulo *Syntactic type* se pueden apreciar la mayoría de los tipos de colocación de la tipología de Hausmann (1984). Si elegimos el tipo *Noun – Prep – Noun*, el resultado es el que muestra la imagen 3.

Imagen 3: Resultado de la extracción de colocaciones del tipo “sustantivo + preposición + sustantivo”.





Como puede observarse en la imagen 3, las colaciones del tipo *sustantivo + preposición + sustantivo* están ordenadas de mayor a menor frecuencia. En el cuarto lugar se encuentra la colocación *cuota de mercado* que aparece en el corpus 193 veces. Con la flechas hacia delante y hacia atrás se pueden ver cada uno de los contextos en los que encontramos esta colocación. En esta imagen mostramos la tercera vez de las 193 veces que aparece en total en el corpus. En rojo se muestra la oración en la que aparece. La colocación aparece en color azul y el resto del texto en color negro.

En la ventana con el rótulo *Collocation Extraction from Multiple Files* se puede ver el corpus del que se han extraído las colocaciones, que en este caso es el corpus español de CorMaC.

## 5. Historia de la fraseodidáctica

Cada vez son más los docentes de lenguas extranjeras que reconocen la necesidad de incluir las unidades fraseológicas en el material didáctico de los cursos de lenguas extranjeras, sobre todo para los niveles avanzados. Pero no se trata de un fenómeno reciente. Según González Rey (2012) el primer intento de introducir la fraseología en la didáctica de las lenguas se sitúa en 1900, pero no es hasta la segunda mitad del siglo XX cuando este impulso cobra fuerza. En las primeras fases, solamente se hablaba de la *didáctica de la fraseología*, pero a finales del siglo XX aparece por primera vez el término de *fraseodidáctica*. Concretamente en sus trabajos de los años 1985 y 1987, el lingüista alemán Peter Kühn acuña el término *Phraseodidaktik* apelando a impulsar una concepción nueva en la creación de material didáctico en este ámbito. De hecho, en su trabajo “Deutsch als Fremdsprache im phraseodidaktischen Dornröschenschlaf. Vorschläge für eine Neukonzeption phraseodidaktischer Hilfsmittel” (1987) establece una comparación entre esta disciplina y la bella durmiente, y hace un llamamiento a despertarla. Según González Rey, la diferencia entre estas dos disciplinas es que “la Fraseodidáctica se define no sólo (sic.) como la didáctica de la fraseología de una lengua sino también como la didáctica de toda una lengua a través de su fraseología” (2012: 67). Vemos pues, que hay un cambio de paradigma para darle más importancia, si cabe, a la fraseodidáctica, incluso por encima de la glotodidáctica (didáctica de lenguas), defendiendo “las posibilidades de empleo de la fraseología como una útil herramienta de enseñanza de nuestra lengua a aprendices extranjeros” (Aguilar Ruiz, 2013: 2).

Es también Kühn (1992 y 1994) quien establece el modelo fraseodidáctico de los tres pasos a la hora de aprender unidades fraseológicas: la descodificación, la comprensión y la producción. Al modelo de los tres pasos de Kühn le sigue el modelo de los cuatro pasos de Lüger (1997). Este autor defiende que entre los pasos 2 y 3, es decir, entre el paso de la comprensión y el paso de la producción, debe hallarse el paso de la consolidación. A raíz de la publicación de estos dos modelos, se

puede observar una proliferación en la producción de libros de texto dedicados a la fraseodidáctica para *DaF* (*Deutsch als Fremdsprache*), es decir, alemán como lengua extranjera, que impulsaron la investigación en este ámbito. En este contexto Ettinger (2007) publica sus recomendaciones para enseñar y aprender unidades fraseológicas. Estas son las tres más importantes para este estudio:

- Para la didáctica de las unidades fraseológicas se deben emplear ejemplos extraídos de textos auténticos.
- Las unidades fraseológicas deben tratarse en niveles avanzados.
- Las unidades fraseológicas se deben dominar primeramente de forma pasiva y en una segunda fase de forma activa.

El primer autor en usar el término específico de la fraseodidáctica en España es Juan Pablo Larreta Zulategui (2001). Este autor realiza un análisis contrastivo entre el alemán y el español para comparar unidades fraseológicas usando un corpus bilingüe. Desde entonces han sido muchos los autores que le han dedicado artículos de investigación y/o libros de texto a la fraseodidáctica (Ruiz Gurillo, 2000; Olímpio, Penadés y Ruiz, 2006; Aguilar Ruiz, 2013; Detry, 2015; Mejía, 2021; etc), pero casi todos ellos se centran en las locuciones. Las unidades fraseológicas que no pertenecen a este grupo han quedado relegadas a un segundo plano, especialmente las colocaciones.

El uso correcto de colocaciones marca en muchas ocasiones la diferencia entre los hablantes nativos de una lengua y los aprendientes de esa misma lengua de un nivel avanzado. A menudo los aprendientes de una lengua traducen literalmente las colocaciones de su lengua materna a la lengua que están aprendiendo con un resultado gramaticalmente correcto, pero estilísticamente deficiente. Aún así las colocaciones no son tratadas en los manuales didácticos con el detalle que se merecen. Por lo tanto, compilar un corpus grande propio y realizar una extracción automática de colocaciones para crear unidades didácticas y actividades que traten este grupo de unidades fraseológicas resulta una medida eficaz para compensar este déficit.

## **6. Propuesta de unidad didáctica**

La compilación de un corpus grande de textos especializados y la extracción de colocaciones, tal y como se ha descrito en este artículo, puede resultar una tarea laboriosa, pero merece la pena teniendo en cuenta la calidad y cantidad de colocaciones que se pueden extraer y la variedad de actividades que se pueden diseñar con ellas.

En este trabajo presentamos una unidad didáctica que consta de una introducción general al tema de las unidades fraseológicas y concretamente de las colocaciones, y de cinco actividades que van de menor a mayor grado de dificultad para que los estudiantes aprendan a usar correctamente las colocaciones españolas del lenguaje del *marketing*

en un contexto académico. Para ello, nos basamos en los cuatro pasos de la fraseodidáctica propuestas por Kühn (1992, 1994) y Lüger (1997) que se mencionaron en el apartado anterior: la descodificación, la comprensión del significado, la consolidación y la producción del fraseologismo, en este caso, de la colocación.

Durante la fase de la descodificación, los estudiantes deben saber identificar las colocaciones en textos auténticos. En esta primera fase se pueden realizar actividades que ayuden a los estudiantes a identificar los campos semánticos, los registros textuales y los ámbitos de especialización de las colocaciones. Así, entramos también de lleno en la fase de la comprensión, puesto que, mediante el ejercicio de la identificación de los campos semánticos y grados de especialización, nos aseguramos de que los estudiantes hayan comprendido el significado de las colocaciones. En la fase de la consolidación, es imprescindible incluir ejercicios repetitivos en los que los estudiantes trabajen con colocaciones que ya conocen de antemano para poder asimilarlas. La fase de la producción es la última y la que requiere mayor esfuerzo por parte de los estudiantes, pues se trata de que reproduzcan adecuadamente en contextos variados las colocaciones que acaban de aprender.

Como se ha mencionado, la unidad didáctica que proponemos a continuación consta de una introducción y de cinco actividades, en las que se combinan las fases propuestas por Kühn y Lüger. No hay ninguna actividad que esté dedicada exclusivamente a una de estas fases puesto que se complementan unas a otras.

### **Introducción general**

Se puede dar por hecho que en los niveles avanzados de ELE los estudiantes han escuchado hablar de las unidades fraseológicas. La mayoría de los manuales dedicados a los niveles A1, A2 y B1 incluyen ejercicios dedicados a este fenómeno lingüístico. Sin embargo, también se puede dar por hecho que la mayoría de ellos no sabe distinguir entre los diferentes tipos de unidades fraseológicas y que, en muchos casos, no las reconocen cuando se topan con una de ellas en un texto. Con una introducción amplia sobre este tema, el/la docente ofrece a los estudiantes los fundamentos primordiales para construir unas estrategias sólidas que les ayuden a la hora de descodificar, comprender, consolidar y reproducir las unidades fraseológicas. En este caso se trata de familiarizar a los estudiantes con el concepto de colocación en un sentido amplio. Es recomendable hacer llegar estos conocimientos mediante el uso de una presentación PowerPoint con elementos interactivos que puede ponerse a disposición de los aprendientes en un LMS (Learning Management System), ya sea Moodle u otro sistema parecido. De este modo los estudiantes siempre tendrán acceso a los contenidos y podrán repasarlos tantas veces como sea necesario para su comprensión.

En cuanto al contenido de esta presentación, para la introducción general, proponemos tratar los siguientes aspectos:

- ¿Qué son unidades fraseológicas? Importancia de las unidades fraseológicas para la comunicación fluida entre hablantes de una lengua.
- Diferencias entre unidades fraseológicas. Concretamente, trasladar a los estudiantes las diferencias entre las colocaciones, las locuciones y las proverbiales, haciendo especial hincapié sobre las colocaciones.
- Explicar las diferencias entre las combinaciones libres de palabras y las colocaciones.
- Ofrecer una breve explicación de las actividades de las que consta la unidad didáctica y sus objetivos.

La lingüística contrastiva, es decir, las comparaciones entre la lengua materna de los estudiantes y el castellano, o incluso, entre el inglés y el castellano, sirve, a menudo, para que puedan ver las diferencias en las unidades fraseológicas de las distintas lenguas. Una práctica interesante consiste en traducir las unidades fraseológicas del castellano a la lengua materna, para hacer conscientes a los estudiantes del efecto que van a conseguir si traducen las unidades fraseológicas de su lengua materna al castellano. En este sentido, tanto la lingüística contrastiva, como la traducción son útiles instrumentos didácticos y se pueden incluir en la presentación a modo de elementos interactivos.

Dado que la unidad didáctica que estamos presentando se ubica en las clases de niveles muy avanzados (B2-C1), resulta adecuado introducir conceptos lingüísticos y teóricos sobre qué es una unidad fraseológica. No resulta problemático adoptar un enfoque metalingüístico, ya que se trata de hacer conscientes a los estudiantes del funcionamiento de la lengua.

### **Primera actividad: Toma de contacto con CorMaC**

*Tabla 1: Ficha técnica de la primera actividad*

Título	Primera toma de contacto con CorMaC: Clasificación de colocaciones según ámbitos de especialización
Nivel	B2-C1
Duración	Entre 25 y 30 minutos
Grado de dificultad	Bajo
Fases	Descodificación, comprensión del significado y producción (en el suplemento)
Objetivo principal	- Clasificar colocaciones del corpus según su pertenencia al lenguaje general, especializado del <i>marketing</i> o al lenguaje académico.

Objetivo secundario	- Hacer conscientes a los estudiantes de las diferencias entre el lenguaje general, especializado del <i>marketing</i> y el lenguaje académico.
Material	CorMaC

La tabla 1 representa la ficha didáctica de la primera actividad. Como puede observarse, aunque el objetivo principal es el aprendizaje de las colocaciones, la actividad cuenta también con objetivos secundarios que se derivan del objetivo principal.

*Desarrollo de la actividad:*

Se les presenta a los estudiantes CorMaC haciendo hincapié en el género textual, el número de palabras, las lenguas representadas y el periodo de publicación de los artículos.

Se les entrega a los estudiantes una lista con las 100 primeras colocaciones del tipo *verbo + sustantivo (objeto directo)* de CorMaC y se les pide que las clasifiquen según pertenezcan al lenguaje general, al lenguaje específico del *marketing* o al lenguaje académico. Para ello, deben rellenar la tabla 1.1.

*Tabla 1.1: Ejercicio para la primera actividad*

Lenguaje general	Lenguaje científico	Lenguaje del <i>marketing</i>

Los estudiantes deben rellenar la tabla trabajando individualmente o en parejas. La lista de colocaciones extraídas de CorMac y proporcionada por el docente garantiza el éxito de esta actividad, puesto que los estudiantes solamente tienen que concentrarse en realizar la clasificación correctamente sin preocuparse de aspectos prácticos como aplicar el software sobre el corpus o realizar la extracción de colocaciones ellos mismos, lo cual dificultaría mucho la tarea. Los resultados a los que suelen llegar los estudiantes quedan reflejados en la tabla 1.2:

*Tabla 1.2: Resultados del ejercicio para la primera actividad*

<b>Lenguaje general</b>	<b>Lenguaje científico</b>	<b>Lenguaje del <i>marketing</i></b>
Tener efecto	Contrastar hipótesis	Realizar compra
Formar parte	Extraer varianza	Satisfacer necesidad
Tomar decisión	Ver tabla/cuadro	Promocionar producto
Establecer relación	Realizar análisis	Adquirir producto
Hacer referencia	Aceptar/rechazar hip.	Promocionar marca
Dar respuesta	Utilizar método	Lanzar producto
Hacer uso	Aportar evidencia	Retener cliente
Cometer error	Seguir procedimiento	Segmentar mercado
Desempeñar papel	Adoptar enfoque	Proyectar imagen

Se entiende que no todos los estudiantes llegan a los mismos resultados. Algunos tienen dificultades a la hora de diferenciar entre el lenguaje general y el científico, o entre el lenguaje científico y el específico del *marketing*. Por este motivo, es importante corregir los resultados en clase y explicar las diferencias entre géneros textuales y lenguajes de especialización. De esta manera garantizamos que se alcance el objetivo secundario.

En esta actividad también se les puede preguntar a los estudiantes qué tipo de colocación piensan ellos que son las más frecuentes en un corpus compuesto por artículos publicados en revistas especializadas en *marketing*: las del lenguaje general, las del lenguaje científico o las del lenguaje del *marketing*. Tras unos minutos de reflexión los estudiantes suelen responder que las colocaciones del lenguaje del *marketing* deben ser las más frecuentes en CorMaC. Sin embargo, después de un tiempo trabajando con CorMaC se percatan de que son las colocaciones del lenguaje general las más frecuentes. Llegados a este punto, es el momento ideal para explicar a los estudiantes que los lenguajes especializados todos toman como base al lenguaje general, no solamente en lo referente a elementos como la sintaxis, sino también en lo que atañe a la fraseología.

También es recomendable comentar la tabla mencionando, por ejemplo, que hay colocaciones que son aptas para aparecer en más de una categoría. Por ejemplo, FipsCo nos propone “realizar compra” como una de las colocaciones más frecuentes. Sin embargo, es una colocación que puede considerarse del lenguaje general al igual que del lenguaje del *marketing*.

En la columna con las colocaciones del lenguaje del *marketing* también se ve el grado de especialización de la colocación. En este sentido, se puede observar que la colocación *segmentar mercado* pertenece a un grado de especialización más alto que la colocación *realizar compra*.

**Suplemento:**

Como actividad complementaria a esta, se puede pedir a los estudiantes que redacten un resumen con las colocaciones del lenguaje científico. Para ello, pueden usar las colocaciones de la tabla con la que acaban de trabajar, pero también todas las demás que aparecen en CorMaC y que también son típicas del lenguaje científico. El *abstract* debe tener entre 150 y 200 palabras de longitud y no tiene que pertenecer necesariamente al ámbito del *marketing*. Puede ser, por ejemplo, un resumen del trabajo fin de carrera en el caso de que los estudiantes se encuentren en cursos avanzados.

Con este suplemento, se añade a esta primera actividad la fase del aprendizaje de la producción de las colocaciones. Se trata de que usen correctamente colocaciones como *adoptar enfoque*, *aceptar* o *rechazar hipótesis*, etc.

**Segunda actividad: Registro de las colocaciones***Tabla 2: Ficha técnica de la segunda actividad*

Título	Registro o nivel de formalidad de las colocaciones
Nivel	B2-C1
Duración	Entre 30 y 35 minutos
Grado de dificultad	Bajo-medio
Fases	Comprensión del significado y producción
Objetivo principal	- Hacer conscientes a los estudiantes de las diferencias en el registro o el nivel de formalidad de las colocaciones.
Objetivo secundario	- Hacer conscientes a los estudiantes de las diferencias en el registro o nivel de formalidad de los géneros textuales.
Material	CorMaC y otros corpus de la lengua española, como CREA.

**Desarrollo de la actividad:**

La segunda actividad que proponemos es muy útil para estudiar el registro o nivel de formalidad de las colocaciones. Así como para observar que hay algunos verbos que tienen mucha capacidad para formar colocaciones, pero, a su vez, son colocaciones que tienen un homólogo en un registro más alto. Por eso, esta actividad consiste en que los estudiantes busquen con la misma base otra colocación con el mismo significado, pero de un registro más alto.

Se trata de sensibilizar a los estudiantes para que vean las diferencias en el registro. Por ejemplo, la diferencia entre *tener tiempo* o

*disponer de tiempo, tener interés o mostrar interés, tener efecto o surtir efecto.* Para ello, deben rellenar la tabla 2.1:

*Tabla 2.1: Ejercicio para la segunda actividad*

<b>Informal</b>	<b>Formal</b>	<b>Informal</b>	<b>Formal</b>
Tener efecto		Tener tiempo	
Tener experiencia		Tener ventaja	
Tener impacto		Tener peso	
Tener actitud		Tener consecuencia	
Tener relación		Tener oportunidad	
Tener interés		Tener repercusión	
Tener propensión		Tener contacto	
Tener influencia		Tener impacto	

Los estudiantes deben rellenar la tabla trabajando individualmente o en parejas. Los resultados a los que suelen llegar quedan reflejados en la tabla 2.2:

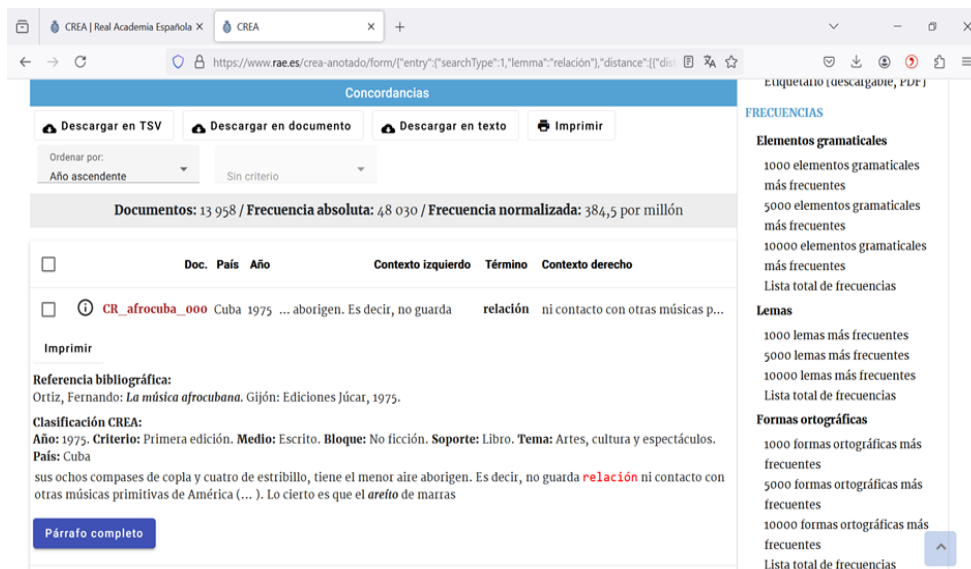
*Tabla 2.2: Resultados del ejercicio para la segunda actividad*

<b>Informal</b>	<b>Formal</b>	<b>Informal</b>	<b>Formal</b>
Tener efecto	Surtir efecto	Tener tiempo	Disponer de tiempo
Tener experiencia	Poseer experiencia	Tener ventaja	Poseer ventaja
Tener impacto	Dejar impacto	Tener peso	
Tener actitud	Mostrar actitud	Tener consecuencia	Traer consecuencia
Tener relación	Guardar relación	Tener oportunidad	
Tener interés	Mostrar interés	Tener repercusión	
Tener propensión	Mostrar propensión	Tener contacto	Mantener contacto
Tener influencia	Ejercer influencia	Tener impacto	Ejercer impacto

Las colocaciones con *tener* son todas extraídos de CorMaC mediante el uso de FipsCo. Pero para buscar las colocaciones sinónimas les proponemos a los estudiantes que usen además otros corpus, como por ejemplo CREA (Corpus de Referencia del Español Actual) de la Real Academia Española. Con este fin, se les ha explicado anteriormente a los estudiantes qué es la RAE y cuáles son sus funciones. De igual manera, se les ha enseñado a hacer búsquedas en el CREA usando su motor de búsqueda. El ejercicio consiste en que el estudiante busque la base de la colocación en CREA, por ejemplo *relación*, y compruebe mediante la función KWIC cuáles son los verbos que pueden aparacer con esta base. En la imagen 4 se observa uno de los resultados de la búsqueda en CREA de la palabra *relación*, en este caso con el verbo *guardar*.



*Imagen 4: Resultado de la búsqueda en CREA de la colocación “guardar relación”.*



Se puede observar que en la tabla elaborada por los estudiantes han quedado algunas celdas libres. Esto se debe a que no se han encontrado equivalencias formales para esas colocaciones con *tener*.

### Tercera actividad: Las colocaciones como unidades fijas

*Tabla 3: Ficha técnica de la tercera actividad*

Título	Las colocaciones como unidades fijas
Nivel	B2-C1
Duración	Entre 20 y 30 minutos
Grado de dificultad	Medio
Fase	Comprensión, consolidación y producción
Objetivo principal	- Hacer conscientes a los estudiantes del aspecto de la fijación en las colocaciones.
Objetivo secundario	- Repaso de la conjugación de los verbos.
Material	CorMaC

#### *Desarrollo de la actividad:*

La tercera actividad de dificultad media no podía faltar. Se trata de un ejercicio para rellenar huecos. El objetivo es hacer conscientes a los alumnos de la fijación de las colocaciones. Este ejercicio es fácil de ejecutar, pues no necesitamos realizar una extracción de colocaciones con FipsCo. Simplemente, con la función de búsqueda, buscamos una palabra como “relación” y sacamos una oración con esa palabra, le quitamos en

este caso el verbo que lleva y buscamos un par de sinónimos de ese verbo. Les pedimos a los estudiantes que rellenen el hueco con el verbo correcto en la tabla 3.1.

Tabla 3.1: Ejercicio para la tercera actividad

Verbos	Oraciones
<i>guardar/mantener/ contar con</i>	Esta variable _____ <i>relación</i> con la hipótesis formulada.
<i>agarrar/coger/tomar</i>	La incertidumbre del consumidor y su miedo a _____ una <i>decisión</i> equivocada son factores a tener en cuenta.
<i>realizar/hacer/cometer</i>	Algunas empresas _____ <i>graves errores</i> a la hora de fijar el precio de sus servicios.
<i>Entregar/dar/otorgar</i>	Para _____ <i>respuesta</i> a estos objetivos, se ha estructurado el trabajo de investigación de la siguiente manera.
<i>Hacer/perpetrar/efectuar</i>	El cliente es incapaz de evaluarlo incluso después de _____ <i>uso</i> de él.
<i>Partir/fraccionar/segmentar</i>	Cuando la variable utilizada para _____ el <i>mercado</i> es la edad, ...
<i>Adquirir/comprar/hacerse con</i>	Puede contribuir también a paliar la reticencia de los consumidores a _____ <i>productos</i> procedentes de países con una imagen desfavorable.

Los estudiantes deben rellenar la tabla trabajando individualmente. Los resultados a los que suelen llegar quedan reflejados en la tabla 3.2:

Tabla 3.2: Resultados del ejercicio para la tercera actividad

Verbos	Oraciones
<i>guardar/mantener/ contar con</i>	Esta variable <i>guarda relación</i> con la hipótesis formulada.
<i>agarrar/coger/tomar</i>	La incertidumbre del consumidor y su miedo a <i>tomar una decisión</i> equivocada son factores a tener en cuenta.
<i>realizar/hacer/cometer</i>	Algunas empresas <i>cometen graves errores</i> a la hora de fijar el precio de sus servicios.
<i>Entregar/dar/otorgar</i>	Para <i>dar respuesta</i> a estos objetivos, se ha estructurado el trabajo de investigación de la siguiente manera.
<i>Hacer/perpetrar/efectuar</i>	El cliente es incapaz de evaluarlo incluso después de <i>hacer uso</i> de él.
<i>Partir/fraccionar/segmentar</i>	Cuando la variable utilizada para <i>segmentar el mercado</i> es la edad, ...
<i>Adquirir/comprar/hacerse con</i>	Puede contribuir también a paliar la reticencia de los consumidores a <i>adquirir/comprar/hacerse con productos</i> procedentes de países con una imagen desfavorable.

Con este ejercicio se subrayada el aspecto de la fijación de las colocaciones, puesto que se hace hincapié sobre el hecho de que la base y el colocador tienen una relación que va más allá de la adecuación semántica, dado que, en la mayoría de los casos, ninguno de los dos se puede sustituir por un sinónimo.

Sin embargo, el hecho de incluir en la actividad un ejemplo (el último), en el que uno de los componentes de la colocación se puede sustituir por un sinónimo, abre la puerta a una discusión sobre el grado de fijación de las colocaciones, incluso a preguntarnos si realmente se trata de una colocación.

Dado que en la tabla 3.1 aparecen los verbos en el infinitivo, los estudiantes practican, además, la conjugación de los verbos.

En las imágenes 5 y 6 se muestra el resultado de la extracción de la colocación *tomar decisión* con algunos ejemplos en contexto.

Imagen 5: Ejemplo número 6 de la colocación “tomar decisión” en contexto

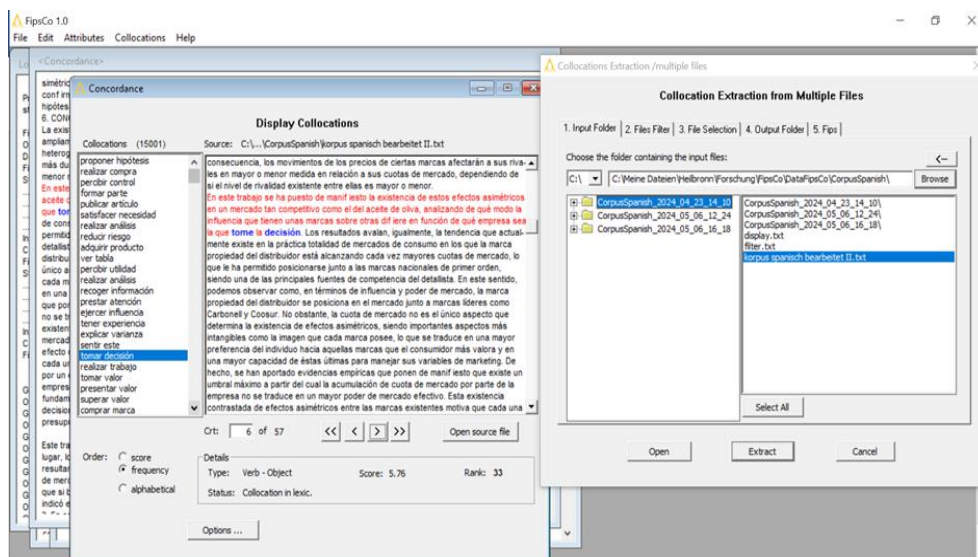
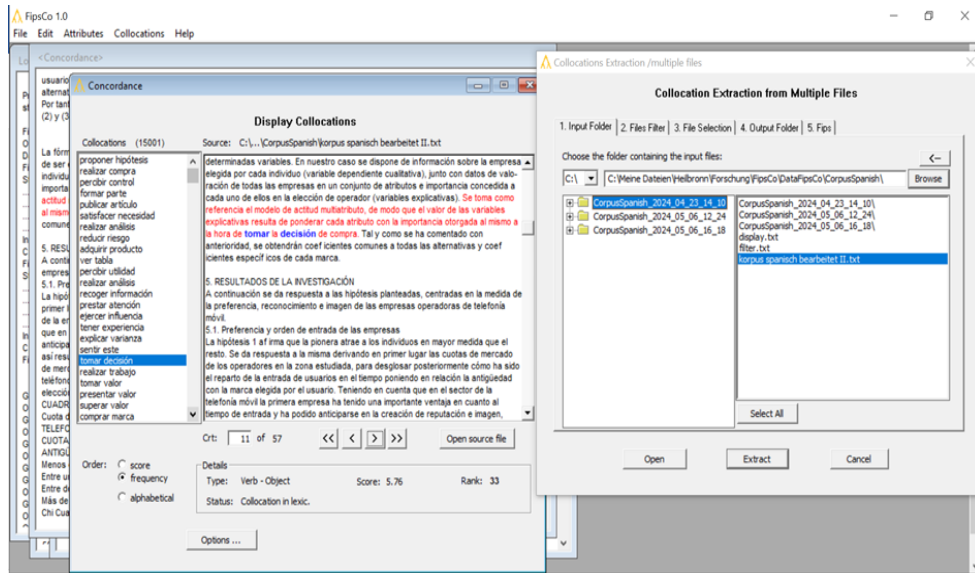


Imagen 6: Ejemplo número 11 de la colación “tomar decisión” en contexto



**Cuarta actividad: Adivina la base**

Tabla 4: Ficha técnica de la cuarta actividad

Título	Adivina la base
Nivel	B2-C1
Duración	Entre 30 y 40 minutos
Grado de dificultad	Medio-alto.
Fase	Comprensión, consolidación y producción (tanto en el desarrollo de la actividad, como en el complemento)
Objetivo principal	- Adivinar la base de una colocación.
Objetivo secundario	- Ampliación del vocabulario.
Material	CorMaC

El grado de dificultad de esta actividad es medio-alto, pero, aun así, es una actividad que les gusta mucho a los estudiantes, porque tiene un componente lúdico. Para su creación, nos hemos basado en el manual *Uso interactivo del vocabulario y sus combinaciones más frecuentes* de las autoras Prada, Salazar y Molero (2012).

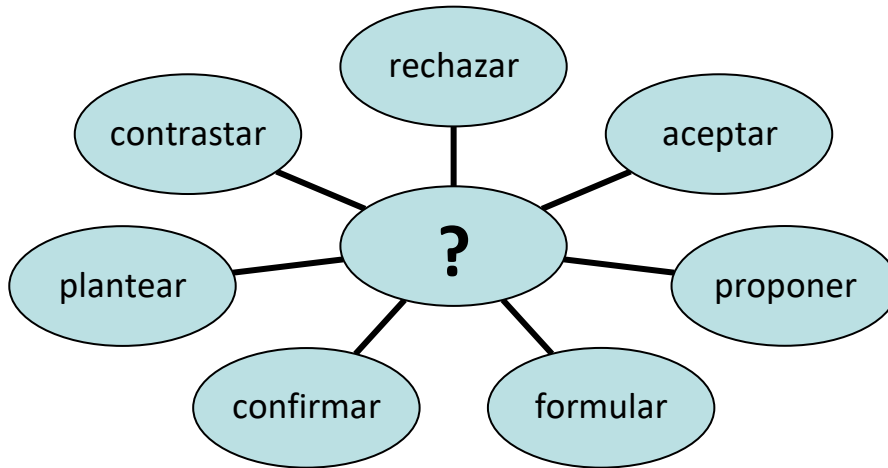
*Desarrollo de la actividad:*

Se les presenta a los estudiantes un diagrama donde se les revela solo un componente de la colocación, en este caso el colocador, y ellos tienen que adivinar la base. Es una actividad que se suele hacer al final de la unidad

didáctica cuando los aprendientes ya están familiarizados con las colocaciones en un ámbito determinado.

Figura 1 para la cuarta actividad<sup>5</sup>

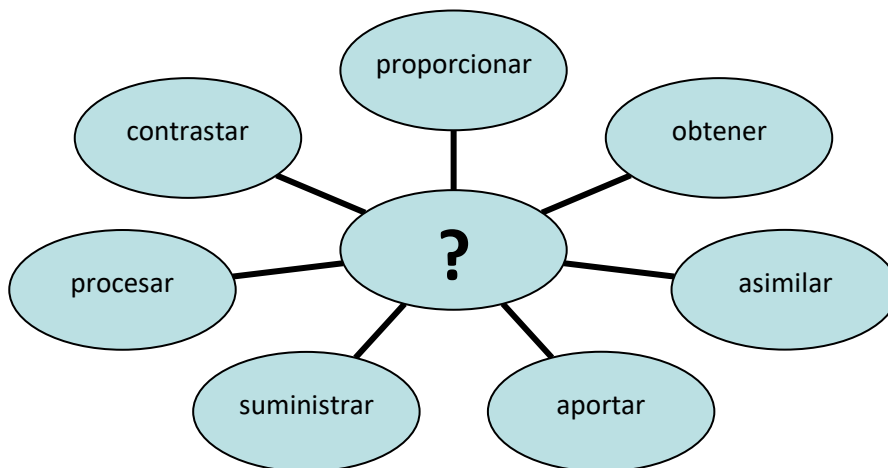
¿Qué base forma colocaciones con estos colocadores?



Siguiente ejemplo:

Figura 2 para la cuarta actividad<sup>6</sup>

¿Qué base forma colocaciones con estos colocadores?



En las imágenes 7 y 8 se pueden observar los resultados con FipsCo de la búsqueda de colocaciones del tipo *verbo + objeto directo* (*contrastar hipótesis* y *plantear hipótesis*).

<sup>5</sup> En este caso la base es la palabra "hipótesis".

<sup>6</sup> En este caso la base es la palabra "información".

Imagen 7: Resultado de la extracción “contrastar hipótesis” (Verbo – Objeto directo)

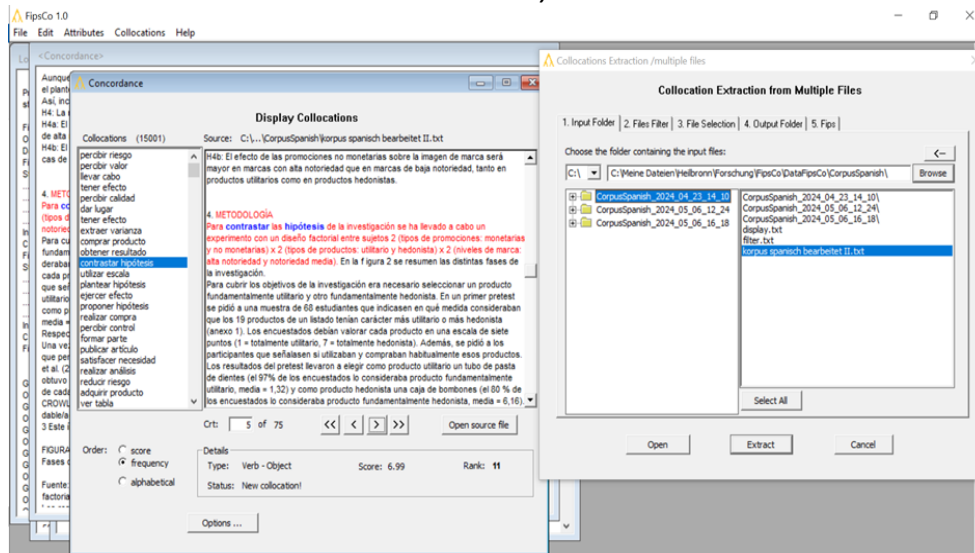
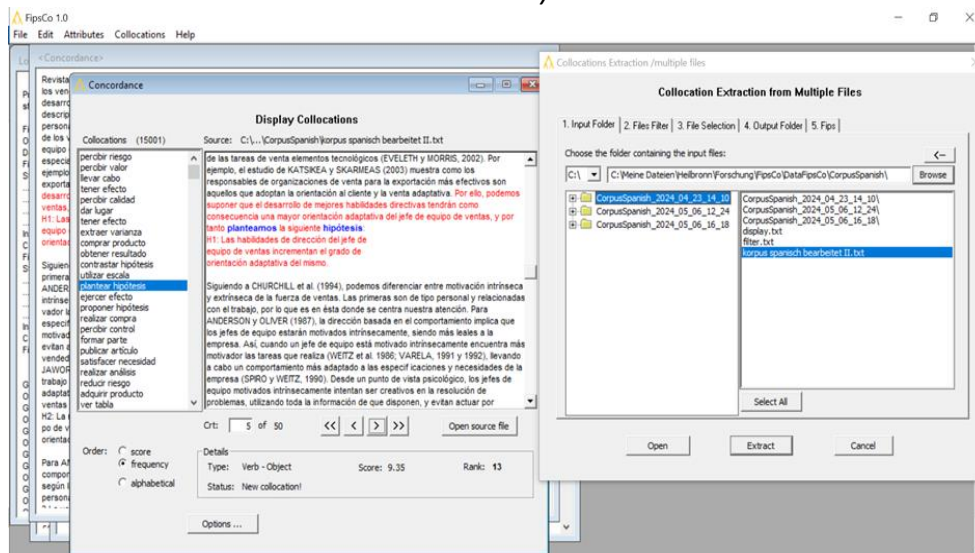


Imagen 8: Resultado de la extracción “plantear hipótesis” (Verbo – Objeto directo)



### Suplemento

Como actividad complementaria a esta, se puede pedir a los estudiantes que escriban una oración con cada una de las colocaciones que acaban de descubrir. Con este suplemento se refuerza la fase del aprendizaje de la producción de las colocaciones en contexto.

### Quinta actividad: Estructuras sintácticas de las colocaciones

*Tabla 5: Ficha técnica de la quinta actividad*

Título	Estructuras sintácticas de las colocaciones
Nivel	B2-C1
Duración	Entre 20 y 30 minutos
Grado de dificultad	Alto.
Fase	Decodificación, comprensión y producción
Objetivo principal	- Hacer conscientes a los estudiantes de las estructuras gramaticales de las colocaciones.
Objetivo secundario	- Repaso de la posición del adjetivo.
Material	CorMaC

La quinta actividad tiene el grado más elevado de dificultad y es el ejercicio perfecto para que los estudiantes sean conscientes de las estructuras de las colocaciones, según Hausmann (1984). Con esta actividad se cierra la unidad didáctica.

*Desarrollo de la actividad:*

Se les entrega a los estudiantes la tabla 5.1 y se les indica que busquen en el corpus colocaciones con estas estructuras.

*Tabla 5.1: Ejercicio para la quinta actividad*

Estructura de la colocación	Ejemplos
Verbo + sustantivo (objeto directo)	
Adjetivo + sustantivo/sustantivo + adjetivo	
Sustantivo + preposición + sustantivo	
Adverbio + adjetivo	
Verbo + adverbio/adverbio + verbo	

Los estudiantes deben rellenar la tabla trabajando en grupos o parejas. La tabla 5.2 muestra un ejemplo de cómo puede quedar la tabla con los resultados:

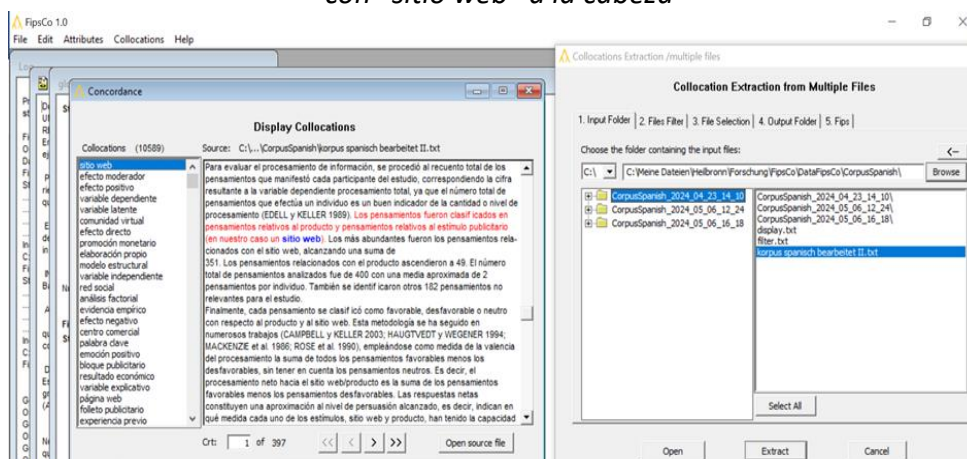
Tabla 5.2: Resultados del ejercicio para la quinta actividad

Estructura de la colocación	Ejemplos
Verbo + sustantivo (objeto directo)	Hacer uso, obtener resultado, tomar decisión, tener lugar, etc.
Adjetivo + sustantivo/sustantivo + adjetivo	red social, palabra clave, sector turístico, etc.
Sustantivo + preposición + sustantivo	base de datos, fuente de información, proveedor de servicio, canal de distribución, objeto de estudio, corriente de investigación, etc.
Adverbio + adjetivo	altamente innovador, ampliamente aceptado, sustancialmente diferente, etc.
Verbo + adverbio/adverbio + verbo	examinar detalladamente, implicar necesariamente, conocer exactamente, etc.

En el caso de la colocación con la estructura *nombre + adjetivo*, se puede hacer un repaso de la posición del adjetivo tal y como sugiere el objetivo secundario de esta actividad.

En la imagen 9 se puede observar el resultado de búsqueda de colocaciones del tipo *nombre + adjetivo* en CorMaC con Fipsco.

Imagen 9: Resultado de la búsqueda de las colocaciones *nombre + adjetivo* con “sitio web” a la cabeza



## 7. Conclusiones

Este trabajo permite elucidar todo el proceso de la creación de un corpus lingüístico propio para su uso en la fraseodidáctica, concretamente en la enseñanza de colocaciones en clase de los niveles avanzados de español para fines específicos (EFE). Se ha expuesto desde la compilación de los



textos que componen el corpus y la extracción de las colocaciones con ayuda de FipsCo mediante un potente analizador sintáctico (*parser*), hasta la creación de una unidad didáctica con introducción y cinco actividades. La propuesta resulta muy útil, especialmente teniendo en cuenta la escasez de material didáctico disponible para la didáctica de EFE en niveles avanzados.

En el caso de CorMaC se trata de un corpus de textos científicos en el ámbito del *marketing*. Sin embargo, este tipo de procedimiento se puede aplicar a cualquier disciplina. Se pueden compilar corpus en el ámbito de los recursos humanos, las finanzas, la medicina, la psicología, el derecho, etc. Además, las posibilidades que ofrecen los corpus como material didáctico son infinitas. Se pueden usar para la enseñanza de la funcionalidad de las preposiciones, de los tiempos del pasado, de los conectores, de las estructuras sintácticas, etc.

Las actividades que componen la unidad didáctica presentada fueron diseñadas en su totalidad por la autora de este artículo sin ayuda de inteligencia artificial. Su implementación en clase tuvo una buena acogida por parte de los estudiantes que se mostraron muy receptivos al tema de las colocaciones, puesto que muchos de ellos no eran conscientes de las consecuencias de un mal uso de estas. La única desventaja de este método es que resulta muy laborioso para el docente. La falta de tiempo muchas veces impide al docente dar rienda suelta a su creatividad diseñando y poniendo en práctica material didáctico de creación propia. Por este motivo, una posible línea de investigación para el futuro podría consistir en comprobar hasta qué punto los chatbots basados en inteligencia artificial tales como Gemini, CoPilot, ChatGPT 4.0 o la plataforma Nolej son capaces de crear unidades didácticas a partir de corpus lingüísticos. De momento, sabemos que son capaces de generar actividades a partir de imágenes, textos, audios e, incluso, de vídeos de youtube. Sin embargo, faltan estudios que brinden información sobre la creación de material didáctico para clases de ELE y/o de EFE usando una combinación de programas de inteligencia artificial y lingüística de corpus.

Lo que queda claro de momento, es que la inteligencia artificial está revolucionando este campo, por lo tanto, apremia sondear las posibilidades que nos brindan.

## 8. BIBLIOGRAFÍA

- AGUILAR RUIZ, Manuel José (2013): "Notas sobre las posibilidades de aprendizaje de español mediante unidades fraseológicas". *MarcoELE, Revista de Didáctica ELE*, 17, pp. 1-23.
- BARTSCH, Sabine (2004): *Structural and functional properties of collocations in English, A corpus study of lexical and pragmatic constraints on lexical co-occurrence*. Tübingen: Narr.
- BURGER, Harald (2003): *Phraseologie: Eine Einführung am Beispiel des Deutschen*. Berlin: Erich Schmidt Verlag.
- CORPAS PASTOR, Gloria (1996): *Manual de fraseología Española*. Madrid: Gredos.

- COUNCIL OF EUROPE (2020): *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Strasbourg, Council of Europe Publishing.
- DETRY, Florence (2015): “Retos iniciales en la comprensión y memorización de las expresiones idiomáticas en clase de ELE”. *MarcoELE. Revista de didáctica ELE*, 21, pp. 1-18.
- DURRANT, Philip y SCHMITT, Norbert (2009): “To what extent do native and non native writers make use of collocations?”. *International Review of Applied Linguistics*, 47, pp. 157-177.
- ETTINGER, Stefan (2007): “Phraseme im Fremdsprachenunterricht”. En Harald Burger et. al. (eds.), *Phraseologie./Phraseology. Ein internationales Handbuch zeitgenössischer Forschung/An International Handbook of Contemporary Research*, 2. Halbband/Volume 2, Berlin, New York: Walter de Gruyter, pp. 893–908.
- EVERT, Stefan (2009): “Corpora and collocations”. En Anke Lüdeling y Merja Kytö (eds.), *Corpus linguistics: an international handbook*. Berlin: De Gruyter, pp. 1212-1248.
- FIRTH, John Rupert (1957): “A Synopsis of Linguistic Theory 1930-55”. *Studies in Linguistic Analysis*. Oxford: The Philological Society, pp. 1-32.
- GOLDMAN, Jean-Phillippe y WEHRLI, Eric (2001). FipsCo: A Syntax-Based System for Terminology Extraction.
- GONZÁLEZ REY, María Isabel (2012): “De la didáctica de la fraseología a la fraseodidáctica”. *Paremia*, 21, pp. 67-84.
- HAUSMANN, Franz Josef (1984): „Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen“. *Praxis des neusprachlichen Unterrichts*, 31, pp. 395-406.
- HAUSMANN, Franz Josef (1989): „Praktische Einführung in den Gebrauch des Student’s Dictionary of Collocations“. *Student’s Dictionary of Collocations*, Evelyn Benson, Morton Benson y Robert Ilson (eds.), Cornelsen, pp. iv-xviii.
- HUNSTON, Susan (2009): “Collection strategies and design decisions”. En Anke Lüdeling y Merja Kytö (eds.), *Corpus linguistics: an international handbook*. Berlin: De Gruyter, pp. 154-168.
- KILGARIFF, Adamy RUNDELL, Michael (2002): “Lexical Profiling Software and its Lexicographic Applications – a Case Study”. *Proceedings of the Tenth EURALEX International Congress*. Copenhagen: Københavns Universitet, pp. 807-818.
- KÜHN, Peter (1985): “Phraseologismen und ihr semantischer Mehrwert. Jemandem auf die Finger gucken in einer Bundestagsrede“. *Sprache und Literatur in Wissenschaft und Unterricht*, 16, pp. 37–46.
- KÜHN, Peter (1987): “Deutsch als Fremdsprache im phraseodidaktischen Dornröschenschlaf. Vorschläge für eine Neukonzeption phraseodidaktischer Hilfsmittel». *Fremdsprachen Lehren und Lernen*, 16, pp. 62-79.
- KÜHN, Peter (1992): “Phraseodidaktik. Entwicklungen, Probleme und Überlegungen für den Mutterspracheunterricht un den Unterricht Daf”. *Fremdsprachen Lehren und Lernen*, 21, pp. 169-189.
- KÜHN, Peter (1994): “Pragmatische Phraseologie: Konsequenzen für Phraseographie und Phraseodidaktik” en Barbara Sandig (ed.): *Europhras 92. Tendenzen der Phraseologieforschung*. Bochum: Universitätsverlag Dr. N. Brockmeyer, pp. 411-428.
- LARRETA ZULATEGUI, Juan Pablo (2001): *Fraseología contrastiva del alemán y el español. Teoría y práctica a partir de un corpus bilingüe de somatismos*. Frankfurt am Main: Peter Lang.

- LÜGER, Heinz-Helmut (1997): "Anregungen zur Phraseodidaktik". *Beiträge zur Fremdsprachenvermittlung*, 31, pp. 69–120.
- MEJÍA, Eva (2021): "Análisis del tratamiento de unidades fraseológicas en los recursos didácticos para estudiantes de ELE". *Doblele: revista de lengua y literatura*, 7, pp. 117-133.
- OLÍMPIO DE OLIVEIRA SILVA, María Eugenia, Inmaculada PENADÉS MARTÍNEZ Y Ana María RUIZ MARTÍNEZ (2006): *Ni da igual, ni da lo mismo: para conocer y usar las locuciones verbales en el aula de español: Nivel avanzado B2-Nivel superior C1*. Madrid: Edinumen.
- PAMIES, Antonio y PAZOS, José Manuel (2005): "Extracción automática de colocaciones y modismos" en Juan de Dios Luque Durán und Antonio Pamies Bertrán (eds.), *La creatividad del lenguaje: colocaciones idiomáticas y fraseología*. Granada: Granada Lingvistica.
- PRADA, Marisa; SALAZAR, Dánica Y MOLERO, Clara María (2012): *Uso interactivo del vocabulario y sus combinaciones más frecuentes*. Madrid: Edelsa.
- REDER, Anna (2006): *Kollokationen in der Wortschaftzarbeit*. Wien: Praesens Verlag.
- ROTHKEGEL, Annely (1994): "Kollokationsbildung und Textbildung", *EUROPHRAS 92 Tendenzen der Phraseologieforschung*, en Barbara Sandig (ed.), Bochum: Brockmeyer, pp. 499-523
- RUIZ GURILLO, Leonor (2000): "Un enfoque didáctico de la fraseología española para extranjeros" en María José Coperías Aguilar, Jorge Redondo Sánchez y Julia Sanmartín Sáez (eds.), *Aprendizaje y enseñanza de una segunda lengua*, Valencia: Universidad de Valencia, pp. 259-274.
- RUIZ YEPES, Guadalupe y KRISHNAMURTHY, Ramesh (2010): "Corpus Linguistics and Second Language Acquisition - the use of ACORN in the teaching of Spanish Grammar", *Lebende Sprachen. Zeitschrift für fremde Sprachen in Wissenschaft und Praxis*. 55(1), pp. 108-122.
- RUIZ YEPES, Guadalupe (2017): "Kollokationen in der Fachsprache des Marketings: Entwicklung und Anwendung eines Extraktionsmodells", *Fachsprache. International Journal of Specialized Communication*, 39 (3-4), pp. 139-157.