

# Uso de corpus lingüísticos por y para profesores de español como lengua extranjera

## Uses of corpora in Spanish Language Teaching

**Magdalena Abad Castelló**

**Profesora titular y formadora. Instituto Cervantes Mánchester  
Doctoranda. Universidad Internacional Iberoamericana**



*Magdalena Abad Castelló es profesora titular del Instituto Cervantes de Mánchester desde 2001. Ha trabajado para instituciones británicas como la Open University o la Universidad de Mánchester. Es licenciada en Filología Inglesa por la Universidad de Murcia en 1992 y Máster en Traducción por la Universidad de Salford en 1997. En 2006 obtuvo su Máster en Formación de Profesores de Español por la Fundación Universitaria Iberoamericana. Sus temas de interés e investigación son la formación de profesores, los contenidos socioculturales y afectivos en el aula de lenguas extranjeras y la enseñanza y aprendizaje de léxico. En este momento está terminando su trabajo de campo sobre el uso de corpus lingüísticos en el aprendizaje y enseñanza de combinaciones léxicas*

### Resumen

Si bien los docentes de lenguas extranjeras reconocen el potencial del uso de corpus lingüísticos en la enseñanza y aprendizaje de lenguas, en el campo de la enseñanza del español como segunda lengua (ELE) sigue siendo una práctica marginal reducida en la mayoría de los casos a la consulta esporádica de dudas lingüísticas. Este artículo explora las posibles causas de este desuso y propone maneras concretas de integrar esta herramienta en la práctica docente que no presuponen un entrenamiento exhaustivo en la materia.

En la primera sección se presentan los resultados de una pequeña encuesta a profesores de español como lengua extranjera del Reino Unido que manifiestan el escaso uso de esta herramienta por parte de los docentes. En la segunda sección se muestran ejemplos concretos de uso de corpus lingüísticos en la clase de ELE tanto con materiales impresos como mediante acceso directo.

### Abstract

While language teachers acknowledge the great potential of corpora in language learning and teaching, the use of corpora is still a minority practice among teachers of

Spanish as a second language with the exception of sporadic corpus consultation to clarify linguistic questions. The present paper explores possible reasons for this underuse and presents concrete ways of using corpora that do not require comprehensive training and can be integrated with everyday language teaching.

The first section presents the results of a small survey of teachers of Spanish as a foreign language in the United Kingdom, which exposes the limited use of this tool. The second section shows concrete examples of the use of language corpus in the SFL class with both teacher-prepared materials and direct corpus searches by learners.

### **Palabras clave**

corpus lingüístico; TIC; formación del profesorado; Aprendizaje basado en datos (ABD); español como lengua extranjera (ELE)

### **Keywords**

Corpus; ICT; teachers development; Data Driven Learning (DDL); Spanish as a foreign language

## **Introducción**

En las últimas dos décadas se han ido difundiendo las grandes ventajas que ofrecen los corpus lingüísticos en diferentes áreas de las humanidades entre las cuales destaca la lingüística aplicada y, en concreto, la enseñanza y aprendizaje de lenguas extranjeras. Para los profesores de lenguas extranjeras los corpus son ante todo una herramienta de consulta para cuestiones que no aparecen en los libros de gramática o no pueden comprobarse en diccionarios. Son particularmente útiles para responder in situ a preguntas de los estudiantes pues sirven para corroborar creencias intuitivas o para rechazarlas y, en este sentido, son extremadamente útiles para el profesor no nativo (Tsui, 2004). Los corpus son asimismo una fuente de ejemplos auténticos por lo que sirven para ilustrar, para crear ejercicios y para elaborar materiales didácticos.

Tim Johns acuñó en 1990 el término “aprendizaje basado en datos”<sup>1</sup>(ABD), también llamado “aprendizaje guiado por datos”, a un enfoque pedagógico en el que el aprendiente se convierte en el investigador o “detective” y el profesor es un mero facilitador. Es el alumno el que examina datos provenientes de corpus lingüísticos y el que extrae sus conclusiones de los mismos. Para Bernardini (2016) los corpus permiten a los estudiantes observar lo que se dice típicamente en unas circunstancias concretas y cómo se dicen y relacionan los dos, significado y forma. Y la adquisición solo se producirá si el aprendiz hace esta conexión entre forma y significado (Asención-Delaney et al. (2015).

Entre las ventajas de uso de los corpus Vyatkina y Boulton (2017) destacan que proporcionan a los aprendices ejemplos de uso evidenciados, les ayudan a desarrollar destrezas analíticas y de resolución de problemas y promueven la autonomía en el aprendizaje. Balunda (2009) y Boulton (2009) defienden que el ABD se adhiere al enfoque comunicativo dado que promociona el aprendizaje inductivo de lengua en contexto, la enseñanza centrada en el alumno y la autonomía en el aprendizaje (capacitan para trabajar de manera autónoma). Además, como sostiene Bernardini (2016), si se fomentan las actividades de investigación en grupo para solucionar problemas lingüísticos se crea un contexto social, colaborativo en el que se construye el

---

<sup>1</sup> Data Driven Learning (DDL), en inglés.

conocimiento. Otra virtud, siguiendo a Mauranen (2004) y Boulton (2009), es que los corpus proporcionan ejemplos genuinos, a menudo auténticos, en contraste con muchos ejemplos de los manuales o las improvisaciones, a menudo simplistas, del profesor. De estos ejemplos el estudiante va a extraer unas conclusiones o descripciones sobre la lengua, que puede que no satisfagan a un lingüista pero pueden serle útiles. Y lo que es más, tal y como sostiene Mauranen (2004), hay que tener en cuenta que la gramática pedagógica imperante hoy día en las aulas de lenguas extranjeras se sirve a menudo de reglas simplificadas y aproximadas.

En resumidas cuentas, el aprendizaje basado en datos acerca al estudiante de lenguas al conocimiento de un nativo. En palabras de Pérez-Paredes y Zapata-Ros:

La utilización de actividades ABD pueden permitir al aprendiz de lenguas acceder a un conocimiento que de forma intuitiva puede llegar a poseer un hablante nativo de la lengua en al menos algunos de los numerosos registros de uso de la lengua. (Pérez-Paredes y Zapata-Ros, 2018: 7).

Pese a todas estas virtudes, lamentablemente, muchos profesores, tanto veteranos como nuevos profesores o profesores en formación, no están familiarizados con el uso pedagógico de corpus lingüísticos. Para Lenko-Szymanska (2014) la dificultad radica en que para implementar aprendizaje basado en datos se necesita tener competencia técnica, conocer bien los usos de los corpus y también saber cómo implementar este conocimiento en el aula, cómo diseñar materiales adecuados y que se complementen con las otras actividades y técnicas de la clase de lengua. Vyatkina y Boulton (2017), a su vez, exponen los siguientes problemas y limitaciones:

- Falta de transparencia en la interfaz de muchos de los corpus disponibles. Se necesita una formación especializada y/o un periodo de entrenamiento para poder percibir mucha de la información que puede ofrecer un corpus.
- Autenticidad. Los ejemplos de los corpus, si bien son genuinos, puede que no sean auténticos para estudiantes de lenguas (Widdowson, 1978) ya que están extraídos de un contexto muy amplio y a menudo proceden de textos que poseen muy poco interés para ellos.
- El tiempo de preparación. Muchas actividades del ABD requieren del profesor una considerable inversión de tiempo y son muy contadas las actividades didácticas listas para usar.

## Usos y funciones de los corpus

Desde un punto de vista didáctico, los corpus pueden ser útiles para actividades de comprensión lectora y auditiva. Ayudan a entender palabras, frases o estructuras al proporcionar un gran número de ejemplos y al poder explorar el cotexto y contexto de estas expresiones. Durante las actividades de producción (sobre todo en la preparación para tareas orales o durante las actividades de expresión o interacción escritas) sirven para comprobar cuestiones sintácticas como régimen de palabras u orden de palabras y para consultar dudas sobre vocabulario como la variedad sociolingüística (diatópica, diastrática, diafásica), connotaciones o colocaciones habituales.

En lo que respecta a las fases de una unidad didáctica los corpus se utilizan durante las actividades posibilitadoras y durante las actividades comunicativas escritas. Y esta herramienta puede ser muy útil en la fase de la postarea (aplicando un enfoque por tareas), en particular durante la retroalimentación, cuando se evalúan los objetivos

lingüísticos y comunicativos y se atiende a las dificultades surgidas. Aquí el profesor puede usar un corpus para ilustrar usos sintácticos o precisar significados, tanto para corregir errores o sugerir mejoras como para responder a dudas de los estudiantes.

Las dos aplicaciones más explotadas de los corpus en la enseñanza de lenguas son los listados de frecuencias y las líneas de concordancia que, en palabras de Cruz Piñol (2012, p. 43) “presentan la palabra buscada (*keyword*) en su contexto de modo que proporcionan mucha información sobre el funcionamiento de ese término en la lengua real y en el nivel textual”. La palabra clave o nódulo aparece en el centro del listado rodeada de las palabras que surgieron junto a ella en el texto. Otra aplicación es el texto contextualizado, es decir, usar el fragmento de donde proviene la línea de concordancia. Por último, los corpus cuentan hoy día con diferentes maneras de presentar estadísticas sobre diferentes variables como región geográfica, periodo o género de texto, información que puede ser invaluable para comprender el uso de una unidad léxica o lexicogramatical.

Existen dos maneras para trabajar con el corpus en el aula de lenguas: acceso directo (los estudiantes consultan directamente el corpus y las líneas de concordancia) o con mediación del profesor, que proporciona a sus estudiantes una selección de datos o concordancias para que las examinen. El acceso directo hace que los alumnos sean más autónomos (Daskalovska, 2015) pero a veces puede abrumar a los estudiantes cuando se enfrentan a las numerosas páginas de concordancias de algunas búsquedas. O’Keeffe y Farr (2003) sugieren empezar con impresiones e ir gradualmente usando el ordenador.

Por lo que se refiere a los aspectos o contenidos investigados con los corpus destacan los contenidos gramaticales y léxicos y, muy frecuentemente, los contenidos lexicogramaticales. Por ejemplo, son frecuentes los estudios de colocaciones preposicionales o adverbiales. En este sentido, es particularmente adecuado el uso de líneas de concordancia pues, además de confirmar significado en contexto (densidad léxica), también nos ayudan a identificar formas y a inferir y/ o corroborar restricciones y reglas sintácticas.

Otras áreas explotadas con ABD son la corrección de tareas o la escritura creativa. Menos frecuentes son las investigaciones en ABD sobre contenidos fonológicos (estudiados con corpus orales) o pragmáticos. También escasean los estudios sociolingüísticos sobre variedades dialectales de la lengua.

En el campo de la enseñanza de ELE los corpus (muy frecuentemente el CREA<sup>2</sup>) se han utilizado principalmente para contenidos léxicos como colocaciones, expresiones idiomáticas, falsos cognados, cuasi-sinónimos y lexicogramaticales<sup>3</sup>. Asimismo se encuentran experiencias didácticas sobre aspectos gramaticales como uso de conjunciones, pronombres o uso de tiempos<sup>4</sup>.

Recientemente se han empezado a explorar en español como L2 otros aspectos como las variedades diatópicas del español<sup>5</sup> o cuestiones relacionadas con la lengua oral. En concreto, se ha utilizado un corpus discursivo Corpus Val.Es.Co.<sup>6</sup>, para estudiar contenidos funcionales, prosodia, partículas discursivas o construcciones suspendidas en la lengua oral<sup>7</sup>.

---

<sup>2</sup> Corpus de Referencia del Español Actual de la Real Academia Española.

<sup>3</sup> Pérez Avila (2006), González Pellicari (2006), Bruno Galván y González Pellicari (2009), Cruz Piñol (2012) y Sánchez Rufat y Jiménez Calderón (2013).

<sup>4</sup> Da Sívla (2006) y Paz (2013).

<sup>5</sup> Bailini (2014).

<sup>6</sup> Corpus Val.Es.Co. (Valencia Español Coloquial) de la Universitat de Valencia.

<sup>7</sup> Nica (2012), Secci (2014) y Pérez Giménez (2015).

## Nuestra encuesta

A pesar de todo el potencial pedagógico de los corpus, su utilización no parece haberse hecho realidad en la práctica diaria entre los docentes de español como lengua extranjera. Solo muy recientemente los manuales de ELE incluyen actividades concretas de uso de corpus y existen muy pocos materiales para usar en el aula.

Para explorar más el uso real de los corpus lingüísticos en el aula de ELE decidimos hacer una pequeña encuesta entre compañeros docentes. Para ello, se elaboró un cuestionario en línea, con *Google forms*, y se envió el enlace a profesores de español de universidades, *colleges*<sup>8</sup> y centros del Instituto Cervantes en el Reino Unido (Ver Cuestionario en Anexo 1).

Las preguntas de indagación eran:

*¿Están familiarizados los profesores con los corpus lingüísticos en español?*

*¿Utilizan en su práctica docente los corpus lingüísticos a su disposición?*

*¿Qué razones aducen los profesores para no utilizar los corpus?*

*¿Para qué utilizan los corpus los profesores?*

Se obtuvo un número reducido de respuestas, 22, y este dato ya puede ser revelador. Se puede aventurar que para muchos profesores el tema no era de su interés y para otros era un tema desconocido en sus clases por lo que no se sentían preparados para responder y elegían no participar. Aunque, por otro lado, son muchos los cuestionarios enviados a profesores de ELE sobre diferentes áreas de investigación por lo que existe una cierta saturación en el gremio.

16 de los participantes son profesores en universidades, cinco son profesores del Instituto Cervantes y tan solo un profesor de *college* respondió a la encuesta. La mayor parte de los participantes son profesores veteranos con más de 12 años de experiencia (12), entre 7 y 12 años de experiencia (7) y solo tres de los encuestados tienen menos de 7 años de experiencia, entre 4 y 6 años de experiencia. En cuanto a la edad, el grueso de los participantes (15) tienen entre 41 y 55 años, 6 están en la franja entre 26 y 40 años y solo un encuestado es mayor de 55 años.

En cuanto al tipo de alumnado, predominan lógicamente los estudiantes universitarios, especialistas (licenciatura en lenguas y/o traducción) y no especialistas (alumnos de centros de lenguas universitarios), seguidos por un tercio de profesores con alumnos adultos, presumiblemente alumnos de centros del Instituto Cervantes y del *college* mencionado anteriormente.

*¿Están familiarizados los profesores con los corpus lingüísticos en español?*

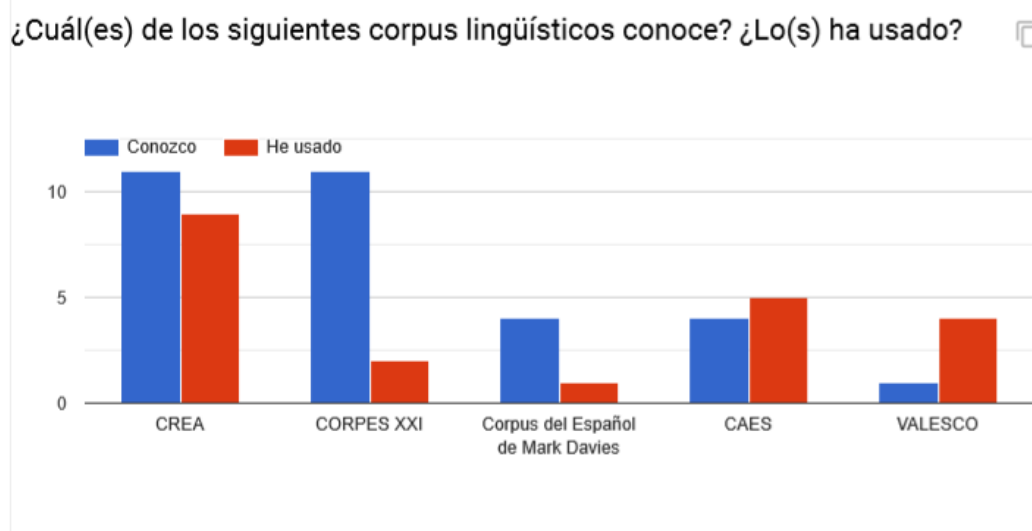
Tal y como se muestra en el gráfico siguiente, los profesores parecen conocer los corpus de la Real Academia, sobre todo el CREA (20) y algunos (13) el CORPES XXI<sup>9</sup>. Nueve profesores conocen el CAES<sup>10</sup>. Solo cinco conocen el Corpus del Español de Mark Davies y otros cinco conocen el Corpus Val.Es.Co.

<sup>8</sup> Con *college* nos referimos a College of Further Education, Further Education se refiere a cualquier estudio a partir de la enseñanza secundaria, que no sea estudio universitario.

<sup>9</sup> Corpus del Español del Siglo XXI de la Real Academia Española.

<sup>10</sup> Corpus de aprendices de español como lengua extranjera, del Instituto Cervantes y de la Universidad de Santiago de Compostela.

Sin embargo, solo nueve profesores usan el CREA, cuatro el CAES, dos el CORPES XXI y un profesor el Corpus del Español. Curiosamente 4 de 5 profesores que conocen el Corpus Val.Es.Co. lo han utilizado.

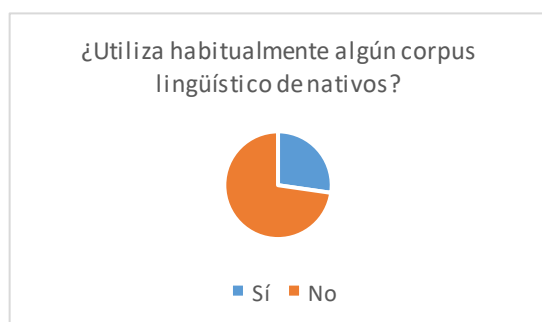


**Gráfico 1.** Conocimiento y utilización de corpus lingüísticos hispanos

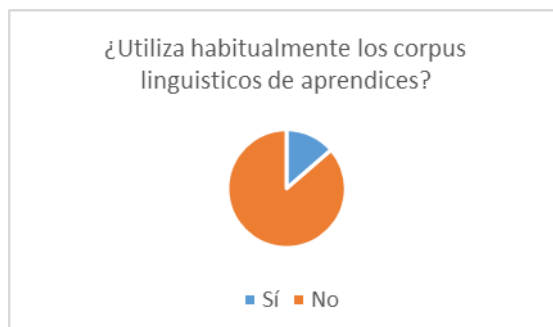
De estos datos se extrae que los profesores están familiarizados con algunos corpus lingüísticos (especialmente el CREA) pero menos de la mitad los utiliza habitualmente.

*¿Utilizan en su práctica docente los corpus lingüísticos a su disposición?*

Como se puede observar en los dos gráficos siguientes, la gran mayoría de profesores (16 de 22) no utiliza habitualmente un corpus lingüístico de nativos (CREA, etc.).



**Gráfico 2.** Uso de los corpus de nativos



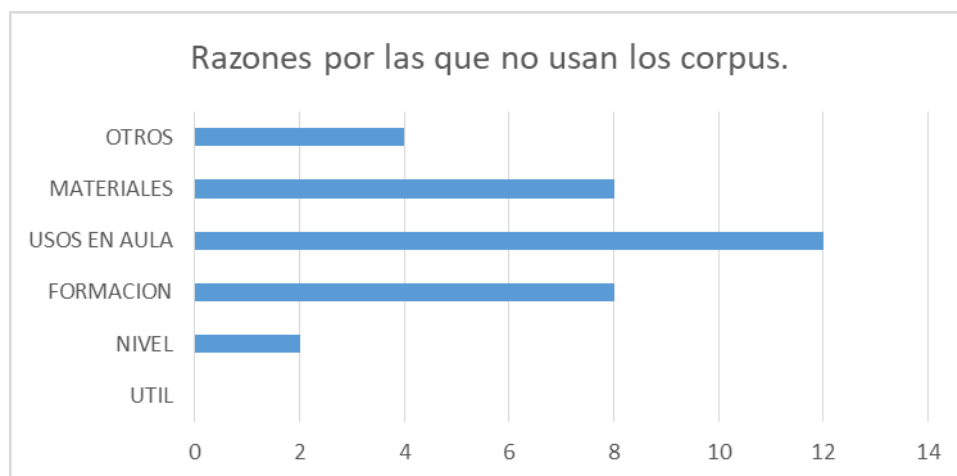
**Gráfico 3.** Uso de los corpus de aprendices

Tan solo 3 de los encuestados utilizan habitualmente algún corpus lingüístico de aprendices, como CAES.

*¿Qué razones aducen los profesores para no utilizar los corpus?*

Como se aprecia en el siguiente gráfico las razones fundamentales por las que la mayoría de profesores participantes no usan los corpus son la falta de formación especializada y de conocimiento sobre usos de los corpus en el aula, seguidas de una carestía de materiales disponibles. Como dato positivo, ninguno de los profesores encuestados considera que los corpus no son útiles. Y, entre otras razones para no

usarlos alegan que no es relevante para sus estudiantes y falta de tiempo para organizar su uso.



**Gráfico 4.** Razones de la no utilización de los corpus

**MATERIALES** No encuentro manuales o materiales con ejemplos prácticos.

**USOS EN AULA** Me falta información sobre usos pedagógicos.

**FORMACIÓN** No tengo la suficiente formación para usarlos.

**NIVEL** El nivel es demasiado avanzado para mis estudiantes.

**UTIL** No creo que sea útil.

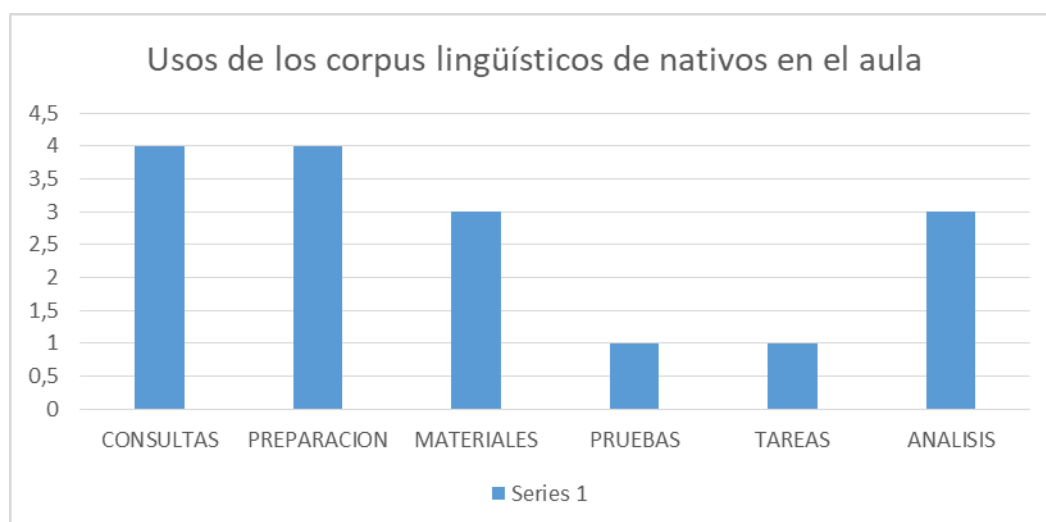
**OTROS**

*¿Para qué utilizan los corpus los profesores?*

A los profesores que respondieron que sí usaban los corpus en su práctica docente, se les preguntó cuáles eran sus usos de los corpus. Como se aprecia en el siguiente gráfico, los usos más frecuentes son las consultas espontáneas y la preparación de tareas. También son frecuentes la creación de materiales y el análisis directo del corpus por los estudiantes.

En cuanto a los corpus de aprendices, de uso aún más minoritario, se mencionan los siguientes usos: creación de materiales, formación de profesores de español, dirección de trabajos de investigación, investigación propia y clases avanzadas de lengua.





**Gráfico 5.** Usos de los corpus de nativos

CONSULTAS Consultas espontáneas

PREPARACION Preparación de actividades

MATERIALES Creación de materiales

PRUEBAS Creación de pruebas

TAREAS Puesta en común de tareas.

ANALISIS Actividades de análisis directo de corpus por los estudiantes (*Data Driven Learning* /Aprendizaje Basado en Datos)

En cuanto a las funciones, los profesores utilizan las cuatro citadas en el cuestionario (textos contextualizados, concordancias, estadísticas, listas de frecuencia) más ejemplos de interlengua, en el caso del corpus de aprendices.

Finalmente se preguntó a los profesores si habían compilado alguna vez un corpus y 5 de los encuestados lo habían hecho para diferentes fines. Citaban proyectos de investigación (incluida una tesis doctoral) y de innovación docente en el ámbito de la enseñanza de lenguas (pragmática), traducir o interpretar, para trabajar un tema específico como ecología, o mundo laboral, para corrección de errores y también como muestras de lengua producidas por no nativos.

De las respuestas a nuestro cuestionario podemos extraer unas primeras conclusiones. En primer lugar se observa que los encuestados conocen algunos corpus lingüísticos pero, con excepciones, no están familiarizados con su utilización pedagógica lo que explica ese uso escaso. De la encuesta se desprende asimismo un uso mayor de los corpus en las universidades (cinco de los 16 docentes) que en el IC (uno de los cinco encuestados). Como cabía esperar en ese contexto, varios de los profesores universitarios han utilizado los corpus en proyectos de investigación.

Por último, el análisis de todos estos datos revela que aquellos docentes que sí utilizan esta herramienta saben sacarle partido para diferentes usos y se sirven de diferentes funciones.

### Ejemplos concretos de uso de corpus lingüísticos en ELE

En esta sección se presentan usos de esta herramienta por parte de la autora del artículo. Todos los ejemplos son genuinos de su práctica docente o de su preparación de pruebas de evaluación, actividades de clase o materiales didácticos.



Como se observa en las respuestas de los cuestionarios, uno de los usos más frecuentes de los corpus por parte de los profesores es la **consulta espontánea en clase**. A menudo los docentes los utilizan para obtener ejemplos genuinos de estructuras gramaticales y evitar improvisaciones que pueden ser simplificadas y no mostrar adecuadamente el uso. Por ejemplo, se puede hacer una búsqueda de líneas de concordancia con “después de que” más Indicativo. Esto nos dará ejemplos reales y también información sobre variedades regionales como que es más frecuente su uso con Indicativo en Hispanoamérica.

6/23/2019 CORPES  
 Corpus del Español del Siglo XXI (CORPES) Santander  
 Listado de concordancias Página: 1

Nº	CONCORDANCIA	después
1	enseñanza. Me dio a conocer los tres reinos de los animales y saber todos los nombres.	de que pasan los treinta días a uno le hace crecer esta uña como de diablo. Entonces
2	Y después vi cómo la luz, segundos	de que el coche o el camión de transporte hubiera pasado por aquel lugar, se volvía
3	un intento de transmitir sus propias voces interiores, luchando por ser rockero	de que la moda del yeah-yeah-yeah pasara. Los Tropicalistas lo veían como un bicho
4	marcadas a fuego en el plástico. Fue mientras miraba la palabra "Esperanza" terminada,	de que Gloria repitiera, paso a paso, la tragedia de la fábrica militar de Río Tercero
5	para ver a Ricky en el mega recital en la 9 de Julio. Además venía muy bajoneada	de que me rechazaran para la escuela de modelos de Anamí Ferreira y para el precasting
6	mariposas. No nos faltaba nada. Yo, personalmente, no la extrañaba en lo más mínimo;	de que se murió le cerré la puerta de mis pensamientos, me quedé con la llave y
7	... y así las papilas quedan como una mesa	de que hubo una comida alborotada.
8	disfruta muchísimo leyéndola, y disfruta aún más cuando no la lee; quiero decir,	de que ya la ha leído, recordando tal o cual momento.
9	«Pero... claro: todo no se puede...», pensó	de que los visitantes se retraron.
10	de movimiento entre el torso y el bulto de las telas, que se desfilen un instante	de que ella hubiese fijado la mirada en tu cara.
11	ahogaba, tal vez la presión o un infarto, por supuesto que el médico llegó horas	de que muriera. Y en el velorio M'bill desaparecía cada tanto y volvía con un olor
12	eso que ahora antes de comer soplo el silbato. Aunque no logro que mi canario coma	de que el Señor feudal se hubiera dignado a estrechársela desde el otro lado de
13	vasaños, como esos pobres diablos que no se lavaban la mano derecha por un mes	de que dos compañeros de clase levantaran la mano (para no dejar en evidencia mi
14	Pidió nuevos voluntarios para su comunidad de los hermanos Dominicos y yo me ofrecí.	de que terminé la universidad para maestra, emigré a Siberia. Ahora que lo pienso
15	-Yo era tan romántica cuando joven...-suspira lina pinchando un tomate.	Después de que un día aventó la puerta y me dijo adiós, yo me pregunto ¿qué más quería
16		de que a su padre, en medio de una desgraciada cacería, lo asesinara la patrulla
17	instrucción primaria comenzó a trabajar en el taller mecánico de Steve Toll, poco tiempo	de que El Negro abandonó mi cuarto, fui por el pasillo que daba al jardín. La oscuridad
18	oportunidad en que la luna se ocultaba tras los nubamones de un cercano temporal.	de que El Negro consumara su venganza. No cumplí, las lágrimas recoman mis mejillas
19	confusión impidieron que tapara el cuerpo. Recordé que me había jurado no llorar	de que Mamá Audolia empezó a vomitar sangre, tuvieron que internarla en un sanatorio
20		Después

Figura 1. Búsqueda de “después de que” en el CORPES XXI

Otra consulta frecuente es la investigación léxica. Ante la pregunta de un estudiante de nivel A2 sobre cuándo usar “coger” o “tomar” se puede trabajar con colocaciones, “tomar un baño funciona” pero no “coger un baño” a no ser que nos refiramos a trasladar un sanitario. Pero también puede ayudar buscar la frecuencia de los dos verbos. En el Corpus del Español de Mark Davies “coger” tiene una frecuencia de 51.839 casos mientras que “tomar” es muchísimo más frecuente, con 1.049.908 ocurrencias. De nuevo se puede buscar la variedad lingüística y se puede observar que “coger” tiene un uso más reducido en América Latina, lo que sería una posible explicación de esa gran diferencia en la frecuencia.

6/23/2019 NOW Corpus: Spanish

Corpus del Español: NOW

SEARCH FREQUENCY CONTEXT OVERVIEW

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

	CONTEXT	ALL FORMS (SAMPLE)	FREQ
1	COGER	51839	

0.750 seconds

Figura 2. Búsqueda de “coger” en el Corpus del Español de Mark Davies

6/23/2019 NOW Corpus: Spanish

Corpus del Español: NOW

SEARCH FREQUENCY CONTEXT OVERVIEW

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

	CONTEXTO	ALL FORMS	FREQ
1	TOMAR	(SAMPLE): 100 200 500	1049908

Figura 3. Búsqueda de “tomar” en el Corpus del Español de Mark Davies

En la **puesta en común de tareas** podemos consultar un corpus de manera espontánea o podemos planificar consultas cuando anticipamos posibles dificultades en la realización de un ejercicio o actividad.

Por ejemplo, en una actividad de un manual de nivel B1 aparece la pregunta *¿Dónde es la fiesta?* El profesor prevé que algún estudiante le va a preguntar si es una errata y, si no lo es, va a preguntar por qué. En lugar de proporcionar directamente una explicación, el profesor busca en el Corpus XXI “¿Dónde es”<sup>11</sup> :

3/21/2019 Concoordancias

REAL ACADEMIA ESPAÑOLA

Corpus del Español del Siglo XXI (CORPES) Versión beta (0.91)

Concordancias | Coapariciones | Configuración | Ayuda | Modo de cita | Sugerencias

Lema Forma ¿Dónde es Clase de palabra (Todos) Grafía original - Subcorpus + Proximidad

Subcorpus Limpiar

Título Autor Fecha de clasificación

Origen (Todos) América España Zonas lingüísticas (Todos) España

Medio (Todos) Escrito Oral Bloque (Todos) Ficción No ficción Soporte (Todos) Internet Libro Tema (Todos) Actualidad, ocio y vida cotidiana Artes, cultura y espectáculos Tipología (Todos) Académico Biografía memoria

Marcas  Cursiva  Negrita  Subrayado  Versalitas

17 casos en 16 documentos.

REF. (Clasificación, país)	CONCORDANCIA	Ordenar por:
1 2001 Esp.	— ¿Dónde es?	Año ascendente   sin criterio
2 2003 Esp.	Ulises. — ¿Dónde es "ahí"?	
3 2004 Esp.	MAMATONTA. — ¿Dónde es por ahí?	
4 2007 Esp.	, me tomó del brazo una viejecita que no logré identificar. «Nena -me preguntó- ¿dónde es el aperitivo?».	
5 2007 Esp.	(¿Dónde es "aquí"?)	
6 2007 Esp.	AINA: ¿Aquí? ¿Dónde es "aquí"? No lo sé. Estaba en mi casa, he abierto la nevera y "pam".	
7 2007 Esp.	mira / a mí me dan estas paranoias mentales / y sé // ¿dónde es el festival de cine // en Franca // tan conocido // que está al lado de Mónaco	
8 2007 Esp.	¿dónde es ahora el decanato?	
9 2009 Esp.	¿dónde es?	
10 2010 Esp.	pero ¿dónde es? bueno da igual	
11 2011 Esp.	-Tu hermano. ¿Dónde es, tío?. pillamos el metro y vamos para allá. Los libros los llevamos nosotros	
12 2012 Esp.	— Lo mismo de cada año. Confiar en nuestras posibilidades y estar arriba. ¿Dónde es arriba para nosotros? Pues ahora no lo sé, pero confío en que volverá a	
13 2015 Esp.	—Pero eso ¿dónde es? ¿En el barrio?	
14 2015 Esp.	— ¿Dónde es la cena? —pregunta ella al cabo de un rato.	
15 2015 Esp.	—¿Allí? ¿Qué lugar es ese? ¿Dónde es allí? —Indaga Larousse, reacio a abrir cauces inesperados a sus particulares	
16 2015 Esp.	¿dónde es la gasolina?	
17 2016 Esp.	—Muy bien. Y eso, ¿dónde es? No tiene que moverse del sitio. Ya iré yo.	

1 - 17 | Imprimir | Exportar | Exportar TGV | 1 de 1 | Ir a página: | Ir

Referencia bibliográfica:

Figura 4. Búsqueda de “¿Dónde es” en el CORPES XXI

Como resultado de la búsqueda se encuentran los ejemplos “¿Dónde es el aperitivo?” “¿Dónde es el festival de cine?” y “¿Dónde es la cena?” Entonces se puede preguntar a los estudiantes qué tienen en común los tres ejemplos. Entre todos se puede llegar a una regla pedagógica, se usa “Dónde es” cuando se sitúa un acontecimiento o actividad en el espacio.

<sup>11</sup> Nótese que en las búsquedas en el CORPES XXI es necesario dejar un espacio entre el signo de interrogación o exclamación y la primera palabra de la pregunta o exclamación.

Los corpus lingüísticos pueden ser una fuente inestimable para **crear materiales didácticos**. Se pueden encontrar textos con ejemplos de un tema o con una colocación precisas. Para empezar el tema del tiempo atmosférico para un nivel B2 /C1 se encontró el siguiente fragmento con varias colocaciones relacionadas con el clima.

Hace un par de años, en **una tarde muy tormentosa del otoño neoyorquino**, caminando a casa desde la universidad me encontré con una larga cola de esperanzados jugadores que aguardaban en la calle su turno para comprar un billete de lotería de la multimillonaria megaloto. Soportaba una **lluvia torrencial salpicada de rayos y truenos** delante de la pequeña tienda de la calle 35. Por casualidad, me acerqué a una pareja que esperaba al final de la cola, divertidos y empapados bajo un diminuto paraguas, y les pregunté amablemente si sabían que la probabilidad estadística de que les tocara el gordo era menor que la de que les **cayera un rayo**. Aunque un tanto sorprendidos por mi pregunta, dos me respondieron al unísono sonrientes que “en teoría, sí”, pero no les preocupaba porque se sentían mucho más cerca del golpe de buena suerte que de la **chispa eléctrica**. Imagino que el antropólogo Lionel Tiger captó este comportamiento cuando dijo aquello de que “calcular con optimismo las probabilidades es una tendencia humana tan básica como buscar comida cuando se tiene hambre”.

**Figura 5.** Texto extraído del CORPES XXI, (*La fuerza del optimismo*, Luis Rojas Marcos)

En los corpus también se pueden hallar audios y transcripciones de audios o se pueden transformar líneas de concordancias en frases para un ejercicio. Por ejemplo, el siguiente ejercicio para practicar “Aunque” o “si bien” está extraído de líneas de concordancia del CORPES XXI. (También para estudiantes de nivel B2 o C1)

a. ¿Aunque? ¿Si bien? ¿O los dos son posibles?

1. \_\_\_\_\_ abandonó la c, nunca pudo evitar verse inmersa en la política insular.
2. La Universidad debe situarse en el centro, \_\_\_ existan otras instituciones de enseñanza media-general y superior.
3. España tendrá que reducir el déficit público, \_\_\_\_\_ podrá discutirse cómo recortarlo.
4. El desayuno es la comida más importante del día \_\_\_ hay muchas personas que se lo saltan.
5. En las grandes ciudades de nuestro entorno, \_\_\_\_\_ se cuenta con mayor acceso a los servicios públicos, la calidad y el hacinamiento son un problema grave.
6. \_\_\_\_\_ estoy a dieta, de vez en cuando me como un helado.
7. Pero \_\_\_ es cierto que esa bonanza ha permitido contar con cuantiosos recursos económicos, los mismos han tenido un efecto perverso sobre la inflación acumulada en Venezuela
8. Las causas de la endometriosis todavía son desconocidas \_\_\_ abundan muchas teorías.

**Figura 6.** Ejercicio creado a partir de concordancias del CORPES XXI

De la misma manera, los datos de los corpus son un recurso excelente para **crear pruebas o exámenes** como ítems de pruebas de selección múltiple o rellenahuecos. Los siguientes ítems se crearon a partir de líneas de concordancia del CORPES XXI.

1. El jugador portugués aguantó el \_\_\_\_\_ de objetos lanzados desde la grada.  
aguacero    chaparrón    huracán    palo de agua
2. Se emocionó ante el \_\_\_\_\_ recibimiento de la gente del pueblo.  
caluroso    caliente    caldeado    acalorado

**Figura 7.** Ítems de una prueba de selección múltiple con datos del CORPES XXI y del CREA

Los ejemplos vistos hasta ahora muestran el uso directo del corpus por parte del profesor. Ahora bien, el objetivo final del Aprendizaje Basado en Datos es que los estudiantes trabajen con esta herramienta de forma autónoma, por ejemplo, en la **preparación de las tareas**. Así, como paso previo a una tarea escrita en la que deben redactar una conversación informal, se puede pedir a los estudiantes que investiguen posibles marcadores conversacionales en un corpus oral. Asimismo se les puede sugerir que comprueben frases y colocaciones de las que no estén muy seguros en un corpus.

Por último, el profesor puede sugerir **tareas de análisis** en las que los estudiantes examinan unidades léxicas con el fin de comprender mejor su significado, forma y función. Se puede proporcionar material impreso como la siguiente tabla comparativa de las combinaciones con “hacerse” y “volverse” (a partir de B2).

HACERSE		VOLVERSE	
	mayor		viral
	fuerte		loco
	fuertes		popular
	gorda		famoso
	sueco		loca
	más		famosa
	mayores		insostenible
	mucho		común
	propia		importante
	visible		difícil
	famoso		agresivo
	efectivo		imposible
	difícil		violenta
	autónomo		tensa
	buen		adicto
	nuevo		fuerte
	regulares		incontrolable
	visibles		insoportable
	anuales		violento
	fácil		complicado

**Figura 8.** Tabla comparativa, datos del Corpus del Español

Los estudiantes pueden subrayar las combinaciones que conocen y/o les parecen lógicas. Pueden comentar aquello que les sorprenda. Para ello pueden servirse de connotaciones negativas o positivas o rasgos concretos de significado como cambios voluntarios o involuntarios.

Una tarea de análisis de contenido gramatical (de nuevo de nivel B2 o C1) puede ser analizar los tiempos verbales de las concordancias de “llevar + gerundio” en la siguiente imagen:

6/27/2019 CORPES



Corpus del Español del Siglo XXI  
(CORPES)

Listado de concordancias



27/6/2019 Página: 1

Nº	CONCORDANCIA	
1	amistades que iban llenando los mortuorios a lo largo de la mañana. El doctor, se decía <b>llevaba</b>	haciendo lo mismo desde el momento de su prematura y siempre anhelada jubilación
2	voces de esa media horilla de corruptela a un sobrino, el joven Juanillo, que se <b>llevaba</b>	encargando desde entonces de la operación con no menor eficacia y hasta auténtico
3	Lo que <b>llevaba</b>	colgando el joven que acababa de embocar la calle de la panadería era una cámara
4	pensaba que la tecla del remarcado se había inventado para su hermano Antonio, al que <b>llevaba</b>	llamando desde que se había ido Teresa, y volvió a pulsarla por rutina, mientras
5	aquí cinco años, soy fotógrafo”, aclaró señalando con el dedo la enorme bolsa que <b>llevaba</b>	colgando, “ando siempre a la caza de cosas bellas y... ya ve usted. Puedo pasar
6	montarle un... espectáculo porque había coincidido semejante día con la luna llena: “ <b>Llevo</b>	soñando algo así por lo menos desde que vivo con él. Se llama Raúl y es fotógrafo
7	atravesando no fuera salvajemente aplastante, pensaron todos de una forma u otra: “Yo lo <b>llevo</b>	viendo ya hace tiempo por las esquelas: hace cinco años, cinco, no más, los conocía
8	“Ya estamos con lo de la salud y el tabaco. Yo <b>llevo</b>	fumando desde siempre y no pienso morirme todavía, Lucía.
9	nada desease más, por fin que se atrevía, que decirle adiós a esa desgracia que <b>llevaba</b>	arrastrando ya tantos años a su lado, a ese loco que no sentía más placer que el
10	cuando recibía el móvil”, sonrió, y dejó por fin recado en el contestador: “Mi vida, <b>llevo</b>	buscándote toda la tarde. Tengo una cena que presumo larga y voy a mantener el
11	sombra”, se le escapó entre dientes, y la bella polaca, que advirtió la queja y <b>llevaba</b>	viéndolo desde el principio contrariado, quiso animarlo: “¿Me vas a dejar que un
12	Qué manía con lo de... don Jesús. ¡Si nos conocemos de toda la vida, Antonio! ¡Si <b>llevo</b>	viniendo aquí... Yo qué sé cuánto llevo viniendo aquí!
13	conocemos de toda la vida, Antonio! ¡Si llevo viniendo aquí... Yo qué sé cuánto <b>llevo</b>	viniendo aquí!
14	en su amo, pensó, y sintió pena, pena por la situación general de Antonio, al que <b>llevaba</b>	viendo muy descentrado desde hacía tiempo, y especial pena por la situación en
15	devolverle la copa, pero estaba cansada y más valía acabar de una vez por todas: “ <b>Llevo</b>	danzando desde las seis y media y estoy más muerta que viva”, le sonrió con cuanta
16	fuera cierto cuando, misterios de la mente, se le vino a la cabeza el apellido que <b>llevaba</b>	buscando más de medianoche: “Romero, Romero, Francisco Romero”, se repitió para
17	por aún insatisfecho y de una profunda euforia por haber realizado las fotos que <b>llevaba</b>	persiguiendo desde los tiempos de la facultad, que anhelaba desde que su antigua
18	No lo mató, papá, pensó Tomás sin ganas de responder a una cuestión sobre la que <b>llevaban</b>	discutiendo más de veinte años, ni lo mató ni aprobó su muerte, don Antonio, nadie
19	mamá, tú has tenido a tu marido y a tus hijos, has sido la señora de la casa, yo <b>llevo</b>	llorando mi único pecado desde que era una chiquilla, durante muchos años me pareció
20	obnubilados, avergonzados el uno del otro, fue la primera vez que se aflojó el cepo que <b>llevaba</b>	torturándome ocho o diez años, aunque nunca se aflojó del todo, aunque siempre

Figura 9. Búsqueda de “llevar + gerundio” en el CORPES XXI

Los estudiantes van a descubrir que los tiempos más frecuentes son el presente y el imperfecto y no van a encontrar el indefinido o perfecto simple. Un error de muchos estudiantes de español es el usar esta perífrasis con tiempos perfectos. Estos ejemplos les pueden ayudar a evitar esos tiempos verbales.

Otra tarea de análisis, de nivel B2 en adelante, puede ser eliminar la unidad léxica buscada de unas líneas de concordancia e intentar descubrir toda la información posible. En el ejemplo siguiente se ha eliminado un marcador conversacional, “y punto”, de todas esas líneas de concordancia.

1. no nos hace falta / él colaboró con la Policía / le damos las gracias públicamente \_\_\_\_\_
2. como nos pidieron / como nos exigieron a / conversar con el Barcelona // lo hemos hecho \_\_\_\_\_ / nada más / después ya es una cuestión de los clubes
3. extradiciones solicitadas por el juez Garzón // usted mencionó las leyes de obediencia debida \_\_\_\_\_ final // quiero saber si el Gobierno español está en conocimiento de que
4. tienen razón / porque sí / sales más tarde / lo que pasa que vas un poco mejor arreglado \_\_\_\_\_ / pero yo qué sé / me gusta verlo / como algo diferente
5. mañana puede salir una ley autonómica que diga en la hostelería no se puede fumar \_\_\_\_\_ y sin embargo según la ley estatal / eeh en algunos casos / tú podrías decidi
6. música que no era en directo / pues no me empanaba mucho de qué iba / yo andaba a lo mío \_\_\_\_\_ pero tenía mucho en directo sobre todo el reggae // había mucho reggae // y
7. por fin Huancayo / punto final de nuestro viaje en tren / \_\_\_\_\_ de arranque del viaje al corazón de la cultura peruana en Huancayo casi todos
8. bien y si no voy también / y yo ahora llamo por teléfono y no voy a dormir / y no voy \_\_\_\_\_ // y entonces / no tengo ¿sabes? que dar / ninguna explicación / las doy porque
9. ves que bueno no / no es culpa tuya / tú no puedes hacer nada / tú haces lo que puedes \_\_\_\_\_
10. conteste no no volváis / si si veis clara una pregunta / no la penséis más / contestarla \_\_\_\_\_ // es preferible // y efectivamente / yo tengo el recuerdo de una compañera / de

Figura 10. El marcador escondido, líneas de concordancia del CORPES XXI

Los estudiantes tienen que investigar la función del marcador, el registro, la posición. Pueden pensar qué marcador usarían en su lengua. Aunque el objetivo no es descubrir cuál es el marcador, también cuentan con algunas pistas por dos usos en los que “y punto” no es un marcador sino una frase, indicios que les ayudarían a rastrearlo.

Por último, el profesor puede diseñar tareas para un uso directo de los corpus. Una tarea fácil (para todos los niveles del MCER) es la búsqueda de la variedad diatópica de una unidad léxica. Por ejemplo, se puede buscar “chévere” en estadísticas.

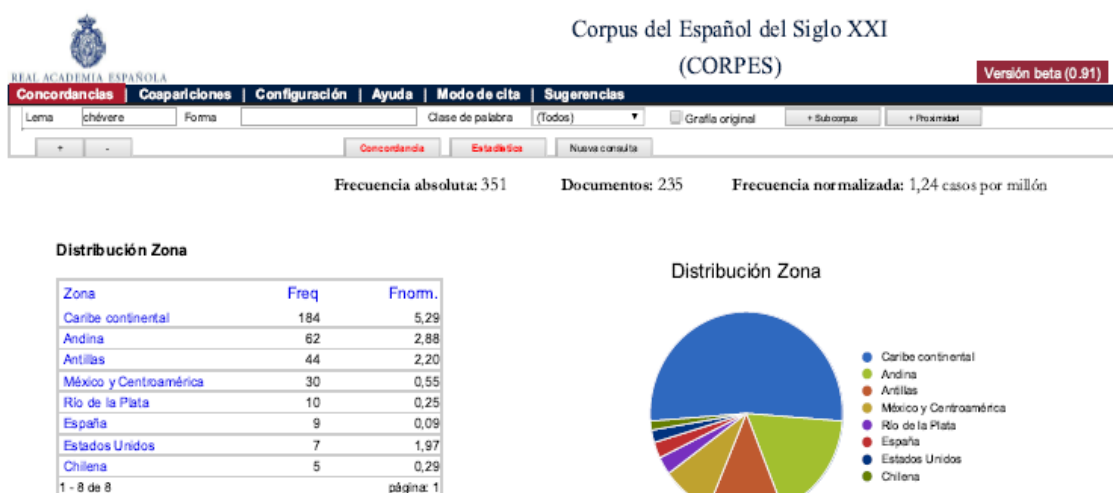


Figura 11. Búsqueda de “chévere” en Estadísticas del CORPES XXI

Este resultado ayudará al estudiante a identificar la variedad geográfica donde la unidad léxica es más usada.

Los corpus pueden dar ejemplos e información sobre expresiones idiomáticas. Por ejemplo, se les puede pedir a un grupo de estudiantes que investiguen seis expresiones idiomáticas relacionadas con la religión. La tarea (recomendada para nivel B2 o en adelante) se realiza en tres grupos, cada equipo busca dos de las frases y las analiza con la siguiente tabla. Después comparten la información con los miembros de los otros grupos.

	(No ser) santo de mi devoción	En santiamén un	Acabar como el rosario de la aurora	Para más inri	Como Dios manda	Quedarse para vestir santos
Número de ocurrencias						
Número de documentos						
Significado ¿Se confirma vuestra hipótesis?						
Las tres áreas geográficas de mayor frecuencia						
Los tres países de mayor frecuencia						
El periodo con mayor frecuencia						
Otros						

Figura 12. Plantilla de análisis de expresiones idiomáticas relacionadas con la religión

Una vez los estudiantes están familiarizados con esta herramienta se les puede pedir que trabajen en casa directamente con un corpus lingüístico. En la figura siguiente se puede ver un ejemplo de uso del CORPES XXI para el análisis de marcadores discursivos.

Vamos a trabajar con tres conectores contraargumentativos de nivel C1. Busca en el CORPES XXI “por el contrario”, “ahora bien” y “con todo”, analiza los ejemplos e intenta responder a las preguntas de la tabla siguiente:

	Por el contrario	Ahora bien	Con todo
Expresa un contraste entre dos miembros			
Introduce un argumento contrario, normalmente fuerte			
Presenta una conclusión contraria			
¿Posición inicial, final o intermedia?			
¿Registro formal o informal? ¿Escrito u oral?			
¿Otra información?			

Figura 13. Plantilla de análisis de conectores contraargumentativos



## Conclusiones

En este artículo se ha explorado el uso de los corpus lingüísticos por parte de profesores de español como lengua extranjera en el Reino Unido. Tras analizar los resultados de una pequeña encuesta, se puede extrapolar que los profesores de español como lengua extranjera conocen la existencia de los corpus lingüísticos, sobre todo de los corpus de la Real Academia Española. Sin embargo, muchos menos profesores los han usado y muy pocos docentes de ELE los usan habitualmente. Las razones de este desuso son fundamentalmente la falta de conocimiento y/o formación específica sobre el uso pedagógico de esta herramienta.

Ahora bien, los profesores que utilizan habitualmente los corpus muestran un gran dominio de la herramienta. Usan los corpus (tanto de nativos como de aprendientes) para diferentes objetivos, se sirven de todas las funciones y algunos han compilado corpus para investigación y otros fines. Por ello, parece evidente que a mayor conocimiento de la aplicación pedagógica de esta herramienta lexicográfica, mayor es la utilización de la misma.

Se necesita, por tanto, una mayor divulgación de buenas prácticas tanto en la formación de profesores como en la publicación de materiales y recursos. En la última sección del artículo se han mostrado algunos ejemplos de uso de los corpus tanto por el profesor como directamente por el alumno. Esperamos que este artículo y estas muestras contribuyan a dar a conocer el gran potencial de esta herramienta lexicográfica entre los profesionales de español como lengua segunda o extranjera.

## Referencias bibliográficas

ASENCIÓN-DELANEY, Y., Collentine, J. G., Collentine, K., Colmenares, J., & Plonsky, L. (2015). El potencial de la enseñanza del vocabulario basada en corpus: optimismo con precaución. *Journal of Spanish Language Teaching*, 2(2), 140-151. doi:<https://doi.org/10.1080/23247797.2015.1105516>

BAILINI, S. (2014). Los corpus como recursos didácticos para la enseñanza de las variedades diatópicas del español. *FIAPE. V Congreso internacional: ¿Qué español enseñar y cómo? Variedades del español y su*, (págs. 25-28). Cuenca. Obtenido de: <https://www.mecd.gob.es/dam/jcr:9879b92d-9541-4b3f-b2bd-67a58a86f26e/3--los-corpus-como-recursos-didacticos-para-la-ensenanza-de-las-variedades-diatopicas-del-espanol--bailinisonia-pdf.pdf>

BALUNDA, S. A. (2009). *Teaching academic vocabulary with corpora: students perceptios of data-driven learning*. Memoria de máster. Indiana University. Obtenido de <https://scholarworks.iupui.edu/bitstream/handle/1805/2049/Balunda%20MA%20Thesis%20Teaching%20Academic%20Vocabulary%20with%20Corpora.pdf>

BERNARDINI, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. En J. Sinclair (Ed.), *How to Use Corpora in Language Teaching*. Amsterdam: Benjamins. doi:<https://doi.org/10.1075/scl.12.05ber>

BERNARDINI, S. (Enero - Junio de 2016). Discovery Learning in the language-for-translation classroom: corpora as learning aids. *Cad. Trad.*, 36(nº especial 1), 14-35. Obtenido de [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2175-79682016000500014](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2175-79682016000500014)

BOERS, F., & LINDSTROMBERG, S. (2009). *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Basingtoke: Palgrave Macmillan.

- BOULTON, A. (2009). Data-driven learning: reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, 35(1), 81-106. Obtenido de: <https://hal.archives-ouvertes.fr/hal-00326990v2>
- BRUNO GALVÁN, C., & GONZÁLEZ PELLIZARI, M. C. (2009). Corpus lingüístico y la enseñanza del español como lengua extranjera. En A. Lluc, & M. C. González Pellizari (Ed.), *Las nuevas metodologías y sus aplicaciones didácticas. Actas del XVII Seminario de dificultades para la enseñanza del español a lusoparlantes.*, (págs. 145-179). Obtenido de: <https://docslide.net/education/xvii-seminario-de-dificultades-especificas-de-la-ensenanza-de-espanol-a.html>
- CHAMBERS, A. (Mayo de 2005). Integrating corpus consultation in language studies. *Language Learning & Technology*, 9(2), 115-125. Obtenido de: <http://llt.msu.edu/vol9num2/chambers/>
- CRUZ PIÑOL, M. (2012). *Lingüística de corpus y enseñanza del español como segunda lengua*. Madrid: Arco Libros.
- DA SILVA, I. M. (2006). La utilización de la lingüística del corpus en la enseñanza del español: usos de pero y sino. En P. Benítez Pérez (Ed.), *Actas del III simposio internacional José Carlos Lisboa de didáctica del español como lengua extranjera del Instituto Cervantes de Río de Janeiro* (págs. 219-228). Instituto Cervantes. Obtenido de: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/publicaciones\\_centros/rio\\_2006.htm](https://cvc.cervantes.es/ensenanza/biblioteca_ele/publicaciones_centros/rio_2006.htm)
- DASKALOVSKA, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 130-144. Obtenido de <https://www.tandfonline.com/doi/abs/10.1080/09588221.2013.803982>
- GONZÁLEZ PELLIZARI ALONSO, M. (2006). Corpus lingüístico y la adquisición de los falsos cognados en ELE. En A. López Ramírez, & M. González Pellizari (Ed.), *Tienes la palabra: aprendizaje del léxico en ELE Actas del XIV Seminario de Dificultades Específicas de la Enseñanza del Español a Lusohablantes*, (págs. 45-73).
- GÜZEL, M. (2014). The Effect of Data Driven Learning (DDL) on Students' Learning Outcomes. *Humanising Language Teaching*, 16(5). Obtenido de <http://www.hltmag.co.uk/oct14/idea02.htm>
- JOHNS, T. (1990). From printout to handout: Grammar and vocabulary teaching in the context of datadriven. *CALL Austria*, 10, 14-34.
- LENKO-SZYMANSKA, A. (2014). Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the effectiveness of a teacher-training course on the. *ReCALL*, 26(2), 260-278. doi:doi:10.1017/S095834401400010X
- MAURANEN, A. (2004). Spoken - general: Spoken corpus for an ordinary learner. En J. Sinclair (Ed.), *How to use corpora in language teaching*. Amsterdam: Benjamins.
- MCCARTHY, M., & CARTER, R. (1995). Spoken grammar: what is it and how can we teach it? *ELT Journal Volume*, 49(3), 207-218.
- MEYER, C. F. (2006). Corpus Linguistics, the World Wide Web, and English Language Teaching. *IBÉRICA*, 12, 9-21. Obtenido de: [https://www.lancaster.ac.uk/fass/projects/corpus/ZJU/xpapers/Meyer\\_Web\\_corpus.pdf](https://www.lancaster.ac.uk/fass/projects/corpus/ZJU/xpapers/Meyer_Web_corpus.pdf)
- NATION, P. (2001). Research into practice: Vocabulary. *Lang.Teach.*, 44(4), 529-539. doi:10.1017/S0261444811000267

NICA, V. (2012). Los corpus discursivos orales como herramienta didáctica para los niveles iniciales. *Foro de profesores de E/LE*, 8. Obtenido de: <https://ojs.uv.es/index.php/foroele/article/view/6617>

O'KEEFFE, A., & FARR, F. (2003). Using Language Corpora in Initial Teacher Education: Pedagogic Issues and Practical Applications. *TESOL Quarterly*, 37(3), 389-418. doi:[https://ulir.ul.ie/bitstream/handle/10344/4591/Farr\\_2003\\_using.pdf?sequence=1](https://ulir.ul.ie/bitstream/handle/10344/4591/Farr_2003_using.pdf?sequence=1)

PAZ, Y. (2013). Aplicación pedagógica de la lingüística de corpus a una clase de gramática avanzada de español como segunda lengua. *MarcoELE Revista de Didáctica ELE*, 6. Obtenido de: <https://marcoele.com/linguistica-de-corpus/>

PÉREZ GIMÉNEZ, M. (2015). Propuestas didácticas para estudiar la sintaxis coloquial en Bachillerato: las “construcciones suspendidas”. *Tejuelo*, 22, 215-234. Obtenido de: [http://dehesa.unex.es/bitstream/handle/10662/4901/1988-8430\\_22\\_215.pdf?sequence=1&isAllowed=y](http://dehesa.unex.es/bitstream/handle/10662/4901/1988-8430_22_215.pdf?sequence=1&isAllowed=y)

PÉREZ-PAREDES, P., & ZAPATA-ROS, M. (2018). Patrones de Pensamiento Computacional y corpus lingüísticos: el aprendizaje de lenguas con datos lingüísticos. *RED. El aprendizaje en la Sociedad del Conocimiento*. Obtenido de: <http://eprints.rclis.org/32209/>

SÁNCHEZ RUFAT, A., & JIMÉNEZ CALDERÓN, F. (2013). Combinatoria léxica y corpus como input. *Language Design*(14), 61-81.

SECCI, D. (2014). Los corpus discursivos orales como recurso didáctico: la oralidad a través del texto en contexto. *Foro de profesores de E/LE*, 10. Obtenido de: <https://ojs.uv.es/index.php/foroele/article/view/6673>

SIDOTI, R. (2011). Las colocaciones en el aula de E/LE. *redELE*(21). Obtenido de <https://www.mecd.gob.es/dam/jcr:592583e9-7f6d-49f6-b7f4-95bbd827c4f1/2011-redele-21-07sidoti-pdf.pdf>

TSUI, A. (2004). What Teachers have always wanted to know - and how corpora can help. En J. Sinclair (Ed.), *How to use Corpora in Language Teaching*. Amsterdam: Benjamins.

URIEL DOMÍNGUEZ, M. (2015). Las colocaciones en un corpus de aprendices valón y flamenco. En Y. Morimoto, M. V. Pavón Lucero, & R. Santamaría Martínez (Ed.), *La enseñanza de ELE centrada en el alumno. Asociación para la Enseñanza del Español como Lengua Extranjera*, (págs. 977-986). Obtenido de: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/25/25\\_0977.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/25/25_0977.pdf)

VINCZE, O. (2015). *Learning multiword expressions from corpora and dictionaries*. Tesis doctoral. Universade da Coruña.

VYATKINA, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1-8. Obtenido de: [https://scholarspace.manoa.hawaii.edu/bitstream/10125/44750/1/21\\_03\\_commentary.pdf](https://scholarspace.manoa.hawaii.edu/bitstream/10125/44750/1/21_03_commentary.pdf)

WIDDOWSON, H. (1978). *Teaching Language as Communication*. Oxford: Oxford University Press.

## Apéndice

### ENCUESTA SOBRE USO DE CORPUS LINGÜÍSTICOS EN LA ENSEÑANZA DE ESPAÑOL

*Este cuestionario forma parte de un proyecto de investigación sobre el uso de corpus lingüísticos por parte de profesores de español en universidades y en centros de lenguas del Reino Unido. Consta de ocho apartados y completarlo le llevará alrededor de 5 minutos.*

*Garantizamos la confidencialidad de los datos y le damos las gracias por su inestimable colaboración.*

#### 1. ¿Cuánto tiempo lleva enseñando español?

- 1-3    4-6    7-12    Más de 12 años

#### 2. ¿Cuál es su centro de enseñanza? Marque todas las respuestas que correspondan.

- College  
 Instituto Cervantes  
 Universidad  
 Otro:

#### 3. ¿Qué tipo de estudiantes de español tiene? Marque todas las respuestas que correspondan.

- Estudiantes adultos (enseñanza no reglada)  
 Estudiantes mayores de 16 años y/o adultos (*colleges*)  
 Estudiantes universitarios especialistas (grados o licenciaturas en lenguas)  
 Estudiantes universitarios especialistas en traducción/interpretación (grados o posgrados)  
 Profesores de lenguas en formación (PGCE u otros)  
 Estudiantes universitarios no especialistas (Centro de idiomas, grados y licenciaturas de otras especialidades)

#### 4. ¿Cuál(es) de los siguientes corpus lingüísticos conoce? ¿Lo(s) ha usado?

	Conozco	He usado
CREA		
CORPES XXI		
Corpus del Español de Mark Davies		
CORDE		
CEMC		
Val.Es.Co		
CAES		

**Si conoce otros corpus lingüísticos, por favor, especifique cuál(es):**

#### 5. ¿Utiliza habitualmente algún corpus lingüístico de nativos (CREA, Corpes XXI)?

- Sí    No

Si ha contestado no, ¿por qué razón? Marque todas las respuestas que correspondan.

- No creo que sea útil.  
 El nivel es demasiado avanzado para mis estudiantes.  
 No tengo la suficiente formación para usarlos.  
 Me falta información sobre usos pedagógicos.  
 No encuentro manuales o materiales con ejemplos prácticos.  
 Otros:

Si ha contestado sí, ¿para qué usos? Marque todas las respuestas que correspondan.

- Consultas espontáneas en clase.
- Consultas en la preparación de actividades o en la corrección de tareas escritas.
- Creación de materiales de enseñanza.
- Creación de pruebas o exámenes.
- Corrección /Puesta en común de tareas.
- Actividades de análisis directo de corpus por los estudiantes (*Data Driven Learning* /Aprendizaje Basado en Datos)
- Otros:

Si ha contestado sí, ¿qué aplicaciones o funciones usa? Marque todas las respuestas que correspondan.

- Textos contextualizados
- Concordancias
- Estadísticas
- Listas de frecuencia
- Otros

#### 6. ¿Utiliza habitualmente los corpus lingüísticos de aprendices como el CAES?

- Sí     No

Si ha contestado sí, ¿para qué usos?

Si ha contestado sí, ¿qué funciones usa?

#### 7. ¿Ha compilado alguna vez su propio corpus?

- Sí     No

Si ha contestado sí, ¿para qué uso?

#### 8. Información general

Nombre \_\_\_\_\_ Apellido(s) \_\_\_\_\_

Sexo: M / F

Edad:         menos de 25     de 26-40     41-55     Más de 55 años

Niveles del MCER que enseña:

Centro de enseñanza:

Lengua(s) materna(s):

**¡Muchas gracias!**