

# Aprendizaje de lengua extranjera a través de herramientas digitales de corpus. Actitudes y autoeficacia en estudiantes universitarios del grado de Educación Primaria

Aprendizagem de línguas estrangeiras através de ferramentas de corpus digital. Atitudes e autoeficácia em estudantes de licenciatura em Educação Primária  
Foreign language learning through digital corpus tools. Attitudes and self-efficacy in university students of Primary Education teacher training

Gema Alcaraz Mármol \*<sup>1</sup>

<sup>1</sup> Universidad de Castilla-La Mancha, Departamento de Filología Moderna, Toledo, España.

## Resumen

El presente estudio tiene como objetivo explorar la autoeficacia y actitudes sobre el uso de dos herramientas digitales de corpus por parte de un grupo de estudiantes de inglés como lengua extranjera que cursa el Grado en Educación Primaria en la Universidad de Castilla-La Mancha. Los participantes realizaron una serie de actividades en el contexto de la metodología basada en datos utilizando como recurso dos herramientas de corpus: SKELL y BNC. Tras la realización de las actividades y la manipulación de las herramientas de corpus, los participantes respondieron un cuestionario donde debían expresar su grado de acuerdo con cada una de las afirmaciones que componían dicho cuestionario. El cuestionario hace referencia a aspectos sobre la dificultad de uso, el nivel de utilidad, o la aplicación fuera del aula de las herramientas digitales de corpus; también se pregunta a los participantes por las actividades de aprendizaje basado en datos para la adquisición de distintos aspectos de la lengua inglesa. Los resultados apuntan a una actitud positiva y un nivel considerable de autoeficacia hacia el corpus. Finalmente, se anima a adoptar este tipo de metodología además de a una formación específica del profesorado para su correcta y óptima aplicación.

*Palabras clave:* Autoeficacia. Aprendizaje basado en datos. Corpus. Actitudes. Lengua extranjera.

## Resumo

O presente estudo tem como objetivo explorar a autoeficácia e as atitudes face à utilização de duas ferramentas de *corpus* digital por parte de um grupo de estudantes de inglês como língua estrangeira que frequentam uma licenciatura em Educação Primária na Universidade de Castilla-La Mancha. Os participantes realizaram uma série de atividades no contexto da metodologia orientada para os dados, utilizando duas ferramentas de corpus: SKELL e BNC. Após a realização das atividades e a manipulação das ferramentas de *corpus*, os participantes responderam a um questionário no qual tinham de expressar o seu grau de concordância com cada uma das afirmações que compunham o questionário. O questionário refere-se a aspectos de dificuldade de utilização, nível de utilidade ou aplicação fora da sala de aula das ferramentas de *corpus* digital; os participantes são também questionados sobre atividades de aprendizagem baseadas em dados para a aquisição de diferentes aspectos da língua inglesa. Os resultados apontam para uma atitude positiva e um nível considerável de autoeficácia em relação ao corpus. Por último, incentiva-se a adoção deste tipo de metodologia, bem como a formação específica de professores para a sua aplicação correta e otimizada.

*Palavras-chave:* Autoeficácia. Aprendizagem baseada em dados. *Corpus*. Actitudes. Língua estrangeira.

## Abstract

The present study aims to explore the self-efficacy and attitudes towards the use of two digital corpus tools by a group of students of English as a foreign language studying for a degree in Primary Education at the University of Castilla-La Mancha. The participants carried out a series of activities in the context of the data-driven methodology using two corpus tools: SKELL and BNC. After carrying out the activities and manipulating the corpus tools, the participants answered a questionnaire in which they had to express their

*Textolivre*  
Linguagem e Tecnologia

DOI: 10.1590/1983-3652.2024.51693

Sección:  
Artículos

Autor correspondiente:  
Gema Alcaraz Mármol

Editor de sección:  
Hugo Heredia Ponce  
Editor de maquetación:  
João Mesquita

Recibido el:  
20 de marzo de 2024  
Aceptado el:  
4 de mayo de 2024  
Publicado el:  
29 de julio de 2024

Esta obra está bajo una  
licencia «CC BY 4.0».



\*Email: gema.alcaraz@uclm.es

degree of agreement with each of the statements that made up the questionnaire. The questionnaire refers to aspects such as difficulty of use, level of usefulness, or application outside the classroom of the digital corpus tools; participants are also asked about data-driven learning activities for the acquisition of different aspects of the English language. The results point to a positive attitude and a considerable level of self-efficacy towards the corpus. Finally, the adoption of this type of methodology is encouraged as well as specific teacher training for its correct and optimal application.

*Keywords:* Self-efficacy. Data-driven learning. Corpus. Attitudes. Foreign language.

---

## 1 Introducción

Las nuevas tecnologías van adquiriendo un papel cada vez más relevante en la enseñanza de lenguas. La era digital ha calado también en las aulas a través de dispositivos electrónicos como las pizarras digitales, las tabletas electrónicas o incluso los dispositivos móviles y las redes sociales. Pese a que la cultura digital ya es un hecho, existen aún ciertos recursos y enfoques relacionados con ella que se resisten a formar parte de la enseñanza de lenguas. Este es el caso de las herramientas digitales de corpus y de la metodología basada en datos derivada de dichas herramientas. A este respecto, precisamente, se presenta este trabajo donde se investiga la aplicación de tales recursos digitales para el aprendizaje de inglés como lengua extranjera en el contexto universitario español.

## 2 Marco Teórico

### 2.1 La lingüística de corpus en la enseñanza de lenguas extranjeras

La lingüística de corpus está adoptando un papel cada vez más relevante en la enseñanza y aprendizaje de lenguas extranjeras (Sinclair, 2004; O'Keeffe; McCarthy y Carter, 2007; Römer, 2009; Szudarski, 2017). Así, ofrece herramientas y enfoques pedagógicos que permiten al estudiantado manejar grandes cantidades de texto auténtico en formato digital, obteniendo información útil sobre el significado de palabras clave y determinadas estructuras gramaticales partiendo de su contexto de uso (Hanks, 2013). McEnery y Xiao (2011) diferencian entre usos indirectos y directos del corpus en el marco de la enseñanza de lenguas.

Tradicionalmente, la lingüística de corpus ha tenido un papel indirecto en el aprendizaje, siendo fuente para el diseño de materiales didácticos para lengua extranjera o para la confección de diccionarios de aprendices donde, entre otros, se ofrece información concreta sobre la frecuencia y contexto de uso de las palabras. Podemos encontrar otras aplicaciones indirectas en el ámbito de las lenguas para fines específicos, donde son comunes los glosarios derivados de corpus de especialidad, por ejemplo. Además, y en la línea de las aplicaciones indirectas, no podemos olvidar los corpus de aprendices, cuyo análisis ofrece esclarecedores datos derivados de errores comunes, y que abren una valiosa ventana a la que asomarse a la hora de investigar el proceso de adquisición de lenguas extranjeras (Meunier, 2002).

Por otro lado, respecto a la aplicación directa de los corpus, también denominado aprendizaje basado en datos, autores como Gabrielatos (2005) y Bernardini (2004) destacan sus efectos positivos en la adquisición de una lengua extranjera. En el aprendizaje basado en datos los aprendices manejan herramientas de corpus, las cuales les permiten explorar el comportamiento de la lengua meta en su contexto de uso real. Así, observan la frecuencia de uso de las palabras y sus concordancias, entre otros. Estos procedimientos desarrollan la capacidad del estudiantado para identificar patrones lingüísticos, hacer generalizaciones fundamentadas en la lengua que están aprendiendo y examinar en detalle cierto léxico y estructuras morfosintácticas, contribuyendo así a su competencia metalingüística. El proceso de enseñanza y aprendizaje tiende así a convertirse en una experiencia más motivadora y auténtica, puesto que el estudiante consolida conocimientos que ya posee, a la vez que descubre e integra nuevos (Boulton, 2012).

Zapata-Ros y Pérez-Paredes (2018) destacan el papel que tienen las herramientas digitales de corpus y el aprendizaje basado en datos en el desarrollo de la capacidad reflexiva y de razonamiento del alumnado, que se materializa en distintas destrezas cognitivas como analizar, hacer inferencias, interpretar o verificar. Como bien apunta Boulton (2009), el aprendizaje basado en datos complementa

el enfoque comunicativo ya que fomenta, por un lado, el aprendizaje inductivo, y por otro la autonomía del propio estudiantado. A este respecto, Flowerdew (2015) menciona tres teorías de aprendizaje que fundamentan esta metodología: la llamada Noticing Hypothesis o Hipótesis de la Captación, el Constructivismo y la Teoría Sociocultural de Vygotsky (1978). Dichas teorías están presentes en el aprendizaje basado en datos y el uso de corpus, puesto que mediante estos se construye conocimiento a partir de la búsqueda y la observación, se aprende por notoriedad de patrones y se refuerza la negociación de significado.

Al interactuar con los corpus, el estudiantado aplica destrezas de pensamiento crítico que implican la observación, la conciencia lingüística, la formulación, la comprobación de hipótesis, y la sensibilidad a la variación lingüística. En una tarea típica de aprendizaje basado en datos, el estudiantado examina un nodo (una palabra o una cadena de palabras) y su contexto inmediato, tanto a la izquierda como a la derecha, así como, si lo desea, el contexto más amplio del párrafo. Cada nodo aparece en una denominada línea de concordancia, y a partir de ahí se exploran las pruebas lingüísticas y se construyen hipótesis sobre la naturaleza del uso de la lengua.

Se considera, por tanto, un enfoque constructivista, inductivo y centrado en el alumnado, que fomenta la autonomía y el aprendizaje por descubrimiento. En este contexto de enseñanza, el profesorado actúa como asesor y guía, facilitando, más que transmitiendo, conocimientos lingüísticos. Se presentan, por tanto, una combinación de innovaciones tecnológicas, educativas y varios recursos en línea que promueven la inclusión de diferentes perspectivas pedagógicas dentro y fuera del aula. Así, contribuye a que el estudiantado universitario, actualmente en su mayoría nativos digitales, aprecien los enfoques basados en corpus y los adapten a su aprendizaje de lengua extranjera. Así, a través de este tipo de aprendizaje el alumnado puede llegar a sus propias conclusiones, con una retención más sólida y prolongada y un mayor conocimiento de la lengua meta. En base a estas premisas, se podría afirmar que el estudiantado es menos dependiente del profesorado, y aprende directamente de la lengua en contexto real en lugar de hacerlo de recursos tradicionales como los libros de texto, las gramáticas o los diccionarios (Thomas, 2015).

Sin embargo, y a pesar de los numerosos estudios que defienden las bondades del aprendizaje basado en datos, este enfoque no parece haber recalado aún en las aulas de lengua extranjera. Callies (2016) sugiere que el profesorado requiere de una serie de competencias y conocimientos para poder aplicar con éxito esta metodología: deben estar familiarizados con la lingüística de corpus y las herramientas digitales, además de tener una base de conocimientos lingüísticos en general. Además, los profesores también deben poseer competencias pedagógicas que les permitan aplicar los corpus en el proceso de enseñanza, diseñando materiales adecuados y pedagógicamente sólidos, combinándolos bien con otras técnicas de enseñanza e incorporándolas al contexto educativo. Se puede afirmar que es necesario formar al profesorado futuro y en activo para que desarrollen todas estas competencias mencionadas arriba.

La reticencia de los docentes de lengua extranjera a aplicar el corpus en sus clases ha sido bien documentada en los últimos años por autores como Römer (2009) o Tribble (2015). Estos autores realizaron encuestas a un representativo número de docentes. Los resultados de sus investigaciones apuntan, principalmente, a la falta de formación en este tipo de metodología, aunque también se mencionan la falta de recursos y herramientas de corpus accesibles, intuitivas y de bajo coste. La mayoría de los docentes afirmaban desconocer las diferentes formas en que los corpus se pueden explotar en el aula y las habilidades necesarias para la aplicación de estos conocimientos. También es interesante mencionar que el nivel educativo al que pertenezcan los docentes influye en su visión sobre este enfoque. En un reciente estudio llevado a cabo por Crosthwaite, Luciana y Schweinberger (2021), docentes de inglés de primaria y secundaria en Indonesia participaron en un taller formativo sobre corpus y aprendizaje basado en datos. Tras preguntar a los participantes por sus impresiones, los investigadores detectaron una clara diferencia entre las opiniones de docentes de distintas etapas educativas. Mientras que los profesores de secundaria veían cierto potencial en el enfoque, aunque con ciertas reticencias, los de primaria se mostraban mucho menos atraídos por este tipo de enseñanza en general.

No obstante, la investigación sobre el corpus y el aprendizaje basado en datos ha dado resultados

satisfactorios en cuanto a su acogida por parte de alumnado. Geluso y Yamaguchi (2014) estudiaron las opiniones de un grupo de estudiantes de inglés como lengua extranjera que utilizaron herramientas de corpus para trabajar sus destrezas orales. Tras realizar varias actividades, respondieron a un cuestionario tipo Likert donde se les preguntaba sobre su experiencia. La mayoría de ellos consideraban dichas herramientas como muy útiles pese a la novedad y el tiempo empleado en saber utilizarlas adecuadamente. En la misma línea, los estudios de Aşık, Vural y Akpınar (2016) y Ming Huei Lin (2016) destacan la buena actitud demostrada por el estudiantado tras haber experimentado la metodología de aprendizaje basado en datos. En el primer estudio, además de un cuestionario, los participantes se sometieron a una entrevista en grupo, donde mostraron una actitud muy positiva hacia esta metodología, especialmente a la hora de trabajar con sinónimos y colocaciones.

En Ming Huei Lin (2016), tanto estudiantes de grado como de posgrado afirmaron sentirse más autónomos y protagonistas de su proceso de aprendizaje. En esta investigación se aplicaron tres acciones formativas: en la primera, únicamente se trabajó con herramientas de corpus, en la segunda y tercera acción formativa se combinó el aprendizaje basado en datos con una metodología tradicional en distintas proporciones para cada grupo. Si bien la actitud positiva se detectó en las tres acciones formativas, fue en el primer grupo donde se obtuvieron las actitudes más favorables. Las herramientas de corpus se utilizaron como base para el aprendizaje de lengua extranjera a través de aplicaciones móviles. En este sentido, Pérez-Paredes *et al.* (2019) quisieron saber si el uso de herramientas de corpus en este tipo de dispositivo tenían la misma acogida que en contextos de aula. Los resultados mostraron que, si bien los usuarios sugirieron mejoras referentes al diseño de algunas actividades, quedaron muy satisfechos con este tipo de aprendizaje.

## 2.2 Autoeficacia en el aprendizaje de lenguas extranjeras

Uno de los conceptos relacionados con las creencias y actitudes respecto al aprendizaje es la autoeficacia. La literatura de los últimos veinte años ha resaltado la importancia de las diferencias individuales a la hora de abordar el proceso de enseñanza y aprendizaje de una lengua extranjera (Mónica, 2021; Dörnyei, 2019; Muñoz, 2019; You y Dörnyei, 2014). Autores como Dörnyei (2019) sugieren que, además de las aptitudes y capacidades cognitivas de los aprendices, factores como la motivación, las creencias y otras individualidades tienen un papel fundamental para que un aprendiz tenga éxito en su adquisición de una lengua extranjera. No obstante, el concepto de autoeficacia ha sido de estudio relativamente reciente, comparado con otros factores individuales.

La autoeficacia comenzó a despertar interés a finales de los años 80 con estudios en el marco de la teoría sociocognitiva de la mano de investigaciones como las de Bandura (1986). Sus investigaciones se centran en explorar cómo lo que los individuos piensan de sí mismos y del entorno influye en su manera de comportarse y de afrontar la realización de determinadas tareas Bandura (1986). En una reciente definición Lu-Chun Lin (2014) concibe la autoeficacia como la confianza que una persona tiene sobre ella misma respecto a su propia capacidad para realizar una tarea, determinando así la cantidad de esfuerzo y tiempo que está dispuesta a dedicar a dicha tarea. Investigaciones como la de Linnenbrink y Pintrich (2003) han observado cómo la autoeficacia ayuda de manera significativa en los procesos de aprendizaje a nivel tanto cognitivo como motivacional.

La autoeficacia se ha relacionado con el uso de estrategias de aprendizaje. Li y Wang (2010), Anam y Stracke (2016) o Saito (2020) observaron que los aprendices con mayores niveles de autoeficacia hacían uso de más y mejores estrategias de aprendizaje. De la misma manera, Mar Haro Soler (2021) explora los beneficios de la autoeficacia para alumnado de traducción. Más concretamente en el campo de las lenguas extranjeras, la relación entre la autoeficacia y el uso de estrategias es una de las líneas más exploradas. Entre los estudios más recientes encontramos a Teng, Wang y Wu (2023) o Bai y Wang (2023). El primero de los estudios es especialmente significativo dado el número de estudiantes que formaron parte del mismo. Más de 500 estudiantes de inglés de origen chino realizaron una encuesta donde se les preguntaba por diferentes estrategias y aspectos motivacionales. Los resultados revelaron que la autoeficacia predecía especialmente el uso de estrategias metacognitivas y aumentaba la motivación por aprender en una época particularmente complicada para la docencia como fue la pandemia por covid-19. En un estudio incluso más reciente, Bai y Wang (2023) comprobaron que

no solo las estrategias como la planificación o la automonitorización se veían favorecidas por la autoeficacia, sino también la motivación intrínseca de los participantes a la hora de realizar distintas tareas en lengua inglesa.

También encontramos investigaciones que muestran cómo la autoeficacia es un buen predictor para determinar el éxito del estudiantado a la hora de realizar ciertas actividades. Por ejemplo, Waddington (2019) exploró qué tareas comunicativas estaban más afectadas por un bajo nivel de autoeficacia. Llegó a la conclusión de que tanto las destrezas receptivas como las productivas se vieron mermadas por esa falta de autoeficacia. Por su parte, Sardegna, Lee y Kusey (2018) observaron que un alto nivel de autoeficacia mejoraba los resultados de aprendizaje en términos de pronunciación. Es la destreza de producción escrita la que ha sido más explorada en este contexto. Como ejemplos podemos mencionar a Woodrow (2011), Villalón, Mateos y Cuevas (2013) o Teng (2024). El primero observó que la mayoría de los participantes presentaba cierto grado de ansiedad a la hora de enfrentarse a una tarea escrita en inglés. Los resultados del estudio apuntaban a que esa relación estaba condicionada por la autoeficacia. Es decir, aquellos estudiantes con más nivel de autoeficacia manejaban mejor sus niveles de ansiedad a la hora de realizar la producción escrita. Por su parte, Villalón, Mateos y Cuevas (2013) introdujeron la variable del género, queriendo averiguar si había diferencias significativas entre estudiantes del género femenino y masculino a la hora de cómo lidiaban con sus tareas de producción escrita según una mayor o menor autoeficacia. Los investigadores no encontraron diferencias significativas entre ambos géneros, si bien confirmaban estudios anteriores respecto al papel de la autoeficacia en la producción escrita en lengua extranjera. En la línea de los dos anteriores, el trabajo de Teng (2024) muestra que el nivel de autoeficacia está íntimamente ligado tanto a estrategias cognitivas como metacognitivas utilizadas en las tareas escritas.

Como se ha discutido arriba, los estudios que se refieren a la autoeficacia en el contexto de enseñanza y aprendizaje de una lengua extranjera se centran en la relación de esta con ciertas destrezas comunicativas y con otras diferencias individuales como son la ansiedad o la motivación. Sin embargo, y a pesar del interés que despierta dicha autoeficacia actualmente en el aprendizaje de lenguas extranjeras, faltan estudios que exploren su papel en el contexto de distintas metodologías de enseñanza como puede ser las que adoptan herramientas digitales de corpus.

### 3 Objetivo de la Investigación

La presente investigación tiene como objetivo conocer el nivel de autoeficacia y actitudes de un grupo de estudiantes del grado de Educación Primaria, respecto a la metodología basada en datos y el uso de las herramientas digitales de corpus.

La novedad de nuestro estudio radica, por tanto, en dos aspectos. Por un lado, explora la autoeficacia en un contexto de aprendizaje basado en datos. Esta metodología de implantación aún incipiente en el contexto de enseñanza de lengua extranjera hace uso de herramientas de corpus digitales como recurso de aula manipulable para el alumnado.

Por otro lado, el perfil de los participantes es concreto. La muestra está formada por estudiantes universitarios, futuros docentes de lengua extranjera en la etapa de educación primaria. Si bien es cierto que los estudios tanto de autoeficacia como de aprendizaje basado en datos se dan en contextos de enseñanza superior, el hecho de que sean futuros docentes de lengua extranjera puede aportar una visión particular, la cual no pueden proporcionar otras investigaciones cuya muestra es estudiantado de otros grados universitarios. Es decir, hablamos de docentes de lengua extranjera en formación, que podrán aplicar la metodología basada en datos y las herramientas de corpus en un futuro cercano. La opinión y actitudes que tengan sobre dicha metodología y recursos puede determinar el papel de los mismos en las aulas de lengua extranjera.

## 4 Metodología

### 4.1 Participantes

En el estudio participaron 33 estudiantes, de las que 20 eran mujeres (60,6 %) y 13 eran hombres (39,4 %) de último año del grado de Educación Primaria de la Universidad de Castilla-La Mancha. Todos ellos cursaban la mención en inglés como lengua extranjera, que los habilitaba para impartir

dicha asignatura en centros escolares de educación básica en España. Sus edades comprendían entre los 20 y los 25 años, excepto uno de los participantes que tenía más de 25 años. Todos ellos contaban con un nivel medio-alto o avanzado de inglés certificado, correspondiente a B2-C1 según el Consejo de Europa (2020). Todos eran hablantes nativos de la lengua española y no conocían ninguna lengua adicional a excepción del inglés. Ninguno de ellos contaba con experiencia previa en el manejo de herramientas digitales de corpus.

## 4.2 Herramientas de corpus y actividades

Para llevar a cabo el estudio se utilizaron dos herramientas de corpus: el British National Corpus (BNC) (2002) y el Sketch Engine for Language Learning (SKELL) (2018). El primero fue recopilado por la Oxford University Press entre los años ochenta y principios de los noventa. Contiene 100 millones de palabras de una amplia gama de géneros, como el lenguaje hablado, la ficción, las revistas, los medios de comunicación escritos y diversos textos académicos. Es de libre acceso, y sólo requiere registrarse en la herramienta. Sin embargo, la versión gratuita sólo permite 20 búsquedas al día y no permite comparar textos. No obstante, estas limitaciones no afectaron al desarrollo de esta investigación, puesto que las acciones formativas que se llevaron a cabo suponían un número de búsquedas menor al límite establecido por la herramienta.

En cuanto a SKELL, es una herramienta especialmente diseñada para el aprendizaje de lenguas extranjeras. Además del inglés, trata el alemán y el italiano, entre otros. Permite verificar fácilmente la búsqueda de palabras y expresiones. Para acceder a esta herramienta no fue necesario registrarse ni pagar. Contiene más de mil millones de palabras y 57 millones de frases. Así pues, el corpus no contiene documentos completos, sino frases ordenadas según la calidad del texto.

Ambas herramientas digitales ofrecen un buen ejemplo de cómo se utiliza el inglés en distintos contextos como puede ser el informal, el formal o el académico. Así, al teclear una palabra o expresión, estas herramientas ofrecen tres opciones: ejemplos reales en los que se ha utilizado esa palabra, colocaciones de esa palabra en distintas posiciones oracionales y sinónimos. Dichos corpus suponen un enorme potencial didáctico para el aula de lengua extranjera.

La Tabla 1 muestra las características de las dos herramientas de corpus utilizadas en esta investigación. Ambas herramientas son muy intuitivas y fáciles de manejar, especialmente para nativos digitales y personas con una competencia digital básica.

**Tabla 1.** Características principales de las herramientas de corpus.

Herramientas	Número de palabras	Tipo	Acceso
<i>British National Corpus</i> (BNC) <a href="http://www.english-corpora.org/bnc">www.english-corpora.org/bnc</a>	100 millones	Inglés general Varios géneros textuales	Se requiere registro Limitación diaria de búsquedas No permite comparación de textos
<i>Sketch Engine for Language Learning</i> (SKELL) <a href="https://skell.sketchengine.eu">https://skell.sketchengine.eu</a>	1000 millones	Inglés general	Sin registro Sin limitación de búsqueda

Fuente: Elaboración propia.

Los participantes utilizaron estas herramientas de corpus en cuatro actividades. En la primera actividad, a los participantes, de manera individual, se les pedía acceder a SKELL y buscar distintas colocaciones con los verbos *make* y *do*. Tras la búsqueda, se les animaba a pensar en el tipo de palabras que acompañan a ambos verbos y la función de esas palabras en una oración. La segunda actividad consistía en familiarizarse con una serie de expresiones idiomáticas relacionadas con partes del cuerpo. Se trabajaron cinco expresiones: *all ears*, *cold feet*, *cold shoulder*, *eagle eye* y *eye to eye*. En este caso los participantes realizaron la actividad en parejas. Se les pidió que accedieran a la herramienta del BNC y que hicieran la búsqueda de cada una de las expresiones idiomáticas, marcando la casilla *Context*. Las expresiones idiomáticas en el BNC aparecen contextualizadas en oraciones reales, y los



participantes tuvieron que inferir el significado de cada una de ellas atendiendo al contexto en el que aparecen. En segundo lugar, completaron una tabla donde debían escribir ejemplos de esas expresiones. Como último paso en esta actividad, se les pidió que usaran esas expresiones en un breve diálogo con su compañero de trabajo. Se estimó una duración de 20 minutos para completarla.

La tercera actividad versó sobre el estilo indirecto o *reported speech*, estaba diseñada para realizarse en 15 minutos, y está inspirada en una propuesta de Brown (2023). Los participantes utilizaron SKELL para llevar a cabo la búsqueda de *I told* y *I said* en la sección *Examples*. Tras observar el comportamiento de los dos verbos, en parejas, completaron un texto con algunas reglas y ejemplos de uso de esas expresiones en el estilo indirecto. Por último, la cuarta actividad consistió en diferenciar el significado de adjetivos/participios terminados en *-ed* o *-ing*. Esta actividad es una adaptación de una propuesta previa de Sutherland (2023) y se llevó a cabo a través de la herramienta del BNC, con una duración de 20 minutos. Se partió de la distinción entre *bored* y *boring*. Tras discutir el significado de ambas, se les pidió a los participantes que utilizaran el BNC para buscar ambos términos y observar qué sustantivos acompañaban a cada una de ellas. Tras esto categorizaron dichos sustantivos en “humanos”, “no-humanos” y “pronombres”, y discutieron con el resto de participantes qué tipo de elementos acompañan a *boring* y *bored* atendiendo a su clasificación. Como parte final de esta actividad, se les presentaron una serie de adjetivos/participios, (*annoying/annoyed*, *confusing/confused*, *exciting/excited*, *disappointing/disappointed*, entre otros), cuyo significado deberían debatir y reflexionar sobre ello, del mismo modo que lo habían hecho con *boring* y *bored*. La Tabla 2 recoge los datos principales de las cuatro actividades.

Tabla 2. Actividades realizadas.

Actividad	Herramienta	Descripción	Duración
Colocaciones con <i>make</i> y <i>do</i>	SKELL	Aprendizaje de distintas colocaciones con los verbos <i>make</i> y <i>do</i> y posterior reflexión sobre ciertas categorías de palabras	10 minutos
Expresiones idiomáticas	BNC	Aprendizaje de expresiones idiomáticas que contienen partes del cuerpo e inferencia de significado	20 minutos
Estilo indirecto	SKELL	Distinción y uso de <i>tell</i> y <i>say</i> en el estilo indirecto	15 minutos
-ed vs. -ing	BNC	Aprendizaje de la diferencia entre las terminaciones -ed e -ing en adjetivos/participios	20 minutos

Fuente: Elaboración propia.

### 4.3 Instrumento y procedimiento

Previamente a la intervención didáctica, se explicó a los participantes, brevemente y de manera clara, qué es un corpus, qué es la metodología ABD, y su utilidad en el aprendizaje de idiomas. También se presentaron las herramientas digitales de corpus de las que se iba a hacer uso, y unas pequeñas indicaciones de cómo manejarlas. Una vez los participantes se familiarizaron con las herramientas, se procedió a realizar las tareas propuestas. La intervención didáctica duró cuatro semanas. Cada semana los participantes trabajaron una actividad. Esta actividad se realizaba dentro del aula de lengua extranjera, y se integraba en la lección de ese día.

Una vez realizadas las actividades por parte de los participantes, se procedió a la recopilación de datos. Para ello se utilizó un cuestionario diseñado *ad hoc* que constaba de 14 ítems. Se tomaron como referencia los instrumentos utilizados por Chen, Flowerdew y Anthony (2019), Ming Huei Lin (2016) y Yoon e Hirvela (2004). Los participantes debían indicar su grado de acuerdo o desacuerdo en una escala Likert de seis puntos, donde 1 indicaba “muy en desacuerdo” y 6 “muy de acuerdo”. Todos los ítems se presentaron en castellano, puesto que queríamos asegurar que los participantes comprendían lo que se les planteaba.

El cuestionario comienza dando unas sencillas indicaciones y pidiendo a los participantes algunos

datos identificativos como edad, género, años estudiando inglés, y conocimiento de otras lenguas extranjeras. Los ítems están organizados en dos secciones: una relacionada con las actitudes sobre las herramientas de corpus y las actividades de aprendizaje basado en datos, y otra sección donde los ítems pretenden reflejar el nivel de autoeficacia de los participantes respecto a las herramientas de corpus y las actividades de aprendizaje basado en datos. El cuestionario hace referencia a aspectos sobre la dificultad de uso, el nivel de utilidad, o la aplicación fuera del aula de las herramientas digitales de corpus; también se pregunta a los participantes por las actividades de aprendizaje basado en datos para la adquisición de distintos aspectos de la lengua inglesa, además de compararlas con actividades tradicionales. El diseño y la distribución del cuestionario se realizó en formato electrónico a través de *Google Forms*. Además de las instrucciones escritas en el propio cuestionario, se explicó a los participantes oralmente cómo proceder. Se realizó en clase y se les informó de que las respuestas eran anónimas y de que no había límite de tiempo.

## 5 Resultados y discusión

La mayoría de los participantes afirmaron que trabajar con herramientas digitales de corpus y realizar actividades de aprendizaje basado en datos aumentaba sus expectativas respecto a su aprendizaje del inglés. Así, un 63,6 % se mostró, de alguna manera, de acuerdo al afirmar que sentían que las herramientas de corpus les hacían avanzar en su proceso de aprendizaje, mientras que, de manera similar, un porcentaje del 75,7 % comentó lo mismo respecto a las actividades de aprendizaje basado en datos. En general, el 93,9 % recomendaría el uso de herramientas de corpus para aprender inglés, incluso más de un 27,3 % afirmó que estaba no sólo de acuerdo sino muy de acuerdo con esta afirmación. Estas afirmaciones son positivas para el aprendizaje desde el punto de vista motivacional. De hecho Raooft, Bee y Swee (2012) argumentan que el creer que se es capaz de hacer una tarea aumenta significativamente las posibilidades de éxito en dicha tarea. De manera similar, tras un meta-análisis de los estudios más relevantes sobre la relación entre autoeficacia y nivel de lengua extranjera, Wang y Sun (2020) concluyeron que a mayor nivel de autoeficacia, mayor competencia lingüística.

Estas respuestas están en consonancia con el estudio de Ming Huei Lin (2016), donde tras la aplicación de la nueva metodología los participantes afirmaban haber mejorado en cuanto a sus creencias de lo que podían ser capaces de aprender y hacer en la lengua extranjera. Así, no se debe pasar por alto el papel de la autoeficacia y el autoconcepto del estudiante de cara al éxito en su proceso de aprendizaje. Investigaciones como la de Sun y Wang (2020) y Genç, Kuluşaklı y Aydin (2016) muestran cómo este tipo de factores afectivos pueden llegar a ser importantes predictores del desarrollo de la adquisición de una segunda lengua. Por tanto, el hecho de que la metodología basada en datos y las herramientas de corpus parezcan contribuir a aumentar el nivel de autoeficacia del estudiantado nos muestra el potencial de estas como elemento de refuerzo en el aprendizaje de la lengua extranjera.

Al preguntar por su capacidad de aprendizaje de inglés a través de esta nueva metodología y recursos, una amplia mayoría afirmó sentirse capaz de hacerlo y de tener éxito (94 %). De manera más específica, el cuestionario planteó afirmaciones respecto a la adquisición de vocabulario, gramática y expresiones idiomáticas en la lengua inglesa (los tres contenidos que se habían trabajado a través de corpus en el estudio). La mayoría se veía capaz de mejorar los tres aspectos gracias a dichos recursos digitales. Los participantes se vieron especialmente capaces de mejorar el aprendizaje de vocabulario (100 %), seguido de las expresiones idiomáticas (94 %) y por último la gramática (81,8 %). Estudios anteriores están en la línea de los resultados obtenidos aquí. Por ejemplo, Saeadakhtar, Bagerin y Abdi (2020) constataron un cambio de actitud hacia el aprendizaje de colocaciones entre estudiantes universitarios de inglés como lengua extranjera. Tras una intervención didáctica basada en el uso de recursos digitales de corpus, los participantes afirmaron tener más ánimo de seguir estudiando colocaciones a través de dicha metodología. Otras dos investigaciones llevadas a cabo por Yao (2019) y Sinha (2021) se centraron en el uso de herramientas de corpus para el aprendizaje de léxico en segunda lengua. También en estos casos los participantes mostraron una actitud muy positiva hacia este tipo de recursos. Por último, Charles (2014) o Guilquin (2021) realizaron encuestas de satisfacción y actitud tras aplicar herramientas de corpus para el aprendizaje de ciertas construcciones gramaticales.



Descubrieron que no solo la actitud era positiva sino que realmente esta metodología basada en datos y el corpus como recurso didáctico suponían una diferencia significativa en cuanto al aprendizaje de dichas estructuras gramaticales.

Los participantes valoraron ocho afirmaciones sobre la dificultad, utilidad, uso fuera y dentro de clase, entre otras, referidas a las herramientas digitales de corpus y la metodología de aprendizaje basado en datos. En cuanto a la dificultad del uso de las herramientas de corpus, más de la mitad de los participantes se mostraron en desacuerdo (54,4 %) con la afirmación de que las herramientas eran difíciles de utilizar. Nuestros resultados coinciden con estudios anteriores como el de Daskalovska (2015), que observó que los participantes de su estudio habían manejado las herramientas digitales de corpus con una destreza bastante alta, a pesar de que habían recibido una formación muy básica sobre ellas, como ocurre en este estudio. Incluso con ese grado mínimo de familiarización, la autora pudo comprobar que el estudiantado había sacado partido a las herramientas. Una de las razones podría ser, entre otras, la edad de los participantes. Tanto los del estudio de Daskalovska como los nuestros pertenecen a la generación de nativos digitales, los cuales de manera intuitiva se hacen en poco tiempo con todo tipo de recursos digitales.

Preguntados por si consideraban dichas herramientas más útiles que otros recursos que ya utilizaban, el 87,9 % estuvo de acuerdo en que sí. Además, manifestaron su deseo de seguir aprendiendo más sobre estas herramientas digitales de corpus el 90,9 %, de los cuales un 27,3 % respondió estar “muy de acuerdo” con esta afirmación. En cuanto a su opinión sobre las actividades de aprendizaje basado en datos, el 69,7 % afirmó no encontrarlas aburridas, además de preferirlas a actividades más tradicionales que se solían hacer en clase con una amplia mayoría del 87,8 %.

Por otro lado, deseaban seguir utilizando dichas herramientas tanto dentro como fuera de clase en la mayor parte de los casos. Preguntados sobre hasta qué punto estaban de acuerdo en utilizar de forma regular las herramientas de corpus en el aula, el 63,6 % respondió estar de acuerdo e incluso muy de acuerdo. En el caso de si les gustaría incorporar dichas herramientas a su aprendizaje fuera del aula, la respuesta también fue muy positiva con el 57,5 % de participantes respondiendo que estaban de acuerdo o muy de acuerdo con ello. Los datos aquí recabados son especialmente relevantes en el sentido de que estos mismos participantes van a ser docentes de inglés en un futuro próximo. Por tanto, el haber tenido una buena experiencia con las herramientas de corpus los puede animar a querer aplicarlas con sus propios estudiantes, dando la oportunidad a esos futuros alumnos a beneficiarse de los buenos resultados que la investigación comentada arriba ha constatado. Así, se puede contribuir a despejar las posibles dudas y recelos que este tipo de recurso y enfoque ha despertado entre los docentes, según estudios anteriores como los de Boulton y Cobb (2017), Çalışkan y Gönen (2018) o Poole (2022).

## 6 Conclusión

El presente estudio explora el nivel de autoeficacia y las actitudes de un grupo de profesorado en formación, que estudia inglés como lengua extranjera en el grado de Educación Primaria con respecto a dos herramientas digitales de corpus y a la metodología de aprendizaje basado en datos. Los resultados extraídos del cuestionario apuntan a una buena acogida tanto del corpus como de la metodología. Además, las respuestas aportadas muestran que el aprendizaje basado en datos y las herramientas digitales de corpus favorecen la autoeficacia en los estudiantes, demostrando predisposición, confianza y motivación en su proceso de aprendizaje de la lengua extranjera. Los participantes encontraron el corpus interesante e intuitivo, afirmando que aumentaba su grado de autonomía y seguridad, reconociendo su utilidad, y la facilidad en su manejo para seguir aplicándolos tanto en el aula como fuera de ella en el estudio autónomo.

Pese a su desconocimiento previo, el interés y la buena actitud expresados por los participantes hacia la metodología basada en corpus resulta alentador para el profesorado que desea incorporarla a sus clases de lengua extranjera. Por otro lado, si bien parece que los participantes se han adaptado a esta metodología, sí que es importante que el profesorado también esté preparado para ello. De esta manera, para que la metodología basada en datos y el corpus como recurso didáctico funcionen, debe existir una formación específica del profesorado, además de un diseño minucioso de las actividades

con un adecuado andamiaje. Esto solamente sería posible si el profesorado que opte por ello esté debidamente entrenado (Smart, 2014).

De cara a futuras investigaciones, se deberían abordar algunas limitaciones que presenta nuestro estudio. Por un lado, se contó con un grupo reducido de participantes que impiden realizar generalizaciones respecto a los resultados obtenidos. Para próximos trabajos se espera contar con un número mayor de participantes, donde se incluyan, además, distintitos niveles de competencia lingüística y otros contextos educativos como la educación básica. Por otro lado, también se debe tener en cuenta la corta duración de la intervención didáctica, si bien se decidió así para evitar el efecto de posibles factores como el abandono de algunos participantes.

## Referencias

- ANAM, Syafi'ul y STRACKE, Elke. Language learning strategies of Indonesian primary school students: In relation to self-efficacy beliefs. *System*, v. 60, p. 1-10, 2016. ISSN 0346-251x. DOI: <https://doi.org/10.1016/j.system.2016.05.001>. Disponible en: <https://doi.org/10.1016/j.system.2016.05.001>. Acceso en: 13 enero 2024.
- AŞIK, Asuman; VURAL, Arzu Şarlanoğlu y AKPINAR, Kadriye Dilek. Lexical awareness and development through data driven learning: Attitudes and beliefs of EFL learners. *Journal of Education and Training Studies*, v. 4, n. 3, p. 87-96, 2016. DOI: <https://doi.org/10.11114/jets.v4i3.1223>.
- BAI, Barry y WANG, Jing. The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, v. 27, n. 1, p. 207-228, 2023. DOI: [10.1177/1362168820933190](https://doi.org/10.1177/1362168820933190). Disponible en: <https://doi.org/10.1177/1362168820933190>. Acceso en: 10 enero 2024.
- BANDURA, Albert. *Social foundations of thought and action: A social cognitive theory*. Nueva Jersey: Prentice Hall, 1986.
- BERNARDINI, Silvia. Corpora in the classroom: An overview and some reflections on future developments. In: SINCLAIR, John (ed.). *How to Use Corpora in Language Teaching*. Amsterdam: Benjamins, 2004. p. 124-138.
- BOULTON, Alex. Testing the limits of data-driven learning: language proficiency and training. *ReCALL*, Cambridge University Press (CUP), v. 21, n. 11, p. 37-51, 2009.
- BOULTON, Alex. What data for data-driven learning? *Eurocall Review*, v. 20, n. 1, p. 23-27, 2012. Disponible en: <https://hal.science/hal-00734993>.
- BOULTON, Alex y COBB, Tom. Corpus use in language learning: A meta-analysis. *Language Learning*, v. 67, n. 2, p. 348-393, 2017. DOI: [10.1111/lang.12224](https://doi.org/10.1111/lang.12224). Disponible en: <https://doi.org/10.1111/lang.12224>. Acceso en: 5 enero 2024.
- BRITISH NATIONAL CORPUS (BNC). *British National Corpus: A Web-based interface to the British National Corpus*. 2002. Disponible en: <http://homepage.mac.com/bncweb/home.html>. Acceso en: 13 enero 2024.
- BROWN, Michael. Using concordance lines to teach participial adjectives. In: VIANA, Vander (ed.). *Teaching English with corpora: a resource book*. Londres: Routledge, 2023. p. 65-68.
- ÇALIŞKAN, Gamze y GÖNEN, Safiye İpek Kuru. Training teachers on corpus-based language pedagogy: Perceptions on vocabulary instruction. *Journal of Language and Linguistics Studies*, v. 14, n. 4, p. 190-210, 2018.
- CALLIES, Marcus. Towards corpus literacy in foreign language teacher education: Using corpora to examine the variability of reporting verbs in English. In: KREYER, Rolf; SCHAUB, Steffen y GÜLKERNING, Barbara Ann (ed.). *Angewandte Linguistik in Schule und Hochschule*. Frankfurt: Peter Lang, 2016. p. 391-415.

CHARLES, Maggie. Getting the corpus habit: EAP students' long-term use of personal corpora. *English for Specific Purposes*, v. 35, p. 30-40, 2014. ISSN 0889-4906. DOI: 10.1016/j.esp.2013.11.004. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0889490613000951>. Acceso en: 5 enero 2024.

CHEN, Meilin; FLOWERDEW, John y ANTHONY, Laurence. Introducing in-service English language teachers to data-driven learning for academic writing. *System*, v. 87, p. 1-11, 2019.

CONSEJO DE EUROPA. *Volumen complementario del Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación*. 2020. Disponible en: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco\\_complementario/mcer\\_volumen-complementario.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco_complementario/mcer_volumen-complementario.pdf). Acceso en: 3 enero 2024.

CROSTHWAITE, Peter; LUCIANA y SCHWEINBERGER, Martin. Voices from the periphery: Perceptions of Indonesian primary vs secondary pre-service teacher trainees about corpora and data-driven learning in the L2 English classroom. *Applied Corpus Linguistics*, v. 1, n. 1, pág. 100003, 2021. Consultado el: 10 de enero de 2024. ISSN 2666-7991. DOI: 10.1016/j.acorp.2021.100003. Disponible en: <https://doi.org/10.1016/j.acorp.2021.100003>.

DASKALOVSKA, Nina. Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, Routledge, v. 28, n. 2, p. 130-144, 2015.

DÖRNYEI, Zoltán. Towards a better understanding of the L2 learning experience, the Cinderella of the L2 motivational self system. *Studies in Second Language Learning and Teaching*, v. 9, n. 1, p. 19-30, 2019. DOI: 10.14746/sslt.2019.9.1.2. Disponible en: <http://dx.doi.org/10.14746/sslt.2019.9.1.2>. Acceso en: 3 enero 2024.

FLOWERDEW, Lynne. Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, v. 48, n. 1, p. 99-116, 2015. DOI: 10.1017/S0261444813000037. Disponible en: <http://doi.org.10.1017/S0261444813000037>. Acceso en: 15 enero 2024.

GABRIELATOS, Costas. Corpora and Language Teaching: Just a Fling or Wedding Bells? *Test-ej*, v. 8, n. 4, p. 1-35, 2005.

GELUSO, Joe y YAMAGUCHI, Atsumi. Discovering formulaic language through data driven learning: Student attitudes and efficacy. *ReCALL*, v. 26, n. 2, p. 225-242, 2014. DOI: 10.1017/S0958344014000044. Disponible en: <https://doi.org/10.1017/S0958344014000044>.

GENÇ, Gülten; KULUŞAKLI, Emine y AYDIN, Savas. Exploring EFL learners' perceived self-efficacy and beliefs on English language learning. *Australian Journal of Teacher Education*, v. 41, n. 2, p. 53-68, 2016. DOI: 10.14221/ajte.2016v41n2.4. Disponible en: <http://dx.doi.org/10.14221/ajte.2016v41n2.4>. Acceso en: 13 enero 2024.

GUILQUIN, Gaëtanelle. Using corpora to foster L2 construction learning: A data-driven learning experiment. *International Journal of Applied Linguistics*, v. 31, n. 2, p. 229-247, 2021. DOI: 10.1111/ijal.12317. Disponible en: <http://doi.org/10.1111/ijal.12317>. Acceso en: 10 enero 2024.

HANKS, Patrick. *Lexical analysis: Norms and exploitations*. Cambridge: MIT Press, 2013.

LI, Yusheng y WANG, Chuang. An empirical study of reading self-efficacy and the use of reading strategies in the Chinese EFL context. *Asian EFL Journal*, v. 12, n. 2, p. 144-162, 2010.

LIN, Lu-Chun. Learning word meanings from teachers' repeated story read-aloud in EFL primary classrooms. *English Language Teaching*, v. 7, n. 7, p. 68-81, 2014. DOI: 10.5539/elt.v7n7p68. Disponible en: <https://doi.org/10.5539/elt.v7n7p68>. Acceso en: 20 enero 2024.

LIN, Ming Huei. Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly*, v. 50, n. 4, p. 871-893, 2016.

LINNENBRINK, Elizabeth y PINTRICH, Paul. The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly*, Routledge, v. 19, n. 2, p. 119-137, 2003. DOI:

10.1080/10573560308223. Disponible en: <http://dx.doi.org/10.1080/10573560308223>. Acceso en: 2 enero 2024.

MAR HARO SOLER, María del. *¿Feedback y autorregulación del aprendizaje? Su impacto en las creencias de autoeficacia del estudiantado de traducción*. Edición: Chelo Vargas Sierra y Ana Belén Martínez López. Granada: Comares, 2021. p. 67-81.

MCENERY, Tony y XIAO, Richard. What corpora can offer in language teaching and learning. In: HINKEL, Eli (ed.). *Handbook of research in second language teaching and learning*. Reino Unido: Routledge, 2011. p. 364-380.

MEUNIER, Fanny. The pedagogical value of native and learner corpora in EFL grammar teaching. In: COMPUTER learner corpora, second language acquisition and foreign language teaching. Ámsterdam: John Benjamins, 2002. v. 12. p. 119-132. DOI: 10.1075/llt.6.10meu. Disponible en: <https://doi.org/10.1075/llt.6.10meu>. Acceso en: 12 enero 2024.

MÓNICA, chamorro Mejía. Actitudes lingüísticas, revisión teórica y repercusiones en la enseñanza de lenguas extranjeras. *Cuadernos de Lingüística Hispánica*, v. 37, p. 1-15, 2021. ISSN 0121-053x.

MUÑOZ, Mónica. Las actitudes y creencias lingüísticas de estudiantes universitarios. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, v. 9, n. 18, p. 1-20, 2019. DOI: 10.23913/ride.v9i18.444.

O'KEEFFE, Anne; MCCARTHY, Michael y CARTER, Ronald. *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press, 2007.

PÉREZ-PAREDES, Pascual; ORDOÑANA GUILLAMÓN, Carlos; VAN DE VYVER, Julie; MEURICE, Alice; AGUADO JIMÉNEZ, Pilar; CONOLE, Grainne y SÁNCHEZ HERNÁNDEZ, Purificación. Mobile data-driven language learning: Affordances and learners' perception. *System*, v. 84, p. 145-159, 2019.

POOLE, Robert. "Corpus can be tricky": revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, Routledge, v. 35, n. 7, p. 1620-1641, 2022. DOI: 10.1080/09588221.2020.1825095. Disponible en: <https://doi.org/10.1080/09588221.2020.1825095>. Acceso en: 10 enero 2024.

RAOOFI, Saeid; BEE, Tan y SWEE, Chan. Self-efficacy in Second/Foreign language learning contexts. *English Language Teaching*, v. 5, n. 11, p. 60-73, 2012.

RÖMER, Ute. Corpus research and practice: What help do teachers need and what can we offer? In: AIJMER, Karin (ed.). *Corpora and language teaching*. Ámsterdam: John Benjamins, 2009. p. 83-98.

SAEEDAKHTAR, Afsaneh; BAGERIN, Maryam y ABDI, Reza. The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System*, v. 91, p. 1-14, 2020. DOI: 10.1016/j.system.2020.102268. Disponible en: <https://doi.org/10.1016/j.system.2020.102268>. Acceso en: 12 enero 2024.

SAITO, Akihiro. Strategy Use, Self-efficacy Beliefs, and Self-regulatedness in Adult Foreign Language Learning. *Australian Journal of Applied Linguistics*, v. 3, n. 2, p. 152-167, 2020. DOI: 10.29140/ajal.v3n2.282. Disponible en: <https://doi.org/10.29140/ajal.v3n2.282>. Acceso en: 9 enero 2024.

SARDEGNA, Veronica; LEE, Juhee y KUSEY, Crystal. Self-efficacy, attitudes, and choice of strategies for English pronunciation learning. *Language Learning*, v. 68, n. 1, p. 83-114, 2018. DOI: 10.1111/lang.12263. Disponible en: <https://doi.org/10.1111/lang.12263>.

SINCLAIR, John. *How to use corpora in language teaching*. Ámsterdam: John Benjamins, 2004.

SINHA, Tara. EFL learners' perception of and attitude to corpus as a vocabulary learning tool. *The Reading Matrix: An International Online Journal*, v. 21, n. 2, p. 106-119, 2021.

SKETCH ENGINE FOR LANGUAGE LEARNING (SKELL). *Lexical Computing Ltd*. 2018. Disponible en: <https://www.sketchengine.eu/skell/>. Acceso en: 13 enero 2024.

- SMART, Jonathan. The role of guided induction in paper-based data-driven learning. *ReCALL*, v. 26, n. 2, p. 184-201, 2014.
- SUN, Ting y WANG, Chuang. College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, v. 90, p. 1-17, 2020. ISSN 0346-251x. DOI: 10.1016/j.system.2020.102221. Disponible en: <https://doi.org/10.1016/j.system.2020.102221>. Acceso en: 10 enero 2024.
- SZUDARSKI, Paweł. *Corpus linguistics for vocabulary: A guide for research*. Londres: Routledge, 2017.
- TENG, Lin Sophie. Individual differences in self-regulated learning: Exploring the nexus of motivational beliefs, self-efficacy, and SRL strategies in EFL writing. *Language Teaching Research*, v. 28, n. 2, p. 366-388, 2024. DOI: 10.1177/13621688211006881. Disponible en: <https://doi.org/10.1177/13621688211006881>. Acceso en: 8 enero 2024.
- TENG, Mark Feng; WANG, Chuang y WU, Junjie Gavin. Metacognitive Strategies, Language Learning Motivation, Self-Efficacy Belief, and English Achievement During Remote Learning: A Structural Equation Modelling Approach. *RELC Journal*, v. 54, n. 3, p. 648-666, 2023. DOI: 10.1177/00336882211040268. Disponible en: <https://doi.org/10.1177/00336882211040268>. Acceso en: 8 enero 2024.
- THOMAS, James Edward. Stealing a march on collocation: Deriving extended collocations from full text for student analysis and synthesis. In: LEŃKO-SZYMAŃSKA, Agnieszka y BOULTON, Alex (ed.). *Multiple affordances of language corpora for data-driven learning*. ámsterdam: John Benjamins, 2015. p. 85-108.
- TRIBBLE, Christopher. Teaching and language corpora: Perspectives from a personal journey. In: LEŃKO-SZYMAŃSKA, Agnieszka y BOULTON, Alex (ed.). *Multiple affordances of language corpora for data-driven learning*. ámsterdam: John Benjamins, 2015. p. 37-62.
- VILLALÓN, Ruth; MATEOS, Mar y CUEVAS, Isabel. High school students' beliefs and learning strategies in the EFL classroom: Are they affected by educational level? *Porta Linguarum*, v. 35, n. 6, p. 653-674, 2013.
- VYGOTSKY, Lev Semenovich. *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press, 1978.
- WADDINGTON, Julie. Developing primary school students' foreign language learner self-concept. *System*, v. 82, p. 39-49, 2019. ISSN 0346-251x. Disponible en: <https://doi.org/10.1016/j.system.2019.02.012>. Acceso en: 13 enero 2024.
- WANG, Chuang y SUN, Ting. Relationship between self-efficacy and language proficiency: A meta-analysis. *System*, v. 95, pág. 102366, 2020. ISSN 0346-251x. DOI: <https://doi.org/10.1016/j.system.2020.102366>. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0346251X20307260>. Acceso en: 5 enero 2024.
- WOODROW, Lindy. College English writing affect: Self-efficacy and anxiety. *System*, v. 39, n. 4, p. 510-522, 2011. ISSN 0346-251x. DOI: 10.1016/j.system.2011.10.017. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0346251X11001278>. Acceso en: 12 enero 2024.
- YAO, Gang. Vocabulary learning through data-driven learning in the context of Spanish as a foreign language. *Research in Corpus Linguistics*, v. 7, p. 18-46, 2019. DOI: 10.32714/ricl.07.02. Disponible en: <https://ricl.aelinco.es/index.php/ricl/article/view/98>. Acceso en: 5 enero 2024.
- YOON, Hyunsook e HIRVELA, Alan. ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, v. 13, p. 257-283, 2004. ISSN 1060-3743. DOI: 10.1016/j.jslw.2004.06.002. Disponible en: <https://www.sciencedirect.com/science/article/pii/S1060374304000190>. Acceso en: 15 enero 2024.
- YOU, Chenjing (Julia) y DÖRNYEI, Zoltán. Language learning motivation in China: Results of a large-scale stratified survey. *Applied Linguistics*, v. 37, n. 4, p. 495-516, 2014. ISSN 0142-6001. DOI: 10.1093/applin/amu046. Disponible en: <https://doi.org/10.1093/applin/amu046>. Acceso en: 9 enero 2024.
- ZAPATA-ROS, Miguel y PÉREZ-PAREDES, Pascual. *El pensamiento computacional, análisis de una competencia clave*. Scotts Valley: Create Space Publishing, 2018.