

EL USO DE CORPUS LINGÜÍSTICOS CON FINES PEDAGÓGICOS: EL CASO DE SEAH PROJECT¹

USING CORPORA IN LANGUAGE TEACHING: THE CASE OF SEAH PROJECT

CLAUDIA COLANTINO

Università degli Studi di Roma “La Sapienza”, Italia.

claudia.colantonio@uniroma1.it

Orcid: 0000-0003-4649-1266

RESUMEN

La lingüística de corpus aplicada a la enseñanza de lenguas extranjeras es un enfoque pedagógico que ha ganado terreno en las últimas décadas: el artículo se plantea analizar la aplicación de los corpus en la enseñanza de una lengua de especialidad, concretamente en el sector de la Arquitectura y de la Construcción. La falta de competencia en el lenguaje académico frecuentemente representa un obstáculo para la movilidad de estudiantes; además, puede imposibilitar la adquisición de conocimientos técnicos y profesionales durante su estancia al extranjero en el proyecto Erasmus. El artículo quiere integrar las aplicaciones prácticas de los corpus en el aprendizaje de la lengua de la Arquitectura y la Construcción, además de presentar el proyecto *Sharing European Architecture Heritage (SEAH)* y la creación de una unidad didáctica en el lenguaje académico-profesional en español.

Palabras clave: Corpus, arquitectura, construcción, unidad didáctica.

ABSTRACT

The corpus-based approach to linguistics and language education has gained prominence over the past decades: this article reports on applying the findings of corpus-based linguistics especially to the learning of a language for specific purposes (LSP). The lack of proficiency in academic language for the study of the specialist sector often constitutes an

¹ Este estudio se circunscribe al proyecto *Sharing European Architectural Heritage: innovative language teaching tools for academic and professional mobility in Architecture and Construction (SEAH)*, financiado por el Programa Erasmus + KA203 y bajo la coordinación del grupo de la Università degli Studi “D’Annunzio” de Chieti-Pescara (Italia) formado por Mariapia D’Angelo, Maria Chiara Ferro y Sara Piccioni.

obstacle to student mobility and may limit the acquisition of technical and professional knowledge during the period of the Erasmus stay in the host institution. The article outlines practical considerations for the integration of language corpora in the learning of the LSP in the field of Architecture and Construction (AC), presenting not only the *Sharing European Architecture Heritage (SEAH)* project, but also an academic-professional language module based on Spanish.

Keywords: Corpora, architecture, construction, module.

*Recibido:*10/09/2022. *Aceptado:* 10/05/2023.

1. INTRODUCCIÓN

La bibliografía más actualizada ya ha destacado la rentabilidad de los corpus en la enseñanza de lenguas extranjeras, como herramientas de investigación que pueden redundar en beneficio de la didáctica y el diseño de materiales (manuales, diccionarios, vocabularios) y como materiales didácticos por el amplio abanico de aportaciones pedagógicas que se pueden colegir, como comprender las combinaciones de la lengua (fraseología y colocaciones), descubrir patrones léxico-gramaticales o reconocer los valores de los términos polisémicos (Albelda Marco, 2011: 85).

El recurso a textos producidos en situaciones naturales y a datos auténticos ofrece a los aprendices una nueva manera de acercarse a la lengua extranjera y a la lengua de especialidad logrando beneficiarse del material recogido y anotado en el corpus construido en función del ámbito de especialidad elegido. En lo que sigue, se introduce el marco teórico del estudio, ilustrando la aplicación de los corpus a la enseñanza de lenguas extranjeras y las peculiaridades del español de la Arquitectura y de la Construcción; sigue una breve descripción del proyecto *SEAH*, en el que el presente estudio se enmarca con su arquitectura, antes de presentar la propuesta didáctica que se ha diseñado para poner de relieve algunas potencialidades de la aplicación del corpus en el aprendizaje del español del sector científico-profesional de AC.

2. MARCO TEÓRICO

2.1. Lingüística de corpus y su aplicación en la enseñanza/aprendizaje de una lengua extranjera

Las lenguas de especialidad constituyen un campo de investigación prometedor al tener un carácter abierto y fértil en el terreno del estudio de la lengua en sí y en

el terreno de la traducción. Un ámbito interesante es el estudio de un tecnolecto combinado con el empleo de corpus lingüísticos: el resultado puede ser realmente motivador para estudiantes de una lengua extranjera que quieran especializarse en un determinado campo del saber.

Antes de profundizar en el tema, resulta conveniente observar que los corpus cuentan ya con una sólida y secular tradición de uso en algunos ámbitos, como las ciencias jurídicas (Ezeiza Ramos, 2011) o los estudios literarios. Ahora bien, los estudios más recientes sobre la didáctica de las lenguas destacan una explotación didáctica cada vez más extendida en cuanto al empleo de corpus y el consecuente cambio que ha supuesto su uso en el aula, gracias a los avances tecnológicos. Efectivamente, el empleo de un corpus puede brindar el acceso a datos lingüísticos reales y auténticos de cara al desarrollo de la conciencia metalingüística de los aprendientes; por otro lado, el profesor puede diseñar actividades y ejercicios que puedan concurrir en la asimilación de los mecanismos de funcionamiento del sistema morfosintáctico y léxico-semántico de una lengua extranjera.

La lingüística de corpus no representa una rama de la lingüística, como la fonética o la sintaxis: para definirla recurrimos a dos citas del académico chileno Parodi que ponen de relieve sus rasgos definitorios y que brindan las peculiaridades de la aplicación de esa metodología. En primer lugar, se pone de relieve su “sustento a la investigación de la lengua en uso a partir de corpus lingüísticos con sustrato en tecnología computacional y programas informáticos ad hoc” (Parodi, 2008: 95).

[...] una metodología para la investigación de las lenguas y del lenguaje, la cual permite llevar a cabo investigaciones empíricas en contextos auténticos [...] Además presenta oportunidades revolucionarias para la descripción, análisis, y enseñanza de discursos de todo tipo. También brinda una base empírica para el desarrollo de materiales educativos y metodológicos de diversa índole [...] (Parodi, 2010: 14-15).

De ahí se desprende que la aplicación de corpus para la enseñanza y el aprendizaje de lenguas para fines específicos ha cobrado protagonismo en los últimos años en el ámbito de la investigación lingüística por su potencial pedagógico como recurso para observar las unidades terminológicas atendiendo a diversos parámetros, como los listados de frecuencia, las ocurrencias o las colocaciones.

Para desarrollar actividades con corpus para las clases de segundas lenguas, una de las metodologías más conocidas es el *Data Driven Learning* (DDL), a saber el aprendizaje guiado con datos (Johns, 1991). En inglés esta metodología cuenta con una sólida tradición (Bennett, 2010; Evison, 2010 por citar unos ejemplos); en español, se podrían mencionar Benavides (2015) y Cruz Piñol (2017), con lo cual la creación de unidades didácticas del proyecto SEAH resulta de gran envergadura en lo que a avances metodológicos y aplicativos se refiere. De todas formas,

el aprendizaje guiado con datos es una metodología realmente novedosa por “[...] placing the learner’s own discovery of grammar at the centre of language-learning, and by making it possible for that discovery to be based on evidence from authentic language use” (Johns, 1991: 3).

El empleo de corpus en el ámbito de la enseñanza y del aprendizaje de una lengua extranjera puede concurrir en el aumento de los conocimientos gramaticales y de la conciencia metalingüística de los aprendientes (Benavides, 2015; Boulton y Cobb, 2017), además de fomentar su autonomía en el aprendizaje. Ahora bien, se trata de un avance de los últimos años aunque su aplicación en el aprendizaje de lenguas extranjeras es mucho más reciente, tal y como apuntan Pitkowski y Vásquez Gamarra (2008: 32).

Para la enseñanza de lenguas, uno de los principales aportes de los corpus ha sido su contribución en el desarrollo de materiales (Marcos Miguel, 2020). Con el corpus se dispone de las estructuras comunicativas y gramaticales más empleadas por los hablantes (Cruz Piñol, 2017; O’Keeffe et al., 2007): efectivamente, contrariamente a los diccionarios, el corpus no ofrece información prescriptiva, por lo tanto se puede colegir otro tipo de información lingüística, es decir, categorías gramaticales específicas, como el uso de conjunciones, las colocaciones, los tiempos verbales, las locuciones preposicionales o los sintagmas nominales a partir de datos auténticos y reales (Ferreira et al., 2020: 77).

Una vez aclarado en qué consiste esta metodología, resulta importante también reflexionar sobre la esencia de un corpus lingüístico y sobre sus características principales: los corpus ofrecen al usuario (investigador, estudiante o cualquier usuario) la posibilidad de tener acceso a un conjunto amplio de textos (escritos y/u orales) digitales ya procesados, representativos de una lengua o de una variedad lingüística y que se han codificado para ser analizados posteriormente de manera científica (Rojo, 2021: 1; Sinclair, 1996).

Además, resulta necesario añadir a este cuadro algunos rasgos definitorios: en un corpus prima la selección de textos auténticos, según criterios establecidos previamente (a modo de ejemplificación, se menciona la modalidad de la lengua, la naturaleza física de los datos, la cobertura y la temática, el número de lenguas, la disponibilidad y el modo de recopilación), para otorgar un uso de la lengua que sea real, auténtico y actual. En términos generales, un corpus lingüístico no es un mero listado de palabras decontextualizadas; más bien recoge datos auténticos procedentes de textos que están almacenados en formato electrónico, lo que, por un lado, acelera la búsqueda de informaciones y, por otro, permite una actualización y una inserción constantes de los materiales. Supuestamente, el desarrollo en formato digital permite gozar de una gestión más autónoma por parte del usuario y un trabajo mejor organizado.

Asimismo, otro rasgo diferenciador es la representatividad, ya que en términos generales un corpus refleja una variedad lingüística o una lengua de especia-

lidad, según la finalidad que tenga: por ejemplo, el *Corpus del Español del Siglo XXI* (CORPES XXI) es un banco de datos general que representa el español del siglo XXI, es decir, la lengua en un eje sincrónico y diatópico ya que cuenta con un conjunto amplio de textos representativos del mundo hispánico en un preciso momento histórico. Otro ejemplo que recoge una variedad lingüística podría ser el *Corpus Diacrónico del Español* (CORDE) que está diseñado para extraer informaciones y datos sobre el español hablado desde sus inicios hasta el año 1974 y recoge textos de diferentes géneros (narrativos, religiosos, jurídicos entre muchos). Por otro lado, *WebLesp. Corpus de comunicación digital especializada en español* es un corpus que corre a cargo de Sara Piccioni y de Gianluca Pontandolfo (2021) y que reúne textos publicados en la web y representativos de la comunicación especializada y divulgativa en español en cuatro macroámbitos de conocimiento especializado: ciencia (el tema del cambio climático), derecho (divorcio en España), economía (macroeconomía española), medicina (cáncer de mama).

Al margen de lo arriba mencionado, es interesante notar que un corpus representa una herramienta de valiosa ayuda y a disposición de los estudiantes de una lengua extranjera o de una lengua de especialidad que les otorga también la autonomía de elegir por sí mismos qué aprender y cómo aprenderlo (Pérez-Ávila, 2007: 24), además de un cierto nivel de descubrimiento y de curiosidad ya que, tal y como señalan Gilquin y Granger (2010: 1), “learners are encouraged to observe corpus data, make hypotheses and formulate rules in order to gain insights into language (inductive approach) or check the validity of rules from their grammar or textbook (deductive approach). They become more involved, more active and more autonomous in the learning process”.

Un corpus tiene distintas funcionalidades: en primer lugar, la concordancia permite realizar búsquedas de palabras, etiquetas, frases, tipos de texto y propone los resultados contextualizados; la lista de palabras resulta útil a la hora de generar listas de frecuencia de varios tipos, como sustantivos, adjetivos o adverbios, y a la hora de generar listas de palabras que contienen determinados caracteres; el parámetro de las palabras clave permite extraer lemas que definen el contenido o el tema de un texto o de un corpus o, según las exigencias de los usuarios, comparar textos y corpus. De todas estas posibilidades se colige que aplicar el trabajo de los corpus a la didáctica de lenguas extranjeras puede redundar en beneficio de la adquisición de contenidos gramaticales, léxicos e históricos y de la ampliación del bagaje terminológico (Gabrielatos, 2005: 10), entre muchos. Ya McCullough (2001) destacaba la importancia de la aplicación de los corpus en el aula de ELE por su capacidad de completar las explicaciones de tipo cultural, léxico o gramatical con ejemplos reales de la lengua, por el aprendizaje de los contextos de uso de los lemas o expresiones, o por el estudio de variedades estilísticas y temáticas en las clases de literatura. Asimismo, O’Sullivan (2007:

277) destaca habilidades que define cruciales en lo que a DDL se refiere, como “predicting, observing, noticing, thinking, reasoning, analysing, interpreting, reflecting, exploring [...]”.

Por último, otra característica de los corpus es la finalidad concreta debido a criterios de diseño: los distintos tipos de corpus se diferencian por el tamaño o por la procedencia de los datos recogidos. En este sentido, por un lado se pueden destacar los corpus en bruto, formados casi exclusivamente por textos, y, por otro, los corpus anotados, como el *Corpus Oral Didáctico Anotado Lingüísticamente* (C-Or-DiAL), enriquecidos por informaciones lingüísticas (lema o categoría sintáctica, estructura lingüística) e informaciones no lingüísticas (autor, procedencia, año, argumento) (Rojo, 2021: 3).

Dependiendo de la finalidad o del tipo de textos analizados se pueden crear corpus con un diseño especial: generales, especializados, monolingües, bilingües o multilingües, escritos u orales. A modo de ilustración, el *Corpus Básico del Español de Chile* y el *Corpus Histórico del Español de México* (CHEM) que son representativos del español de Chile y de México (Pitkowski y Vásquez Gamarra, 2008: 33-34). Entre los muchos, en este artículo destacamos los corpus especializados, como *SEAH* (del cual se hablará detenidamente a continuación), que permiten obtener listas de vocabulario a partir de la frecuencia de uso, acercarse a una lengua de especialidad o incluso ayudar para la elaboración y la confección de materiales didácticos y programación de cursos.

2.2. Hacia una delimitación del español de la Arquitectura y de la Construcción

La arquitectura se concreta en la lengua de la Arquitectura y de la Construcción que, como cualquier tecnolecto o lengua de especialidad, comparte el sistema morfosintáctico y el léxico con la lengua común, aunque cabe observar que cuenta con rasgos diferenciadores propios. Ahora bien, esta variación interna depende de tres factores que Garofalo (2021: 104) identifica, en el ámbito del lenguaje jurídico, en la temática, en la situación de uso y en los rasgos lingüísticos diferenciadores, y que bien se podrían aplicar a cualquier lengua de especialidad.

El componente temático consta de los contenidos y los conocimientos propios de un campo del saber, en otras palabras, el conjunto de términos, lemas y palabras especializados y marcados por su univocidad, monorreferencialidad y monosemia, y representaría lo que Cortelazzo (1990) y Calvi (2006: 9-13) definen eje o dimensión horizontal. En nuestro caso se trataría de la perspectiva disciplinaria desde que se aborda el contenido de la Arquitectura y de la Construcción (diseño urbanístico, historia de la arquitectura, arquitectura verde, por poner algunos ejemplos). Por otro lado, la situación de uso denotaría el eje vertical e implicaría

una reflexión en torno a los factores pragmáticos y comunicativos de una intervención, a la relación entre los participantes y al contexto: en términos sencillos, el para quién o destinatario de la comunicación y el nivel de conocimiento del que tendría que disponer; el quién o rasgos del emisor que usa la lengua de especialidad y su nivel de especialización (especialista, semiespecialista, público lego); el dónde o contexto situacional en el que se emplea la lengua de AC.

Por último, los rasgos lingüísticos diferenciadores atañen al nivel léxico, al nivel morfosintáctico y al nivel estilístico de un campo del saber. A modo de ejemplificación del nivel estilístico, mencionamos el ámbito profesional y el ámbito académico dentro del sector de AC que concretan tipologías textuales con una estructuración y unos rasgos estilísticos propios: en cuanto al primero, destacamos dictámenes periciales, normativas y leyes, artículos periodísticos, redacción del proyecto básico y de ejecución; por lo que se refiere al ámbito académico sobresalen trabajos de fin de grado o de fin de máster, memorias, normativas, revistas y libros. Por lo general, es posible afirmar que estas tipologías textuales con carácter científico-técnico están destinadas a la comunicación entre especialistas del campo de AC, así como al intercambio de información con semiespecialistas o público lego en la materia.

En lo referente a la morfología, abundan rasgos que ‘repercuten en la articulación discursiva, puesto que en el discurso científico-técnico la construcción de un texto posibilita la transmisión de contenidos científicos y técnicos mediante una lengua especializada bajo un prisma léxico, sintáctico y de la configuración textual completa’ (Trovato, 2021: 92).

Atendiendo al ámbito léxico, el campo del saber de AC destaca por su interdisciplinariedad, puesto que su léxico es el resultado de varias influencias y de la combinación con disciplinas de índole distinta. Algunas expresiones proceden de diferentes especialidades (ingeniería civil, edificación, mecánica, química, informática, aeronáutica, telecomunicación, por citar unos ejemplos): el español de AC se apropia de términos de la medicina (se habla de *patología* en edificaciones para referirse a las lesiones que una sobrecarga podría producir en algún elemento de la edificación), de la biología (*intersticios*, *células de madera*, *hibridación* que alude a la puesta en crisis de parámetros arquitectónicos preestablecidos) o de la filosofía (epistemología del diseño arquitectónico, arquitectura holística, hermenéutica).

En términos generales, las palabras relacionadas con AC entran en la lengua española como préstamos; la mayoría pertenecen a la arquitectura clásica y proceden del latín (*claustró*, *arco*, *capilla*, *muro*, *pináculo*, *paramento*, *zócalo*, *torre*, *estípite*, *ventana*) y del griego (ábside, cornisa, nártex). Otras palabras derivan del árabe y se motivan por la duradera permanencia de los musulmanes en la Península Ibérica, como *azotea*, *zaguán*, *ajimez*, *mocárabe*, *tabique*, *cenefa*, entre muchas.

3. SEAH PROJECT: EL CORPUS PARA FINES PEDAGÓGICOS

Aclaradas algunas de las características y de las tipologías de los corpus, en este apartado se pasa a comentar el proyecto *Sharing European Architectural Heritage: innovative language teaching tools for academic and professional mobility in Architecture and Construction (SEAH)*, que nace con el propósito de abordar la lengua de especialidad de AC en un entorno didáctico y pedagógico.

Se trata de un proyecto que se propone tres objetivos: en primer lugar, la creación de un corpus lingüístico compuesto por textos representativos de la comunicación profesional y académica en el ámbito de la AC en cinco idiomas (alemán, español, francés, italiano y ruso); la segunda fase consiste en desarrollar módulos didácticos que puedan facilitar la adquisición de los principales aspectos gramaticales y léxicos de la lengua de AC. La última parte del proyecto prevé que el corpus y los módulos/cursos temáticos sean disponibles en línea y accesibles a través de una *Open Educational Resources Platform*, es decir, una plataforma abierta de interrogación que utiliza el programa *Sketch Engine*, lo cual posibilita el autoaprendizaje y la autonomía en el uso del corpus para resolver dudas concretas, profundizar en ellas o explorar fenómenos lingüísticos de interés por medio del acceso a materiales auténticos y reales.

El proyecto nace a partir de las necesidades de formación lingüística para desenvolverse en la lengua de especialidad del sector de la AC de los estudiantes que cursan estudios en una Universidad extranjera en el ámbito de un programa de intercambio (Piccioni, 2022), aunque su uso se podría extender también a profesores de lenguas extranjeras y a profesionales de la AC.

A continuación se presenta solo el apartado del corpus SEAH relacionado con la lengua española y se remite a Piccioni *et al.* (2021) y a D'Angelo M. (2021) para profundizar en los demás apartados que conforman el corpus, ya que se excederían los límites de espacio de la presente investigación.

El proyecto prevé la colaboración de Internet Web Solution, empresa española de programación web y posicionamiento orgánico en Google y distintos polos universitarios: Università degli Studi “Gabriele D’Annunzio” de Chieti – Pescara (Italia), Université Bordeaux Montaigne (Francia), Masaryk University (República Checa), Polotsk State University (Bielorrusia) y Universidad Politécnica de Madrid (España). Cada una de las universidades se ha encargado y ocupado de una parte del corpus en la lengua del lugar que hospeda la universidad (italiano, francés, alemán, ruso y español).

3.1. El corpus ES_SEAH

El corpus SEAH relacionado con la lengua española (ES_SEAH), exactamente como las secciones de las demás lenguas, se compone de textos auténticos y representativos de la AC seleccionados a través de una labor previa de *needs analysis* que ha establecido los ámbitos léxicos, temáticos y disciplinarios, además de los géneros textuales y los dominios (*domain* en la interfaz gráfica del corpus) a incluir. A modo de ilustración, los dominios elegidos son cuatro: ACCAD, es decir, el dominio académico que pone en relación investigadores y/o estudiantes con la comunidad científica; DID es el dominio que reúne textos pensados por investigadores o profesores para estudiantes; PROF aglutina textos creados por profesionales del sector científico-profesional de AC para otros profesionales y, finalmente, DIV reúne textos producidos por investigadores, profesores o profesionales para el público lego.

El foco del corpus ES_SEAH radica en cinco áreas temáticas que corresponden a las áreas léxicas que son objeto de los módulos lingüísticos desarrollados en la segunda parte del proyecto, a saber, renovación, edificios y obra civil, edificios no residenciales, edificios residenciales y materiales de construcción.

4. PROPUESTA DIDÁCTICA

4.1. Introducción

En este artículo se presenta una ejemplificación de unidad didáctica cuyos destinatarios serían estudiantes de español como lengua extranjera que quieran profundizar o especializarse en la lengua específica de la Arquitectura y de la Construcción.

La secuenciación que existe en la totalidad de las actividades que compone la siguiente unidad didáctica motiva al estudiante y fomenta su aprendizaje de la lengua extranjera, puesto que la reflexión meta e interlingüística cobran una especial relevancia (Úbeda Mansilla y Escribano Ortega, 2019: 144). Al mismo tiempo, la progresión de las actividades implica poner en práctica algunas de las destrezas y habilidades lingüísticas (a saber, la comprensión lectora y la comprensión auditiva), a la vez que se deben aclarar conceptos y terminología, lo que según Aguirre Beltrán (1998: 20) permite mejorar las capacidades y fomentar la comunicación de los aprendientes. Por último, cabe destacar que la resolución de ejercicios se basa en el progresivo empleo del corpus SEAH, que se convierte en una herramienta más a la hora de favorecer el conocimiento y la activación de mecanismos cognitivos, como observaciones sobre la lengua de la AC en el plano léxico (sinónimos, expresiones fijas, colocaciones, etc.) y en el plano morfosintáctico (D'Angelo, 2022: 14).

4.2. Descripción de la unidad didáctica

A partir de estas premisas, a modo de ejemplificación en la tabla que sigue se presenta la organización de la Unidad 1, Viviendas sostenibles. El objetivo del módulo es enriquecer el vocabulario técnico con ejercicios en línea sobre los eco-materiales y las propiedades de los mismos, además de reflexionar sobre la concordancia de género y número.

Tabla I. Estructuración de la primera unidad didáctica.

Unidad 1: Viviendas sostenibles (disponible en: https://www.seahproject.eu/see-training.php?id=10)	
Estructura	Contenido
1. ¿Qué es una vivienda sostenible?	
1.1. Para empezar: actividad con Quizlet.	Actividad lúdica o motivadora, como sopa de letras, puzzle, etc.
1.2. ¿Qué es una vivienda sostenible?	Comprensión lectora sobre las características de la vivienda, con autocorrección.
1.3. Responde a las siguientes preguntas	
1.4. Aquí tenemos cinco instalaciones básicas, tanto en una vivienda sostenible como en una convencional.	El ejercicio prevé colocar las imágenes en el orden correcto para que cada una coincida con la definición de la derecha.
2. ¿Qué es una casa pasiva?	
2.1. Conocemos lo que nos aporta una vivienda pasiva. Escucha el siguiente video y responde a las preguntas.	Comprensión oral.
2.2. Contesta a las siguientes preguntas con verdadero (V) o falso (F).	
3. Vocabulario y gramática.	
3.1. Usa el corpus ES_SEAH para observar con qué sustantivos se combinan las formas indicadas abajo.	Aplicación del corpus para averiguar cuáles son las concordancias con un lema.
3.2. Como demuestran los ejemplos sacados, en español los adjetivos concuerdan con el sustantivo al que acompañan, tanto en su género (masculino o femenino) como en su número (singular o plural).	Ejercicio gramatical sobre la concordancia de género y número de los adjetivos.
4. Juguemos con las palabras.	

Continuación Tabla I.

4.1. Juguemos con las palabras.	Actividad lúdica con crucigrama.
5. Comunicar.	Actividades relacionadas con las expresiones metafóricas del lenguaje de los arquitectos.
6. ¿Sabías que...?	Apartado de las curiosidades de la lengua, como las tres raíces latinas que componen la palabra sostenible o los 17 Objetivos del Desarrollo Sostenible.
7. Mi glosario.	Núcleo del módulo con un apartado de autoevaluación.
7.1. Ecomateriales: el diagrama muestra algunos de los ecomateriales más antiguos en la construcción. ¿Puedes señalar en el siguiente detalle constructivo algunos de estos materiales?	Ejercicio para relacionar las imágenes con las denominaciones.
7.2. ¿Conoces las palabras correspondientes a estos términos en tu idioma? Añádelas en la tercera columna.	Búsqueda de correspondencias en el idioma materno del usuario.
7.3. Usa el corpus SEAH para enriquecer tu vocabulario.	Actividades que prevén el empleo del corpus para descubrir más sobre materiales naturales y viviendas sostenibles.
7.4. ¿Cuántos tipos de piedra conoces?	Uso del corpus para comprobar las combinaciones con un lema seleccionado.
7.5. ¿Y qué propiedad tienen los demás materiales?	Se extiende el empleo del corpus a otros lemas (ej.: aluminio, ladrillo, vidrio).
7.6. ¿Cuántos tipos de madera conoces?	Se aplica otro criterio de interrogación para analizar otras concordancias.
7.7. ¿Cuántos tipos de paneles conoces?	Búsqueda con otro lema y correspondencia en el idioma materno del usuario.
8. ¿Qué has aprendido?	Se resumen los objetivos, las actividades complementarias, las palabras clave y las webs recomendadas.

Relativamente a la explotación didáctica del corpus ES_SEAH, a continuación se describen algunas actividades que cuentan con su aplicación según los principios del *data driven learning* (DDL), y que están pensadas tanto para estudiantes de una lengua extranjera de nivel intermedio/avanzado B1-B2, según las directrices del Marco Común Europeo de Referencia (MCER), como para profesores y profesionales del sector.

A modo de ejemplificación, en la unidad que se acaba de presentar, a los estudiantes se les pide, a través de un análisis guiado de concordancias (ejercicio 3.1., Figura 1), los sustantivos que se combinan con la palabra *visto* y también que los resultados se ordenen alfabéticamente según las palabras que la anteceden (a saber, el contexto de izquierda, o *left context*). Esta primera toma de contacto con el corpus resulta útil para la siguiente actividad que consiste en buscar las concordancias con los adjetivos *tosco* y *ligero*.

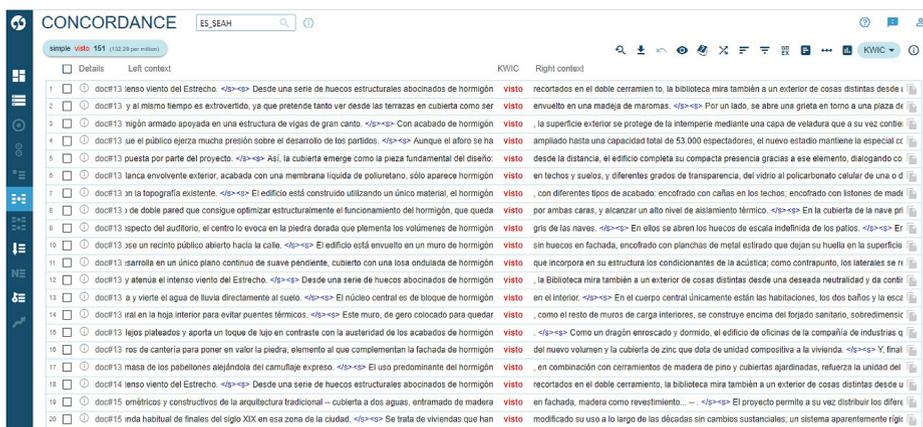


Figura 1. Interrogación para la palabra *visto* con la función *concordance*.

En otra actividad de la unidad didáctica (ejercicio 7.4., Figura 2) encaminada a enriquecer el vocabulario, a los estudiantes se les pide emplear la función *concordance* en su opción *advanced* y con la sintaxis CQL, que es un lenguaje de interrogación que permite analizar estructuras gramaticales y léxicas más complejas. Para que los estudiantes familiaricen progresivamente con el corpus, se les proporciona la cadena de búsqueda (o *query*) que les permite visualizar las concordancias del sustantivo *piedra* seguido de varios adjetivos, como *desigual*, *arenisca*, *caliza*, *blanca*, etc. Además, una vez que hayan terminado la primera tarea, deberán cambiar la palabra *piedra* por *ladrillo*, *vidrio* y *aluminio*, de manera que puedan ampliar el léxico relacionado con el área semántica de la unidad.

En otra actividad (ejercicio 7.6., Figura 3) se les pide a los estudiantes recuperar la secuencia *madera de* para descubrir cuáles son los árboles que producen madera: en este caso se pide emplear la sección de las concordancias que ofrece, entre muchas, la opción de búsqueda simple y una concordancia en orden alfabético según las palabras que siguen la secuencia (contexto de derecha o *right context*). Los resultados muestran distintos tipos de árboles, como *pino*, *roble*, *nogal*, *castaño*, etc.

seahproject.eu/see-training.php?id=10

1. PLEASE HELP US IMPROVE THIS SEAH TRAINING MODULE

Please rank the following aspects of the training materials on a scale from 1 to 5	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. The unit topic is relevant to me	<input type="radio"/>				
2. The training materials fitted my language level	<input type="radio"/>				
3. The assignments are clear	<input type="radio"/>				
4. The use of the SEAH corpus was useful for my learning	<input type="radio"/>				
5. I will keep using the SEAH corpus as a learning aide in the future	<input type="radio"/>				

PLEASE REPORT ANY ISSUES YOU FIND IN THE UNIT AND SPECIFY IN WHICH EXERCISE

[Send feedback](#)

Figura 4. Sección de *feedback* de la Unidad 1.

seahproject.eu/see-training.php?id=10

OTROS MATERIALES:

- Sobre edificación sostenible:
<https://blog.structuralia.com/caracteristicas-de-una-edificacion-sostenible-que-debes-conocer>
- En el siguiente vídeo puedes ver el ejemplo de Casa Martina, una casa autosuficiente desconectada de la red. Reflexiona sobre si es lo mismo una casa pasiva y una casa autosuficiente. <https://www.youtube.com/watch?v=KCTIakzjYI>
- Sobre los 17 Objetivos de Desarrollo Sostenible:
<https://www.un.org/sustainabledevelopment/es/sustainable-development-goals/>
- Sobre género y número en español, ejercicios:
<https://aprenderespanol.org/gramatica/masculino-femenino.html>
<https://aprenderespanol.org/gramatica/singular-plural.html>
- Transcripciones con lenguaje especializado de los dos vídeos de esta unidad (¿Qué es una casa pasiva? y Casa Martina):
<https://docs.google.com/document/d/15ILBKYiit4LDNvReadmciT7Bj9wml-UIP2h8dXKPS8/edit?usp=sharing>

En este documento, puedes consultar la terminología del fotovoltaico en español, inglés, italiano, catalán, rumano, francés, gallego y portugués: http://www.realter.net/wp-content/uploads/2013/06/DEF_Lessico_panlatino-2014-1.pdf

BIBLIOGRAFÍA:

- Diccionario Bilingüe de Metáforas y Metonimias Científico-Técnicas. Routledge. London & New York. 510 páginas. ISBN 978-1-138-86005-6 (hbk) ISBN: 978-1-315-71681-7 (ebk). Georgina Cuadrado Esclápez, Irina Argüelles Álvarez, Pilar Duran Escribano, María-José Gómez-Ortiz, Silvia Molina Plaza, Joana Pierce McMahon, María del Mar Robisco Martín, Ana Roldán Riejos, Paloma Úbeda Mansilla.
- Corpus SEAH <https://corpora.unich.it/seah/#open>

Figura 5. Materiales complementarios de la Unidad 1.

5. CONCLUSIONES

A tenor de lo anterior, las actividades que se han ilustrado permiten desprender las ventajas de la aplicación del corpus SEAH a la enseñanza de las lenguas extranjeras porque se compone de material auténtico y real de los hablantes nativos, no creado artificialmente y representativo de la comunicación profesional y académica en el ámbito de la Arquitectura y de la Construcción (Gilquin y Granger, 2010: 1); favorece el autoaprendizaje y la autonomía de los usuarios que, a través de interrogaciones distintas según sus exigencias, construyen el bagaje nocional necesario para comunicar en lengua extranjera; tomar en consideración el significado de las palabras no de forma aislada, sino en relación a su contexto de uso en textos reales; finalmente, favorece la indagación de los rasgos contextuales de lemas y expresiones en relación a un tema específico.

A esto, hay que añadir que la aplicación de los corpus en la enseñanza de una LE contribuye al desarrollo de la conciencia reflexiva acerca del funcionamiento lingüístico de la lengua misma (Mendikoetxea, 2014: 12), de manera que el aprendiente pueda interiorizar el comportamiento de cada componente del plano morfosintáctico (a saber, las reglas del funcionamiento del sistema gramatical de la lengua extranjera) y léxico-semántico (a modo de ejemplificación, el conocimiento de los campos semántico y léxico de algunas voces léxicas que tienen semejanzas y se prestan a otras posibilidades, como la polisemia, la antonimia o la sinonimia).

El presente artículo se enmarca dentro del ámbito pedagógico de la lingüística de corpus aplicada a la enseñanza de lenguas extranjeras, que representa un área de investigación novedosa, innovadora y fructuosa en el terreno de los estudios de la lengua. En la primera parte hemos esbozado el marco teórico y conceptual de la lingüística de corpus y posteriormente hemos procedido a ilustrar las características diferenciales de un corpus, como ser el contenedor de expresiones y palabras pertenecientes a un ámbito comunicativo y que se extraen de recursos digitales. Además, hemos propuesto una panorámica holística de las tipologías principales de corpus, como los generales, los multilingües, y se ha hecho hincapié en los especializados, como el corpus *SEAH*.

En la segunda parte nos hemos propuesto ofrecer una serie de propuestas y actividades enfocadas a la didáctica de la lengua de la Arquitectura y de la Construcción que pueden contribuir a desplegar las potencialidades de los corpus lingüísticos en la enseñanza de las lenguas extranjeras en general, y en la enseñanza del español como lengua extranjera más en concreto. Efectivamente, estamos en medio de un marco didáctico en ebullición, de una nueva línea de investigación que puede aportar notables beneficios que permitan mejorar la calidad de la enseñanza de una lengua de especialidad.

REFERENCIAS

- Aguirre Beltrán, Blanca. 1998. Enfoque, metodología y orientaciones didácticas de la enseñanza del español con fines específicos. *Carabela*, 44, 5-29.
- Albelda Marco, Marta. 2011. Rentabilidad de los corpus discursivos en la didáctica de lenguas extranjeras. En F. J. de Santiago-Guervós, H. Bongaerts, J. J. Sánchez Iglesias & M. Seseña Gómez, (coords.). *Del texto a la lengua: la aplicación de los textos a la enseñanza-aprendizaje del español L2-LE* (83-96). Alicante, España: Biblioteca Virtual Miguel de Cervantes.
- Benavides, Carlos. 2015. Using a Corpus in a 300-Level Spanish Grammar Course. *Foreign Language Annals*, 48(2), 218-235.
- Bennett, Gena. 2010. *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. Michigan, Estados Unidos: University of Michigan Press.
- Boulton, Alex y Cobb, Tom. 2017. Corpus Use in Language Learning. *Language Teaching*, 50(4), 348-393.
- Calvi, Maria Vittoria. 2006. *Lengua y comunicación en el español del turismo*. Madrid, España: Arco/Libros.
- Castillo Fadić, María Natalia. 2012. *Corpus básico del Español de Chile*.
- Cortelazzo, Michele. 1990. *Lingue speciali. La dimensione verticale*. Padova, Italia: Unipress.
- CORDE = Real Academia Española: Banco de datos. *Corpus diacrónico del español*. Disponible en <https://www.rae.es/banco-de-datos/corde>
- CORPES XXI = Real Academia Española: Banco de datos. *Corpus del Español del Siglo XXI*. Disponible en <https://www.rae.es/banco-de-datos/corpes-xxi>
- Cruz Piñol, Mar. 2017. *Lingüística de corpus y enseñanza del español como 2/L*, Madrid, España: Arco/Libros.
- D'Angelo, Mariapia. 2021. Discorso accademico-disciplinare in prospettiva glottodidattica: una proposta multilingue *corpus-based* per il linguaggio dell'architettura e delle costruzioni. En M. D'Angelo y M. Ožbot (eds.), *Lingue, testi e discorsi: Studi in onore di Paola Desideri* (pp. 347-361). Firenze, Italia: Franco Cesati Editore.
- D'Angelo, Mariapia. 2022. Risorse linguistiche *corpus based* per l'italiano dell'architettura e delle costruzioni, *Italiano a stranieri*, 32, 12-16.
- DLE/ASALE = Real Academia Española: *Diccionario de la lengua española*. Disponible en <https://dle.rae.es/>
- Evison, Jane. 2010. CARPER Corpus tutorial. Disponible en: <https://calper.la.psu.edu/publications-cpt/corpus-tutorial/>
- Ezeiza Ramos, Joseba. 2011. Plataforma Garalex: infraestructura tecnológica para la investigación y la didáctica de lenguaje del ámbito de las ciencias jurídicas. En M. L. Carrió Pastor & M. Á. Candel Moral (eds.), *Actas del III Congreso*

- Internacional de lingüística de corpus. Las tecnologías de la información y las comunicaciones: presente y futuro en el análisis de corpus* (pp. 683-694), Valencia, España: Editorial Universitat Politècnica de València.
- Ferreira, Anita, Elejalde, Jessica y Blanco, Lorena. 2020. Análisis de errores en el Corpus CAELE: estudio sobre la concordancia gramatical en el verbo “ser” en aprendientes francófonos y anglosajones. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 14(29), 76-99.
- Gabrielatos, Costas. 2005. Corpora and language teaching: just a fling or wedding bells?. *TESL-EJ. Teaching English as a Second or Foreign Language*, 8(4), 1-35.
- Garofalo, Giovanni. 2021. El lenguaje jurídico. En F. San Vicente & G. Bazzocchi (eds.), *Leti. Lengua española para traducir e interpretar* (pp. 103-126). Bologna, Italia: CLEUB.
- Gilquin, Gaëtanelle y Granger, Sylviane. 2010. How can data-driven learning be used in language teaching?. En A. O’Keeffe & M. McCarthy (eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359-370), London/New York, Inglaterra/Estados Unidos: Routledge.
- Johns, Tim. 1991. Should be persuaded: Two examples of data-driven learning. En T. Johns & P. King (eds.), *Classroom concordancing. English Language Research Journal* (pp. 1-16). Birmingham, England: Centre for English Language Studies, University of Birmingham.
- Marcos Miguel, Nausica. 2020. Aplicaciones para la didáctica de ELE del *Corpus del español, E-SEDLL*, 3, 32-44.
- McCullough, James. 2001. Los usos de los corpóra de textos en la enseñanza de lenguas. En M. Trenchs Parera (ed.), *Nuevas tecnologías para el autoaprendizaje de lenguas* (pp. 125-149). Lleida, España: Milenio.
- Mendikoetxea, Amaya. 2014. Corpus-based Research in Second Language Spanish. En Geeslin, k. L. (ed.), *The Handbook of Spanish Language Acquisition* (pp. 11-29), Oxford, Inglaterra: Wiley/Blackwell.
- O’Keeffe, Anne, McCarthy, Micheal y Carter, Ronald. 2007. *From corpus to classroom: Language use and language teaching*. Cambridge, England: Cambridge University Press.
- O’Sullivan, Íde. 2007. Enhancing a process-oriented approach to literacy and language learning: the role of corpus consultation literacy, *ReCALL*, 19(3), 1-35.
- Parodi, Giovanni. 2008. Lingüística de corpus: una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada*, 46(1), 93-119.
- Parodi, Giovanni. 2010. *Lingüística de corpus: De la teoría a la empiria*. Madrid/Frankfurt, España/Alemania: Iberoamericana/Vervuert.
- Pérez-Ávila, Elena A. 2007. El corpus lingüístico en la didáctica del léxico del español como LE. *Boletín de la Asociación para la Enseñanza del Español como Lengua Extranjera*, 37, 11-32.
- Piccioni, Sara. 2022. [En prensa] Planificación curricular basada en corpus para el

- aprendizaje del español de la Arquitectura y Construcción: el caso de las locuciones preposicionales en el corpus ES_SEAH, *Rassegna Italiana di Linguistica Applicata*.
- Piccioni, Sara y Pontrandolfo, Gianluca. 2021. WebLesp: corpus de comunicación digital especializada en español. Arquitectura, compilación y usos. *Revista de Lingüística Teórica y Aplicada (RLA)*, 59(1), 13-37.
- Piccioni, Sara, D'Angelo, Mariapia y Ferro, Maria Chiara. 2021. I Corpora SEAH di comunicazione specializzata nel settore dell'Architettura e delle Costruzioni. Struttura, compilazione e usi, *Linguistica*, 61(2), 99-125.
- Pitkowski, Elena Fabiana y Vásquez Gamarra, Javier. 2008. El uso de los corpus lingüísticos como herramienta pedagógica para la enseñanza y aprendizaje de ELE, *Tinkuy: Boletín de investigación y debate*, 11, 31-51.
- Pontrandolfo, Gianluca y Piccioni, Sara. 2022. *Comunicación especializada y divulgación en la red: aproximaciones basadas en corpus*. London/New York, Inglaterra/Estados Unidos: Routledge.
- Real Academia Española = *Diccionario panhispánico de dudas*. Disponible en <https://www.rae.es/dpd/>
- Rojo, Guillermo. 2021. *Introducción a la lingüística de corpus en español*. London/New York, Inglaterra/Estados Unidos: Routledge.
- Sinclair, John. 1996. *EAGLES. Preliminary Recommendations on Corpus Typology*. Disponible en <http://www.ilc.cnr.it/EAGLES96/corpusstyp/corpusstyp.html>
- Trovato, Giuseppe. 2021. El lenguaje científico-técnico. En F. San Vicente & G. Bazzocchi (eds.), *Leti. Lengua española para traducir e interpretar* (pp. 87-102). Bologna: Italia, CLEUB.
- Úbeda Mansilla, Paloma y Escribano Ortega, María Luisa. 2019. Activamos y adquirimos vocabulario, mejorando la pronunciación: actividades con alumnos de ELE en el ámbito de la arquitectura y la construcción, JEFEEVi-1. En P. Pérez Cañizares, L. Magnacco, I. Pérez Torres & J. Schnitzer (coords.), *Contribuciones seleccionadas de las Primeras Jornadas de Español con Fines Específicos* (pp. 141-155). Viena, Austria: Centro de Publicaciones. Ministerio de Educación y Formación Profesional.
- Universidad Nacional Autónoma de México. *Corpus Histórico del Español en México* (CHEM). Disponible en: <http://www.corpus.unam.mx/chem>
- Università di Firenze. *Corpus Oral Didáctico Anotado Lingüísticamente* (C-OrDiAL). Disponible en: <http://lablita.it/app/cordial/corpus.php>