

# LA PRAXIS DE LA ENSEÑANZA DE LA LENGUA A PARTIR DE LOS CORPUS TEXTUALES, ELEMENTO UNIFICADOR ENTRE TRADICIÓN E INNOVACIÓN EDUCATIVA<sup>1</sup>

---

ALONSO CAMPO, MARÍA ARACELI  
*Universidad de Valladolid*

FERNÁNDEZ TIJERO, MARÍA CARMEN  
*Universidad de Valladolid*

## 1. INTRODUCCIÓN

En la actualidad, las diversas sociedades exigen que la formación de los profesionales se base en la aplicación práctica de los saberes, pero no solo eso, sino que el ritmo frenético de avance y cambio que acontece pone como objetivo primordial de la educación actual la capacidad de adaptación, la versatilidad en todas las áreas de desarrollo humano. Esta necesidad de adaptabilidad, formación en destrezas y habilidades estratégicas es a lo que el sistema educativo intenta dar respuesta en sus continuas revisiones, especialmente en la última Ley Orgánica 3/2020, la LOMLOE, aprobado el 29 de diciembre de 2020, que supone un nuevo cambio de perspectiva del currículo: se centra en el desarrollo de las competencias y no en los contenidos —«saberes básicos», conforme a la LOMLOE.

---

<sup>1</sup> El presente trabajo se ha realizado en el marco del GIR Trans-REAL Lab, de la Facultad de Educación de Palencia, Universidad de Valladolid, al cual pertenecen ambas autoras.

Los nuevos requerimientos dan lugar a un nuevo enfoque por competencias, que va inevitablemente unido a un cambio metodológico, donde los contenidos no son el centro del diseño curricular. Ahora el centro de la programación lo constituye el desarrollo competencial global del estudiante. Establece las «situaciones de aprendizaje» como base del programa educativo, que hace del estudiante el agente de su propia formación, al gestionar los contenidos y habilidades de distintas áreas de forma conjunta y global para la resolución de problemas que se plantean como retos educativos.

De esta forma, el sistema educativo debe adaptar los materiales, formas de trabajo, equipamiento y formación de sus profesionales a la pragmática de esos saberes teórico-prácticos. Uno de los primeros escollos que los estudiantes deben superar es el manejo de fuentes de diversa índole en todas las áreas. El currículo, por otra parte, no contempla una formación específica para esta habilidad en ninguna de las áreas, por lo que se trata de formación estratégica que se debe abordar de forma interdisciplinar.

El área de Lengua castellana y Literatura, en concreto, se presenta como un espacio óptimo para el manejo de diverso tipo de herramientas de análisis textual que, aun sin suponer una formación específica destinada al manejo de fuentes, sí propicia una formación para el desarrollo de la competencia estratégica a la vez que se potencia la adquisición de las competencias lingüística, discursiva y sociocultural desde un enfoque comunicativo a partir del estudio de textos auténticos que permiten la observación del uso de la lengua en contextos reales de comunicación.

Específicamente, la Lingüística de corpus ofrece una serie de métodos y técnicas para la adquisición no solo de la competencia en comunicación lingüística sino de las otras siete competencias clave estipuladas en la LOMLOE.

## 2. OBJETIVOS

El presente trabajo tiene como objetivo principal mostrar las posibilidades que nos brinda la Lingüística de corpus para su aplicación a la enseñanza de la lengua española en entornos escolares.

### 2.1. OBJETIVO GENERAL

- La aplicación de corpus textuales en el aula de Educación Primaria como herramienta para el aprendizaje de la lengua española. El uso de corpus en los estudios lingüísticos ha sido tradicionalmente aplicado a identificar irregularidades, excepciones, cambios de uso..., en definitiva, un estudio de corte más sociolingüístico. Lo que se propone es que sean los propios estudiantes quienes manejen esos corpus integrándolos como una estrategia más de aprendizaje que podrán aplicar no solo a los contenidos del área de lengua, sino de forma transversal a todas las áreas de conocimiento.

### 2.2. OBJETIVOS ESPECÍFICOS

- Compilar un corpus de textos adecuados para los estudiantes de Educación Primaria. El corpus textual CORDALES\_Primaria será digital, multitextual, con diversas tipologías textuales, y multimodal, con múltiples funcionalidades que permita al estudiante analizar los datos desde diferentes perspectivas.
- Analizar el corpus con el fin de diseñar materiales didácticos enfocados al entorno escolar de Primaria, que faciliten un aprendizaje de la lengua a partir de textos reales, acorde, por tanto, con el enfoque comunicativo de enseñanza de las lenguas promovido desde la UE.
- Aplicar el uso de corpus en entornos escolares para el aprendizaje de la lengua, y en el caso específico del presente estudio, en el aspecto de Lengua y variación, de forma que los estudiantes puedan apreciar la evolución de la lengua, cómo van variando los términos, en función de qué factores, etc.

### 3. METODOLOGÍA

La aplicación de esta metodología basada en la compilación de corpus con fines didácticos y en entornos escolares se ha aplicado, con distintas funcionalidades, en apenas algún corpus como el desarrollado en el marco del proyecto BNClab – *Corpus for Schools. Corpus resources for A-level English Language and English Language Teaching*<sup>2</sup>, dirigido por la Dra. Dana Gablasova en el *ESRC-funded Centre for Corpus Approaches to Social Sciences* de la Universidad de Lancaster.

Este estudio propone la aplicación de este tipo de metodología al aula de Lengua castellana y Literatura de Educación Primaria con el objetivo de que el alumnado indague sobre cómo varía el lenguaje en función del lugar o región en la cual vivimos, el grupo social o la propia situación comunicativa, creando así también una conciencia social y ciudadana con base en los propios testimonios que encontramos en las fuentes de primera mano.

El Corpus didáctico de la lengua española en entornos escolares, *CORDALES\_Primaria*, incluye textos escritos de diferentes tipologías y género, los cuales se han agrupado conforme al género en tres subcorpus diferentes:

1. Manuales de texto de Educación Primaria
2. Textos literarios
3. Miscelánea: recetas, artículos de prensa, etc. Otros textos de tipo comunicativo susceptibles de su aplicación en Educación Primaria.

Tanto para la compilación como para la propia explotación del corpus se ha utilizado Sketch Engine<sup>3</sup> (Kilgarriff et ál. 2004 y 2014), una de las herramientas más populares actualmente para la consulta, creación y análisis de corpus, debido a su gran versatilidad, multifuncionalidad y fácil manejo para un usuario no experto<sup>4</sup>. Tal como se describe en la propia página de Sketch Engine, se trata de una aplicación informática

---

<sup>2</sup> Para más información, véase <https://wp.lancs.ac.uk/corpusforschools/bnclab/>.

<sup>3</sup> © Copyright - Lexical Computing CZ s.r.o.













<sup>4</sup> La herramienta Sketch Engine está disponible en <https://www.sketchengine.eu/>.

que permite tanto la consulta de más de 700 corpus existentes en más de 100 idiomas diferentes, entre ellos el español, como la compilación de corpus textuales *ad hoc* en una o varias lenguas, marcados y anotados morfológicamente de manera automática.

El «marcado» y «anotación morfológica» de un corpus hace referencia a la codificación de información adicional en el corpus mediante un proceso de análisis y etiquetación, que puede ser de naturaleza no lingüística o lingüística. Por «marcado» se entiende la codificación estructural del texto mientras que la codificación lingüística se conoce como «anotación». A su vez, las anotaciones lingüísticas pueden realizarse conforme a diferentes niveles: nivel morfológico (asociando lema y categoría a las formas), nivel sintáctico, nivel léxico-semántico, entre otros. El hecho de poder marcar y anotar morfológicamente de manera automática el corpus, lo que normalmente se conoce como «compilación del corpus», constituye una gran ventaja y aumenta enormemente las diferentes posibilidades de búsqueda en el corpus y, como consecuencia, de la explotación de los datos extraídos del mismo.

Una vez compilado el corpus, Sketch Engine integra múltiples herramientas para poder observar diferentes aspectos sobre el uso de la lengua, tal como se muestra en la siguiente pantalla.

**FIGURA 1.** Funciones integradas en Sketch Engine para el análisis del corpus.

 <b>Word Sketch</b> Collocations and word combinations	 <b>Word Sketch Difference</b> Compare collocations of two words
 <b>Thesaurus</b> Synonyms and similar words	 <b>Concordance</b> Examples of use in context
 <b>Parallel Concordance</b> Translation search	 <b>Wordlist</b> Frequency list
 <b>N-grams</b> Multword expressions (MWEs)	 <b>Keywords</b> Terminology extraction
 <b>Trends</b> Diachronic analysis, neologisms	 <b>Text type analysis</b> Statistics of the whole corpus
 <b>OneClick Dictionary</b> Automatic dictionary drafting	 <b>Bilingual terms</b> Bilingual terminology extraction

El objetivo, como ya se avanzó en el Capítulo 2, no se limita a obtener un resultado al realizar determinadas búsquedas, sino que se trata de un enfoque *corpus-driven*, en términos de Tognini-Bonelli (2001) en contraposición con el enfoque *corpus-based*. En este último caso, el usuario utiliza el corpus para comprobar o validar su teoría o presuposición mientras que desde una perspectiva *corpus-driven*, el usuario «se deja llevar» por el corpus. Conforme a la aclaración de Martín Herrero (2009), es el enfoque *corpus-driven* el que nos permitirá aproximarnos al estudio de la lengua participando activamente el proceso de aprendizaje:

De acuerdo con Tognini-Bonelli, el cometido del lingüista es acceder a la integridad de los datos del corpus como un todo, las descripciones serán el resultado de las evidencias dadas en el corpus. Es el corpus de origen, puesto que las teorías partirán de los datos del corpus, para después procesarse y constituirse gracias a él. Este es el punto de partida de la lingüística de corpus: el uso del corpus como un “*corpus-driven*”, puesto que es la forma de que el corpus participe activamente.

Todos los lingüistas de corpus insisten en la idea de la “mirada limpia” como única forma de trabajar “honestamente” con el corpus y recalcan que si aplicamos las teorías antes de ver el corpus no dejaremos que este nos dé, gracias a la aplicación de un software específico, toda la información que nos puede ofrecer (p. 1024).

Este enfoque «dirigido», en el que el alumnado también se involucra en el proceso de búsqueda, definiendo, redefiniendo, orientando su búsqueda, etc., coincide con el enfoque *data-driven learning*, conocido como DDL, propuesto precisamente desde la aplicación de la lingüística de corpus a la enseñanza de lenguas, sobre todo, de las lenguas extranjeras (Boulton, 2017 y 2021). Boulton y Tyne (2013) proponen una definición amplia de DDL, basándose en Gilquin y Granger (2010, p. 359, como se cita en Boulton y Tyne, 2013): «“*data-driven learning* (DDL) consists in using the tools and techniques of corpus linguistics for pedagogical purposes” – whatever the corpus linguistic tools and techniques, and whatever the pedagogical purposes.» (p. 98).

En nuestro caso, partimos de la base de un enfoque DDL para la elaboración de propuestas didácticas donde el alumno sea el actor de su propio proceso de aprendizaje y se deje dirigir por el propio corpus para fomentar un aprendizaje por descubrimiento de diferentes aspectos sobre su lengua y su entorno.

**FIGURA 2.** Fases para la elaboración de las situaciones de aprendizaje.



Fuente: elaboración propia mediante la herramienta Canva®

Se han definido en principio siete módulos diferentes para focalizar el aprendizaje sobre aspectos diferentes, como son:

- Lengua y evolución
- Lengua y variación
- Lengua y cultura
- Lengua y sociedad
- Lengua e ideología
- Lengua y literatura

- Otros: Lengua y ciencia, por ejemplo.

Para cada uno de estos módulos se pueden elaborar sencillas propuestas didácticas, organizadas en 5 fases:

1. Introspección: reflexión, el alumnado recapacita sobre un término de una forma guiada.
2. Descubrimiento: a partir de la propia búsqueda en el corpus el alumnado descubre por sí mismo posibles nuevos aspectos.
3. Análisis: ese descubrimiento le permite desarrollar el razonamiento metalingüístico.
4. Creación: puesta en acción de esos nuevos conocimientos que están adquiriendo acerca del uso y evolución de la lengua.
5. Reflexión: la reflexión final que ayuda a fijar nuevos conocimientos y adquirir las habilidades relativas a nuevas metodologías de trabajo de investigación-acción.

La idea reside en guiar al alumnado hacia un aprendizaje por descubrimiento, colaborativo y *corpus-driven* y, donde el estudiante aprenda diferentes destrezas y habilidades que faciliten tanto el desarrollo de competencias estratégicas e informacionales como de las competencias clave, prestando especial atención a la competencia en comunicación lingüística, desde una perspectiva interdisciplinar.

#### 4. RESULTADOS

Uno de los aspectos en los que se pone especial énfasis en la LOMLOE es el tema de la variación lingüística y el conocimiento de las variedades de la lengua. Este es uno de los motivos por los que se ha comenzado por este módulo, que es en el que nos basaremos para mostrar una propuesta didáctica a modo de ejemplo.

Teniendo en cuenta las fases especificadas en el apartado de Metodología, se organiza una secuencia didáctica en torno al módulo de Lengua y variación con la siguiente estructura, dividida en 5 bloques o pasos:

1. Bloque 1: Lluvia de ideas.
2. Bloque 2: Búsquedas.
3. Bloque 3: Análisis.



4. Bloque 4: Creación.
5. Bloque 5: Reflexión.

Si ponemos en relación las fases definidas en el apartado de Metodología para organizar la propuesta didáctica con los bloques o pasos que conforman la estructura de la secuencia para trabajar la lengua y su variación, la correspondencia quedaría de la siguiente manera: el bloque 01, «Lluvia de ideas», se corresponde con la fase de Introspección; los bloques 02 y 03, «Búsqueda» y «Análisis», con la fase de Descubrimiento; el bloque 04, con la fase de Creación; y, por último, el bloque 5, con la fase de Reflexión.

La secuencia didáctica propuesta se centra en la variación léxica y, específicamente, en el campo léxico de la alimentación. La organización por campos léxicos y no por campos semánticos se debe al hecho de tener en consideración diferentes categorías de palabras. Al igual que del Morral Barrigüete (2008, p. 188), aunque esta autora lo aplica a la enseñanza de lenguas extranjeras, nos interesan aquellas unidades léxicas no sólo relacionadas con el campo semántico al que pertenecen (la comida, por ejemplo), sino también el uso de las palabras que acompañan a cada unidad; es decir, sus estructuras gramaticales, su combinatoria, las expresiones o frases hechas, etc., que precisamente podremos observar a través del uso del corpus.

El primer bloque o paso es el punto de partida o detonador de la situación de aprendizaje. En primer lugar, se le preguntará al alumno sobre nombres de fruta que se digan de diferente manera en otra región o país a modo de «lluvia de ideas». Se les entregará un texto con nombres de fruta que podrían ser desconocidos para el alumnado, de manera que el estudiante tenga que identificar la variante e intentar deducir a qué fruta podría corresponder. El propio corpus nos puede servir como fuente para buscar un texto adecuado.

Una vez se pone en común el conocimiento previo que puedan tener sobre el léxico, se pasaría a la fase de Descubrimiento. Para ello, el alumno ha de utilizar Sketch Engine para poder realizar consultas en el corpus a partir de las diferentes funcionalidades que la herramienta nos ofrece, mostradas anteriormente en la Figura 1. En este bloque iremos

guiando al alumno para que conozca las herramientas al mismo tiempo que se les va haciendo preguntas sobre el uso de la palabra en contexto. Esta tarea puede desarrollarse por grupos, lo cual potenciaría el trabajo colaborativo, aunque se podría adaptar para el trabajo individual, dependiendo de las características del grupo-clase.

Por ejemplo, la primera búsqueda podría realizarse a través de la herramienta *Concordance* (concordancia, en español):

**FIGURA 3.** *Concordancias de la unidad léxica «durazno».*

o de banana 30 Licuado de manzana y **durazno** 31 Pastel de calabaza 32 Pollo frutado con pur  
 1 cucharada sopera de mermelada de **durazno** o naranja.</s><s>Preparación " Pelar la banan  
 de manzana Papilla yde pollo y zapallo **durazno** " a partir de "\*\*\*\*\* \*\*\*\*\* \*\*\*\*\* 7 MESES 6 Nunca  
 MINAS Y MINERALES Ingredientes " 1 **durazno** chico.</s><s>" 1 manzana chica.</s><s>" 1 cu  
 y retirar las semillas.</s><s>" Pelar el **durazno** y retirar el carozo.</s><s>" Introducir las frutas  
 \* \* \* \* FRUTAS Manzana Banana Pera **Durazno** Palta Mamón Todas las frutas \* \* \* \* \* \* \* \* página  
 le pimienta inglesa, opcional Postre de **duraznos** y jengibre Me gusta 1.</s><s>Vierte 2/3 taza d  
 era con el disco para rebanar.</s><s>2 **duraznos** maduros medianos, pelados y sin hueso 2 cucl  
 la tapa.</s><s>2.</s><s>R ebana los **duraznos** al vaporizador.</s><s>Retira la tapa y el disco  
 a el aspa para moler/picar y agrega los **duraznos** , la avena y el jengibre (si se desea) al tazón.</  
 ediana madura, pelada y sin semillas 1 **durazno** mediano maduro, pelado y sin hueso 1.</s><s>  
 >R ebana los chabacanos, la pera y el **durazno** en la canasta vaporera.</s><s>Retira la tapa y  
 r y agrega los chabacanos, la pera y el **durazno** al tazón.</s><s>Mezcla agregando suficiente li  
 ñ R E C E T A S ( E T A P A 3) Me gusta 2 **duraznos** medianos maduros, pelados y sin hueso ¼ de t  
 aporera.</s><s>2.</s><s>R ebana los **duraznos** en la canasta vaporera.</s><s>Retira la tapa y  
 a el aspa para moler/picar y agrega los **duraznos** y Ricotta al tazón.</s><s>Mezcla agregando si  
 sas de uvas, frutos secos, orejones de **durazno** o frutas brillantadas.</s><s>Se puede sustitui

Fuente: www.sketchengine.eu (© Copyright – Lexical Computing CZ s.r.o.)

Esta aplicación permite visualizar ejemplos de la unidad léxica en su contexto, en formato KWIC —*Keyword in Context*—, de manera que se muestra las palabras que van a la izquierda y a la derecha de la unidad léxica buscada. Si clicamos en la unidad léxica también se podría visualizar un mayor contexto, en caso de necesidad. La versatilidad de la

herramienta ofrece múltiples posibilidades para la visualización de los datos.

Asimismo, se puede guiar al alumnado para que utilicen la aplicación Wordlist (listado de palabras, en español), que extrae un listado de las palabras más frecuentes del corpus.

**FIGURA 4.** *Listado de palabras ordenadas por frecuencia.*

Lemma	Lemma	Lemma	Lemma
1 ingredientes ...	14 kcal ...	27 trocear ...	40 desinfectar ...
2 rebanada ...	15 vaporera ...	28 ecorecetario ...	41 uras ...
3 cucharadita ...	16 proteínas ...	29 tazón ...	42 rebanar ...
4 minipimer ...	17 cucharada ...	30 zanahoria ...	43 socioeducativa ...
5 porciones ...	18 gr ...	31 recetas ...	44 licuar ...
6 frijoles ...	19 verd ...	32 calabacín ...	45 tortilla ...
7 sopero ...	20 procedimiento ...	33 espinaca ...	46 acelga ...
8 braun ...	21 puré ...	34 cacerola ...	47 pimiento ...
9 jitomate ...	22 cebolla ...	35 perejil ...	48 hervir ...
10 g ...	23 frijol ...	36 rehogar ...	49 batata ...
11 hidratos ...	24 ml ...	37 batidor ...	50 ejote ...
12 lípidos ...	25 licue ...	38 rallar ...	
13 taza ...	26 lave ...	39 ajo ...	

Fuente: [www.sketchengine.eu](http://www.sketchengine.eu) (© Copyright – Lexical Computing CZ s.r.o.)

El listado nos abre la puerta a indagar sobre otros alimentos que podrían no ser reconocidos por el estudiante como, por ejemplo, el caso de «jitomate», «ejote» o «batata», cuyo uso podría ser desconocido para un alumnado de la variante peninsular del español. También podemos detectar unidades léxicas referentes a utensilios para cocinar, como «vaporera» o «minipimer», sustantivos acotadores como «rebanada» o «cucharada», o verbos para indicar acciones que se realizan con los alimentos, como «rebanar», «trocear» o «rallar». Este último nos puede llevar a aspectos ortográficos como la diferencia de «rallar» y «rayar». Igualmente, podemos ver usos morfológicos para distinguir medidas o

cantidades como «taza» o «tazón». El corpus nos dirige hacia diversidad de usos y la combinación de la aplicación de Listado de palabras junto con las Concordancias ofrece un enriquecimiento sustancial sobre el uso de las unidades léxicas que sería difícil de conseguir mediante la introspección.

Otra aplicación parecida a WordList es Keywords (palabras clave, en español), la cual extrae tanto las unidades simples como las unidades sintagmáticas (multipalabras) principales, es decir, aquellas nucleares y que se corresponden con categorías léxicas (sustantivos, adjetivos, verbos y adverbios, principalmente).

Si queremos ir más allá, la aplicación de Word Sketch nos facilita la observación de las colocaciones y la combinatoria léxica de una palabra. Word Sketch podría decirse que es una de las funcionalidades que distingue la herramienta de Sketch Engine de otras aplicaciones de compilación y explotación de corpus, ya que nos extrae las colocaciones no solo más frecuentes sino las más significativas, agrupadas por la función gramatical de la unidad léxica en relación con el colocado, entendiendo por «colocado» la palabra que acompaña a la unidad léxica nuclear que conforma la colocación. Por ejemplo, si buscamos los *word sketches* de la unidad léxica «durazno», tal como se muestra en la Figura 5, podemos observar qué adjetivos modifican protípicamente a la unidad «durazno», por ejemplo, «maduro» o «mediano», qué verbos acompañan a «durazno» como objeto, como, por ejemplo, «rebanar» o «pelar», qué unidades léxicas van con «durazno» en la construcción [durazno+y/o+sustantivo], como, por ejemplo, «durazno y Ricotta», «mango o durazno», qué unidades léxicas van con «durazno» en la construcción [sustantivo+de+durazno], como, por ejemplo, «avena de durazno» o «orejón de durazno», etc.

FIGURA 5. Word sketches de «durazno».

modifiers of "durazno"	verbs with "durazno" as object	"durazno" and/or ...	... de "durazno"
<b>maduro</b> ... duraznos maduros	<b>rebanar</b> ... duraznos rebanados	<b>Ricotta</b> ... duraznos y Ricotta	<b>Avena</b> ... Avena de durazno
<b>mediano</b> ... duraznos maduros medianos	<b>pelar</b> ... Pelar el durazno	<b>mango</b> ... mango o durazno	<b>orejón</b> ... orejones de durazno
	<b>agregar</b> ... picar y agrega los duraznos	<b>jengibre</b> ... duraznos y jengibre	<b>Postre</b> ... Postre de duraznos
	<b>cortar</b> ... cortar los duraznos	<b>pera</b> ... chabacanos , la pera y el durazno	<b>mermelada</b> ... mermelada de durazno
	<b>poner</b> ... Pon los duraznos	<b>jugo</b> ... duraznos y el jugo	
		<b>naranja</b> ... durazno o naranja	
		<b>manzana</b> ... manzana y durazno	
		<b>fruto</b> ... durazno o frutas	

Fuente: [www.sketchengine.eu](http://www.sketchengine.eu) (© Copyright – Lexical Computing CZ s.r.o.)

Los Word Sketches también se pueden observar de manera más visual mediante un gráfico como el que se muestra en la siguiente Figura.

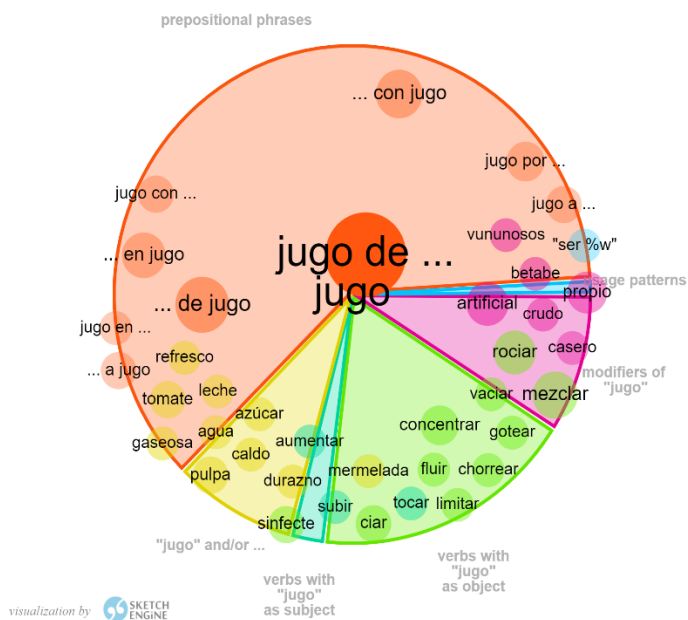
FIGURA 6. Gráfico con los word sketches de «durazno».



Fuente: [www.sketchengine.eu](http://www.sketchengine.eu) (© Copyright – Lexical Computing CZ s.r.o.)

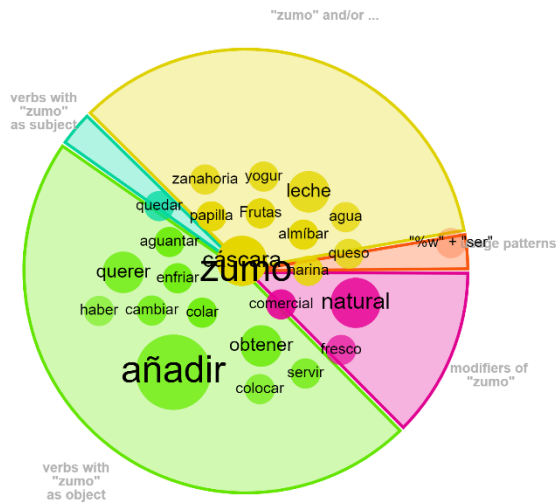
Entre las unidades léxicas que acompañan a «durazno» en la construcción [sustantivo+de+durazno], podemos guiar a los alumnos para que se fijen en la unidad léxica «jugo» y que descubran la diferencia, por ejemplo, entre «jugo» y «zum».

**FIGURA 7.** Gráfico con los word sketches de «jugo» en Sketch Engine.



Fuente: [www.sketchengine.eu](http://www.sketchengine.eu) (© Copyright – Lexical Computing CZ s.r.o.)

**FIGURA 8.** Gráfico con los word sketches de «zumo».



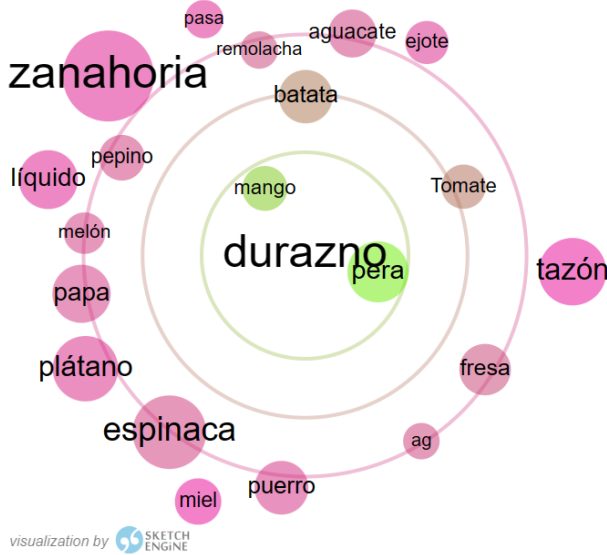
visualization by  SKETCH ENGINE

Fuente: [www.sketchengine.eu](http://www.sketchengine.eu) (© Copyright – Lexical Computing CZ s.r.o.)

La observación de «durazno» nos conduce a la aplicación Thesaurus (tesaruo, en español), la cual nos permite detectar otras unidades léxicas que funcionan de manera similar a «durazno».



FIGURA 8. Palabras relacionadas con «durazno».



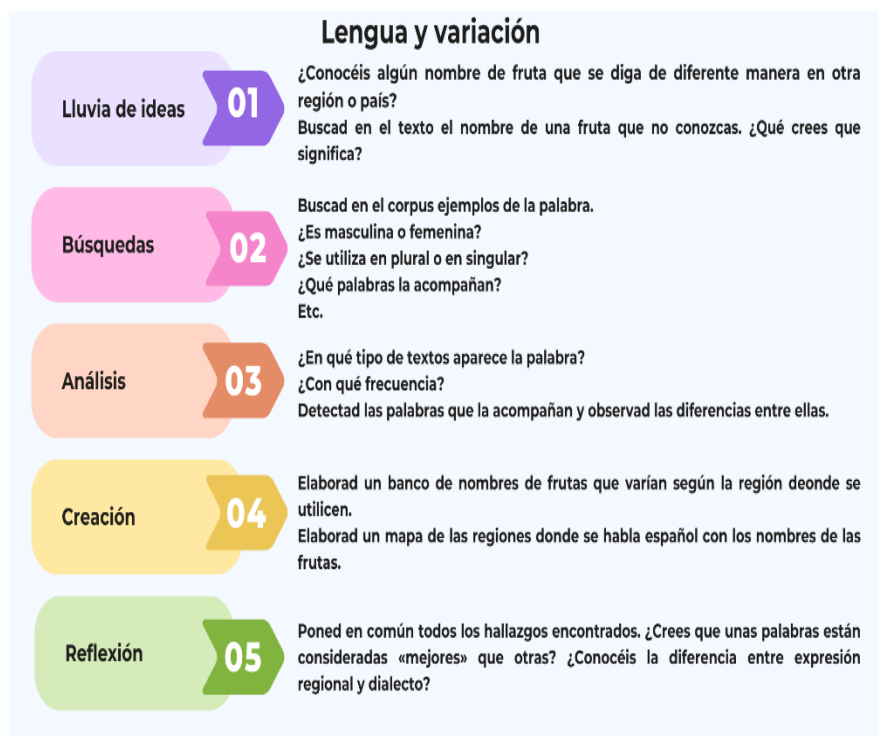
Fuente: [www.sketchengine.eu](http://www.sketchengine.eu) (© Copyright – Lexical Computing CZ s.r.o.)

La exploración del corpus nos lleva totalmente a un aprendizaje *data-driven*, donde el alumnado por grupos va descubriendo el entorno de las unidades léxicas no basándonos solo en aspectos gramaticales sino también en culturales, sociales, etc.

Una vez que finalizamos con los bloques o pasos 2 y 3 llega el momento de llevar a la acción todo lo aprendido. En el bloque o paso 4, el de Creación, podemos sugerir alguna actividad donde los estudiantes hayan de generar o crear un recurso, como, por ejemplo, un banco de nombres de frutas o un mapa de las regiones donde se habla español con los nombres de las frutas.

Por último, es importante terminar mediante una fase de Reflexión, el bloque o paso 5, con el fin de poner en común los hallazgos encontrados y autoevaluar su proceso de aprendizaje. En la Figura 9 se muestra a modo de resumen la secuencia didáctica planteada.

**FIGURA 9.** Ejemplo de secuencia didáctica correspondiente al módulo de Lengua y variación.



Fuente: elaboración propia mediante la herramienta Canva®

Se ha de tener en cuenta que el enfoque *corpus-driven* no permite conocer de antemano todos los resultados de las posibles búsquedas, pues estas son infinitas, por lo que el aprendizaje es un descubrimiento tanto para el alumno como para el propio profesor. Esto implica que el profesor ha de ir guiando, sugiriendo al alumno posibilidades, pero es el alumno el que dirige su propio descubrimiento sobre el uso. No obstante, sí es aconsejable que el docente realice búsquedas previstas para poder lanzar preguntas, sugerir vías o caminos para la indagación, etc.

## 5. DISCUSIÓN

La Lingüística de corpus con fines didácticos no supone un ámbito totalmente nuevo, pues la aplicación de los corpus textuales al

aprendizaje de lenguas se viene desarrollando desde hace ya tiempo, aunque la mayoría de los estudios se centran en el aprendizaje de lenguas extranjeras. En el ámbito de la didáctica de lenguas extranjeras se han desarrollado principalmente corpus que recopilan textos escritos u orales de producciones de los estudiantes con el fin de detectar y analizar los errores frecuentemente cometidos. A estos corpus se les ha denominado «corpus de aprendizaje», como equivalente del término inglés *learner corpora*. Otros autores (Núñez Nogueroles, 2019) se refieren a ellos como «corpus de aprendices» y se han aplicado de cara a la enseñanza de ELE.

En entornos escolares su aplicación es más reciente, y son escasos los corpus con fines didácticos para su aplicación en edades más tempranas. La propuesta que presentamos se aplica al desarrollo de materiales didácticos, al igual que en el caso de los corpus de aprendizaje, pero la diferencia radica en las fuentes que alimentan el corpus. En nuestro caso, el corpus se configura a partir de textos reales que utilizan los estudiantes en un entorno escolar y no a partir de sus propias producciones. Es por ello que creemos necesaria una clara distinción entre los corpus escolares, como sería el caso del corpus *CORDALES\_Primary* y los corpus de aprendizaje.

Tal como hemos pretendido mostrar en estudio, que constituye una primera aproximación a la aplicación de la Lingüística de corpus en entornos escolares y, específicamente, en Educación Primaria, la adaptación de los corpus textuales para que el alumnado pueda indagar sobre el funcionamiento del lenguaje y cómo este nos permite categorizar una realidad, cambiarla o moldearla permite diseñar verdaderas propuestas didácticas enfocadas al descubrimiento de la lengua y de su uso en situaciones de comunicación real<sup>5</sup>.

A través de la herramienta Sketch Engine, los estudiantes pueden recapacitar sobre diversas cuestiones relacionadas no solo con el significado de las palabras, su variación en el uso, sino las construcciones gramaticales, contextos o variaciones diastráticas..., lo que potencia su

---

<sup>5</sup> Cf. McEnery y Wilson. 2001.

reflexión metalingüística y favorece, por ende, el manejo de las construcciones lingüísticas y mejora el pensamiento crítico.

Estas búsquedas requieren en todo momento de la acción de guía por parte del docente, cumpliendo así los objetivos de autogestión del aprendizaje y autonomía en el proceso de formación que las nuevas disposiciones reguladoras educativas requieren.

Otro de los aspectos positivos que resultan de la utilización de corpus escolares es la alta capacidad de atracción de atención que es inherente a estas herramientas. En nuestro caso, Sketch Engine proporciona gráficos y cuadros sinópticos que hacen que resulte más sencillo gestionar y comprender conceptos con múltiples variables.

Si bien las ventajas de los corpus para fines didácticos de cara a un mayor desarrollo de la competencia en comunicación del hablante se hacen evidentes, se ha de tener en cuenta que también se ha de hacer frente a ciertas dificultades o limitaciones que pueden surgir. El alumnado de Primaria está en proceso de desarrollo de la competencia digital y una herramienta como Sketch Engine no está adaptada al entorno escolar. Quizás en un futuro se podría implementar esta adaptación, pero se ha de tener en cuenta el apoyo constante que necesitará el alumnado para conocer el uso de las diferentes funcionalidades.

Por otro lado, el enfoque DDL supone un reto para el docente, pues no tiene un control previo del camino que va a seguir el alumno en sus búsquedas, por lo que necesita un conocimiento óptimo no solo de la aplicación sino de los métodos y técnicas de la Lingüística de corpus como metodología de indagación lingüística.

Por último, hemos de destacar que la propia versatilidad y funcionalidad de una metodología *corpus-driven* permite la adaptación de las propuestas didácticas a diferentes escenarios y niveles, aunque por los aspectos comentados previamente sería difícil de aplicar en un nivel inicial de Primaria.

## 6. CONCLUSIONES

Nuestro estudio tenía como objeto de estudio el uso de un corpus di-  
dático escolar para el aprendizaje de las competencias clave especifica-  
das en la LOMLOE, prestando atención a la competencia comunica-  
tiva.

La LOMLOE insiste en el desarrollo de la competencia comunicativa,  
enlazando con el desarrollo de la competencia social y ciudadana y la  
competencia digital, en el marco de acciones de la Agenda 2030:

[...]la educación para el desarrollo sostenible y la ciudadanía mundial  
ha de incardinarse en los planes y programas educativos de la totalidad  
de la enseñanza obligatoria, incorporando los conocimientos, capacida-  
des, valores y actitudes que necesitan todas las personas para vivir una  
vida fructífera, adoptar decisiones fundamentadas y asumir un papel  
activo –tanto en el ámbito local como mundial– a la hora de afrontar y  
resolver los problemas comunes a todos los ciudadanos del mundo.

[...]El desarrollo de la competencia digital no supone solamente el do-  
minio de los diferentes dispositivos y aplicaciones. El mundo digital es  
un nuevo hábitat en el que la infancia y la juventud viven cada vez más:  
en él aprenden, se relacionan, consumen, disfrutan de su tiempo libre.  
Con el objetivo de que el sistema educativo adopte el lugar que le co-  
rresponde en el cambio digital, se incluye la atención al desarrollo de  
la competencia digital de los y las estudiantes de todas las etapas edu-  
cativas, tanto a través de contenidos específicos como en una perspec-  
tiva transversal, y haciendo hincapié en la brecha digital de género.

El trabajo con este tipo de corpus aúna el trabajo de la competencia  
digital, por lo que supone la búsqueda e interpretación de resultados en  
la propia herramienta; la competencia social y ciudadana se ve refor-  
zada por esa conciencia de diversidad cultural que le proporciona el  
conocimiento de otros entornos lingüísticos; y más evidente aún si cabe  
es el desarrollo de la habilidad estratégica que la sociedad actual tanto  
requiere en nuestros egresados y profesionales de todas las áreas.

## 8. REFERENCIAS

- Boulton, A. (2021). Research in data-driven learning. En P. Pérez-Paredes y G. Mark, *Beyond Concordance Lines: Corpora in language education*. Studies in Corpus Linguistics, 102 (pp. 9-34). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.102.01bou>
- Boulton, A. (2017). Data-driven learning and language pedagogy. En S. Thorne y S. May (Eds.), *Language, Education and Technology: Enciclopedia of Language and Education*. Springer Cham. [https://doi.org/10.1007/978-3-319-02237-6\\_15](https://doi.org/10.1007/978-3-319-02237-6_15)
- Boulton, A. y Tyne, H. (2013). Corpus linguistics and data-driven learning: a critical overview. *Bulletin suisse de linguistique appliquée*, No. 97, 97-118.
- del Moral Barrigüete, C. (2008). La enseñanza-aprendizaje del vocabulario en ELE con el método *Collage*. *Publicaciones* 38, 183-193.
- Martín Herrero, C. (2009). Aproximación a ciertas perspectivas en Lingüística de Corpus. En P. Cantos Gómez y A. Sánchez Pérez (Eds.), *A survey of corpus-based research* (pp. 1020-1032). Asociación Española de Lingüística de Corpus.
- McEnery, T. y Wilson, A. (2001). *Corpus Linguistics: An Introduction*. Edinburgh University Press.
- Núñez-Nogueroles, E. E. (2019). Pasado, presente y futuro de los corpus de aprendices de ELE. Una revisión bibliográfica. *ReiDoCrea- Monográfico sobre Perspectivas transnacionales en la enseñanza de lenguas*, 8(3), 170-190.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Studies in Corpus Linguistics, 6. John Benjamins Publishing Company. <https://doi.org/10.1075/scl.6>
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. y Suchomel, Vít (2014). The Sketch Engine: en years on. *Lexicography* 1, 7-36.
- Kilgarriff, A., Rychlý, P., Smrž, P. y Tugwell, D. (2004). The Sketch Engine. Proceedings of the 11th EURALEX International Congress (pp.105-116). Euralex, Université de Bretagne-Sud.