

Raquel Taranilla. Reseña de [Cruz, Mar. 2012. *Lingüística de corpus y enseñanza del español como 2/L*. Madrid: Arco Libros. Infoling 3.19 \(2013\)](#)
<<http://infoling.org/informacion/Review171.html>>

En su cerca de medio siglo de historia, la Lingüística de Corpus (LC) ha supuesto una auténtica revolución metodológica en la investigación lingüística (Hunston 2002: 1), puesto que ha permitido al estudio de las lenguas manejar grandes cantidades de producciones lingüísticas reales a partir de las cuales extraer generalidades o corroborar hipótesis. De forma progresiva, el empleo de corpus ha ido alcanzando prácticamente todas las ramas de la disciplina, desde aquellas que tienen un fin estrictamente teórico hasta las de carácter aplicado. Este es el caso de la enseñanza de español 2/L, cuyos profesionales están cada vez más interesados en las aplicaciones de la LC.

El propósito del libro *Lingüística de corpus y enseñanza del español como 2/L*, de la profesora Mar Cruz Piñol, satisface justamente ese interés. Con la finalidad fundamental de ser útil a los profesores de español 2/L, este libro se ocupa de cómo las aportaciones de la LC pueden enriquecer la docencia de la lengua. Su objetivo, más allá de realizar una mera descripción de las posibilidades de la explotación de corpus lingüísticos, es ser una puerta de entrada a un campo de investigación y desarrollo de recursos docentes con posibilidades infinitas. De hecho, una de las virtudes del libro es que se esfuerza por implicar al lector en el descubrimiento de nuevos modos de incorporar el uso de corpus textuales en el aula de español. Es por ello que las explicaciones van acompañadas de numerosos ejemplos y de referencias constantes a distintos recursos. Muchos de ellos están disponibles en línea, de modo que el lector puede complementar el contenido del libro e innovarlo según sus propias necesidades. En ese sentido, este libro responde a la perfección al espíritu de la colección *Manuales de formación de profesores de español 2/L* (editorial Arco/Libros), en la que está editado, y que se caracteriza por ser una serie esencialmente práctica, útil y actualizada.

El libro está dividido en dos partes. La primera, que reúne del primer al séptimo capítulo, contiene una presentación general de la LC, pensada específicamente para acercar a los profesionales del español 2/L sus conceptos básicos y sus herramientas. Una vez asentados tales rudimentos, la segunda parte del libro, que se extiende del capítulo 8 al 14, se ocupa de las aplicaciones de la LC a la enseñanza del español 2/L. Cada parte se completa con una serie de preguntas y actividades con las que el lector puede evaluar el conocimiento que va adquiriendo. Asimismo, a fin de que el trabajo no pierda actualidad, la autora recurre a una dirección web (www.delicious.com/MarCruzPinol/corpus) en la que se pueden localizar fácilmente tanto los materiales citados como otros recursos sobre el mismo tema. Por último, se incluye un completo listado de recursos en línea sobre LC, que permitirá al lector ampliar la explicación en función de sus motivaciones.

2. Los conceptos fundamentales de la LC

Los primeros capítulos del libro abordan los fundamentos de la LC, expuestos de manera clara y concisa, e ilustrados con ejemplos pertinentes para los profesionales del español 2/L. Aunque la autora no pretende realizar una panorámica exhaustiva de esta área de la Lingüística, sino más bien abrir una ventana al tema específicamente pensada para los docentes de la lengua, lo cierto es que este primer bloque proporciona un buen planteamiento general sobre el empleo de corpus en el estudio de la lengua, complementado con numerosas referencias bibliográficas que permiten al lector conocer los textos capitales de la LC. Destaca, asimismo, que desde las primeras páginas se recurra a imágenes de consultas a corpus en lengua española para hacer más comprensible la explicación.

Con el propósito de contextualizar la LC, los dos primeros capítulos abordan su origen y su relación con las corrientes lingüísticas del momento. El libro se abre con la definición

de “corpus” propuesta por Sinclair (2004), según la cual un corpus es una recopilación de textos en soporte electrónico, seleccionada de acuerdo con criterios objetivos para representar una lengua o una variedad de lengua, y que tiene el propósito de proporcionar datos empíricos a la investigación lingüística. Seguidamente, se describen los antecedentes de la LC y la posición sobre el empleo de corpus que han mantenido las escuelas lingüísticas dominantes en las últimas décadas (a saber, el estructuralismo, el generativismo y el cognitivism lingüístico-social). Para terminar, y en consonancia con la voluntad de actualidad que inspira todo el trabajo, la autora describe cuál es la situación en el momento presente, en el que identifica una “tendencia hacia el eclecticismo” (pág. 32), es decir, un interés por combinar el empleo de datos empíricos con el conocimiento que proviene de la introspección y de la intuición del propio investigador.

El capítulo 3 comienza justificando la utilidad del empleo de corpus. En palabras de la autora,

Entre los muchos beneficios del manejo de estos bancos de datos podríamos mencionar que permiten disponer de muestras reales de la lengua en uso, poner a prueba y comprobar las teorías lingüísticas o las intuiciones, comprobar la gramaticalidad de las oraciones dentro de su contexto, extraer reglas o descripciones lingüísticas a partir de la observación de la lengua, obtener información cuantitativa y, por lo tanto, verificable, y manejar un gran volumen de información en poco tiempo (pág. 35).

Dicho capítulo también se ocupa de uno de los conceptos cruciales —pero también más problemáticos— de la LC: el de la representatividad de un corpus. Se trata de saber cuándo un corpus es representativo de una variedad de lengua y, por tanto, puede ofrecer conclusiones válidas sobre esa variedad. Según apunta la autora, la LC aún no ha podido ofrecer criterios claros sobre cuándo un corpus puede considerarse representativo. En cualquier caso, dado que este problema no impide en nada que el empleo de corpus sea rentable en la enseñanza de español 2/L, no interesa insistir más en esta cuestión.

En el capítulo 4 se presentan algunas maneras de utilizar un corpus. En particular, se abordan las dos formas básicas del análisis lexicométrico: los listados de frecuencias (esto es, listas que permiten conocer la frecuencia en que se usa una palabra en una determinada variedad de lengua) y las concordancias (que muestran cuáles son los contextos habituales de una palabra). En la línea general del libro, prima el deseo de mostrar la utilidad de conocer la frecuencia de uso de las palabras y sus concordancias recurrentes en la enseñanza de la lengua. En concreto, se destaca su conveniencia a la hora de elaborar materiales didácticos, gramáticas y diccionarios. Además, la autora ofrece una valiosa relación de diccionarios de frecuencias y de corpus del español que permiten consultar frecuencias, concordancias y colocaciones. Toda la explicación se ilustra con imágenes de búsquedas reales en distintos corpus, lo que sin duda facilita la lectura y familiariza al lector con los recursos de la LC y las diversas formas de búsqueda. Por último, se dedica un apartado a los programas informáticos que permiten realizar análisis lexicométricos, lo que sin duda constituye una herramienta muy apreciable y poco conocida entre los profesionales del español 2/L.

El capítulo 5 está dedicado a la anotación de corpus, que consiste en incorporar en el texto crudo información de tipo lingüístico, lo que aumenta las posibilidades de que un corpus pueda ser rentabilizado. Por una parte, el capítulo desglosa de forma clara los distintos niveles de anotación a los que se puede someter un corpus: a saber, la lematización y la anotación morfológica, sintáctica y semántica, principalmente. Por otra parte, se enumeran diversos corpus del español disponibles en línea y que permiten realizar consultas por lema, por categoría gramatical, por función sintáctica y por relación semántica. Como se ha dicho arriba, la finalidad de todo ello es asentar las bases que permitirán, en la segunda parte del libro, dar cuenta de las distintas aplicaciones de los corpus al aula de español 2/L.

El capítulo 6, a pesar de que es probablemente el más descriptivo de la obra, ya que se ocupa de la tipología de los corpus lingüísticos, no deja por ello de implicar al lector y de empujarle a experimentar por sí mismo con los muchos corpus que se mencionan. El

capítulo clasifica los corpus según los criterios siguientes: el canal de producción lingüístico, el grado de representatividad, la especificidad de los textos, el porcentaje y la distribución de los textos, los idiomas incluidos y, por último, el nivel de análisis. Todo ello, como es común en el resto del manual, está ejemplificado a partir de imágenes de distintos corpus, desde los que son más conocidos hasta los que, sin serlo tanto, resultan sin duda recursos interesantes.

Finalmente, el primer bloque se cierra con el capítulo 7, dedicado a las aplicaciones de la LC y, en especial, a las aplicaciones en el estudio del español. Su mayor valor es que relata casos concretos que permiten al lector comprender las potencialidades del uso de corpus. Así, para empezar, la autora da cuenta de la rentabilidad de los corpus en la elaboración de gramáticas, diccionarios y repertorios léxicos del español: es el caso, entre otros que se refieren, de la vigésima segunda edición del *Diccionario de la lengua española*, que tuvo en consideración el corpus CREA, o del *Diccionario del Castellano del siglo XV en la Corona de Aragón*, que está basado en un corpus del castellano de ese siglo. Además, el empleo de corpus también es ventajoso en estudios de neología, disponibilidad léxica, dialectología, fraseología y normativa.

3. Aplicaciones de la LC a la enseñanza de español 2/L

La segunda parte del libro está dedicada al uso de corpus en la enseñanza de español 2/L. El capítulo 8, de carácter introductorio, aborda algunas implicaciones didácticas del empleo de corpus en la docencia de una lengua. En particular, destaca una consideración que el profesor del español 2/L que desee emplear corpus en sus clases debe tener en cuenta: si bien la aplicación de corpus a la enseñanza de segundas lenguas se ha vinculado a un aprendizaje proactivo y autónomo, no por ello el profesor puede quedarse al margen, sino que ha de aprender a intervenir de otra manera. Según advierte la autora, en ese contexto docente la implicación del profesor sigue siendo crucial; este no puede desentenderse del aprendizaje a través de corpus, sino que ha de guiar y controlar todo el proceso.

El capítulo 9 está dedicado a la manera más sencilla de emplear un corpus en el aprendizaje de lenguas: la consulta de ocurrencias y de concordancias. Este tipo de consulta, que es especialmente útil para la enseñanza de vocabulario, puede ser explotado de muy diversos modos. La autora sugiere algunas actividades y da herramientas para idear muchas otras. Así, entre otros recursos, propone usos concretos que corpus como el CREA o el CorpusEye pueden tener en el aula; igualmente, presenta las plantillas contenidas en el manual en línea ICT4LT (*Using concordance programs in the Modern Foreign Language classroom*), invitando a plantear, por ejemplo, actividades para trabajar los falsos amigos.

Una de las cualidades más destacadas de este libro y, de forma muy particular, de su segunda parte es la gran cantidad de experiencias reales que se recogen. Véase, como muestra, este fragmento:

También en relación con el uso de la prensa en línea como corpus, McCullough (2001: 132-133) describe el caso de un profesor de inglés que tomó una noticia de actualidad de un diario en línea (una resolución judicial muy esperada), utilizó un programa de procesamiento de corpus para obtener el listado de todas las palabras, escogió las que pertenecían al lenguaje jurídico, buscó con el mismo programa las concordancias de esos términos en el texto de la noticia, seleccionó los casos más interesantes y, a partir de ahí, preparó un ejercicio de lenguaje específico, en el que los alumnos tenían que colocar el término jurídico correcto en diez frases. A continuación, buscó los pronombres relativos seguidos o no de coma y, siguiendo el mismo procedimiento, preparó un ejercicio sobre oraciones subordinadas explicativas y especificativas (pág. 120-121).

Para la enseñanza del léxico también se pueden emplear listas de frecuencias, que son el tema del capítulo 10, y que han sido tradicionalmente utilizadas para confeccionar glosarios y diccionarios, además de en el diseño de actividades para aprender vocabulario. Aunque, según explica la autora, su uso para la docencia ha sido objeto de algunas críticas, la verdad es que bien pueden ser un recurso más al alcance de los profesores de lenguas.

Más novedosa es, sin duda, la explotación de corpus de aprendices, esto es, corpus de producciones lingüísticas de individuos que están aprendiendo una lengua (ya sea como 1/L o como 2/L). Tal como explica el capítulo 11, existen diversos corpus de aprendices de español 2/L que recopilan textos escritos (como CEDEL2, CATE o SAELE) y que pueden emplearse tanto en la investigación de la adquisición del español como para preparar actividades para la docencia. Por lo que se refiere a la lengua oral, que como bien afirma la autora “es la destreza para la que existen menos recursos” (pág. 135), contamos con corpus que permiten acceder a las grabaciones de las producciones orales de los aprendices. Es el caso de los corpus OLAC, SPLLOC y SPT. En este último corpus, además, se incluyen actividades de léxico, cuestiones gramaticales, fonéticas y pragmáticas, que pueden inspirar a los profesores a diseñar otros materiales y que, por otro lado, pueden servir para que los estudiantes observen y aprendan de errores cometidos por otros aprendices. “Así el corpus se convierte en un recurso para aprender de la experiencia entre iguales” (pág. 137).

El capítulo 12 está dedicado a la aplicación de búsquedas en corpus anotados. La explicación viene acompañada por ejemplos muy claros de cómo emplear las consultas a la hora de preparar materiales docentes o de proponer recursos en el aula. Así, se plantean algunas de las posibilidades que puede tener la búsqueda en corpus lematizados y en corpus que contienen información gramatical, sintáctica o semántica. Como muestra de los muchos ejemplos que van aportándose durante el capítulo, puede verse el fragmento que sigue, en el que se propone una de las posibilidades de la búsqueda de las relaciones semánticas de la palabra “vacaciones” en la aplicación Araknion:

[C]omo se puede ver, aquí, además de los casos de proximidad léxica como *playa* (*vacaciones en la playa*), *verano* (*vacaciones de verano*), *disfrutar* (*disfrutar de las vacaciones*), *turistas* (*turistas de vacaciones*), obtenemos otros casos de palabras relacionadas semánticamente con las vacaciones (como *ocio*, *sueldo*, *subsidio... o recorte*). En el aula de E/2L, este recurso podría servir, por ejemplo, como ayuda para la composición de textos: a partir del entorno léxico-semántico de una palabra (escogida en función de las motivaciones del grupo y del momento), es fácil animar a los estudiantes a imaginar una historia (pág. 144).

El capítulo 13 se enfrenta a uno de los ámbitos de la enseñanza del español con menos trayectoria, pero con una demanda en aumento: la enseñanza de español 2/L para fines específicos. Se trata de un tipo de docencia especializada que puede enriquecerse en gran medida a partir del uso de corpus, dada la complejidad que supone para el profesor añadir al conocimiento de la lengua otro campo de saber específico (la economía, la medicina, el derecho, el turismo, etc.), que posee un discurso con claves propias. Aunque esta es, ciertamente, un área de la docencia del español poco desarrollada todavía, la autora propone diversos ejemplos de usos de corpus para elaborar materiales de español para fines específicos. Uno de ellos es la plataforma ElektraVoc, dedicada al aprendizaje del español de la economía; a partir de un corpus de artículos periodísticos, se han elaborado un glosario y algunas actividades que sirven para enseñar el léxico propio de la economía en contextos de uso reales.

Además de la aplicación de las consultas en corpus a la actividad estrictamente docente, la explotación de corpus puede emplearse también en la elaboración de materiales didácticos. Así, el capítulo final se centra en la confección de diccionarios y manuales a partir de búsquedas en corpus de español. Aparte de justificar la oportunidad de emplear corpus en la confección de obras lexicográficas y recursos didácticos, la gran aportación de este capítulo es que presenta materiales que, en algunos casos, son poco conocidos para los profesionales del español 2/L y que, además, pueden inspirar en la creación de nuevos recursos. En el caso de los diccionarios basados en corpus, merecen ser citados, por ser obras recientes, el *Gran diccionario de uso del español actual* y el *Diccionario de locuciones adverbiales para la enseñanza del español*. En cuanto a los manuales, se mencionan *Cumbre*, *PortaVoces* y *Trampas y pistas*, y como muestra de actividades interactivas en línea, *FrameNet Español*.

4. Valoraciones generales

El libro *Lingüística de corpus y enseñanza del español como 2/L* llena un vacío en la bibliografía de la enseñanza de español 2/L. Tal y como ha ocurrido antes en otras lenguas de nuestro entorno, la aplicación de la LC a la docencia del español es, cada vez más, una demanda de los profesionales del español y este trabajo supone una magnífica iniciación a ese campo, puesto que cumple una doble función. De un lado, realiza una visión general de la LC aplicada a la enseñanza de lenguas clara y rigurosa. De otro lado, da cuenta de una extraordinaria variedad de recursos y propone numerosos ejemplos capaces de sugerir al lector nuevos desarrollos en función de sus propias inquietudes e intereses.

El libro demuestra un conocimiento profundo y muy actualizado de los corpus del español, así como de los recursos docentes generados a partir de las consultas a corpus. Además, desde su inicio, resulta patente que es el resultado de una larga reflexión sobre las necesidades de la enseñanza del español 2/L, así como de la voluntad de innovar y aprovechar las oportunidades que ofrecen las nuevas tecnologías para optimizar el aprendizaje de lenguas. La autora ha combinado una consolidada obra teórica en ese ámbito con su experiencia como docente de español 2/L. Solo desde esa doble vertiente puede producirse un libro útil, completo y tan sugerente, y que ha de constituir un punto de partida ineludible para aquellos profesionales del español interesados en introducir nuevas formas de aprendizaje en el aula.

Referencias bibliográficas

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge UP.

McCullough, J.L. 2001. "Los usos de los corpórea de textos en la enseñanza de lenguas", en Trenchs, M. (ed.) *Nuevas tecnologías para el autoaprendizaje y la enseñanza de lenguas*. Lleida: Milenio, 125-140.

Sinclair, J.M. 2004. *Trust the text: Language, corpus, and discourse*. Londres: Routledge.