

## Book Reviews and Notices

**CRUZ PIÑOL, Mar (2012): *Lingüística de corpus y enseñanza del español como 2/L*. Madrid: Arco/Libros. Manuales de formación de profesores de español 2/L. [ISBN: 978-84-7635-850-4] 192 páginas.**

El auge que experimenta actualmente el aprendizaje de segundas lenguas ha propiciado que otros ámbitos de la Lingüística, en principio ajenos, hayan volcado sus estudios, o parte de ellos, en facilitar la enseñanza de L2. Es el caso de la lingüística de corpus. El propio profesorado ha sido consciente de las ventajas que esta puede presentar en la docencia, de ahí que hayan surgido numerosas propuestas para el empleo de los diferentes corpus lingüísticos en la enseñanza de lenguas y, en nuestro caso, en el español; sin embargo, la mayor parte de los docentes carecen de la formación necesaria para un máximo aprovechamiento de los mismos.

En este contexto, surge la obra de Mar Cruz Piñol, *Lingüística de corpus y enseñanza del español como 2/L*, con la que pretende que el lector sea capaz de manejar algunos conceptos clave de la LC (lingüística de corpus) y que conozca las posibilidades de estos corpus en el aprendizaje de lenguas, con el propósito final de “contribuir a la formación y especialización del profesorado de español como 2/L” (13). Por todo ello, la finalidad pedagógica de la obra es evidente.

Tal objetivo explica igualmente la estructura que la autora determina para su libro. Este aparece dividido en dos partes: la primera de ellas, “La lingüística de corpus: conceptos fundamentales”, se ocupará, como ya hemos comentado, de iniciar al lector en la LC, mientras que, en la segunda, “La lingüística de corpus y el español como 2/L”, se aplican a la enseñanza de segundas lenguas los conceptos vistos en el apartado anterior, mostrando cómo los corpus facilitan la labor del docente. Es por ello por lo que afirmábamos que la finalidad didáctica justifica la propia estructura, puesto que es necesario que el lector sea conocedor de una serie de conceptos generales de la LC para que sea posible su posterior aplicación en la enseñanza del español. Asimismo, Cruz Piñol cierra cada una de estas dos partes con una serie de preguntas de síntesis (y más adelante sus respectivas respuestas), con las que pretende que el lector compruebe por sí mismo si los conceptos se han comprendido verdaderamente. Encontramos también varias actividades que, en cierta

manera, obligan al lector a consultar las webs propuestas y llevar a la práctica los conocimientos previamente asimilados.

Otro de los aspectos que refuerza su carácter práctico y pedagógico son las numerosas imágenes que la autora incluye con la interfaz de distintos corpus. Las explicaciones se vuelven mucho más claras una vez que el lector puede concretar en ejemplos los planteamientos teóricos expuestos. Por otro lado, al hacer alusión a corpus fácilmente accesibles, por estar disponibles en línea, se anima al receptor a consultarlos y a que sea él quien experimente por sí mismo.

La autora comienza proporcionando la definición de corpus, cuyo significado queda restringido al ámbito de la Lingüística, donde se caracteriza por: a) ser una recopilación de textos; b) estos textos deben seleccionarse a partir de unos criterios objetivos y no arbitrariamente; c) es representativo de la lengua; d) constituye la base de los estudios lingüísticos y e) suele estar disponible en línea. Entre todas ellas, la autora incide en su carácter representativo; es más, “la representatividad es, pues, la piedra angular de la LC, pues de ello depende que se puedan extraer conclusiones fiables a partir de los datos estadísticos” (36).

Asimismo, es posible localizar sus precedentes en las líneas de concordancia, que mostraban la palabra en su contexto, y los índices de frecuencia, según la distribución de las letras. Además, el desarrollo de la LC discurre en paralelo a las teorías que diferencian las distintas etapas de la Lingüística: si el estructuralismo despierta el interés por las muestras reales de lengua, el generativismo las enviará a un segundo plano, al centrar su atención en el concepto abstracto de competencia. El cognitivismo incide en la función social del lenguaje. Por ello, mantiene que, para estudiarlo, es necesario recurrir a datos reales, que nos permitan contemplar la lengua en su uso, de ahí que los corpus vuelvan a revalorizarse. Actualmente, se opta por un modelo ecléctico en el que los conceptos teóricos son esenciales para el análisis de las muestras lingüísticas.

En la actualidad, los listados de frecuencias y de concordancias continúan empleándose, si bien su informatización ha permitido grandes ventajas. Dentro de los primeros, señala lo útil que puede resultar el proceso de lematización. Esto es, “relacionar cada FORMA (la palabra tal y como aparece en el texto) con su LEMA (la entrada bajo la que buscaríamos esa palabra en un diccionario)” (53). Las concordancias, por su parte, tienen en cuenta el contexto, con lo que su utilidad para conocer cuáles son las distribuciones más frecuentes en una lengua es evidente. Incluso, señala la existencia de corpus que permiten seleccionar las grafías

que aparecerían en torno a la palabra en la que se centra la búsqueda. En todos estos apartados, como anteriormente señalamos, Cruz Piñol ofrece ejemplos de corpus en línea que animan al lector a investigar por sí mismo.

De la misma forma, encontramos información acerca de cómo se lleva a cabo la lematización a través de un programa informático, lo cual parece alejarse de las pretensiones inmediatas del docente; no obstante, la propia autora justifica la inclusión de este apartado, afirmando que resulta útil para que el profesor pueda conocer mejor el funcionamiento de estos programas y, con ello, explotar al máximo sus posibilidades. En definitiva, la lematización aporta datos acerca de la propia morfología de la lengua, con lo que el valor del corpus se eleva considerablemente, un valor que asciende aún más si se añade, como la autora propone, la categoría gramatical, información sintáctica, semántica y pragmática, con lo que el lector puede ya imaginar la dificultad que esta tarea conlleva.

En cuanto a la tipología de corpus, se establecen varias clasificaciones en función del criterio seleccionado; por ejemplo, desde la perspectiva de la lematización, puede resultar interesante la última clasificación incluida (entre un total de seis), según el nivel de análisis. De esta forma, podemos distinguir entre: a) corpus simples; b) corpus lematizados; c) corpus analizados morfológicamente; d) corpus con información sintáctica superficial; e) corpus analizados sintácticamente y f) corpus etiquetados semánticamente. Así pues, en principio, es lógico pensar que la ausencia de anotación morfológica reducirá las posibilidades en análisis de carácter gramatical; sin embargo, Guillermo Rojo, a partir de su artículo “Sobre codificación y explotación de corpus textuales. Otra comparación del Corpus del Español con el CORDE y el CREA” (2010), problematiza la sencillez aparente de esta clasificación al señalar (refiriéndose al CORDE y al CREA) que “la rica codificación de estos corpus [...] permitirá obtener resultados más fiables” (32). Enumera entre sus ventajas consultar la distribución según los diferentes tipos de texto o según las distintas variedades del español, así como poder observar con detalle la evolución de las formas en periodos concretos de tiempo. En definitiva, la idea que nos interesa de Guillermo Rojo es que no siempre la ausencia de anotación lingüística ha de verse como una deficiencia en los corpus, sino que, en algunas ocasiones, esta puede verse fácilmente suplida por otro tipo de información detallada que nos ayudará igualmente a resolver nuestras dudas. A pesar de que esta tarea se logre a fuerza de aumentar el número de búsquedas, puede resultar muy útil para que el alumnado sea capaz de resolver sus dudas a partir de sus propias investigaciones.

Finalmente, en esta primera parte, Cruz Piñol enumera las aplicaciones de la LC, entre las que, además de la elaboración de gramáticas y diccionarios, los estudios neológicos, fraseología y normativa léxica, se encuentra el aprendizaje de segundas lenguas, con lo que conecta directamente con la segunda parte del libro.

Esta segunda parte se inicia con una serie de consideraciones didácticas, que sintetizan brevemente las distintas posibilidades de los corpus en este ámbito. En ellas, profundizará la autora en los capítulos posteriores. Así pues, la LC puede resultar útil para:

1. determinar si una forma léxica es empleada en la lengua (consulta de ocurrencias);
2. mejorar la elaboración de manuales, sobre todo, los que se encuadran en el enfoque léxico;
3. ofrecer datos reales que faciliten el análisis de la interlengua;
4. poner al alumno en contacto con *input* real, de manera que pueda ser él quien deduzca las estructuras de la lengua a partir de ellas.

Una vez señaladas estas, la autora advierte de un peligro importante, que no deberá perderse de vista: dada la variedad de datos que se pueden encontrar en los corpus y, más aún, si partimos de la consideración de la web como gran corpus textual, el profesor deberá ser especialmente cuidadoso con la información que selecciona. Este problema está muy presente en la primera de las opciones que hemos señalado, las ocurrencias, por lo que el docente deberá estar capacitado para discernir entre información útil y aquella que no lo es.

Por otro lado, las concordancias permiten conocer las combinaciones más habituales de una lengua, lo que facilita en gran medida la elaboración de diverso material didáctico. En esta misma línea, Izquierdo Gil, en “Los corpora electrónicos en la enseñanza del vocabulario español LE: la utilidad pedagógica de las concordancias” (2002), reclama la necesidad de tener en cuenta las colocaciones reveladas por los listados de concordancias a la hora de elaborar un léxico básico, ya que, como ella misma comenta “los índices de frecuencia clásicos constituyen únicamente la base o el punto de partida para el desarrollo de la competencia comunicativa” (209). Señala, pues, entre todas estas colocaciones, la importancia de las de tipo gramatical, que incluyen formas verbales. Dentro de estas, en las construcciones con *verbo soporte*, especialmente frecuentes en el español, los verbos empleados suelen ser verbos deslexicalizados (*dar, tener, hacer...*), que se prestan a una diversa

polisemia, pues el contenido léxico recae en los complementos que los acompañan. Volviendo a Cruz Piñol, propone que, en el caso concreto del español, las concordancias puedan resultar útiles para que el alumno conozca no solo la lengua estándar, sino también algunas de las principales variedades que esta presenta. Según el manual en línea *Using concordance programs in the Modern Foreign Languages classroom* descubrimos que son muchas las actividades que presenta, además de la ya citada por la autora, en la que propone trabajar con los “falsos amigos”. Con ellas, el propio alumno ha de acudir a los programas de concordancias. En la web, sugieren, por ejemplo, que el alumno encuentre la palabra más apropiada para un contexto dado o que este analice el contexto en el que aparecen los adverbios de esa lengua para saber cuál es su posición sintáctica, a través de la búsqueda de una serie de combinaciones proporcionadas por el profesor, lo que podría aplicarse también a otro tipo de estructuras sintácticas.

De la misma forma, las listas de frecuencias también se presentan como fuente para la creación de materiales, si bien presenta cierta complejidad a la hora de adecuar estas listas de palabras a los diferentes niveles y, sobre todo, si tenemos en cuenta que “el aprendizaje de la lengua no siempre discurre en paralelo con la estadística” (127). A esta dificultad, habría que añadir la que entrañan los corpus lematizados que, aunque permiten resolver rápidamente casos de homonimia o conocer restricciones morfosintácticas de un término, no son recomendables hasta un nivel de C1, como comenta la autora a juicio de KrisBuyse.

Para el análisis de la interlengua, resultan verdaderamente útiles los corpus de aprendices, elaborados por los propios estudiantes. El profesor conoce así el estadio de la lengua en el que se sitúa el aprendiz y el tipo de errores, a veces generalizados, que se asocian a ese nivel. En esta línea, Joseba Abaitua en “Tratamiento de corpora bilingües” (2002) alude al *International Corpus of Learner English*, con textos elaborados por estudiantes de inglés con diferente lengua materna, con el objetivo de “comprobar la hipótesis del 'modelo de lengua común' al que se llega en los niveles avanzados” (18). El análisis de este corpus ha permitido observar algunas de las actitudes más habituales que los estudiantes adoptan con respecto a la segunda lengua; por ejemplo, ante la presencia de falsos amigos, lo más frecuente es que el estudiante evite la palabra en su discurso, seleccionando otra que no se perciba como conflictiva. En conclusión, el estudio de la interlengua a partir de estas muestras permitiría la elaboración de materiales más eficaces, e incluso, resultaría

útil para que los alumnos sean conscientes de algunos de los errores frecuentes y los eviten posteriormente en sus producciones.

Los corpus lingüísticos también pueden resultar muy útiles en el español para fines específicos (EFE), lo cual es posible apreciar gracias al largo listado que la autora proporciona de trabajos elaborados a partir de corpus en este ámbito. Los corpus ofrecen aquí datos objetivos para la selección de un determinado léxico. Por último, señala su aplicación en la elaboración de diccionarios bilingües, manuales y ejercicios, para acabar demandando la necesidad de un mayor número de actividades disponibles gratuitamente en línea.

En conclusión, como hemos querido hacer ver desde el inicio de esta reseña, la labor desempeñada por Cruz Piñol no solo contribuye a una mejor formación del profesorado de español como segunda lengua, sino que, debido a su marcado carácter práctico, pretende animar al profesorado a que emplee estos corpus en la confección de sus materiales didácticos, que serán ahora mucho más eficaces. Es decir, más que proporcionar un mero conocimiento teórico, lo que pretende es que las recomendaciones y contenidos que aquí se incluyen sean realmente útiles para la configuración de las clases, donde tendrán ahora una mayor autonomía. Por todo ello, *Lingüística de corpus y enseñanza del español como 2/L* se convierte en una obra verdaderamente útil y necesaria para los docentes de este ámbito.

Marina Castro Cruz  
*Universidad de Córdoba*

\*\*\*\*\*

**BOTMA, Bert; KULA, Nancy C. y NASUKAWA, Kuniya (eds.), *Continuum Companion to Phonology*. London: Continuum International Publishing Group, 2011 [ISBN: 978-0-8264-3423-4] 540 páginas.**

Los profesores Bert Botma, Nancy Kula y Kuniya Nasukawa nos presentan la edición de un manual que agrupa teorías fonológicas que se emplean y se discuten en gran parte de los departamentos de lingüística a comienzos del siglo XXI, desde los modelos generativos de rasgos y reglas de derivación, que llevan cerca de 50 años en desarrollo, hasta los modelos de Fonología de Laboratorio y Fonología basada en la frecuencia de uso de formas lingüísticas, de origen más reciente.