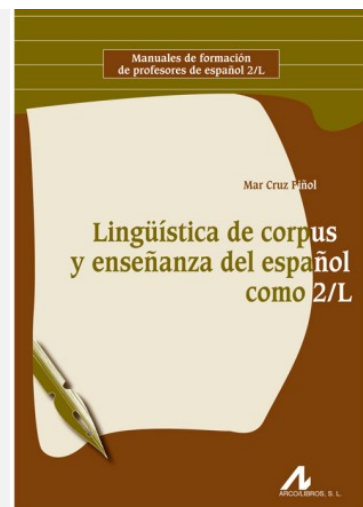


GRAS, PEDRO

UNIVERSITEIT ANTWERPEN, BÉLGICA

MÁS ALLÁ DE LAS GRAMÁTICAS Y LOS DICCIONARIOS,  
LOS CORPUS COMO HERRAMIENTAS EN LA ENSEÑANZA / APRENDIZAJE DE LENGUAS



Cruz Piñol, Mar

*Lingüística de corpus y enseñanza del español como 2/L*

Arco Libros, 2017 (2012)

2ª edición, revisada y actualizada

ISBN: 978-84-7635-850-4

192 páginas

**A**l inicio de un artículo sobre el uso de corpus lingüísticos en la investigación léxica y gramatical, el lingüista estadounidense Charles J. Fillmore (1929-2014), uno de los padres de la Lingüística Cognitiva, opone dos estereotipos de lingüistas: los *lingüistas de sofá* frente a los *lingüistas de corpus*. Según la caricatura de Fillmore, mientras los primeros se sientan a pensar en frases que suenan bien o mal recurriendo a su propia intuición, los segundos dedican sus esfuerzos a calcular las frecuencias de expresiones en colecciones de textos orales o escritos –los corpus lingüísticos. A propósito de esta oposición, el profesor Fillmore afirma lo siguiente:

“Tengo que hacer dos observaciones. La primera es que no creo que haya ningún corpus, por grande que sea, que contenga información sobre todas las áreas de la gramática y el léxico del inglés que quiero explorar; todos los que he visto resultan inadecuados. La segunda observación es que cada corpus que he tenido la ocasión de explorar, por pequeño que fuera, me ha mostrado hechos que no podría haber descubierto de ningún otro modo. Mi conclusión es que ambos tipos de lingüista [de sofá y de corpus] se necesitan mutuamente. O mejor aún, que ambos tipos de lingüistas, en la medida de lo posible, deberían existir en el mismo cuerpo.” (Fillmore 1991: 35, la traducción es mía.)

Aunque seamos hablantes nativos de una lengua, no tenemos en mente todas las posibilidades de combinación de las piezas léxicas y gramaticales de nuestra(s) lengua(s) materna(s). Si le pregunto a un hablante nativo qué verbos pueden seguir a *romper* en una expresión del tipo “rompió a”, es posible que piense en algunos como *reír* o *llorar*, pero muy probablemente olvide otros que también son posibles, como *llover* o *hablar*, y que los corpus

documentan. Al mismo tiempo, es poco probable que un corpus contenga todas las combinaciones posibles del verbo *romper*. Se hace, pues, necesario un equilibrio entre el recurso a datos auténticos –los corpus– y la intuición del hablante, que reconoce matices y ausencias que los corpus por sí solos no pueden suministrar.

En el caso de la enseñanza, aprendizaje e investigación de segundas lenguas, la necesidad del uso de corpus lingüísticos es todavía más evidente. Los aprendices –por definición, hablantes con una competencia parcial en la L2/LE– no pueden confiar en su propia intuición para saber si la combinación “rompió a comer galletas” es una expresión posible, imposible o extraña en español. En este contexto, los corpus lingüísticos pueden constituir herramientas didácticas eficaces para favorecer el desarrollo de la competencia lingüística y comunicativa, así como para compensar las limitaciones propias de una competencia parcial, tal como muestra con numerosos ejemplos el libro *Lingüística de corpus y enseñanza del español como L2* de Mar Cruz Piñol, profesora de la Universitat de Barcelona y reconocida especialista en el uso de las nuevas tecnologías en la enseñanza del español, que acaba de conocer su segunda edición, ampliada y actualizada. En el resto de este texto, voy a describir someramente los contenidos del libro y voy a comentar algunas experiencias de actividades propuestas en el libro en el marco de los cursos de lengua y lingüística españolas que imparto en la Universiteit Antwerpen (Bélgica).

El libro *Lingüística de corpus y enseñanza del español como L2* forma parte de la colección Manuales de formación de profesores de español 2/L2, cuyo objetivo es poner a disposición de los profesores de ELE información actualizada y práctica sobre diferentes áreas relevantes para la enseñanza de idiomas. En

consonancia con este doble carácter de actualización y práctica, el manual se estructura en dos partes: una primera parte destinada a presentar los conceptos básicos de la lingüística de corpus (capítulos 1-7) y una segunda parte que recoge diversas aplicaciones de la lingüística de corpus a la enseñanza del español (capítulos 8-14). Cada parte se cierra con una serie de preguntas cerradas de comprensión (con sus correspondientes soluciones) y una serie de actividades abiertas que invitan a poner en práctica las técnicas expuestas en el manual.

Como decía, la primera parte tiene como objetivo suministrar al lector los conocimientos básicos sobre la lingüística de corpus, con especial atención a los recursos sobre el español. Esta parte se abre con dos capítulos generales, que sitúan la lingüística de corpus en el panorama de la investigación lingüística, tanto en una perspectiva histórica (capítulo 1) como actual (capítulo 2). El resto de capítulos que componen esta parte introducen los conceptos básicos de la disciplina, como *representatividad*, *concordancia*, *frecuencia* o *lematización*, recurriendo en la mayoría de los casos a ejemplos del español. De hecho, la parte teórica se cierra con dos capítulos que ofrecen un panorama completo y actualizado de los corpus disponibles en español, atendiendo a diversos parámetros (oral/escrito, general/específico, monolingüe/plurilingüe, etc.). La obra cuenta con la ventaja de que los enlaces de acceso a los corpus se encuentran disponibles en un gestor de favoritos en línea –en la red Mendeley– que la autora ha puesto a disposición de todos los lectores en la segunda edición, de este modo las referencias pueden actualizarse y ampliarse de forma constante, en consonancia con el dinamismo propio de los recursos en línea. Esta primera parte, por sí misma, es una herramienta útil para el profesor de ELE, puesto que pone a su disposición una selección amplia y actualizada de recursos lingüísticos, que en muchas ocasiones no forman parte de su repertorio de herramientas de trabajo.

Sobre la base de lo expuesto en los primeros capítulos, la segunda parte recoge, de forma organizada, un conjunto de propuestas de explotación de los corpus lingüísticos en el ámbito del ELE, como recurso didáctico en cursos generales o para fines específicos, así como herramienta técnica en la creación de materiales. Esta parte se abre con un capítulo sobre la relación entre el uso de corpus lingüísticos y diferentes aproximaciones a la enseñanza de lenguas extranjeras (capítulo 8). A continuación, se encuentran cuatro capítulos que examinan el potencial pedagógico de diversas herramientas propias de la lingüística de corpus: las *ocurrencias* y *concordancias* (capítulo 9), las *listas de frecuencias* (capítulo 10), los *corpus de aprendices* (capítulo 11) y los *corpus lematizados* o etiquetados con información sintáctico-semántica (capítulo 12). Cierran este apartado dos capítulos que abordan la utilidad del uso de corpus en dos áreas especiales: la enseñanza del español para fines específicos y la creación de materiales didácticos.

Esta segunda parte es, sin duda, la principal contribución de este manual, ya que presenta de forma reflexiva un amplio abanico de propuestas didácticas basadas en el uso de corpus lingüísticos, bien recuperadas de la bibliografía bien propuestas por la propia autora. De este modo, el manual de Mar Cruz viene a llenar un vacío en el ámbito de la enseñanza del español como lengua extranjera. Por una parte, si bien existen obras dedicadas al uso de corpus lingüísticos en la enseñanza de lenguas, así como capítulos dedicados a este tema en volúmenes generales sobre lingüística de corpus (Burnard y McEnery 2000, Sinclair 2004, Aijmer 2009, McEnery y Xiao 2010, Harris y Moreno 2010, entre otros), se trata de obras escritas en inglés y que se centran mayoritariamente en esta lengua. Por otra, existen trabajos dedicados a la aplicación de corpus lingüísticos a la enseñanza del español como lengua extranjera –de hecho, la mayoría referenciados en el manual de

Cruz-; sin embargo, carecen de la visión que conjunto que aporta una monografía, que presenta de forma organizada diversas propuestas pedagógicas agrupadas en función del tipo de herramienta implicada (listas de frecuencias, concordancias, corpus lematizados, etc.).

He tenido la ocasión de llevar algunas de las propuestas de este libro en el marco de mis clases de lengua y lingüística española en la Universiteit Antwerpen, así como en diversos cursos de formación de profesores de ELE. En particular, el recurso a corpus lingüísticos me ha resultado especialmente útil para abordar de forma léxica algunos problemas gramaticales, y de este modo hacerlos más concretos. Voy a comentar algunos ejemplos.

Como es bien sabido, los verbos de cambio (*ponerse, hacerse, volverse, quedarse*, entre otros) suelen plantear problemas a los aprendices de ELE, debido a que en muchas lenguas estos se corresponden con una única unidad léxica (inglés, *become*; francés, *devenir*; neerlandés, *worden*, etc.). Una buena parte de las explicaciones, nocionales o funcionales, ha intentado explicar la combinación de verbos de cambio y predicados nominales (*ponerse contento, hacerse rico, volverse inteligente, quedarse callado*) a partir de una serie de rasgos: cambios permanentes vs. transitorios, positivos vs. negativos, graduales vs. abruptos, etc. El problema de este tipo de acercamientos es que es difícil que den cuenta de todas las combinaciones posibles. Por ejemplo, algunos manuales indican que el verbo *hacerse* selecciona palabras –nombres o adjetivos– que describen la profesión o la ideología, como *hacerse policía* o *independentista*. En cambio, no explican que también se use el verbo *hacerse* –en el sentido de ‘empezar a ser’– en expresiones como *se hizo famoso* –que no es una profesión, de momento– o *se hizo de noche*. Una alternativa consiste en ayudar a los estudiantes a buscar un conjunto de ejemplos de cada uno de estos verbos

–empleando corpus lematizados que permiten la búsqueda por categorías gramaticales, como nombre y adjetivo– y pedirles a continuación que agrupen los ejemplos en función del significado de los atributos. De este modo, pueden llegar a deducir por sí mismos que *hacerse* como verbo de cambio (‘empezar a ser’) se combina preferentemente con atributos que designan ideología (*hacerse independentista*) y profesión (*hacerse policía*), pero también características acumuladas (*hacerse rico, famoso*), paso del tiempo (*hacerse de día, noche, tarde*) o comportamientos fingidos (*hacerse el tonto* o *el despistado*). En definitiva, el recurso a los corpus ayuda –entre otras cosas– a construir generalizaciones partiendo de la consulta de datos reales.

La segunda experiencia que querría reseñar tiene que ver otro aspecto complejo en la adquisición del español como lengua extranjera: la selección de los complementos –y pronombres– con los que se construye un verbo. Como es bien sabido, la mayoría de los verbos pueden ocurrir en distintos entornos sintácticos, de modo que requieren de presencia de diferentes tipos de sintagmas y pronombres, como muestran los siguientes ejemplos del verbo *comer*, extraídos de la Base de datos de Verbos, Alternancias de Diátesis y Esquemas Sintáctico-Semánticos del Español (ADESSE<sup>1</sup>, Universidad de Vigo):

- (1) Dormías siempre en tu estudio y, a veces, incluso COMÍAS allí, a deshora.
- (2) Me sentía destemplado: no HABÍA COMIDO nada en todo el día.
- (3) Piensa por tu cuenta; que nadie te COMA el coco.

---

<sup>1</sup> ADESSE es de acceso libre y gratuito en la siguiente URL: <http://adesse.uvigo.es>

(4) Mientras te aguardaba, recogía las algarrobas que habían caído de los árboles y me las comía.

La información relativa a la construcción de los verbos, muy pertinente para los aprendices de ELE, no queda convenientemente recogida en los diccionarios, ni siquiera en los de aprendizaje, que recogen contextos prototípicos, como (2) en los que comer ocurre en una construcción transitiva, pero frecuentemente dejan de lado otros contextos como la construcción con dativo posesivo (3) o la construcción pronominal (4). La base de datos ADESSE pone a disposición de los profesores y autores de materiales –así como a los alumnos avanzados de ELE, especialmente aquellos que cursan estudios de Letras o se preparan para ser profesores de español– información completa sobre las posibilidades de construcción de un verbo, así como la frecuencia relativa de estas construcciones en un corpus de español contemporáneo, de modo que la frecuencia puede ser tenida en cuenta como indicador de la pertinencia de cada esquema en la programación del proceso de enseñanza del ELE.

Estos dos casos pretenden ilustrar las posibilidades que la consulta de corpus y bases de datos suponen para los profesionales del ELE. El recurso a datos auténticos, en muchos casos enriquecidos por anotación con información gramatical, semántica y pragmática e indicación de frecuencia, ofrecen a profesores y aprendices una nueva vía de acercarse a la realidad gramatical, léxica y discursiva del español.

Me gustaría cerrar este texto, por una parte, agradeciendo a la profesora Mar Cruz que haya puesto a disposición de la comunidad de profesionales del ELE un conjunto de recursos y propuestas metodológicas presentadas de forma clara y coherente, y, por otra, animando a los profesores y demás profesionales del ELE a acercarse

a las posibilidades que ofrece la lingüística de corpus en la enseñanza del español.

## REFERENCIAS

Aijmer, Karin (ed.), 2009: *Corpora and Language Teaching*. Ámsterdam: John Benjamins.

Burnard, Lou; McEnery, Tony (eds.), 2000: *Rethinking language pedagogy from a corpus perspective*. Nueva York: Peter Lang.

Fillmore, Charles, 1991: "Corpus linguistics" or "Computer-aided armchair linguistics", en *Directions in Corpus Linguistics*. Estocolmo: Mouton de Gruyter. 35-66.

Harris, Tony; Moreno Jaén, María, 2010: *Corpus Linguistics in Language Teaching*. Berlín: Peter Lang.

McEnery, Tony; Xiao, Richard, 2010: What corpora can offer in language teaching and learning, en *Handbook of Research in Second Language Teaching and Learning*. Londres: Routledge. 364-380.

Sinclair, John (ed.), 2004: *How to use corpora in language teaching*. Ámsterdam: John Benjamins.

FECHA DE ENVÍO: 19 DE DICIEMBRE DE 2017

FECHA DE ACEPTACIÓN: 16 DE ENERO DE 2018