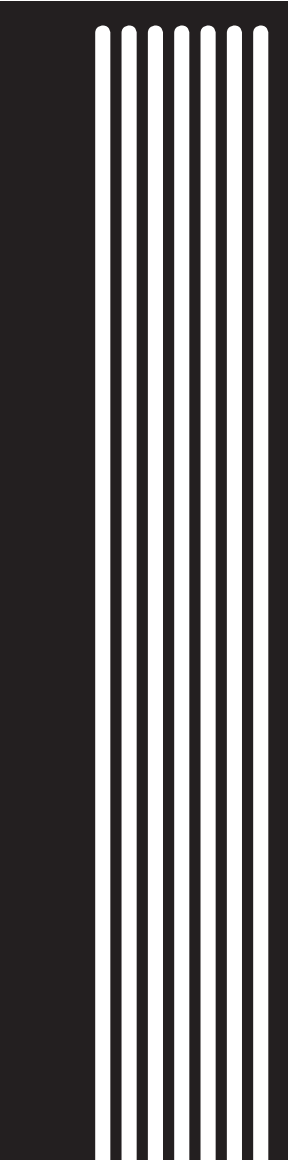




# DOSSIER THÉMATIQUE



**Corpus y diccionarios para  
la investigación en lengua española**



---

## Introducción: Corpus y diccionarios para la investigación en lengua española

### Introduction: Corpus and Dictionaries for Spanish Language Research

---

SARA GÓMEZ SEIBANE [sara.gomezs@unirioja.es]  
Universidad de La Rioja, España

Este volumen de la revista *Études romanes de Brno* reúne una selección cuidadosa de artículos sobre la lengua española a partir de corpus y diccionarios. El lector tiene en sus manos parte de los trabajos presentados en las I Jornadas Filológicas Internacionales-UR (JFI-UR) con el título “Corpus, bases de datos y diccionarios para la investigación de la lengua y la literatura”, celebrado en la Universidad de La Rioja y en Cilengua del 30 de septiembre al 2 de octubre de 2015. La mayor parte de los artículos que se recogen en este volumen fueron presentados en versiones anteriores en el mencionado encuentro, que reunió a especialistas de lo lingüístico y lo literario que utilizan corpus, bases de datos y diccionarios en su investigación, con el fin de mostrar las posibilidades de consulta y explotación que este tipo de herramientas ofrece en diversas áreas, como la literaria, la lingüística y la relacionada con la enseñanza de lenguas.

A partir de este encuentro científico, decidimos recoger una muestra de los resultados de la investigación en lengua española, fundamentalmente en su vertiente histórica pero también sincrónica, y completarla con otras contribuciones solicitadas a expertos en el manejo de corpus y diccionarios como herramientas de trabajo. Con respecto a la sincronía, en “Anglicismos no asimilados en el *DRAE* (23ª ed.)”, García Andrevia revisa los anglicismos del *DRAE* (2014) en dos corpus del propio banco de datos de la institución, el *CREA* y el *CORPES XXI*. El autor comprueba que la nueva edición del diccionario ha incorporado 195 anglicismos en cursiva y señala que, cuando la Academia considera que el préstamo no es la forma preferida, remite a expresiones adaptadas o castizas. Sin embargo, el estudio arroja que para un número importante de extranjerismos, como (*blue*) *jean*, *boy scout*, *camping*, o *top-model*, entre otros, la remisión a voces españolas no responde al criterio de mayor frecuencia de uso. Según García Andrevia, esto se explica por el celo de la institución académica tanto por mantener estable el sistema gráfico-fónico del español, como por evitar el desplazamiento de su léxico por voces foráneas.

Por su parte, García Salido (“La frecuencia de corpus como criterio para nivelar colocaciones léxicas”) mide hasta qué punto la frecuencia de uso de las colocaciones léxicas coincide con la práctica docente del español como segunda lengua, y si responde a los conocimientos de los

aprendientes. Los resultados confirman la validez de la frecuencia de uso como criterio fundamental para nivelar las colocaciones. No obstante, son menos concluyentes los resultados con respecto a la intuición de los profesores sobre la frecuencia colocacional y la correlación entre el nivel asignado a las colocaciones y la dificultad para los aprendientes.

Desde una perspectiva histórica, Molina Sangüesa estudia el tecnolecto matemático renacentista en “La confección del *Glosario de aritmética y álgebra en el Renacimiento hispano* como instrumento para el estudio del léxico matemático del siglo XVI”. A partir del corpus del *Diccionario de la Ciencia y de la Técnica del Renacimiento (DICTER)* y con el apoyo de herramientas informáticas cuyo uso, explotación y dificultades también aborda aquí, esta autora elabora un glosario de especialidad sobre la aritmética y el álgebra del siglo XVI. Sus primeras conclusiones destacan, por un lado, un origen mayoritariamente latino de las voces en la etapa de inicio de la terminología; y por otro lado, el recurso a la derivación como mecanismo neológico interno más productivo.

Siguiendo con las aportaciones léxicas en áreas específicas, en “Recepción de léxico de confitería decimonónica en diccionarios del español” de Torres Martínez se analizan las voces de especialidad compiladas en un tratado culinario decimonónico sobre confitería, pastelería y destilería. Centrada en las voces adscritas al campo de la confitería, la autora observa qué términos se registran en los diccionarios generales del español, su fecha de lematización y el modo de incorporación en los repertorios. Este trabajo supone una aportación al ámbito de la alimentación, una de las áreas de estudio compartidas por la antropología y la lexicología, de gran interés para la historia de la vida social.

Por su parte, “Otra mirada al rincón de la Lexicografía: Cuba” de García Aranda describe cuatro repertorios léxicos publicados entre 1840 y 1893 en Cuba, destinados a la enseñanza de lenguas extranjeras. Se trata del *Vocabulario auxiliar de bolsillo español e inglés* de Juan José Romero (Matanzas, 1840), el *Arte de hablar bien francés* de Pierre Nicolás Chantreau versionado por Pablo Dupuy (La Habana, 1850), *El recreo de la juventud* de Alfredo Carricaburu (La Habana, 1891) y la anónima *Illustrated Guide Book of the Island of Cuba* (La Habana, 1893). La autora analiza este material como documentación que permite conocer mejor la historia del léxico americano, tanto en sus voces autóctonas como en la presencia y adaptación de voces de procedencia española y extranjera.

Con una perspectiva histórica más extendida, Ramírez Luengo, en “Los corpus lingüísticos en la historia del léxico: algunos datos sobre la generalización de los indigenismos antillanos en el español de España”, aborda la evolución de los americanismos a partir de una interpretación de este concepto basada exclusivamente en el uso, y de naturaleza eminentemente dinámica. El autor analiza el empleo de estas voces entre los siglos XVI al XIX en CORDE, considerando tanto su porcentaje de aparición en América y España, su presencia en obras españolas de contenido americano y no americano, así como su extensión por géneros textuales y temáticas variadas. De este modo, Ramírez Luengo consigue explicar el tránsito de ciertos indigenismos antillanos de lo que denomina ‘americanismos puros’ a ‘voces generales’, diatópicamente no marcadas.

El volumen se cierra con la aportación de Vila Carneiro y Faya Cerqueiro, “Fórmulas de despedida de matiz religioso en las cartas del siglo XVII”. Las autoras analizan los actos de habla de despedida en cartas del siglo XVII, de temática variada y de diferentes zonas geográficas. Sus conclusiones subrayan el alto grado de convencionalización de las expresiones con referente

religioso, en lo relativo a las formas verbales empleadas (*guardar* y *prosperar*) y a la divinidad invocada (*Dios* y *Señor*), así como su amplia difusión tanto en lo cronológico como en lo geográfico.

En definitiva, los artículos seleccionados no solo constituyen una aportación esencial para la vertiente histórica del español, como son los aspectos formularios de la correspondencia, el léxico específico o la difusión del léxico americano, sino que también son una contribución fundamental a cuestiones tan actuales como el tratamiento de los anglicismos por parte de la Real Academia Española y la reflexión sobre la enseñanza de las colocaciones léxicas del español.

