

Bases de datos y corpus para la investigación y la enseñanza de la lengua y la literatura: una introducción

Cada vez son más los estudios de carácter lingüístico y literario que se enmarcan o se aproximan a lo que actualmente llamamos Humanidades Digitales. El uso de bases de datos y corpus en la investigación, sin embargo, no es algo nuevo, si bien estos instrumentos se han ido refinando con el tiempo, facilitando así la labor investigadora y mejorando el alcance de esta.

En el presente monográfico se ponen a disposición del lector nueve artículos de naturaleza muy diversa que se ocupan del empleo de estas herramientas no solo para la investigación en lengua y literatura, sino también para su enseñanza. Pretende ser, en suma, una muestra representativa de lo que hay y de lo que queda por hacer que esperamos resulte de gran utilidad a investigadores y docentes de lenguas romances.

La ventaja que la lingüística lleva a los estudios literarios en el uso de las tecnologías (si bien paulatinamente se va acortando) es perceptible también en el número de contribuciones sobre este campo que se incluyen en nuestra publicación. Una de ellas es la de Cazorla y García-Aranda, quienes dan a conocer las múltiples posibilidades de explotación filológica y lingüística que la *Biblioteca Virtual de la Filología Española* nos brinda, un recurso gratuito y de cómodo uso que, sin duda, favorece la labor de quienes nos dedicamos a la investigación. Sobre herramientas también versa el artículo de Rodríguez Puente. Esta investigadora ofrece una visión general de aquellas disponibles para el estudio de la historia de la lengua española y hace hincapié en el hecho de que la lengua oral apenas tiene representación en los corpus diacrónicos del español, en comparación con lo que se observa en la tradición anglosajona. Por su parte, Ivana Lončar, en su estudio de caso vinculado a la elaboración del *Refranero bilingüe español-croata*, analiza la utilidad de diccionarios, bases de datos y corpus para la investigación lingüística y deja constancia de la necesidad de ser cautos en su uso.

La lengua, en este caso la gallega, también es protagonista en la contribución aportada por Dopazo Entenza, que detalla cómo se empezó a dar forma al CORILGA (*Corpus Oral Informatizado da Lingua Galega*): la composición, los criterios de selección, la transcripción... y expone las posibilidades de utilización que tiene, así como en qué estado se encuentra actualmente y cuáles son los proyectos de futuro que el equipo de investigación que está detrás de él contempla. En la misma línea de investigación de Dopazo, se halla el artículo de Nuria Polo, en el cual propone una serie de pautas para la confección de corpus

orales, referidas fundamentalmente a los participantes, las grabaciones y las herramientas.

A medio camino entre la lingüística y la didáctica, se encuentra la contribución de García, Vincze, Orol y Alonso, donde se examina un corpus de lecturas graduadas destinadas a aprendices de español como lengua extranjera con el objetivo de comprobar si se establece alguna relación entre la frecuencia léxica y la secuenciación del vocabulario en materiales ELE.

Más centrado en la didáctica, en este caso de la literatura, está el trabajo de Alicia Vara, quien señala la infinidad de posibilidades que una herramienta como *TESO (Teatro Español de Siglo de Oro)* puede ofrecer en el aula del Grado de Educación Infantil para trabajar los estereotipos de género.

En el ámbito literario, contamos con el trabajo de Álvarez Cifuentes, que pone de relieve la utilidad de las nuevas herramientas para el estudio de las literaturas ibéricas, centrándose especialmente en el nuevo corpus de Literatura Hispano-Portuguesa de la Fundación Biblioteca Virtual Miguel de Cervantes, dirigido por José Miguel Martínez-Torrejón. Por su parte, Blanca Ripoll da cuenta de cómo se gestó la base de datos que recoge las publicaciones del semanario barcelonés *Destino* y reflexiona sobre las diversas problemáticas que la configuración de esta herramienta acarreó y las múltiples posibilidades que esta brinda a la producción científica en estudios culturales.

Todos estos artículos son solo una muestra de la importancia que bases de datos y corpus revisten en la investigación de la lengua y la literatura romances, y con ellos esperamos contribuir a completar bibliográficamente un campo de estudio que se encuentra en alza en la actualidad.

**Zaida Vila Carneiro (ed.)
Universidad de La Rioja**