

Rebeca Lázaro Niso (ed.), *Corpus y bases de datos para la investigación en literatura*, Logroño, Fundación San Millán de la Cogolla, 2017, 183 págs.

DOI: <https://doi.org/10.24197/cel.9.2018.CIII-CVIII>

Rebeca Lázaro Niso, con la edición de *Corpus y bases de datos para la investigación en literatura*, ofrece a la comunidad científica un volumen donde se conjugan los dos polos característicos de las humanidades digitales: el conocimiento humanístico de los investigadores y la aplicación de las nuevas tecnologías a una disciplina humanística; aquí la literatura. Ello se logra mediante la inclusión de trabajos de diferente cariz aunados en unas tipologías definidas por la editora en la “Presentación”:

- los que manifiestan el contraste entre propuestas innovadoras basadas en las nuevas tecnologías y la investigación tradicional,
- los más centrados en los usos de las bases de datos y los corpus en relación con la investigación y
- los orientados a la transferencia de resultados, con énfasis en la difusión y la aplicación didáctica.

La utilidad de los corpus y las bases de datos vertebrará las tres tipologías y el total de las aportaciones.

El primero de los trabajos, presentados en orden alfabético, pertenece a Josefa Badía Herrera: “Bases de datos para el estudio del teatro clásico español. Propuesta de modelo para el análisis de microsecuencias”. Guiada por la noción de que las nuevas tecnologías habilitan novedosos paradigmas para los estudios literarios, la autora articula una contribución encauzada hacia el análisis microsecuencial del teatro del Siglo de Oro; para ello, acude a la base de datos *ARTELOPE*, dirigida por Joan Oleza. Se centra, como muestra, en la comedia inédita *Los infortunios del conde Neracio* (Real Biblioteca, ms. II-460-17); el resultado es contundente sobre la eficacia de *ARTELOPE*. A partir del significativo dato de que la densidad por palabra de *Los infortunios del conde Neracio* (5,04 versos por réplica) supera la media de la comedia palatina (3,5 versos por réplica), género al que se adscribe, Badía Herrera escruta los factores desencadenantes de esta peculiaridad. Tras atender a varios valores (relación y reacciones de los personajes, número de intervenciones, referencia al pasado o al futuro en las intervenciones, etc.), concluye que el desajuste obedece a que se destinan largos parlamentos a la descripción

del modo de actuar de los personajes, con lo cual la acción gana en morosidad.

Shai Cohen, con “Apertura, difusión y accesibilidad: el uso teórico y práctico de una base de datos”, alude a la necesidad de diseñar y fomentar herramientas que recojan y ordenen la inmensidad de datos de la que dispone el investigador. Reclama así la existencia de bases de datos consistentes y de fácil acceso para la comunidad científica. Para evidenciar la importancia de dichas plataformas, desarrolla dos ideas: primero aclara algunos de los tecnicismos propios de las humanidades digitales, así como problemas teóricos imprescindibles; después incide en dos proyectos en que se ha involucrado: uno sobre bilingüismo en el campo de refugiados de Konik, donde colaboró activamente, y otro orientado hacia la blogosfera *Grisosfera*, la cual desde su nacimiento en 2012 ha actuado como vínculo entre investigaciones especializadas y público general en el marco de los estudios auriseculares.

“Apostillas para un estudio del corpus sacramental de Lope de Vega”, de Juan Manuel Escudero Baztán, es el tercer capítulo. El investigador expone las líneas del proyecto en que está inmerso, referente a los autos sacramentales de Lope de Vega, poco estudiados pese a su relevancia en un género tan esencial para el teatro áureo como este y, específicamente, para su configuración en el siglo XVII. Destaca Escudero Baztán lo imperioso de editar y escudriñar el corpus sacramental lopesco, no solo para la comprensión del género, sino también a propósito de Lope y la estética aurisecular; hasta la fecha solo Menéndez Pelayo, aún en el siglo XIX, ha editado el conjunto de estos autos sacramentales. Son obras que, en algún caso, han sido atribuidas a otros autores por error y que requieren ediciones apropiadas que den pie a análisis sólidamente asentados en el texto, de lo cual se está encargando el Grupo de Investigación Siglo de Oro (GRISO), radicado en la Universidad de Navarra. El investigador señala dos fases del proyecto. En primer lugar, dados los inexcusables problemas de autoría de los autos sacramentales supuestamente lopescos, optaron por trabajar con los seguros: entre otros, los cuatro publicados en *El peregrino en su patria* y otros de la recopilación que, en 1644, llevó a cabo el licenciado José Ortiz de Villena. En la nueva fase investigan diez nuevos autos sacramentales de autoría también en firme. Es reseñable el panorama textual de los diez autos sacramentales que traza Escudero Baztán.

Simon Kroll, a través de “Calderón como poeta. Sobre el uso de las humanidades digitales para el estudio métrico”, pone de relieve la faceta de Calderón de la Barca como versificador y como poeta, que habría que

tener en cuenta con la de dramaturgo, que es la que le ha otorgado un lugar en la historia literaria española. Sugiere Kroll que la crítica ha descuidado la veta poética de Calderón, y para demostrar su importancia se sirve de las comedias *El secreto a voces* y *Para vencer amor, querer vencerle*, de donde entresaca ejemplos que resaltan la labor versificadora de Calderón y que, de paso, fundamentan el estudio del panorama de las asonancias calderonianas. Insiste el investigador en proseguir con el estudio de Calderón como poeta mediante las humanidades digitales, que podrían propiciar, por ejemplo, el diseño de una base de datos de tales asonancias con la que avanzar en el conocimiento de la maestría versificadora de Calderón y, más ampliamente, en los fructíferos nexos entre métrica y nuevas tecnologías.

El siguiente trabajo, como algún otro, está firmado por dos autores, Raquel López Sánchez y Borja Navarro Colorado: “Propuesta teórica y metodológica para el desarrollo de un corpus digital representativo del Romancero nuevo”. Desde el principio de que el romancero nuevo surge en el seno de una generación de autores que, hacia 1580, encuentra en el cauce métrico, y formal, del romance una vía para vehicular sus inquietudes estéticas y temáticas, los investigadores convergen en la necesidad de promover un específico corpus digital para él. Para ello, repasan los proyectos realizados hasta la fecha que, con la mediación de las nuevas tecnologías, han intervenido sobre la corriente poética que aquí interesa, a los cuales agregan las principales publicaciones para su estudio bajo el prisma de “la construcción de la cultura literaria hispánica” (p. 86), y que abarcan desde mediados del siglo XX hasta 2014. El trabajo se detiene en el proyecto de corpus del romancero nuevo que está en marcha, con lo que se expone toda una suerte de dificultades inherentes a la tradicional labor de la historiografía literaria que se entreveran, y se acucian, con las propias de las nuevas tecnologías. El resultado, sin embargo, merece el esfuerzo invertido: el corpus permitiría trabajar más fácilmente con tan espinosos asuntos como el de la autoría o la multiplicación de las variantes textuales. Se utilizan técnicas provenientes del procesamiento del lenguaje natural.

María del Rosario Martínez Navarro y Marta Rodríguez-Manzano encierran en “«Al bien hacer jamás le falta premio»: herramientas digitales para la didáctica y el estudio de la lengua y la literatura hispánica” una propuesta pedagógica original. En el capítulo las autoras mencionan y explican una serie de recursos tecnológicos que pueden utilizarse en el aula para motivar al alumnado y potenciar su aprendizaje: *ARTELOPE*,

*BIDISO*, *Biblioteca Virtual Miguel de Cervantes*, *Bieses*, *Canon 60 - TC12*, *Corpus CHARTA*, *PARES*, *PHEBO*, etc. No conciben estas herramientas como privativas del aula universitaria, sino que pueden abrirse, y conviene hacerlo, a otros niveles educativos. Asimismo, incorporan un grupo de propuestas para ejercitarse con las herramientas presentadas, que no deben quedar relegadas al uso exclusivo de los investigadores. La docencia se ve así dotada de nuevos instrumentos especialmente efectivos de cara a la lectura de los textos y su estudio en el aula.

“Visualización del corpus sacramental calderoniano: una línea del tiempo”, a cargo de Davinia Rodríguez Ortega y Álvaro Baraibar, pone de manifiesto los beneficios de situar sobre un eje cronológico los autos sacramentales de Calderón de la Barca y glosarlos con información sobre su título, autoría, datación, testimonios y representaciones acreditadas, entre otras opciones. Se nutre el planteamiento de la investigación centrada en los autos sacramentales calderonianos que viene promoviéndose en el GRISO desde hace más de dos décadas, bajo la dirección de Ignacio Arellano. De este modo, han visto la luz cuidadas ediciones con las cuales se ha avanzado considerablemente en la comprensión de unos textos bastante menos explorados de lo necesario antes de la labor del GRISO. No restan muchos títulos por editar y de esa circunstancia se aprovecha el capítulo: teniendo a mano la mayor parte de los datos, es posible configurar ya una línea cronológica significativa. Una herramienta ajustada a estos parámetros facilitaría la transmisión de la dilatada investigación desarrollada en el GRISO a un público amplio y a niveles educativos no necesariamente universitarios.

“Repertorios y bases de datos sobre poesía catalana de las épocas medieval y moderna: estado de la cuestión”, de Helena Rovira Cerdà, funciona como catálogo de las herramientas digitales principales para el estudio de la poesía catalana medieval y moderna, si bien no se limita a este espectro: menciona, igualmente, recursos generales sobre poesía hispánica medieval. En esta última línea, se recurre a *Philobiblon*, compuesto a su vez de cuatro bibliografías (*BETA*, *BITAGAP*, *BITECA* y *BIPA*), amén de a *MedDB* y *MetCa*, no sin antes aducir unas breves reflexiones sobre la lírica trovadoresca y la lírica catalana medieval. Ahora bien, el foco del trabajo reside en la poesía catalana, cuyos corpus y bases de datos se organizan en torno a la Edad Media y a la Edad Moderna. De la Edad Media se nombran herramientas como *RIALC*, *Cançoners DB* y *The Last Song of the Troubadours*, mientras que de la Edad Moderna se

citan menos: *Nise* y *MCEM*. Rovira Cerdà apunta la necesidad de más corpus, sobre todo de poesía catalana moderna. De este modo, cobra lógica el proyecto que realiza en compañía de Johan Mahiques: el “Cens de poesia catalana de l’Edat Moderna”, que se alimenta de la tesis doctoral de este segundo investigador, en que ha recogido “un censo de autores y poemas en catalán conservados en impresos datados entre 1474 y 1814” (p. 137): el *Repertori d’Obres en Vers*. Desde los parámetros anteriores, Rovira Cerdà y Mahiques aspiran a volcar la información recogida en una base de datos y a complementarla. Aunque en el capítulo aún no se pueda consignar la existencia del censo en línea, que estaba en construcción, hoy en día sí que está listo para su consulta. El repertorio *Cens de Poesia Catalana de l’Edat Moderna* se convierte, así las cosas, en una herramienta imprescindible para todo investigador en poesía catalana medieval y moderna.

Isabel Sainz Bariain, en “Corpus de bibliógrafos para el estudio de autores novohispanos”, muestra un selecto número de referencias bibliográficas mexicanas que penetran en la realidad novohispana, lo que conforma una rica compilación para el conocimiento del período colonial. Sus aportes comprenden el arco que se prolonga desde el siglo XIX hasta el XXI, bien que con predilección por el XIX y el principio del XX. Se apoya en tres cortes cronológicos: a) desde el siglo XIX hasta mediados del XX, b) del siglo XX hasta la actualidad y c) finales del siglo XX y comienzos del XXI. Cada apartado contiene información sobre distintos autores y sus obras y, en consecuencia, sobre cómo interpretaron el pasado colonial.

Como penúltimo trabajo, está “Funcionalidad de un corpus áureo para la investigación literaria: un ejemplo de interrelación”, de Simón Sampedro Pascual. La contribución aborda la emblemática y presta atención a los corpus y proyectos más destacados al hilo de este género: *Corpus de sonetos castellanos del Siglo de Oro*, *Corpus de Textos Espirituales del Siglo de Oro*, *BIDISO*, etc. No sortea Sampedro Pascual los problemas concernientes a la emblemática como género, en todo momento contemplada a la luz de sus propiedades en el auge de la estética barroca y en el siglo XVII, “cuando la empresa emerja como forma artística donde lo abstracto cobra visibilidad” (p. 163). A pesar del despunte barroco, no puede menospreciarse el avance del género a lo largo del XVI, gracias, entre otros, a la labor de dos precursores: Andrea Alciato y Juan de Horozco y Covarrubias. Tras la definición general, se adentra en las empresas políticas, cuyo máximo artífice no fue otro que Diego de

Saavedra Fajardo. Como ejemplo, se recurre a *Ganar por la mano el juego*, comedia de Álvaro Cubillo de Aragón, donde las empresas políticas hacen “constantes llamadas al espectador” (p. 167).

Sara Sánchez Bellido y Nicolás Asensio Jiménez ofrecen una última aproximación en “El Archivo Digital del Romancero: metodología y desarrollo de un proyecto de preservación”. Se muestran el sentido y la utilidad de la digitalización a la cual se están sometiendo los fondos del Archivo del Romancero Menéndez Pidal-Goyri, objeto de un proyecto de investigación orientado a la visibilidad y la accesibilidad del corpus, para lo cual se ha optado por la creación de un repositorio que almacena la versión digitalizada de los documentos y sus metadatos, gracias al sistema Dublin Core. La base de datos pretende así alojar una información fundamental para todo especialista en el romancero tradicional hispánico, de tal suerte que con el acceso podrá disponer de un corpus esencial y, sin embargo, “aún hoy inédito en su mayor parte” (p. 173). Se menciona, entre otras cosas, el descomunal esfuerzo de Ramón Menéndez Pidal por reunir cuantos romances pudo y su deseo de que estos y otros se publicasen en la totalidad de sus variantes, fin inalcanzable hasta ahora. Es el momento de que, mediante el apoyo de la tecnología, la idea de Menéndez Pidal pueda materializarse.

En conclusión, Rebeca Lázaro Niso edita un libro cargado de reflexiones sugestivas, compendios prácticos y proyectos prometedores alrededor de los corpus y las bases de datos en su aplicación al estudio de la literatura. Estas nuevas herramientas, enraizadas en el auge de las humanidades digitales, han traído consigo nuevos problemas a los estudios literarios y nuevas soluciones. Pero sobre todo nuevas maneras de abordar esta realidad tan compleja, y tan humana, que es la literatura.

PEDRO MÁRMOL ÁVILA  
Universidad Autónoma de Madrid  
[pedro.marmol@uam.es](mailto:pedro.marmol@uam.es)