

Corbella, Dolores; Fajardo, Alejandro y Langenbacher-Liebgoth, Jutta (eds.), *Historia del léxico español y Humanidades digitales*, Berlín, Peter Lang, 2018, 498 pp.

El oleaje producido por la revolución digital y la irrupción de las nuevas tecnologías en el ámbito investigador también ha alcanzado las costas de lo lingüístico, generando especial impacto en las disciplinas que se centran en el estudio del léxico, caracterizadas en muchos casos por su estrecha relación con los testimonios “en papel” – físicos –: lexicografía, geografía lingüística, estudios de corte diacrónico, etc. El nacimiento en los últimos años de corpus documentales, atlas lingüísticos en línea, bases de datos o la digitalización de materiales ha permitido al investigador tener acceso inmediato a una enorme cantidad de información antes inimaginable, que, por otro lado, no solo plantea ventajas, sino también nuevos retos y problemas derivados de una cierta falta de sistematicidad o estandarización.

Esta compilación de estudios centrada en el impacto que el futuro digital ejerce en el estudio histórico del léxico se presenta dividida en tres bloques principales. El primero de ellos – que lleva por título *Corpus y recursos actuales* – comprende nueve artículos en los que se presentan y describen algunos de los corpus más destacados, siendo Mar Campos Souto (Universidad de Santiago de Compostela) la que inaugura el bloque con “Las bases documentales del NDHE: Entre la realidad y el deseo”, que desarrolla un análisis de algunas de las fuentes principales de las que se vale el proyecto del *Nuevo Diccionario Histórico del Español* – dirigido por José Antonio Pascual –, entre las que encontramos una intersección entre la tradición – como el *Fichero General* de la Academia – y la innovación digital, faceta que representan los corpus académicos – destacando el papel del propio *Corpus del Diccionario Histórico* (CDH) y los problemas que supone el empleo de este tipo de bases de datos (representatividad, fiabilidad, lematización, etc.) – y la digitalización de fondos documentales. La comparativa entre el *Corpus del Español* y los corpus académicos CORDE, CREA y CORPES XXI constituye el punto central del estudio de Mark Davies (Brigham Young University) “Uso del *Corpus del Español* y los corpus relacionados para la lexicografía histórica española”. A través de ejemplos de búsqueda, Davies presenta su *Corpus del Español* como una herramienta que permite realizar búsquedas mucho más acotadas y precisas que las que pueden ofrecer los corpus académicos, dado que estos presentan una interfaz mucho más limitada, no pudiendo ofrecer datos como la frecuencia de uso o formas de búsqueda de tipo formal o semántico, aspectos que permiten una profundización mayor en el estudio histórico del léxico. Virginia Bertolotti (Universidad de la República, Uruguay) y Concepción Company Company (Universidad Nacional Autónoma de México) nos invitan a cruzar el charco de la mano de “El corpus para América: *CORDIAM*”, capítulo en el que presentan aquellos aspectos principales que caracterizan al *Corpus Diacrónico del Español de*

América — tanto documentales como informáticos —, presentando así una herramienta fundamental en el estudio del español americano que incluye testimonios desde el siglo XVI hasta los siglos XIX-XX y cuyo espectro textual abarca tanto textos literarios como periodísticos, notariales, cronísticos o epistolares. Aún en la orilla americana del Atlántico nos encontramos con el “Tesoro léxico de los americanismos contenidos en los vocabularios hispano-amerindios coloniales (1550-1800) [TELEAM]” en el que Esther Hernández (Instituto de Lengua, Literatura y Antropología, CSIC) presenta este proyecto que persigue desarrollar una herramienta de estudio del léxico histórico americano — especialmente en el ámbito de la neología y la etimología — reivindicando el papel e importancia de la lexicografía misionera, esto es, a partir de vocabularios bilingües “español-lenguas indígenas” que los misioneros elaboraron durante los casi 3 siglos de periodo colonial, un recurso alternativo a las fuentes clásicas como las crónicas. La editora del volumen, Dolores Corbella (Universidad de La Laguna), nos propone un paseo por la historia de los “tesoros lexicográficos” con “Del tesoro lexicográfico analógico al digital”, una labor iniciada por Samuel Gili Gaya y su *Tesoro lexicográfico* que serviría de ejemplo a ambos lados del Atlántico y que ha tenido especial repercusión en el ámbito dialectal, en el que podemos destacar obras como el *Tesoro léxico de las hablas andaluzas*, el *Léxico del leonés actual*, el *Tesoro lexicográfico del español de Puerto Rico* o, ya en el plano digital, proyectos tan monumentales como el *Nuevo Tesoro Lexicográfico de la Lengua Española* de la Real Academia, un recurso que, entre otros aspectos, ofrece una panorámica no solo de la historia del léxico, sino también de la propia evolución de la técnica y labor lexicográficas. En “Una nueva fuente para la historia del léxico español: el DEMel”, Rafael Arnold, Stefan Serafin, Anna-Susan Franke y Jutta Langenbacher-Liebgott (Universität Rostock y Universität Paderborn) — editora del volumen — presentan un proyecto lexicográfico centrado en la digitalización del *Diccionario del español medieval* de Bodo Müller (1987-2005): el *Diccionario del español medieval electrónico*, del que nos presentan su “estado de la cuestión”, esto es, la configuración de las 1700 fichas en papel de las que dispone el proyecto y la apariencia de la interfaz de búsqueda y sus futuras aplicaciones de cara al estudio del léxico histórico desde los primeros testimonios del español hasta el siglo XV. Uno de los grandes trabajos pendientes que el mundo de la filología hispánica aún no ha conseguido tachar de su lista es la culminación del *Atlas Lingüístico de la Península Ibérica* — el ALPI — de Tomás Navarro Tomás. En este antepenúltimo capítulo del primer bloque, “Geolingüística y Humanidades digitales: el *Atlas Lingüístico de la Península Ibérica* (ALPI)”, Pilar García Mouton (Instituto de Lengua, Literatura y Antropología, CSIC) revela el estado actual del proyecto y demuestra la gran utilidad que las nuevas tecnologías pueden tener en un campo como es el de la geolingüística a través de la pregunta 520a, referida a los nombres que recibe el *jabalí*. El “ALPI 2.0” se revela como una utilísima herramienta a la hora de determinar aspectos lingüísticos como “la existencia o la ausencia, la extensión y la vitalidad de una voz” (p. 189). Las materias que pueden dar lugar al nacimiento de un nuevo diccionario son prácticamente infinitas, afirmación que nos demuestra Rolf Eberenz (Université de Lausanne) con su “Hacia un diccionario de la alimentación y la culinaria medievales y renacentistas”, un ambicioso proyecto que tiene como objetivo ofrecer un panorama léxico —aunque

también sociocultural— del vocabulario culinario castellano de los siglos XIII al XVII y del que Eberenz nos proporciona información relativa al corpus base y a algunos aspectos macro y microestructurales. Por último, el cierre del primer bloque les corresponde a José Calvo Tello, Ulrike Henny-Krahmer y Christof Schöch (Universität Würzburg y Universität Rostock), quienes nos presentan “*Textbox: análisis del léxico mediante corpus literarios*”, un conjunto de nueve corpus multilingües de corte literario que permiten un estudio del léxico que combina la perspectiva lingüística con la literaria —frecuencias de palabras y presencia en obras de determinados periodos, distribución de voces, etc.—, amén de ilustrar las ventajas que ofrece la anotación en formato TEI y el empleo de *tokens*.

El segundo bloque, *Crítica de los recursos en línea: el desorden digital*, pretende reflejar el “lado oscuro” del mundo digital, esto es, el desfase existente entre la actualización y creación de nuevos soportes digitales y la falta de un criterio homogéneo a la hora de llevar a cabo tareas como el tratamiento de los textos o la procedencia de los materiales; en palabras de los propios editores “se están creando las autopistas de la información, pero no siempre las rotondas que permitan la optimización de los resultados” (p. 9). Así, el editor del volumen Alejandro Fajardo (Universidad de La Laguna) nos ofrece en “*Lexicografía histórica con corpus y recursos digitales: aspectos metodológicos*” un exhaustivo análisis comparativo de múltiples corpus diacrónicos —generales y especializados— y de algunos recursos en línea tales como bibliotecas y hemerotecas digitales, análisis que permite evidenciar sus diversas *raisons d’être* y la falta de sistematicidad derivada de la inexistencia de modelos comunes en el campo de la digitalización de materiales. Continuando con esta línea de “problemas 2.0”, Francisco Javier Herrero Ruiz de Loizaga (Universidad Complutense de Madrid, Instituto Universitario Menéndez Pidal) se centra en “*Algunos problemas en la aplicación de los corpus informatizados al estudio de la diacronía del español, con especial atención a los procesos de gramaticalización*”. Si bien es cierto que los corpus se constituyen como una herramienta fundamental en el ámbito de los estudios lingüísticos de corte diacrónico, lo cierto es que el investigador debe andarse con ojo, pues no se encuentran exentos de errores como fallos en la lematización o la transcripción, digitalizaciones incorrectas, erratas, etc., o carencias tales como desigualdad del número de fuentes textuales en los distintos periodos, errores que no ofrecen una lectura clara de los datos históricos y que, por ende, dificultan un poco más —si es que era posible— el estudio histórico del léxico.

Por último, el tercer bloque —*Del corpus a los estudios léxicos*— nos ofrece una relación de siete estudios que demuestran las posibilidades reales —prácticas— que el mundo de los corpus digitales puede ofrecer a la disciplina lexicográfica. Miguel Calderón Campos (Universidad de Granada) nos presenta la utilidad del *Corpus diacrónico del español del reino de Granada* en el estudio de los andalucismos: en “*Andalucismos en el corpus del reino de Granada*” se nos presenta el análisis de diversas voces que el *DLE* considera como estándares, si bien *CORDEREGRA* —un corpus con un perfil mucho más específico que los corpus generales— proporciona datos que evidencian su condición de voces dialectales andaluzas y que nos ayudan a

precisar su significado concreto, amén de permitir observar la vitalidad y continuidad de algunos de ellos no solo en Andalucía, sino también en dominios lingüísticos estrechamente relacionados con las variedades andaluzas, como son el español de Canarias o el español de América. El mundo de la documentación notarial – también presenten en *CORDEREGRA* – se revela como una importante fuente en el estudio del léxico, de ahí que Pedro Sánchez-Prieto Borja y Delfina Vázquez Balonga (Universidad de Alcalá) nos ofrezcan en “El léxico en los documentos de la Comunidad de Madrid (ss. XVI-XIX)”, una parte del proyecto del *ALDICAM-CM* que tiene como fin el estudio del léxico de la vida cotidiana madrileña entre los siglos XVI y XIX a partir de archivos municipales de la Comunidad de Madrid de diversa índole: inventarios de bienes, expedientes judiciales, memorias de obras públicas, etc., en los que puede encontrarse un enorme caudal léxico de ítems relativos al cuerpo humano, la indumentaria o la vida doméstica que permiten, entre otros aspectos – como la caracterización de la lengua de los distintos periodos –, dibujar de una manera más precisa la realidad lingüística de la Comunidad de Madrid a lo largo de los tres siglos que comprende el corpus documental base. Es, precisamente, el *Atlas Lingüístico Diacrónico e Interactivo de la Comunidad de Madrid* el que María Jesús Torrens (Instituto de Lengua, Literatura y Antropología, CSIC) emplea como tema principal en el capítulo “El proyecto *ALDICAM-CM* y el ejemplo de los fueros de Alcalá para el estudio de la historia del léxico”, un proyecto centrado en la elaboración de un atlas lingüístico de la Comunidad de Madrid que cuenta con un corpus de documentos archivísticos fechados entre el siglo XIII y el XIX, ofreciendo de este modo una interesante perspectiva de la evolución de la lengua hablada y escrita en Madrid que se postula como una magnífica herramienta en el estudio histórico del léxico – como demuestra la autora a partir de la comparativa del *Viejo fuero de Alcalá* (s. XIII) y el *Nuevo* (s. XVI) –. El proyecto *CorLexIn* también concentra sus esfuerzos en el estudio del léxico de la vida cotidiana a partir de documentos notariales, si bien en este caso el marco temporal se sitúa en el siglo XVII. En “El proyecto *CorLexIn* y la variación diatópica en el léxico del Siglo de Oro”, José Ramón Morala Rodríguez y María Cristina Egido Fernández (Universidad de León) presentan algunas de las múltiples posibilidades que ofrece el *Corpus Léxico de Inventarios*, entre las que destacan la documentación de localismos que, en la mayoría de los casos, aparecen infrarrepresentados en los corpus generales de la Academia o no figuran con la marcación diatópica que les correspondería; la definición de isoglosas léxicas más precisas y definidas que permiten establecer el ámbito de uso de determinadas voces; o el estudio de los préstamos lingüísticos y su grado de integración en el caudal léxico general, aspecto que no siempre se corresponde con los datos ofrecidos por los corpus generales. En esta misma línea, aplicada en este caso al estudio del español medieval, Miguel Ángel Puche Lorenzo (Universidad de Murcia) destaca la importancia del uso de fuentes documentales alternativas como la prosa jurídica. En “Estudio del léxico castellano a través de fuentes medievales murcianas”, Puche se vale de documentación jurídica inédita del siglo XIV para ilustrar el potencial de este tipo de textos, amén de profundizar en el estudio del castellano “periférico” y su contribución en la configuración del léxico español. Casi rozando el final del volumen se encuentra el capítulo dedicado a “Fuentes lexicográficas y estudio del léxico: el *Diccionario de la*

lengua castellana de la Real Academia Española (1817-1852)”, obra de M.^a Ángeles Blanco Izquierdo, Gloria Clavería Nadal y Enrique Jiménez Ríos (Centro de Estudios de la RAE, Universidad Autónoma de Barcelona y Universidad de Salamanca) en la que analizan la evolución lexicográfica que puede apreciarse en las seis ediciones del *Diccionario de la lengua castellana* comprendidas entre 1817 y 1852 –5.^a y 11.^a respectivamente–: influencias de obras académicas y no académicas, cambios macroestructurales, cambios en la definición, presencia y ausencia de marcación, etc.; lo que, claramente, permite una profundización en la historia lexicográfica académica y, por ende, una muestra de la evolución de la técnica lexicográfica hispánica a través de su obra por excelencia. Finalmente, José Ignacio Pérez Pascual (Universidade da Coruña) nos traslada a un periodo “desatendido” desde el punto de vista de la historia del léxico como es la Edad de Plata. A través del mundo de las publicaciones periódicas a las que puede accederse a través de medios como las hemerotecas digitales, “Las publicaciones periódicas y el estudio del léxico de la ‘Edad de Plata’ pretende reivindicar no solo la importancia léxica de ese medio siglo comprendido entre 1885 y 1936 –una nueva realidad que trae consigo un nuevo mundo de voces neológicas y préstamos–, sino también la necesidad de acudir a fuentes externas al ámbito puramente académico para “refinar considerablemente el conocimiento que poseemos acerca de la historia de nuestro vocabulario” (p. 495).

En conclusión, la obra combina perfectamente las tres perspectivas que deben tenerse en cuenta y sobre las que se debe reflexionar a la hora de abordar el estudio del léxico a través del uso de herramientas digitales: en primer lugar, el investigador debe conocer las herramientas de las que dispone y puede valerse para llevar a cabo su tarea, de ahí la importancia de reseñar los proyectos existentes actualmente o que se encuentran en fase de desarrollo en el ámbito de la historia del léxico. En segundo lugar, debe ser consciente de que, a pesar de las múltiples ventajas que dichos recursos digitales suponen frente a las fuentes tradicionales, el mundo “2.0” también adolece de ciertas carencias generadas, en su mayoría, por la falta de estándares y criterios homogéneos en la selección de materiales o el tratamiento de los datos, aspecto de considerable importancia y al que parece no habersele prestado la atención que merece. Es necesaria, por ende, una labor de reflexión y revisión tanto desde el punto de vista teórico como metodológico. Por último, una aplicación práctica de la teoría, esto es, una muestra real de las posibilidades que dichas herramientas digitales ofrecen para la historia del léxico, una muestra del avance que la era digital supone para una disciplina –o un conjunto de disciplinas– que, como indicábamos al principio, parece reacia a separarse de la seguridad que le otorga el papel; pero que, en el fondo, ya se ha convertido en compañera inseparable del mundo digital.

Alejandro Junquera Martínez