

## Introducción

Tradicionalmente, el estudio histórico de la lengua española se ha fundado en una mayoría de textos literarios, pues, desde sus inicios, la Escuela de filología española siguió el criterio de su fundador, Ramón Menéndez Pidal, quien dio prioridad al análisis lingüístico de tales escritos y relegó el de los documentos de archivo únicamente a los periodos desiertos de producción literaria, en particular a los orígenes de la lengua.

La superioridad de los textos literarios frente a los documentos de archivo se defiende por su calidad estética y el valor de sus contenidos, pero casi siempre los escritos literarios o asimilables (pongamos por caso las crónicas medievales) presentan no pocos inconvenientes para los estudios lingüísticos, entre los que destacan la falta de manuscritos autógrafos —pues generalmente los testimonios conservados están separados de los originales por las manos de uno o varios copistas y por años o hasta siglos de distancia— y la artificialidad estilística propia de los discursos literarios. En contraste, el interés filológico de los documentos de archivo reside esencialmente en el hecho de que su localización espacial y temporal suele ser precisa y fiable. Su muy gran diversidad temática y pragmática permite, por otra parte, estudios extremadamente variados.

El cuestionamiento epistemológico de la calidad de los datos sobre los que reposa nuestro conocimiento de la lengua del pasado, sumado a las posibilidades que han ofrecido en estas últimas décadas el desarrollo tecnológico y la financiación pública de proyectos de edición de documentos de diversos archivos españoles y americanos, han conducido a que se vayan constituyendo corpus documentales de diversa envergadura que permiten la exploración de fenómenos lingüísticos de modo sensiblemente diferente ahora en comparación con el siglo pasado. Entre esos corpus, pueden consultarse hoy en línea CODEA (*Corpus de Documentos Españoles Anteriores a 1800*), CORHEN (*Corpus Histórico del Español Norteño*), CORDEREGRRA (*Corpus diacrónico del español del Reino de Granada*), CorLexIn (*Corpus Léxico de Inventarios*), CODEMA (*Corpus*

*Diacrónico de Documentación Malagueña*), CODCAR (*Corpus de Documentos de Cancillería Real*), DITECA (*Diccionario de Textos Concejiles de Andalucía*) o CHEM (*Corpus Histórico del Español en México*).

El año 2005, gracias a la iniciativa de creación de la red internacional CHARTA (*Corpus Hispano y Americano en la Red: Textos Antiguos* <<https://www.redcharta.es/>>) por colegas de las universidades de Alcalá, Deusto, Murcia, Las Palmas de Gran Canaria y Valladolid, así como del Consejo Superior de Investigaciones Científicas, abre una nueva etapa para los investigadores especialistas de la historia de la lengua española, que progresivamente irán disponiendo de cada vez más datos lingüísticos de documentos de múltiples archivos que cubren el amplio arco cronológico que va del siglo XII al XIX.

Son muchos los fenómenos que ahora pueden estudiarse de manera cuantitativamente diferente gracias a los corpus que van recopilándose desde hace décadas y que se han ido poniendo en común en la red CHARTA, y resultan muy notables los avances que aportan algunas de las herramientas incorporadas a algunos de esos corpus, como la del cartografiado, que permite organizar dialectalmente las formas lingüísticas con unos cuantos clics. Todos estos desarrollos contribuyen sin duda a una mejora cualitativa de nuestra comprensión de los fenómenos lingüísticos del español en su devenir temporal.

El presente volumen colectivo reúne un conjunto de trabajos de historiadores de la lengua española que en la mayor parte de los casos ejercen también como editores de documentos de archivo y que manejan corpus documentales para estudiar fenómenos lingüísticos que ya han sido objeto de estudio en el pasado o para explorar otros no tratados hasta aquí, a la luz de nuevos datos a menudo masivos y geolocalizados con precisión. A través de las páginas de este libro, el lector podrá descubrir cómo el conocimiento de la historia de la lengua española se ha enriquecido considerablemente durante estos últimos años y, en ese sentido, cómo se reescribe esta historia gracias a la observación de los datos extraídos de la gran cantidad de documentos de archivo editados a partir del nuevo siglo.

Los dos primeros capítulos de este libro presentan algunos fondos archivísticos inéditos y nos informan de la creación de corpus documentales y su valor para el estudio histórico de la lengua española.

En el primero de ellos, Dolores Corbella («Fuentes archivísticas canarias») nos habla del *Corpus documental de las islas Canarias* (CORDICan),

que su grupo de investigación ha estado utilizando hasta ahora internamente para la elaboración de un diccionario histórico de ámbito regional (*Diccionario histórico del español de Canarias* – DHECan), pero que puede igualmente ser de gran utilidad para el estudio fonético-fonológico y morfosintáctico de la variedad canaria en épocas antiguas (abarca desde finales del siglo xv hasta los últimos años del xix). Este corpus modular diatópico de pequeño dominio está perfectamente datado y etiquetado, y está siendo volcado actualmente en la red para que esté disponible en acceso libre en un futuro próximo.

En el segundo trabajo, «Corpus de documentos castellanos redactados en Cataluña (siglos xviii y xix): inicio de un proyecto», Joan Torruella y Gloria Clavería nos muestran un nuevo corpus que, a pesar de encontrarse en su fase inicial, permite ya el análisis de diferentes aspectos lingüísticos de la modalidad del castellano hablado en tierras catalanas y sus interferencias con la lengua catalana. La finalidad del corpus ESenCAT es proporcionar datos, a través de un portal web, para poder reconstruir la compleja historia de la formación del castellano oral y escrito en Cataluña.

Otras dos contribuciones tienen por objetivo exponer las enseñanzas sobre la lengua que se extraen de dos corpus de documentos no explorados hasta la fecha.

Rolf Eberenz con «Fiestas cortesanas en las crónicas del siglo xv: representación discursiva y significación» analiza cómo se retratan las fiestas nobiliarias en una nómina de catorce obras cronísticas del siglo xv que incluyen episodios festivos referidos al periodo que va desde el reinado de Pedro I hasta el de los Reyes Católicos. Además de proporcionar una tipología de las diversas celebraciones nobiliarias reflejadas en las páginas de esas crónicas, Eberenz se fija como objetivo primordial explorar cómo se ha forjado el léxico asociado a esas fiestas (la sala y su mobiliario, las comidas, los manjares, la danza, la música y el teatro).

Pedro Sánchez-Prieto Borja para su estudio «Madrid y su tierra en la conformación del español moderno» parte de un corpus de cartas de petición de los siglos xviii y xix conservadas en los fondos de beneficencia de Madrid y que forman parte del proyecto ALDICAM, que pretende elaborar un corpus que sirva de base a un atlas lingüístico diacrónico de la Comunidad de Madrid. En estas cartas rastrea los rasgos fonéticos, morfológicos, sintácticos y léxicos y muestra, de un lado, cómo los documentos de tipo administrativo desempeñaron un papel significativo en la conformación de la lengua que hoy hablamos y escribimos, y, de otro lado,

cómo el desarrollo urbano de Madrid como corte y capital pudo también ser decisivo en la forja del español moderno.

En los siguientes tres capítulos se abordan cuestiones de grafemática, fonética y fonología.

Lola Pons Rodríguez («La intervención gráfica en la edición de textos españoles del siglo xv: tendencias y variaciones»), constata el nivel de variación interna dentro de la praxis editorial en un conjunto de ediciones de textos del siglo xv. Las cotas de preservación o de transformación del testimonio del que se parte varían entre editores, incluso cuando existe una comunidad de principios en cuanto al tipo de texto, de destinatario, de edición y de solidez de la práctica filológica en que se asienta el propio proceso ecdótico. En su análisis, Pons Rodríguez mide qué zonas del nivel gráfico se respetan más y cuáles menos en el volcado de la fuente al texto editado, revisando las propuestas que se han hecho para textos españoles y no españoles con relación a los estándares de representación gráfica editorial.

Hiroto Ueda con «Las grafías bilabiales sonoras <u>, <v> y <b> del español en relación con el fonema /f/ y el paradigma sibilante» sigue la huella a los procesos históricos que han recorrido las grafías <u>, <v> y <b> desde el siglo xii hasta el xvii en las distintas regiones de España. Explora el uso de dichas grafías atendiendo a parámetros lingüísticos (condición léxica, condición morfológica, función fonológica, posición dentro de palabra, etc.) que somete a un análisis multivariante. Para su estudio se basa en el corpus CODEA (*Corpus de Documentos Españoles Anteriores a 1800*) y como herramienta de análisis emplea el nuevo sistema LYNEAL (*Letras y Números en Análisis Lingüísticos*).

José Ramón Morala y Hermógenes Perdiguero («La isoglosa de la aspiración de /f/ en el siglo xvii») estudian la distribución diatópica en el Siglo de Oro de los casos en los que se mantiene un resultado aspirado para la /f-/ latina u otros fonemas con los que esta ha confluído. Los datos los extraen del corpus de documentación notarial CorLexIn en el que analizan las confusiones entre las grafías de las antiguas dorsopalatales (<x, j, g>) con la grafía <h>, dibujando así la isoglosa que en el siglo xvii separa las zonas de mantenimiento de la realización aspirada de aquellas en las que se pierde cualquier referencia para la /f/ medieval. Del mismo modo, logran fijar también las zonas no estrictamente castellananas en las que se conserva la /f-/ originaria.

Entre la fonética y el léxico se sitúa el trabajo de Inés Carrasco Cantos y Pilar Carrasco Cantos «La palabra y los textos: acercamiento fonético y

léxico a la documentación andaluza (siglos XVI–XIX)», quienes nos brindan una investigación sobre meridionalismos de estos dos niveles de la lengua presentes en los corpus CODEMA (*Corpus Diacrónico de Documentación Malagueña*) y DITECA (*Diccionario de Textos Concejiles de Andalucía*). En ella exploran en qué medida la RAE, en sus diversas obras, y los estudios dialectológicos los recogen como tales dialectalismos y proporcionan nuevos datos que permiten concretar dataciones e identificar nuevas formas aún no calificadas de diatopismos en las obras especializadas.

Cinco son los capítulos que tratan sobre aspectos de lexicología y semántica desde una perspectiva dialectológica.

Igualmente centrada en Andalucía está la contribución de Pilar López Mora y Livia García Aguiar «De las voces de origen árabe en DITECA: aportación al estudio histórico del arabismo léxico (siglos XIII–XVIII)», en la que abordan desde una doble perspectiva (semántica e histórico-cronológica) el estudio de los arabismos en el *Diccionario de Textos Concejiles de Andalucía*. El interés de analizarlos en el DITECA estriba en que el corpus en el que este diccionario se basa reúne un léxico heterogéneo procedente de las ordenanzas (que legislaban los aspectos más variados de la vida de las ciudades) de distintos municipios de toda la geografía andaluza en un arco cronológico que abarca cinco siglos. Todos los campos designativos que se encuentran en ellas permiten hacer una valoración más exacta de la influencia del elemento árabe en la formación del léxico español.

El objetivo que persigue Emiliana Ramos Remedios en su capítulo «*Venta y renta*: un proceso de variación y cambio léxico», partiendo del CORHEN (*Corpus Histórico del Español Norteño*), es explicar el proceso de variación y cambio que afecta a las diversas lexías empleadas en esos textos para referirse a conceptos como ‘venta’ o ‘renta’ —en contraste con procesos paralelos observados en textos contemporáneos de otras diatopías hispánicas— y que llega hasta fines de la Edad Media con el triunfo en castellano de las palabras *venta* y *renta* sobre otras variantes. La pregunta que subyace a su estudio es a qué obedece la variación léxica que se observa en los textos medievales (variación diatópica, consecuencia de la adaptación de préstamos, tradiciones de escritura locales o usos individuales de escribanos).

Seguidamente, Robert Verdonk nos presenta «La correspondencia de don Fernando Álvarez de Toledo, III Duque de Alba, en tierras flamencas (1567–1573): su valor para la historia del español de Flandes y del

español general». Este corpus de cartas e informes del III Duque de Alba (quien transformó el gobierno de Flandes a su llegada a Bruselas en 1567, proponiendo una organización política y militar, basada en precedentes españoles y llevada por personal español), resulta fundamental para la observación de la situación de lenguas en contacto que surge entre el español y las diferentes lenguas autóctonas de Flandes (el francés y el neerlandés, con sus respectivos dialectos). En esta ocasión, Verdonk se centra en analizar una serie de préstamos del francés y el neerlandés que no habían aparecido en la correspondencia del Duque hasta su llegada a Flandes, preguntándose, además, en qué medida estos lograron integrarse más tarde en el español general.

En «Sobre léxico de la documentación notarial medieval del sur de Ávila: léxico de la vida cotidiana», M<sup>a</sup> Nieves Sánchez González de Herrero y Raquel Sánchez Romo emplean el CODOMSA (*Corpus de Documentación Medieval del Sur de Ávila*) para explorar la parcela del léxico de la vida cotidiana, que recoge las voces relacionadas con algunas categorías onomasiológicas que comprenden el ajuar doméstico, el vestido, los muebles, las herramientas y útiles propios de determinados oficios y algunas otras más que giran en torno a las distintas actividades que ocupan la rutina diaria. Se aprecia en esta investigación el interés de las autoras por cubrir toda la historia del campo léxico al que dedican su estudio, pues cruzan los datos proporcionados por los documentos antiguos con encuestas realizadas a habitantes de Mombeltrán en 2016 y 2017 con el fin de comprobar la continuidad que las formas puedan tener aún hoy.

Igualmente en el estudio del léxico de la vida cotidiana se detiene Del-fina Vázquez Balonga, con «Nuevos datos para la historia del léxico: aportaciones de los inventarios de bienes en la ciudad de Madrid en el siglo XVIII». En este trabajo nos ofrece algunos datos sobre los usos léxicos de términos de tejidos, colores y todo tipo de objetos en diferentes contextos sociales, documentados en inventarios madrileños extraídos de fondos archivísticos inéditos. Estos fondos aportan información nueva sobre la datación de voces registradas y la localización de voces no registradas o hápax, ya que, durante la elaboración de los inventarios se busca el término más preciso y, por tanto, el más empleado en su contexto. Esto permite a Vázquez Balonga encontrar términos con un marcado carácter diatópico y diacrónico.

Los siguientes dos capítulos continúan trabajando con corpus ya mencionados, como el CODOMSA y el CORHEN, pero centrándose esta vez en aspectos morfosintácticos.

Vicente Marcet Rodríguez en «Los tiempos compuestos en el castellano medieval: los documentos del sur de Ávila (siglo xv)» se dedica a un estudio morfológico de una serie de documentos del corpus CODOMSA. En ese conjunto de textos notariales redactados a lo largo del siglo xv, analiza las ocurrencias de tiempos compuestos de indicativo y subjuntivo, fijándose en aspectos como la doble auxiliaridad de estas con *haber* y *ser*, la posibilidad de intercalar algún elemento oracional entre el auxiliar y el participio o la concordancia en género y número entre el participio de la forma verbal compuesta y el complemento directo.

Por su parte, M<sup>a</sup> del Carmen Moral del Hoyo, en «Volver a (y revolver) los orígenes del castellano: el Corpus Histórico del Español Norteño (CORHEN)» describe, en primer lugar, el corpus CORHEN (tipología documental seleccionada, distribución temporal y geográfica), su preparación editorial (criterios de transcripción y edición) y la herramienta informática de libre acceso que permite su consulta, para plantear, en segundo término, los resultados más relevantes de algunos de los fenómenos lingüísticos estudiados hasta el momento. En concreto, Moral del Hoyo se centra en determinados procesos de variación gramatical como la morfología verbal del futuro e imperfecto de subjuntivo, de los perfectos fuertes, los demostrativos, la asimilación de preposición y artículo o el paradigma de los pronombres átonos de 3<sup>a</sup> persona, cuyo análisis permite arrojar luz sobre el papel inter e intradialectal del castellano norteño y central en la selección de variantes.

Por otra parte, contamos en este volumen con dos trabajos dedicados a cuestiones de pragmática.

Cristina Tabernero Sala en «Fórmulas de cortesía y mecanismos de atenuación en la retórica epistolar del siglo xix» aborda el estudio de más de un centenar de cartas del siglo xix procedente de un archivo particular de la localidad navarra de Redín a la búsqueda de las estrategias de cortesía y atenuación especialmente relevantes en este tipo textual que se acerca, por su concepción, a la inmediatez comunicativa propia de la conversación, y que emplea, sin embargo, a partir de un molde prefijado y formulario, el medio material propio de los intercambios de la distancia, la escritura. La autora atiende particularmente a las fórmulas de saludo y despedida, a las formas y pronombres de tratamiento y a los diversos mecanismos de atenuación presentes en este corpus documental.

Precisamente este mismo corpus es el que utiliza Concepción Martínez Pasamar para su trabajo «Fenómenos de covariación lingüística en

cartas del siglo XIX: las estrategias de intensificación y otras dimensiones de la variación». La autora examina en este caso las partes menos formularias de las epístolas para atender a las consideradas constantes discursivas de lo coloquial —saltos temáticos, construcción parcelada del discurso, dislocaciones, presencia de relatos— y, con mayor detalle, a los rasgos de expresividad y estrategias de intensificación del contenido o de la actitud. La combinación de estos fenómenos con las variables diastrática y diafásica muestra la influencia del factor sociocultural, que incluye el grado de dominio y manejo de la escritura, en la aparición de estos rasgos de proximidad en los textos.

Cierra este libro el capítulo de Belén Almeida «La historia de la lengua desde la perspectiva de la edición de textos: reflexiones sobre el valor de los textos escritos por no profesionales» dedicado a explorar cuestiones culturales asociadas a la escritura manuscrita. En él se nos exponen los factores que debemos tener en cuenta al atribuir un determinado grado de cultura y de control escritural al responsable efectivo de un documento y/o a su firmante. La autora ilustra hechos como los rasgos de variación en cuanto a género (posibles rasgos de «escritura de mujer»), los de variación cultural (maestros y responsables de cargos públicos de diversas categorías) o los propiamente reflejados en los documentos (braquigrafía, informatividad de los blancos, etc.) como muestra de los distintos grados de apropiación de técnicas escriturales que pueden detectarse en documentación en proceso de edición.

El conjunto de estos capítulos conforma un volumen en el que filólogos y especialistas en la edición de textos de distintas épocas avanzan en el conocimiento de aspectos de la historia de la lengua española en sus diferentes niveles (fonético-fonológico, grafemático, léxico-semántico, morfosintáctico y pragmático), gracias al examen de datos en corpus documentales que permiten realizar descripciones y análisis sólidamente fundamentados. Es ya un hecho que toda nueva página de investigación sobre la historia de la lengua española no puede prescindir de la información que esos nuevos corpus proporcionan, ya que contienen un caudal de testimonios de los usos lingüísticos en lugares y momentos identificados de modo fiable y son producto, además, de variados perfiles de escritores. La progresiva ampliación de dichos corpus a lo largo de estos años, tanto como el desarrollo de herramientas de búsqueda y análisis para su explotación, están contribuyendo a que los “bad data” con los que William Labov asociaba la lingüística histórica dejen definitivamente de ser tales.



A (re)escribir la historia de nuestra lengua con más y mejores datos desean contribuir las páginas de este volumen en el que, por ello, trasluce constantemente en filigrana el interés de continuar ampliando las colecciones de ediciones filológicas de documentos que nos informen de modo cada vez más fino sobre el pasado del español.

