

Andrés Enrique Arias (Ed.). *Diacronía de las lenguas iberorrománicas: nuevas aportaciones desde la lingüística de corpus*. Madrid: Iberoamericana-Vervuert, 2009, 416 páginas

El presente libro contiene una selección de 21 trabajos presentados en el Coloquio Internacional sobre Corpus Diacrónicos en Lenguas Iberorromances, el cual tuvo lugar en Palma de Mallorca en octubre de 2007. Al brindar una visión panorámica de los trabajos realizados actualmente sobre el estudio histórico de las lenguas iberorromances, este compendio busca profundizar en la metodología de la lingüística de corpus aplicada a la diacronía de dichas lenguas, pues ha habido poco interés en analizar de manera crítica los datos utilizados en las investigaciones de índole histórica.

Tres trabajos se ocupan de aspectos relacionados con cuestiones filológicas del diseño de corpus. Uno de ellos, escrito por Pedro Sánchez-Prieto Borja *et al.* muestra el *CODEA (Corpus de Documentos Españoles Anteriores a 1700)*, proyecto que pretende confrontar los documentos históricos de manera interdisciplinaria con el fin de obtener una comprensión global del texto. El segundo, de Marcela Cabrera de la Red y Miguel Gutiérrez Maté, señala la problemática alrededor del diseño de escritura de los documentos americanos. Beatriz Arias, enfocada hacia la historia del español de América, propone un corpus que contemple factores lingüísticos y extralingüísticos en relación con el origen y desarrollo de la variedad del español de Nueva España.

Por otro lado, José Antonio Pascual y Carlos Domínguez detallan el avance del *Corpus del nuevo diccionario histórico* de la Real Academia Española, el cual intenta compilar 50 millones de palabras para ilustrar la historia léxica y geográfica del español. Joan Torruella presenta el diseño del *Corpus Informatizat del Català Antic (CICA)*, cuyo fin radica en realizar un estudio gramatical del catalán anterior al siglo XVII. Mediante encuestas dirigidas a investigadores, Ana Sttulic-Etchevers y Soufiane Rouissi describen las necesidades principales que se deben tener en mente (como las representaciones digitales, la flexibilidad de consulta y la colaboración en el manejo de documentos) a la hora de elaborar un corpus del judeoespañol.

Otros dos trabajos se encargan de cuestiones de arquitectura informática. Mark Davies, al comparar el *Corpus del español* y el *Corpus do português* con el *CORDE*, señala la escasa utilidad de una colección de textos si no cuenta con las herramientas informáticas adecuadas para la búsqueda de datos. Por su parte, Kim Schulte plantea el uso de *regular expressions* para llevar a cabo complejas búsquedas en corpus sin anotaciones.

El tema de otros recursos electrónicos puestos al servicio de variedades iberorromances es tratado por Xulio Viejo *et al.*, quienes se ocupan del desarrollo del corpus *Eslema (Corpus Xeneral de la Lingua Asturiana)*. En otro trabajo, Ricardo Pichel y Xavier Varela Barreiro

se refieren al corpus *Tesouro Medieval Informatizado da Lingua Galega*. Un tercer recurso electrónico aplicado a lenguas iberorromances es el proyecto de “Edición del fichero manuscrito de léxico judeoespañol de Cynthia Crews”, pormenorizado en el trabajo de Aitor García Moreno.

Enfocados en la caracterización histórica de la variedad regional del español peninsular, específicamente, se ofrecen dos trabajos. Miguel Calderón Campos y María Teresa García Godoy ilustran los criterios utilizados en la selección de los textos del *Corpus diacrónico del español del Reino de Granada (CORDEREGRA)* y analizan lingüísticamente ciertos datos extraídos del corpus. Mercedes Abad Merino se refiere a las dificultades en la selección de un corpus que permita un estudio sociolingüístico histórico y las características que este debería tener. A través de un corpus documental del antiguo Reino de Murcia, la autora expone un análisis sociolingüístico a partir de textos históricos.

Tres estudios se ocupan de la utilización de traducciones bíblicas medievales como corpus. Andrés Enrique Arias investiga las implicaciones de trabajar con el corpus *Biblia medieval*, en cuanto a sus beneficios y limitaciones. El trabajo de Valentina Vincis consiste en el análisis de las estructuras de ilación oracional en textos bíblicos romanceados. Asimismo, se encarga de aspectos metodológicos de la lingüística de corpus en relación con la teoría de las tradiciones discursivas. Por su parte, Lola Pons investiga los nexos concesivos en las Biblias medievales para observar la variación y sustitución de estos elementos.

Los últimos cinco trabajos se centran alrededor del uso de corpus para el estudio histórico del español a partir de nuevas aplicaciones. Robert Blake y Gina Lee utilizan el programa de concordancia *MOCA (Multimedia Online Corpus Analysis)* para crear una base de datos con documentos de los siglos IX al XII procedentes del Monasterio de Sahagún. Josep Alba-Salas parte de los datos del *Corpus del español* para analizar estructuras del tipo *meter miedo* y *poner miedo* desde una perspectiva diacrónica. Dorein Nieuwenhuijsen se encarga de la evolución de los pronombres *conmusco* y *convusco* a partir del *CORDE* y trata la problemática alrededor de las herramientas de búsqueda en diferentes manuscritos. Natalya Stolova, mediante el *Corpus del español* y el *CORDE*, indaga sobre el uso del auxiliar perfectivo *ser* en español postclásico, a la vez que muestra los beneficios de los corpus en formato electrónico. Enrique Pato utiliza el *CODEA*, el *Corpus del español* y el corpus *Biblia medieval* para investigar la evolución del indefinido *alguien* desde sus primeras apariciones hasta el siglo XIII y así corroborar su origen occidental.

Cabe concluir que dicha compilación, por incorporar investigaciones relacionadas con el empleo de corpus históricos desde nuevas perspectivas, se considera como una herramienta indispensable para los interesados en el aspecto metodológico de la investigación diacrónica de las lenguas iberorromances.