

Pervivencia y Literatura:

documentos periféricos al texto literario



Coordinado y editado por
CARMEN F. BLANCO VALDÉS y ELISA BORSARI



cilengua

SAN MILLÁN DE LA COGOLLA
2023

UN PRIMER ACERCAMIENTO AL ANÁLISIS ESTILOMÉTRICO DE LA *GENERAL ESTORLA**

MIGUEL LAS HERAS CALVO

Universidad de La Rioja / University of British Columbia (Canadá)

1. INTRODUCCIÓN

La estilometría con medios informáticos se viene aplicando a las obras literarias españolas desde hace algunos años para esclarecer dudas respecto a la autoría de ciertas obras. Esta está proporcionando información verdaderamente llamativa, pues en muchos casos contradicen las supuestas atribuciones que se han dado¹. Este tipo de análisis también se ha empleado en textos medievales, donde destacan los esfuerzos de Fradejas Rueda. El profesor, entre otras pruebas que ha ido realizando y presentando en sucesivos congresos, planteó «aplicar estas técnicas a textos de un pasado más lejano, tan lejano que a veces el gran problema que subyace en la defectuosa transmisión textual que han sufrido» (Fradejas Rueda 2019a: 50). Es aquí mismo donde, movido por el largo debate sobre la participación de Alfonso X en las obras de su escritorio, analiza estilométricamente los textos del *scriptorium alfonsí*. Los resultados son realmente

* Esta investigación se ha llevado a cabo gracias a las ayudas para la recualificación del sistema universitario español (2021-2023), modalidad Margarita Salas, financiadas por el Ministerio de Universidades (RD 289/2021) y la Universidad de La Rioja (BOR Resolución 1299/2021).

Asimismo, este estudio se enmarca en las actividades del proyecto de investigación *DHumar II: From Middle to Golden Age: Translation & Tradition* (Ref. PY20_00469, financiado por la Consejería de Transformación Económica, Industria, Conocimiento y Universidades de la Junta de Andalucía y por FEDER Una manera de hacer Europa) y del proyecto *The Confluence of Religious Cultures in Medieval Historiography: A Digital Humanities Project* (Ref. 61R01897, financiado por el Social Sciences and Humanities Research Council of Canada [SSHRC]).

1. Algunos de los ejemplos más llamativos son los del *Quijote* de Avellaneda o el *Lazarillo de Tormes*. Para una revisión de estos y otros estudios famosos sobre estilometría, véase, por ejemplo, Cerezo Soler/Calvo Tello (2019: 233-236).

sugere, pues tras el análisis del corpus de textos —tomados de las transcripciones del *Hispanic Seminar of Medieval Studies*—, obtiene dos grandes ramas (Il. 1): por una parte, las obras que son traducciones de textos árabes (rama izquierda) y, por otra, las obras que tienen fuentes latinas (rama derecha).

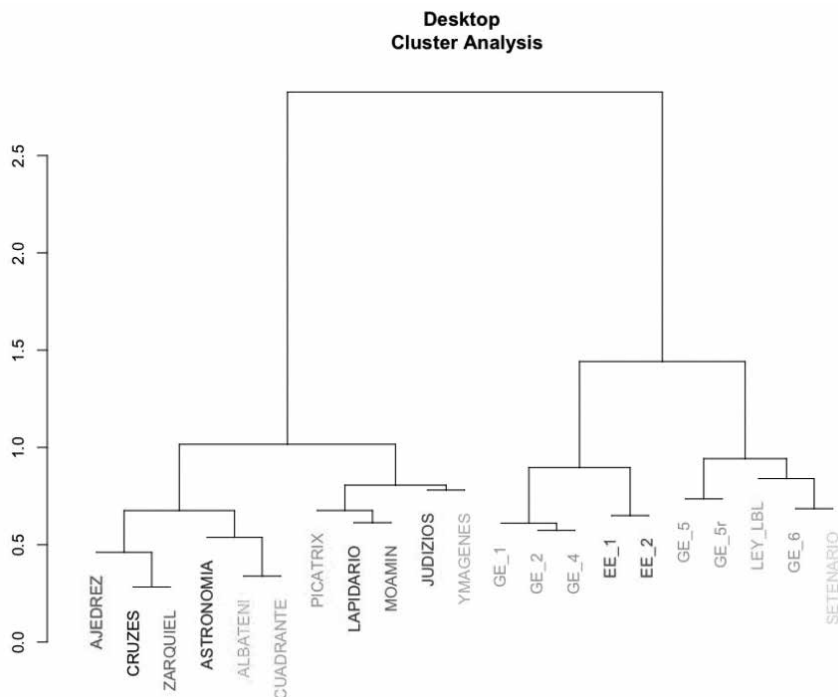


Ilustración 1. Obras del *Scriptorium* alfonsí (3000 MFW, Culled @ 0%, Classic Delta distance) (Fradejas Rueda 2019a: 59)

Respecto a la obra historiográfica que nos ocupa, la *General estoria*, adscrita a la segunda rama, constituye una obra singular dentro de la Edad Media, entre otros motivos, por su extensión, la variedad de fuentes empleadas o la manera en la que se imbrican los libros bíblicos y las historias de los gentiles. Sin lugar a dudas, este proyecto historiográfico debió de contar con un número muy elevado de colaboradores que trabajaron probablemente en sucesivas etapas en la recopilación de las fuentes, su traducción, su compilación y, finalmente, el volcado del texto definitivo en los manuscritos regios².

2. Para más información sobre la colaboración entre los colaboradores alfonsíes, consúltense, entre

Precisamente, con los textos de los manuscritos regios conservados, en la línea del trabajo de Fradejas (2019a), nuestro propósito es aplicar esta técnica al texto de la *General estoria* para observar qué relaciones estadísticas hay entre los diferentes libros que conforman sus códices. Los manuscritos escogidos para este trabajo son los códices regios 816 de la Biblioteca Nacional de España, que contiene la Primera parte de la obra (*GE1*), y el manuscrito 539 Urb. Lat. de la Biblioteca Apostólica Vaticana, en el que se halla la Cuarta parte (*GE4*). Se toman las ediciones semipaleográficas del *Hispanic Seminary of Medieval Studies*, cuyos textos fueron transcritos por Kasten, Nitti y Jonxis-Henkemans, corregidos por Sánchez-Prieto, Díaz Moreno, Trujillo Belso y Gago Jover, y recopilados por este último (Gago Jover 2011a y 2011b).

Respecto a la preparación de los textos para su tratamiento, se han creado ficheros de texto plano con cada uno de los libros que componen estos códices y se han eliminado todos los metadatos y etiquetas irrelevantes para el procesamiento de los textos (número de folios, columnas, rúbricas... [Fradejas Rueda 2019b]). Estos documentos se han denominado de acuerdo con las peculiaridades organizativas de cada códice: a) para los libros que componen la Primera parte, se menciona su pertenencia a *GE1*, el libro bíblico en el que se halla (Génesis, Éxodo, Levítico, Números o Deuteronomio) y, finalmente, el número de libro; b) para los libros de la Cuarta parte, se alude a su adscripción a *GE4*, el número de libro según su posición en el códice y su nombre. Finalmente, para el procesamiento de estos ficheros, se ha empleado, como es habitual en este tipo de análisis, el programa RStudio junto al paquete *stylo* (Eder/Kestmont/Rybacki 2016). Se ha recurrido a un análisis en el que se han seleccionado las 3000 palabras más frecuentes (*MFW*)³ y se ha recurrido a la distancia euclídea⁴.

muchos otros, Menéndez Pidal (1951), Catalán (1963), Sánchez-Prieto (2009a: LXXXVIII y ss.), Fernández-Ordóñez/Orellana (2009: xxv y ss.) o Rubio Tovar (2014).

3. Como es bien sabido, la herramienta establece sus relaciones entre los diferentes textos a través de las palabras más empleadas, cuyo número puede ser seleccionado. Tomo este número de palabras al igual que hicieron Fradejas Rueda (2019a: 59) y Cerezo Soler/Calvo Tello (2019: 240) en sus respectivos estudios. No obstante, se han hecho varias pruebas con 100, 300 y 1000 *MFW*. Los resultados del incremento progresivo del número de palabras desembocan finalmente en los dendrogramas que presentamos a continuación. Todas estas pruebas, así como los resultados, se han realizado siguiendo, entre otras fuentes, la guía de Fradejas Rueda (2023).
4. Utilizo este tipo de distancia, como hace Fradejas Rueda (2023), pero se prevé que en un futuro próximo se realicen también otras pruebas con Classic Delta, muy empleado en estilometría, para ver si los resultados muestran cambios significativos.

2. ANÁLISIS ESTILOMÉTRICO DE LA *GENERAL ESTORIA*

2.1. *Ms. 816 BNE (GE1)*

Como es bien sabido, al igual que se venía haciendo en otras obras historiográficas latinas y romances (Sánchez-Prieto 2009a: lx, 2009b), para la organización de *GE1* se tomaron los libros bíblicos del Pentateuco, que, además de aportar un eje estructural para la narración, dotaban a esta de temas. Sin embargo, los contenidos que en ella se encuentran no son el resultado de una mera traducción de la Vulgata, sino que son numerosísimas las aportaciones que se insertan de otros comentaristas, exégetas o historiadores, que hacen a propósito del texto bíblico, o la adición de historias de gentiles. Como apunta Sánchez-Prieto (2009a: lxxvi), el contenido estrictamente bíblico representa, aproximadamente, un 10 % del total de *GE1*. Cada uno de los libros del Pentateuco queda a su vez subdividido en otros libros: Génesis comprende los libros del 1 al 10; Éxodo, los libros del 11 al 16; Levítico, del 17 al 20; Números, del 21 al 25; y, por último, Deuteronomio abarca los libros del 26 al 29⁵. Tras procesar los textos de estos libros con *stylor*, el dendrograma que se obtiene es el siguiente (Il. 2).

En primer lugar, se observan dos grandes ramas: la primera de ellas comprende, entre otros, todos los libros de Génesis. La mayor parte de ellos se distribuye en el dendrograma agrupados por parejas de libros consecutivos (libros 1 y 2, libros 3 y 4, libros 6 y 7 [al que se le añade cerca el 5] o libros 8 y 9 [al que se le añade cerca el libro 10]). Además, llama la atención que, en esta primera rama, se incluyan, asimismo, otros libros. Por una parte, se añade el undécimo, perteneciente al Éxodo. En este sentido, la obra refleja la estructura de los libros del Pentateuco; sin embargo, se consideraba que la materia tratada en Génesis y Éxodo constituía una única historia y así se plasmaba en el Prólogo de este segundo libro del Pentateuco: «Departa maestre Pedro [...] e diz que la estoria del libro Éxodo e la del primero libro de la Biblia que viene ante d' éste, que es el Génesis, que non an departimiento ninguno entre sí, e que una estoria misma son» (Sánchez-Prieto 2009a: lxv). Esto podría justificar, en cierto modo, que el primer libro del Éxodo tenga gran semejanza estadística con los del Génesis. Y, por otra parte, resulta llamativo que en esta rama se incluyan algunos de los últimos libros del código, que se adscriben a los libros de Números (23) y Deuteronomio (26, 27 y 29).

La segunda gran rama contiene los libros restantes del Pentateuco. Por una parte, se observa que los libros del Éxodo (salvo el aludido antes), de nuevo,

5. No hay que olvidar que el código no dispone de su último cuaderno que, como afirma Fernández-Ordóñez (2002: 45), dispondría, posiblemente, de ocho folios, como los del resto del manuscrito.

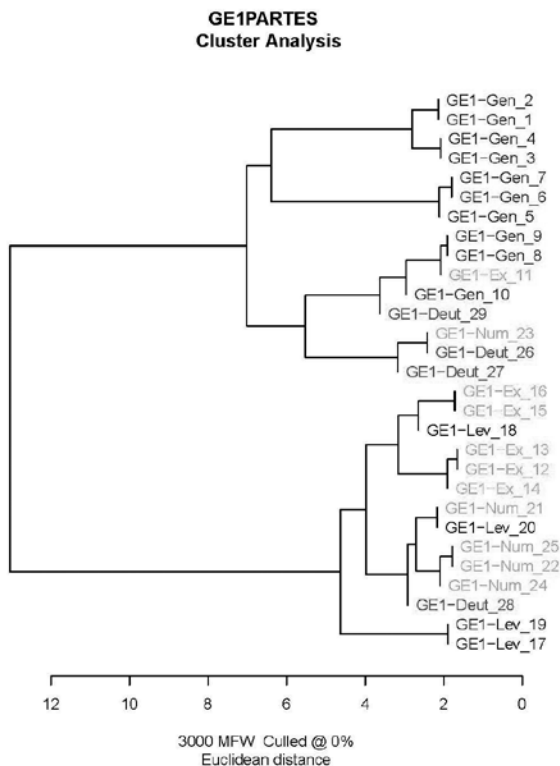


Ilustración 2. Dendrograma de los libros de la Primera parte de la *General estoria*

quedan distribuidos por parejas de libros consecutivos (libros 12 y 13 [a los que les queda cerca el libro 14] o libros 15 y 16); por otra, también se observa que los libros pertenecientes a Números se hallan muy cercanos entre sí. Sin embargo, esta organización tan destacada parece que no se repite de la misma manera con los libros de Levítico y Deuteronomio.

2.2. Ms. 539 Urb. Lat. (GE4)

Por su parte, el manuscrito 539 Urb. Lat. de la Biblioteca Apostólica Vaticana narra el periodo comprendido entre el reinado de Nabucodonosor hasta el comienzo del de Tolomeo Filópator. En *GE4* se produce un cambio relevante en la organización del contenido, pues, como anotan Fernández-Ordóñez y Orellana (2009: xiv), ahora la narración ya no se basa en la historia del pueblo hebreo,

sino que se toman como referencia los reinados de los reyes gentiles. Los libros bíblicos que se insertan en esta parte se presentan como *estorias unadas*. La compilación y posterior redacción parece que no fue sencilla, pues en ocasiones, como ocurre con la redacción de los primeros años de Ciro de Persia, no coinciden las cabeceras con lo enunciado en los prólogos, lo que da a entender que los encargados de confeccionar el código «no supieron entender la estructura que se les proponía» (Fernández-Ordóñez/Orellana 2009: xviii).

Si acudimos al análisis estilométrico para observar cuáles son las relaciones estadísticas entre los libros que componen *GE4*, tras la preparación de los archivos, que suman un total de 30⁶, obtenemos el dendrograma que se puede ver en Il. 3.

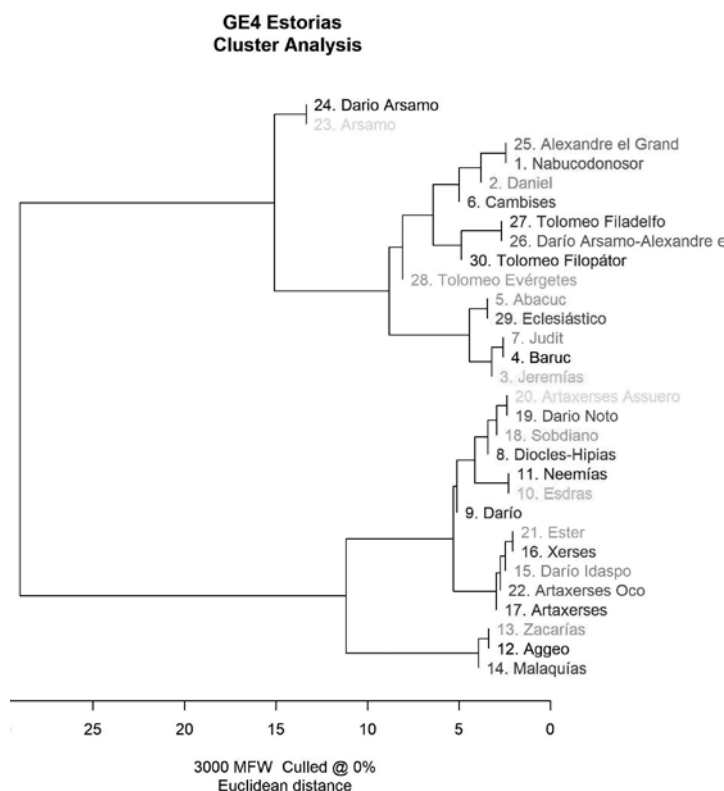


Ilustración 3. Dendrograma de los libros de *GE4*

6. En ocasiones hemos tenido problemas a la hora de fragmentar los textos, como es el caso del libro de Darío Arsamo, al que le sigue el libro de Alexandre el Grand, ya que la historia de este

Respecto a la organización global de la obra, los libros de *GE4* quedan divididos en dos ramas, que comprenden, por una parte, los libros 1 (Nabucodonosor) al 7 (Judit) y los libros 23 (Arsamo) al 30 (Tolomeo Filopátor); y, por otra, los libros 8 (Diocles-Hipias) al 22 (Artaxerses Oco). Pese a que pueda parecer que la repartición de los libros queda entrecortada, lo cierto es que la suma de los libros de cada rama suma un total de quince, justamente la mitad.

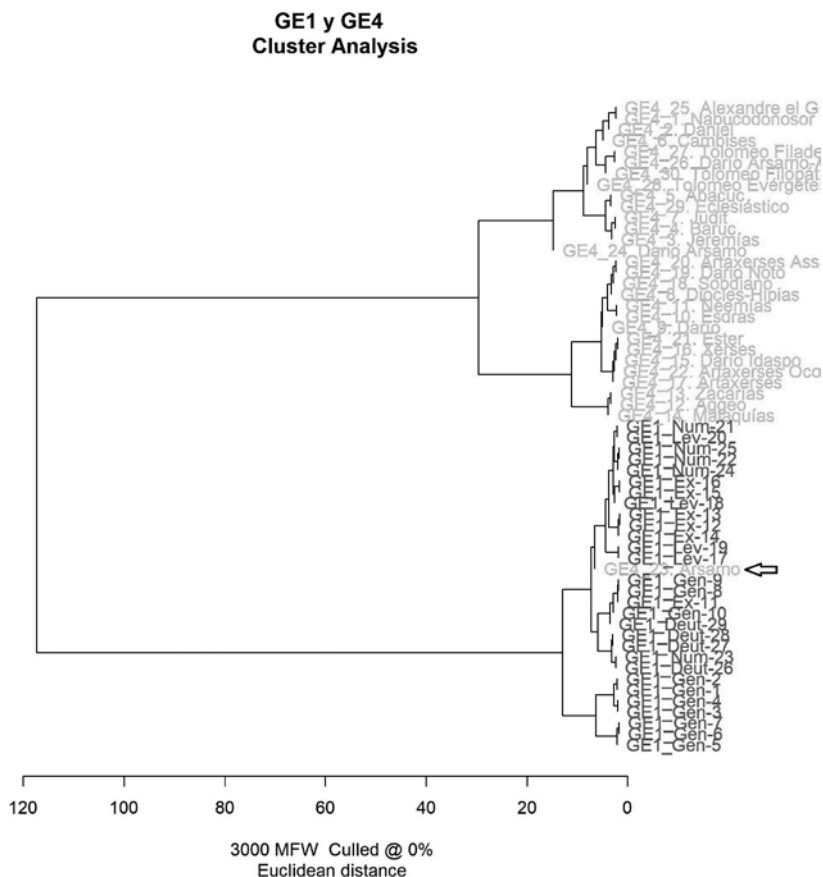
Respecto a la primera rama (rama superior), *a priori*, se observa que en la organización de *G4*, a diferencia de *GE1*, la agrupación de los libros contiguos no se da en demasiadas ocasiones; solamente se observa en los libros de Arsamo y Darío Arsamo (23 y 24) y en los libros 26 y 27 (Darío Arsamo-Alexandre el Grand y Tolomeo Filadelfo).

En la segunda rama la agrupación de los libros sí sigue un orden más acorde con la lógica de la propia estructura de *GE4*, pues en ciertas ocasiones los libros contiguos quedan juntos en la rama. Por ejemplo, me refiero a los libros de Esdras y Neemías (10 y 11); los de Aggeo y Zacarías (12 y 13), cercanos al de Malaquías (14); o los de Darío Noto (19) y Artaxerses Assuero (20), cercanos a Sobdiano (18). Solamente se documentan dos agrupaciones de libros lejanos, como los de Xerses (16) y Ester (21), y Darío Idaspo (15) y Artaxerses Oco (22).

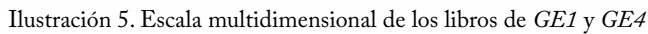
2.3. Comparación de *GE1* y *GE4*

Además de los resultados de cada una de las partes, resulta interesante, aplicando exactamente la misma metodología, comparar los libros de ambos códices, ya que nos hallamos ante manuscritos regios. De esta manera, aplicando el análisis estilométrico a los 59 libros que conforman en conjunto ambas partes, se obtiene el siguiente dendrograma (II. 4).

último se adscribe dentro del reinado de aquel. No obstante, aunque la extensión del libro de Darío Arsamo resulta bastante breve, se ha optado dejar la organización del códice, que también se mantiene en la edición crítica (Fernández-Ordóñez/Orellana 2009).

Ilustración 4. Dendrograma de los libros de *GE1* y *GE4*

El dendrograma queda dividido en dos grandes ramas: por una parte, en la rama superior, se encuentran los libros que pertenecen a *GE4*; y, por otra, los libros de *GE1*. Si bien, como se aprecia en la imagen, el libro 23 de *GE4* (Darío Arsamo), no se localiza en la rama de su correspondiente Parte, sino que se inserta en los libros de la *GE1*, lo cual indica que, estadísticamente, este libro se halla más cercano a determinados libros de *GE1* que de *GE4*. En este mismo sentido, si se aplica un análisis MDS (*Multidimensional Scaling*), que sirve para visualizar en una gráfica el nivel de similitud estadística de los diferentes textos analizados, se obtiene el siguiente gráfico (Il. 5).



3. ALGUNAS CONCLUSIONES Y COMENTARIOS

Como se ve, el análisis estilométrico de los manuscritos regios de la *General estoria* parece resultar bastante fructífero y arroja unos datos que son, cuanto menos, bastante sugerentes. Algunos de los datos que reseñamos son los siguientes:

a) El dendrograma de los libros que componen *GE1* muestra la minuciosa organización que debieron de tener los equipos de redactores alfonsíes. Así, los libros de Génesis quedan todos ellos adscritos a una misma rama y, por lo general, los libros que se agrupan son contiguos. Exactamente lo mismo ocurre con los libros que conforman el Éxodo y Números. Los libros de Levítico y Deuteronomio aparecen un tanto más dispersos.

b) En cuanto al dendrograma de *GE4*, aunque hay varios casos en los que las agrupaciones de libros contiguos muestran estructura y organización, parece que no se repite la meticulosa organización de *GE1*. Así, a través del análisis estilométrico se coincide con la conclusión alcanzada por Fernández-Ordóñez/Orellana (2009: xx) a propósito de la distorsión estructural: «la combinación de la materia bíblica y pagana generó muchos desajustes y contradicciones en la organización del texto, hasta el punto de que no todos los colaboradores alfonsíes entendieron de la misma forma el diseño global de la obra».

c) Por último, respecto a la comparación de los códices regios, el dendrograma obtenido parece señalar que la relación entre los equipos de estas dos partes fue bastante escasa, pues únicamente el libro de Arsamo, perteneciente a *GE4*, se incluía en la rama de los libros de *GE1*.

Además, los dendrogramas no solo nos han dado información sobre las semejanzas existentes entre los diferentes libros, sino que también nos han planteado dudas sobre las que tendremos que profundizar. Así, habrá que ver cuáles son los motivos por los que los libros del final de la Primera parte aparecen junto a los del Génesis; qué relación hay entre los libros agrupados que no son consecutivos; por qué el libro de Arsamo se adscribe a los libros de *GE1*; o a qué se debe la distancia estadística que hay entre el libro de Tolomeo Evérgetes de *GE4* y el resto de los libros de esta parte.

No obstante, este trabajo constituye un primer acercamiento al análisis estilométrico de los códices regios de la *General estoria*, en cuya precisión debemos aún seguir trabajando. No hay que olvidar que, pese a que la división en libros parece resultar productiva, la extensión de estos es muy amplia en la mayor parte de los casos y con seguridad habría numerosas manos copiando cada uno de los libros. Fradejas Rueda (2019b: 74) decía que:

decididamente, los métodos estilométricos (y lexómicos) puede que no nos digan nada nuevo de un texto, aparte de confirmar lo que ya sabíamos (que no es poco), pero su importancia radica en las posibilidades que puede ofrecer para futuras investigaciones.

No cabe duda de que este tipo de análisis aporta pistas de gran valor para el conocimiento de la organización del trabajo en la *General estoria*. Estas pistas ponen

sobre la mesa un amplio abanico de líneas de investigación en las que, junto con la ayuda de otros estudios —como pueden ser análisis de tipo paleográfico, lingüístico, ideológico o de contenido, entre otros—, habrá que seguir profundizando.

REFERENCIAS BIBLIOGRÁFICAS

- CATALÁN, Diego (1963), «El taller historiográfico alfonsí. Métodos y problemas en el trabajo compilarorio», *Romania*, 88, pp. 354-375. DOI: <https://doi.org/10.3406/roma.1963.2906>
- CEREZO SOLER, Juan y CALVO TELLO, José (2019), «Autoría y estilo. Una atribución cervantina desde las humanidades digitales. *El caso de La conquista de Jerusalén*», *Anales cervantinos*, 51, pp. 231-250. DOI: <https://doi.org/10.3989/anacervantinos.2019.011>
- EDER, Maciej; KESTEMONT, Mike y RYBICKI, Jan (2016), «Stylometry with R: A Package for Computational Text Analysis», *The R Journal*, 16/1, pp. 1-15. DOI: <https://doi.org/10.32614/RJ-2016-007>
- FERNÁNDEZ-ORDÓÑEZ, Inés (1992), *Las Estorias de Alfonso el Sabio*. Madrid: Istmo.
- FERNÁNDEZ-ORDÓÑEZ, Inés (1999), «El taller historiográfico alfonsí: la *Estoria de España* y la *General estoria* en el marco de las obras promovidas por Alfonso el Sabio», en Jesús Montoya Martínez y Ana Domínguez Rodríguez (eds.), *Scriptorium alfonsí, de los libros de astrología a las «Cantigas de Santa María»*. Madrid: Editorial Complutense, pp. 105-126.
- FERNÁNDEZ-ORDÓÑEZ, Inés (2001), «El taller de las *Estorias*», en Inés Fernández-Ordóñez (ed.), *Alfonso X el Sabio y las crónicas de España*. Valladolid: Universidad de Valladolid, Servicio de Publicaciones, pp. 60-82.
- FERNÁNDEZ-ORDÓÑEZ, Inés (2002), «*General estoria*», en Carlos Alvar y José Manuel Lucía Megías (coords.), *Diccionario filológico de literatura medieval española. Textos y transmisión*. Madrid: Castalia, pp. 42-54.
- FERNÁNDEZ-ORDÓÑEZ, Inés y ORELLANA, Raúl (eds.) (2009), *General estoria. Cuarta parte*. Madrid: Biblioteca Castro.
- FRADEJAS RUEDA, José Manuel (2019a), «Estilometría y la Edad Media castellana», en Nanette Rißler-Pipka (ed.), *Theorien von Autorschaft und Stil in Bewegung: Stilistik und Stilometrie in der Romania. Romanische Studien*. München: AVM.edition, pp. 49-74.
- FRADEJAS RUEDA, José Manuel (2019b), «Alfonso X y la estilometría». En línea <<https://7partidas.hypotheses.org/1072>> [consulta: 07/02/2023].
- FRADEJAS RUEDA, José Manuel (2023), *Cuentapalabras: Estilometría y análisis de texto con R para filólogos*. Valladolid: Universidad de Valladolid. En línea <<http://www.aic.uva.es/cuentapalabras/>> [consulta: 05/01/2023].

- GAGO JOVER, Francisco (ed.) (2011a), «*General Estoria I*». *Obra en prosa de Alfonso X el sabio. Digital Library of Old Spanish Texts. Hispanic Seminary of Medieval Studies*. En línea <<http://www.hispanicseminary.org/t&c/ac/index.htm>> [consulta: 10/10/2022].
- GAGO JOVER, Francisco (ed.) (2011b), «*General Estoria IV*». *Obra en prosa de Alfonso X el sabio. Digital Library of Old Spanish Texts. Hispanic Seminary of Medieval Studies*. En línea <<http://www.hispanicseminary.org/t&c/ac/index.htm>> [consulta: 10/10/2022].
- MENÉNDEZ PIDAL, Gonzalo (1951), «Cómo trabajaron las escuelas alfonsíes», *Nueva Revista de Filología Hispánica*, 5-4, pp. 363-380. DOI: <https://doi.org/10.24201/nrfh.v5i4.196>
- RUBIO TOVAR, Joaquín (2014), «La traducción en la *General Estoria*», en Jaume Aurell, Rafael Narbona Vizcaíno *et al.*, *La cultura en la Europa del siglo XIII. Emisión, intermediación, audiencia. Actas de la XL Semana de Estudios Medievales de Estella*. Pamplona: Gobierno de Navarra, Departamento de Cultura, Turismo y Relaciones Institucionales, pp. 247-284.
- SÁNCHEZ-PRIETO BORJA, Pedro (2001), «Sobre el concepto de *original* (el caso de la *General estoria* de Alfonso el Sabio)», en Leonardo Funes y José Luis Moure (eds.), *Studia in honorem Germán Orduna*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá, pp. 571-582.
- SÁNCHEZ PRIETO BORJA, Pedro (ed.) (2009a), *General Estoria. Primera parte*. Madrid: Biblioteca Castro.
- SÁNCHEZ-PRIETO BORJA, Pedro (2009b), «Biblia e historiografía en los códices medievales», en Pedro Manuel Cátedra, Eva Belén Carro y Javier Durán (coords.), *Los códices literarios de la Edad Media. Interpretación, historia, técnicas y catalogación*. San Millán de la Cogolla: Cilengua, pp. 71-90.