

Estilometría y la Edad Media castellana

José Manuel Fradejas Rueda (Valladolid)

RESUMEN: En este trabajo se muestra que para los análisis estilométricos con la librería *stylo* se pueden utilizar transcripciones semipaleográficas de textos españoles medievales. Se ha puesto en prueba con las obras completas de don Juan Manuel (regularizadas y semipaleográfica), con las obras que conforman el *Scriptorium* alfonsí (semipaleográficas). Este corpus ha arrojado un resultado curioso, pero aún inexplicable, y es que considerando tan solo las 100 palabras más frecuentes, el sistema ha podido establecer que las obras del *Scriptorium alfonsí* se agrupan en dos grandes bloques: 1) obras traducidas del árabe y 2) obras de fuente latina. Alentados por los resultados, y tras estudiar los trabajos de *lexómica*, se ha analizado el *Libro de la caza de las aves* de Pero López de Ayala y el *Libro de la montería*. Los resultados, aunque conocidos, han sido muy interesantes ya que al dividir los textos en pasajes de 800 o 1000 palabras, el sistema ha sido capaz de detectar la existencia de las distintas fuentes de las que han bebido tanto Ayala como el *Libro de la montería* y agruparlas entre sí.

ZUSAMMENFASSUNG: Der Beitrag zeigt, dass für stilometrische Untersuchungen mit dem *Stylo*-Paket für R auch semi-paläografische Transkriptionen mittelalterlicher Texte genutzt werden können. Verglichen werden dazu die gesammelten Werke Juan Manuels (regularisiert und semi-paläographisch) mit den Werken aus dem *Scriptorium* von Alfonso X (semi-paläographisch). Mit diesem Korpus wurden interessante aber kaum erklärable Ergebnisse erzielt. Beispielsweise kann die Methode schon bei 100 MFW (most frequent words) die Werke des *Scriptorium* von Alfonso X in zwei Blöcke gruppieren: 1. Werke, die aus dem Arabischen übersetzt sind, 2. Werke, die aus dem Lateinischen stammen. Angespornt durch diese Ergebnisse und im Zusammenhang lexometrischer Studien am Wheaton College, wurde die gleiche Methode auf einen Vergleich des *Libro de la caza de aves* von Pero Lopez de Ayala (in einer modernisierten Fassung) und des *Libro de la montería* (in einer semi-paläographischen Fassung) angewendet. Die bereits bekannten Ergebnisse wurden noch spannender, weil durch eine Aufteilung der Texte in Abschnitte von 800 und 1000 Wörtern mit der stilometrischen Methode die unterschiedlichen Quellen in den Texten nachgewiesen werden können.

PALABRAS CLAVE: estilometría; *lexómica*; autoría; fuentes; Edad Media; español medieval

SCHLAGWÖRTER: Stilometrie; *Lexomik*; Autorschaft; Quellen; Mittelalter; Altspanisch

^o Este artículo forma parte de las investigaciones del proyecto *Edición crítica (digital) de las Siete Partidas*, financiado por el MINECO y el FEDER (ref. FFI2016-75014-P AEI-FEDER, EU) que se lleva a cabo en la Universidad de Valladolid. Para más información acerca de este proyecto véase el sitio web <https://7partidas.hypotheses.org/> y el repositorio <https://github.com/7PartidasDigital>.

Introducción

En un trabajo reciente pude comprobar que la estilometría automatizada, ya sea aplicando el *script* de aprendizaje de programación en R publicado por Jockers¹, ya sea con el paquete *stylo*, funciona correctamente a la hora de identificar un texto moderno o contemporáneo escrito en español². También se han realizado algunos ensayos para ver si esas mismas técnicas, y en especial el paquete *stylo*³, son aplicables a textos de etapas anteriores cuya autoría es altamente disputada, como quiénes son (o no son) los autores del llamado *Quijote de Avellaneda* o el del *Lazarillo de Tormes*⁴. Lo cierto que esos análisis parecen contradecir todo lo dicho hasta el momento sobre la autoría de estas dos obras. Ante el éxito de estos ensayos, y alguno más del que tengo noticias, me planteé la posibilidad de aplicar estas técnicas a textos de un pasado más lejano, tan lejano que a veces el gran problema subyace en la defectuosa transmisión textual que han sufrido. Hemos estado haciendo pruebas, no se las puede considerar de ninguna otra manera, con textos de una serie de autores de los que tenemos un relativo buen conocimiento, en algunos muy profundo, desde el punto de vista de la filología más tradicional, y los resultados son los que exponemos a continuación.

Juan Manuel

Es un tema clásico de la historia de la literatura medieval española afirmar que las obras de don Juan Manuel (1282–1348) se pueden dividir en dos grandes bloques: uno de influencia alfonsí y otro de un estilo más personal. En el primer bloque se incluyen la *Crónica abreviada*, el *Libro de la caza* y el *Libro del*

¹ Matthew L. Jockers, *Text Analysis with R for Students of Literature* (Cham: Springer, 2014), doi:10.1007/978-3-319-03164-4.

² José Manuel Fradejas Rueda, “El análisis estilométrico aplicado a la literatura española: las novelas policiacas e históricas”, *Caracteres: estudios culturales y críticos de la esfera digital* 5, n.º 2 (2016): 196–245.

³ Maciej Eder, Jan Rybicki, y Mike Kestemont, “Stylo: a Package for Stylometric Analyses”, 2013.

⁴ Nanette Rissler-Pipka, “Avellaneda y los problemas de la identificación del autor: propuestas para una investigación con nuevas herramientas digitales”, en *El otro Don Quijote: la continuación de Fernández de Avellaneda y sus efectos*, ed. Hanno Ehrlicher (Ausburgo: Institut für Spanien, Portugal- und Lateinamerikastudien, 2016), 27–51; Nanette Rissler-Pipka, “Der falsche Quijote? Autorschafts attribution für spanische Prosa der frühen Neuzeit”, en *DHd 2016 Modellierung, Vernetzung, Visualisierung*, 2016, 212–17; Nanette Rissler-Pipka, “Digital Humanities und die romanische Literaturwissenschaft. Der Autorschaftsstreit um den Lazarillo de Tormes”, *Romanische Forschungen* 128, n.º 3 (2016): 316–42.

caballero e del escudero; y en el segundo, el resto de sus obras. La consideración alfonsina del *Libro del caballero e del escudero* se debe a que se inspiró en el título 21 de la segunda partida; la del *Libro de la caza* se debía a que siempre se ha dicho que lo escribió siguiendo un desaparecido, e ignoto, libro de la caza escrito por el rey Alfonso X⁵, sin embargo, se ha podido demostrar que Juan Manuel sigue el modelo del *De arte venandi cum avibus* de Federico II⁶. Resta una tercera obra de corte alfonsí: la *Crónica abreviada*, que es un resumen de la *Estoria de España* de Alfonso X.

El análisis del texto con el paquete *stylo* se ha realizado sobre la edición normalizada de Carlos Alvar⁷ y nos muestra, como se observa en la figura 1, que la *Crónica abreviada* es una obra que nada tiene que ver con el estilo de don Juan Manuel, que es un elemento haplofílico dentro del dendrograma, o por usar el término estadístico, un dato atípico.

⁵ Esta afirmación se basa en los que dice don Juan Manuel en el prólogo al *Libro de la caza*: “Et el dicho rey don Alfonso deseando el saber, como dicho es, et pagándose de todas las cosas nobles et apuestas et sabrosas et aprehochosas, entendiendo que en la caça ha estas quatro cosas muy complidamente a los que quieren usar della como deven, et non dexar por ella otros fechos mayores, ca los que en otra manera caçassen, aunque guardassen el sabor et la apostura de la caça, non guardarían la nobleza nin el aprovechamiento, por ende mandó fazer muchos libros buenos en que puso muy complidamente toda la arte de la caça, tan bien del caçar, como del benar, como del pescar. Et puso muy complidamente la teórica et la práctica como conviene a esta arte; et tan complidamente lo fizo, que bien cuidan que non podría otro emendar nin eniader ninguna cosa más de lo que él fizo, nin aun fazer tanto nin tan bien como él”, Juan Manuel, *Obras completas*, ed. Carlos Alvar y Sara Finci (Madrid: Fundación José Antonio de Castro, 2007).

⁶ José Manuel Fradejas Rueda, “La influencia del *De arte venandi cum avibus* de Federico II en el *Libro de la caza* de Juan Manuel”, en *Los libros de caza* (Tordesillas y Valladolid: Seminario de Filología Medieval, Universidad de Valladolid, 2005), 41–54; José Manuel Fradejas Rueda, “Juan Manuel y Federico II de Hohenstaufen”, en “*Siempre soy quien ser solía*”: estudios de literatura española medieval en homenaje a Carmen Parrilla, vol. 1 (La Coruña: Universidade da Coruña, Servicio de Publicaciones, 2009), 137–48.

⁷ El mismo editor (Juan Manuel, *Obras completas*) nos facilitó una copia en formato Word de los textos tal y como los entregó a la editorial. Ese fichero se convirtió a texto plano y cada una de las obras se guardó en un fichero independiente que se han identificado como: ABREVIADA.txt (99767 palabras-token; 6943 palabras-tipo), ARMAS.txt (7576 palabras-token; 1395 palabras-tipo), ASUNCION.txt (2862 palabras-token y 638 palabras-tipo), CAB_ESC.txt (26508 palabras-token; 2695 palabras-tipo), CAZA.txt (28647 palabras-token; 2916 palabras-tipo), ENFENIDO.txt (15748 palabras-token; 1856 palabras-tipo), ESTADOS.txt (103243 palabras-token; 5922 palabras-tipo) y LUCANOR.txt (77764 palabras-token; 5717 palabras-tipo) lo que ha supuesto un corpus de 362115 palabras-token y 15821 palabras-tipo. Una vez obtenidos estos materiales, se han procesado con *stylo*.

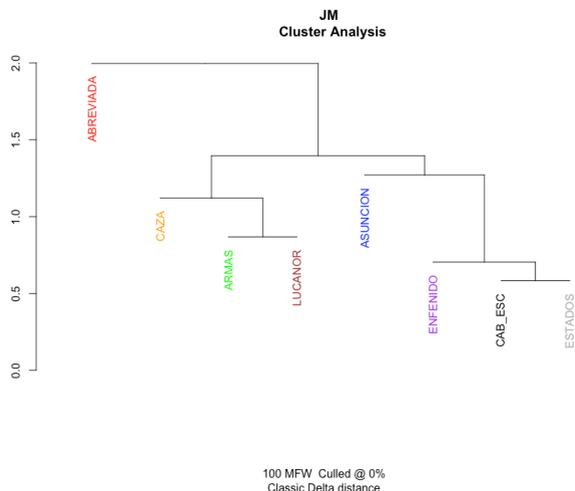


Fig. 1: Obras de don Juan Manuel, según la edición normalizada de Alvar (100 MFV)

Al aumentar el número de palabras que se tienen en cuenta a 1000, la geometría del dendrograma se mantiene, y la *Crónica abreviada* sigue siendo un dato atípico (fig. 2).

El objetivo es comprobar si el hecho de tratarse de versiones normalizadas de un texto medieval y no fieles a un manuscrito, ha podido influir en este resultado puesto que en la edición de Alvar se han regularizado usos gráficos como la alternancia de u/v y i/j/y, que pueden tener valores vocálicos o consonánticos, y se ha acentuado según las reglas ortográficas actuales. Para esta parte del análisis se han manejado las transcripciones semipaleográficas de Ayerbe-Chaux⁸ preparadas para el Dictionary of Old Spanish Language (DOSL n adelante) del Hispanic Seminary of Medieval Studies (HSMS en adelante). Ahora bien, se ha tomado la decisión de convertir la nota tironiana, que representa la conjunción copulativa y, que en estas transcripciones se transcriben con &, se han sustituido por e en la *Crónica Abreviada* ya que es la forma mayoritaria cuando no se usa la nota tironiana, y por et en el resto de las obras ya que es la de mayor uso. Asimismo, se han efectuado otros dos cambios: la conversión de la secuencia <n-> en <ñ> y de <c'> en <ç>, dado que estas dos letras no existen en el juego de caracteres ASCII con el que se

⁸ Reinaldo Ayerbe-Chaux, *Textos y concordancias de la obra completa de Juan Manuel* (Madison: Hispanic Seminary of Medieval Studies, 1986).

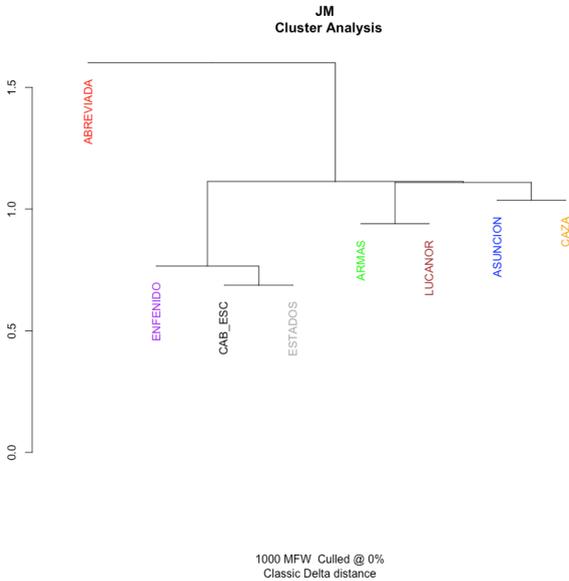


Fig. 2: Obras de don Juan Manuel, según la edición normalizada de Alvar (1000 MFW)

transcribieron originalmente los textos del HSMS⁹. A pesar de estos cambios, no parece que la fidelidad a un testimonio medieval sea de gran valor, como lo demuestran las figuras 3 y 4 en donde se puede ver que la Crónica abreviada sigue siendo un rasgo atípico dentro de la producción textual de don Juan Manuel, tanto si solo se consideran las 100 palabras más frecuentes (fig. 3) como si se tienen en cuenta las 1000 más frecuentes (fig. 4).

Alfonso X

Bajo el nombre de Alfonso X (1221–1284), la historia de la literatura española agrupa un rico corpus textual que comprende obras legales, históricas, científicas y de recreación. Uno de los temas más debatidos es el referente a la intervención del rey en la redacción de las obras y el equipo con el que contaba en el llamado *Scriptorium alfonsí*. En este punto, se examinarán las obras alfonsíes desde un punto de vista estilométrico para observar cómo se agrupaban los distintos textos que constituyen este rico corpus textual.

⁹ Kenneth Buelow y David Mackenzie, *A manual of manuscript transcription for the Dictionary of the Old Spanish language* (Madison Wis.: Hispanic Seminary of Medieval Studies, 1977), 11–2.

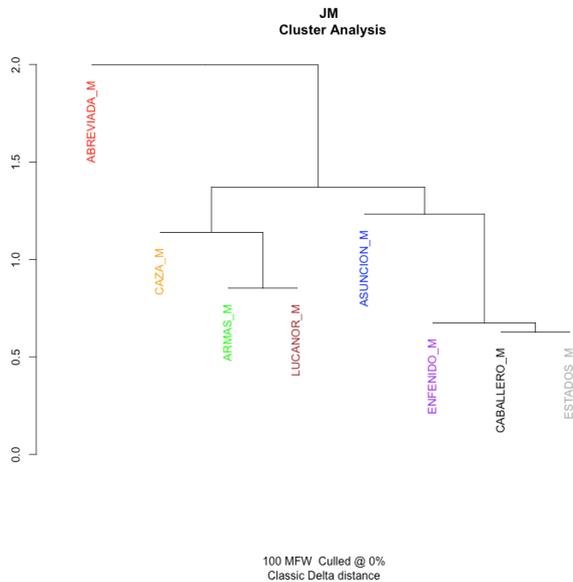


Fig. 3: Obras de don Juan Manuel, según la edición semipaleográfica de Ayerbe-Chaux (100 MFW)

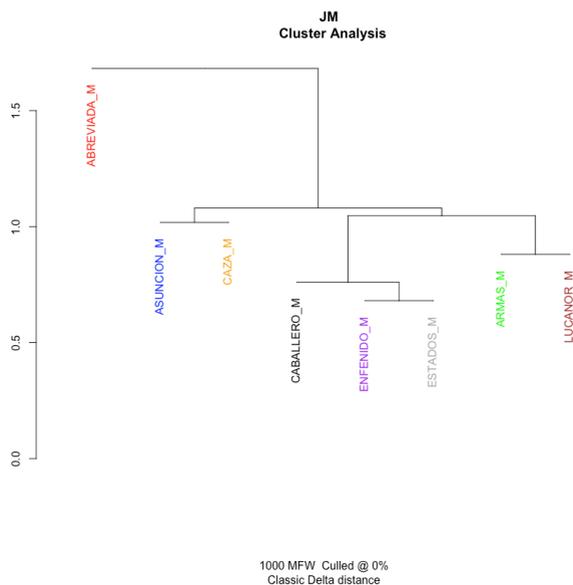


Fig. 4: Obras de don Juan Manuel, según la edición semipaleográfica de Ayerbe-Chaux (1000 MFW)

Para esto se ha contado, de nuevo, con todos los textos del *Scriptorium alfonsí* transcritos para el *Dictionary of Old Spanish Language* del Hispanic Seminary of Medieval Studies, transcripción llevada a cabo por Kasten *et al.*¹⁰ y algunos textos más recopilados por O'Neill¹¹ y reunidos por Gago-Jover¹². Como en el caso de las obras de don Juan Manuel, se ha procedido a sustituir la nota tironiana & bien por *e* bien por *et*, de acuerdo con el uso mayoritario en cada testimonio de una u otra forma (véase Tabla 1), de manera que no se pierda información.

Un sencillo análisis de grupos con el paquete *stylo*, en el que solo se han tenido en cuenta las 100 palabras más frecuentes, ofrece un interesante y curioso dendrograma (fig. 5).

Las obras del rey Sabio aparecen distribuidas en dos grandes bloques: a la derecha, traducciones del árabe y a la izquierda, obras históricas y jurídicas con fuente latina. En casi todos los casos la longitud de los subclados es muy corta, lo que indica una gran homogeneidad entre los textos, especialmente en la serie *Judizios, Ymagenes, Ajedez, Moamin, Lapidario* y *Picatrix*.

¹⁰ Lloyd August Kasten *et al.*, *The electronic texts and concordances of the prose works of Alfonso X, El Sabio* (Madison: Hispanic Seminary of Medieval Studies, 1997).

¹¹ John O'Neill y Hispanic Seminary of Medieval Studies, *Electronic texts and concordances of the Madison Corpus of early Spanish manuscripts and printings* (New York: Hispanic Seminary of Medieval Studies, 1999).

¹² Las obras que se han tenido en cuenta son: *Canones de Albateni, Libros de ajedrez, dados y tablas, Libros del saber de astronomía, Libro de las cruces, Estoria de España I, Estoria de España II, General estoria I, General estoria II, General estoria IV, General estoria V, General estoria VI, General estoria VI, Judizos de las estrellas, Lapidario de Alfonso X, Libro de las leyes, Moamyn-Libro de las animalias, Picatrix de Alfonso X, Libro del cuadrante señero, Libro de las formas y de las imagenes, Tablas de Zarquiel, Setenario*. El *Libro de las animalias que cazan*, erróneamente considerado como el *Libro de Moamín*, es transcripción de Cárdenas – *The text and concordance of Biblioteca Nacional Manuscript RES. 270–217 Libro que es fecho de las animalias que caçan, the Book of Moamin* (Madison [Wis.]: Hispanic Seminary of Medieval Studies, 1987) –, mientras que la del *Setenario* lo es de Bonch-Bruevich (*Electronic text and concordances of the Setenario by Alfonso el Sabio: Biblioteca Capitular de Toledo, MS 43–20* (Madison: Hispanic Seminary of Medieval Studies, 1999)). Todo este corpus textual supone un total de 3.234.321 palabras y se halla reunido en la Biblioteca Digital de Textos del Español Antiguo (Francisco Gago Jover, “Obra en prosa de Alfonso X el Sabio: biblioteca digital de textos del español antiguo”, 2011, <http://www.hispanicseminary.org/t&c/ac/index-es.htm>). El recuento es el ofrecido por el HSMS para cada uno de los textos (véase Tabla 2).

	<i>e</i>	<i>et</i>	<i>&</i>
AJEDREZ	474	157	2014
ALBATENI	7	1403	1714
ASTRONOMIA	130	4942	15826
CRUZES	30	1550	5790
CUADRANTE	20	417	341
EE_1	6282	731	13317
EE_2	429	10181	17457
GE_1	626	4477	38388
GE_2	166	6382	13922
GE_4	270	563	37880
GE_5	1	3691	9457
GE_5r	6	3982	11329
GE_6	2	313	1189
JUDIZIOS	2822	3	19984
LAPIDARIO	8	2621	5420
LEY_LBL	2330	43	4325
MOAMIN	10499	10	96
PICATRIX	17	582	2584
SETENARIO	8	1420	3022
YMAGENES	1	58	1164
ZARQUIEL	0	0	12

Tab. 1: Formas y ocurrencia de la conjunción copulativa en el corpus alfonsí

	palabras-token	palabras-tipo
AJEDREZ	53638	1899
ALBATENI	44386	2115
ASTRONOMIA	275364	6238
CRUZES	130770	2815
CUADRANTE	10210	633
EE_1	258521	14620
EE_2	352894	16385
GE_1	535532	20663
GE_2	243965	1447
GE_4	456311	20520
GE_5	146468	10343
GE_5r	193144	12549
GE_6	17338	2839
JUDIZIOS	225456	7413
LAPIDARIO	103121	5655
LEY_LBL	153664	6833
MOAMIN	107190	6345
PICATRIX	34584	4421
SETENARIO	78729	638
YMAGENES	17203	2104
ZARQUIEL	849	127

Tab. 2: Número de palabras-token y palabras-tipo de cada uno de los textos tenidos en cuenta

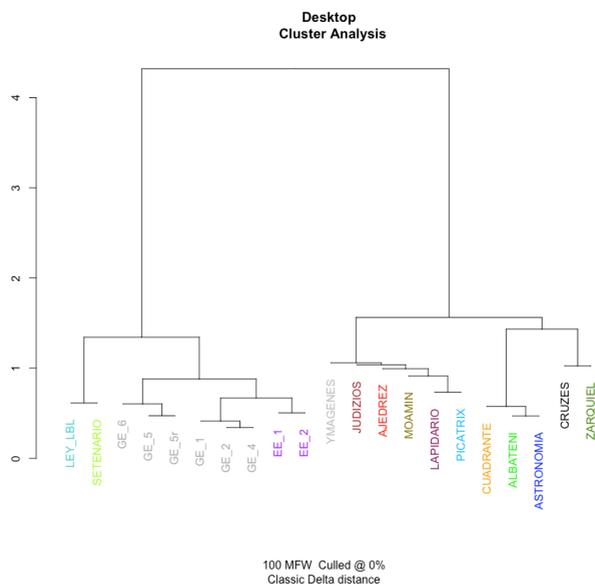


Fig. 5: Obras del Scriptorium alfonsí (100 MFW)

El que el *Libro de los animales que cazan* (o *Moamín*)¹³, obra traducida del árabe cuando aún Alfonso no había accedido al trono, se asiente cómodamente dentro del compacto y homogéneo grupo de obras de origen árabe traducidas a lo largo de 30 años puede indicarnos que la traducción es alfonsí realmente,

¹³ Designar a esta obra como *Moamín* se debe a un error que se remonta a finales de los años 1940, cuando Håkan Tjerneld (“Una Fuente Desconocida Del Libro de La Montaña Del Rey Alfonso El Sabio”, *Studia Neophilologica* 22, n.º 2 (21 de enero de 1949): 171–93, doi:10.1080/00393274908587044) estableció la relación entre la versión castellana del *Kitab al-yawarīh*, hasta entonces conocido como *Libro de cetrería* (El Escorial, ms. V.II.19), y un texto editado por el mismo Tjerneld (*Moamīn et Ghatrif: traités de fauconnerie et des chien de chasse* (Stockholm: Fritze, 1945)) debido a la pluma de un cetrero árabe conocido simplemente como *Moamīn* y que François Viré (“Sur l’identité de Moamyn le fauconnier”, *Comptes Rendus de Séances* 111, n.º 2 (1967): 172–5) trató de identificar con Hunayn ibn Ishaq (808–873). Al acceder a principios de los años 1980 a un nuevo testimonio de la versión castellana (olím Phillipps 11719, BNE RES/270) se pudo establecer que el autor era Muhammad ibn Abd Allah ibn Umar al-Bayzar (Muḥammad ibn Abdallah Al-Bāzyār, *Libro de los animales que cazan* = *Kitāb al-Īwārīh*, ed. José Manuel Fradejas Rueda (Madrid: Editorial Casariego, 1987)), autor del que años más tarde se publicarían partes del texto árabe bajo el título de *Kitab al-Mutawakkil* (véase Muḥammad ibn Abdallah Al-Bāzyār, *Das Falken- und Hundebuch des Kalifen al-Mutawakkil: ein arabischer Traktat aus dem 9. Jahrhundert*, ed. Anna Akasoy y Stefan Georges (Berlin: Akademie Verlag, 2005)).

aunque se realizara en un momento previo a su coronación como rey, por lo que no se sostiene la apreciación de Fernández-Ordóñez de que el *Libro de los animales que cazan* puede pertenecer al entorno de Fernando III, aunque este no aparece “como responsable o instigador, no digamos autor”¹⁴.

Incluso si aumentamos a 3000 palabras, entrando en el resbaladizo ámbito de las palabras semánticas, se observa en la figura 6 que seguimos teniendo dos grandes grupos: uno constituido por las obras traducidas del árabe y el otro conformado por las obras históricas y legales con un fortísimo componente de traducción, o fuentes, desde el latín. Pero la situación de estos dos grupos es inversa a la mostrada anteriormente: el grupo árabe está ahora a la izquierda, mientras que el latino se sitúa a la derecha, lo que es poco o nada significativo, puesto que un dendrograma solo representa los grupos, ya que para una estructura de grupos cualquiera existen 2^{n-1} árboles posibles¹⁵.

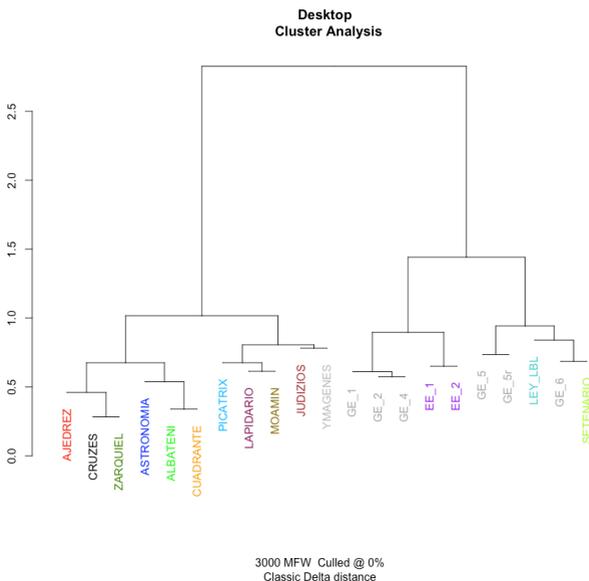


Fig. 6: Obras del *Scriptorium alfonsí* (3000 MFW)

¹⁴ Inés Fernández-Ordóñez, “Alfonso X el Sabio en la historia de español”, en *Historia de la lengua española*, ed. Rafael Cano Aguilar (Barcelona: Ariel, 2004), 386–7.

¹⁵ Hermann Moisl, *Cluster Analysis for Corpus Linguistics* (Berlin: De Gruyter Mouton, 2015), 214.

Otro de los asuntos candentes dentro de los estudios alfonsíes es el de la redacción y finalización de las *Siete Partidas*. En el esquema anterior (fig. 5 y 6), se ha tenido en cuenta el llamado *Libro de las leyes* (British Library, ms. Additional 20787), que corresponde al testimonio más antiguo de la primera partida¹⁶, pero no existe ningún testimonio de época alfonsí de todas las *Partidas*. Es más, ninguno de ellos es anterior al reinado de Sancho IV y, por lo que parece, la gran mayoría de los conservados son copias realizadas después de 1348, cuando Alfonso XI declara en el Ordenamiento de Alcalá 1348 las *Siete Partidas* como derecho supletorio (título 28, ley 1) y ordena hacer dos ejemplares sellados con su sello de oro y plomo¹⁷. A finales del siglo XV, en 1491, se fija un texto que se convertirá en el texto normativo durante los 65 años siguientes: la edición de Alonso Díaz de Montalvo. De esta tenemos una versión electrónica cuidada con los mismos criterios de transcripción que el resto del corpus alfonsí¹⁸. Al incorporarla, se ve que se sitúa junto al *Libro de las leyes*, aunque el *Setenario* se desplaza y se une a GE6, algo que ya habíamos visto en la figura 5. Por otra parte, las dos versiones de las *Siete Partidas* se unen en un grupo absolutamente compacto con una distancia cero, como se puede apreciar en la figura 7, a pesar de que la distancia cronológica entre ambas versiones es de 200 años.

Este tipo de análisis con textos tanto contemporáneos como medievales, como los que acabamos de mostrar, permiten establecer con alta confianza la atribución a un autor de una obra con independencia de si los textos están escritos en lengua actual o en la de tiempos pasados, o si están escritos en forma modernizada o con grafías antiguas. Pero... ¿puede el análisis de grupos, la estilometría, decirnos algo más acerca de una obra? Acabamos de ver que hay algo curioso en los textos alfonsíes: un sistema automatizado separa las obras que son traducción de un texto árabe de las que tienen fuentes latinas, aunque no se puede aventurar que sean traducción directa de un texto latino, ya que cada uno de ellos tiene múltiples fuentes. Uno de los datos relativos de la influencia del árabe en la escritura castellana es la abundantísima coor-

¹⁶ Juan Arias Bonet, *Primera partida según el manuscrito Add. 20.787 del British Museum* (Valladolid: Universidad de Valladolid Secretariado de Publicaciones, 1975).

¹⁷ Alfonso Otero Varela, "Las Partidas y el Ordenamiento de Alcalá en el cambio del ordenamiento medieval", *Anuario de historia del derecho español* 1, n.º 63 (1993): 451-548.

¹⁸ Francisco Gago-Jover, "'Siete partidas': textos legales españoles. Biblioteca digital de textos del español antiguo", 2013, <http://www.hispanicseminary.org/t&c/lex/index-es.htm>. La edición de 1491 incorpora unas adiciones de mano de Montalvo, pero se han eliminado para este estudio, ya que podrían introducir ruido.

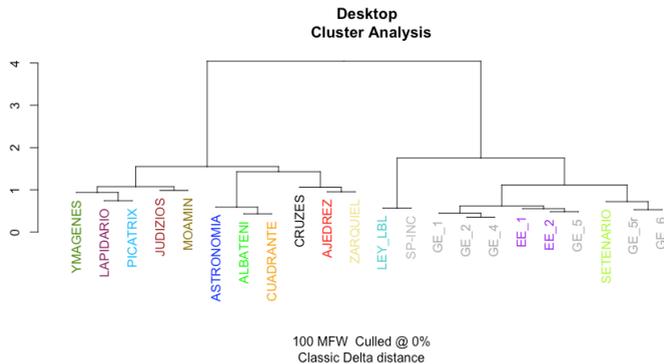


Fig. 7: El corpus alfonsí con las *Siete Partidas* de 1491 (SP-INC) (100 MFW)

dinación copulativa¹⁹. ¿Qué pasaría si se retirara este marcador discursivo de altísima frecuencia en el corpus alfonsí (209009 casos de *et* y 63886 casos de *e*)? El resultado es que al borrarse de los textos la conjunción copulativa, no varía un ápice la geometría básica del dendrograma, esto es, sigue habiendo un grupo de traducciones del árabe y otro, de fuentes latinas como se aprecia en la figura 8.

Lexómica

En el Wheaton College han desarrollado un método estilométrico al que designan *lexómica*. Este nuevo método adopta técnicas de la bioinformática, de la estilometría y del análisis textual tradicional, es decir, de los métodos filológicos de siempre. Lo que hacen es utilizar técnicas de la estadística informatizada para identificar patrones que después interpretan usando los métodos filológicos tradicionales. Al inicio de sus investigaciones, pudieron comprobar que “the computational methods told us where in a text to look”²⁰.

Según exponen, los métodos lexómicos “differ slightly from pioneering stylometrics analyses in two major ways [...] First [...] we include every word in our analyses. Second [...] we divide our texts into segments and analyze the

¹⁹ Rafael Lapesa, *Historia de la lengua española* (Madrid: Gredos, 1981), 150 y Lola Pons Rodríguez, *La lengua de ayer. Manual práctico de historia del español* (Madrid: Arco Libros, 2010), 358, aunque esta autora no le da validez a este rasgo.

²⁰ Phoebe Boyd et al., “Lexomic Analysis of Anglosaxon Prose: Establishing Controls with the Old English Penitential and the Old English Translation of Orosius”, *Selim: Journal of the Spanish Society for Medieval English Language and Literature* = *Revista de la Sociedad Española de Lengua y Literatura Inglesa Medieval*, n.º 19 (2012): 7–58, aquí 9.

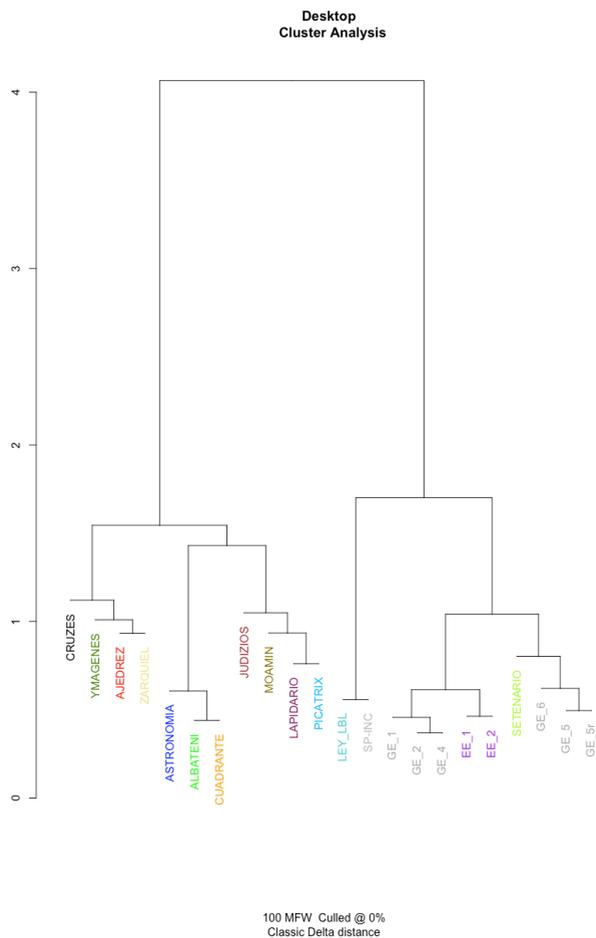


Fig. 8: El corpus alfonsí con las *Siete Partidas* de 1491 (SP-INC) y eliminadas las conjunciones copulativas *e* y *et* (100 MFW)

relationship of these to each other”²¹. Y, aunque reconocen que estos análisis tienen cierta propensión al establecimiento de la autoría, sus objetivos no están “focused primarily on author identification but instead on text’s sources or affinities”²². En nota a pie de página reconocen que las fuentes y las afinidades pueden tener que ver con los problemas de autoría, y de hecho lo han utilizado para mostrar que *Guthlac B*, uno de los poemas anglosajones que narran las leyendas de san Guthlac (673–714), lo escribió Cynewulf²³.

También reconocen que un análisis lexómico puede no decir nada nuevo sobre un texto. Sin embargo, admiten que gracias a la correlación de la geometría de un dendrograma con el conocimiento previo que podemos tener sobre un texto dado, “we can have reasonable confidence in the relationships suggested by our methods when these are applied to texts whose interrelationship is unknown or controversial”²⁴, es decir, examinar con los métodos lexómicos textos que sabemos cómo *funcionan* gracias a los métodos filológicos puede servirnos como *control* para aquellos que no conocemos aún²⁵.

Downey et al.²⁶ insisten repetidamente en que la lexómica muestra una relativa homogeneidad en el vocabulario de los autores examinados. Sin embargo, esto es una falsa apreciación ya que no tienen claro qué se ha de entender por el léxico de un texto, o bien no tienen nada claro qué palabras utilizan los programas para realizar el análisis de grupos, aunque, cuando tratan de explicar que el fragmento 17 del *Planctu naturae* de Alain de Lille (a. 1128–1202) es un dato atípico, una rama *haplofilica* o *simplifolius*, reparan en que las palabras más abundantes son *alii*, *alienus* e *isti*, poco informativas semánticamente²⁷. Puesto que la lexómica, así como la estilometría, se basan en el análisis de grupos, con independencia de la fórmula matemática empleada (Euclidiana, Delta, Manhattan, Camberra, Eder...), sin hacer uso de las palabras semánticas sino de las gramaticales, que suelen ser las más frecuentes, la homogeneidad de estilo a la que se refieren parece que es independiente del contenido temático del texto.

²¹ Boyd et al., “Lexomic Analysis”, 9.

²² Boyd et al., “Lexomic Analysis”.

²³ Michael D C Drout et al., “Of Dendrogrammatology: Lexomic Methods for Analyzing Relationships among Old English Poems”, *Journal of English and Germanic Philology* 110, n.º 3 (2011): 301–36.

²⁴ Drout et al., “Of Dendrogrammatology”, 38, y Sarah Downey et al., “Lexomic Analysis of Medieval Latin Texts”, *The Journal of Medieval Latin*, n.º 24 (2014): 225–75, aquí 232.

²⁵ Boyd et al., “Lexomic Analysis”, 14, y Downey et al., “Lexomic Analysis”, 232.

²⁶ Downey et al., “Lexomic Analysis”.

²⁷ Boyd et al., “Lexomic Analysis”, 240.

También afirman que el “lexomic analysis is based on critical editions”²⁸, debido a que “previous lexomic research suggests cluster analyses based on critical editions are more reliable than those based on diplomatic texts”²⁹. Pero, como acabamos de ver, en el caso de don Juan Manuel, los dendrogramas de las obras de este autor, ya se hayan basado en el análisis de la edición normalizada de Alvar³⁰, ya sea en la edición semipaleográfica de Ayerbe-Chaux³¹ no varían. Lo mismo cabe decir de las transcripciones de las obras de *Scriptorium alfonsí*³², que no parecen afectar en nada los resultados, por lo que no se puede afirmar que el “lexomic analysis goes somewhat against the grain of manuscript-approaches”³³. Es más, ellos mismos reconocen que “the differences between diplomatic and critical editions are relatively invisible to lexomic methods [...] because the most common words in Anglo-Saxon are those least likely to be emended”³⁴.

Partiendo de las ideas y sugerencias del grupo formado por Downey, Boyd y Drout et al.³⁵ y conociendo perfectamente el comportamiento del *Libro de la caza de las aves* de Pero López de Ayala (1332–1407), y del *Libro de la montería* (siglo XIV, c. 1350), hemos realizado un análisis lexómico de ambos textos, pero con el paquete *stylo*, puesto que los procesos que ejecuta el software diseñado por este grupo, Lexos³⁶, también los puede realizar *stylo* sin mayores problemas, ya que las fases del trabajo son las mismas. La única diferencia de importancia es que Lexos se puede encargar de fragmentar los textos, bien por número de palabras, bien por medio de hitos (*milestones*)³⁷, mientras que en

²⁸ Boyd et al., “Lexomic Analysis”, 18.

²⁹ Michael D. C. Drout et al., *Beowulf Unlocked: New Evidence from Lexomic Analysis* (Palgrave Macmillan, 2016), 19.

³⁰ Juan Manuel, *Obras completas*.

³¹ Ayerbe-Chaux, *Textos y concordancias de la obra completa de Juan Manuel*.

³² Gago Jover, “Obra en prosa de Alfonso X el Sabio.”

³³ Boyd et al., “Lexomic Analysis”, 19.

³⁴ Boyd et al., “Lexomic Analysis”, 24–5.

³⁵ Sarah Downey, Michael D C Drout, y Michael J Kahn, “Books Tell Us’: Lexomic and Traditional Evidence for the Sources of Guthlac A California University of Pennsylvania”, *Modern Philology* 110, n.º 2 (noviembre de 2012): 69–86, doi:10.1086/668252; Downey et al., “Lexomic Analysis”; Boyd et al., “Lexomic Analysis”; Drout et al., “Of Dendrogrammatology”; Michael D. C. Drout, *Tradition and Influence in Anglo-Saxon Literature: An Evolutionary, Cognitivist Approach* (New York: Palgrave Macmillan, 2013); Drout et al., *Beowulf Unlocked*.

³⁶ Se puede trabajar en línea (<http://lexos.wheatoncollege.edu/upload>) o se puede instalar el software en máquinas Windows 7 o superior), Mac OSX o Linux (<https://wheatoncollege.edu/academics/special-projects-initiatives/lexomics/lexos-installers>).

³⁷ Estos hitos pueden ser capítulos, folios, etc.

stylo es una tarea previa, aunque poco complicada con unos pequeños *scripts* en R o cualquier otro lenguaje de programación. Las demás tareas son idénticas y lo único que pueden variar son las fórmulas para calcular las distancias. Tras cargar el texto en el programa, este se encarga de convertirlo todo a minúsculas (para evitar que, por ejemplo, la palabra *Rey* y *rey* cuenten como dos palabras diferentes, cuando, en realidad, son una), y eliminar la numeración arábiga y los signos de puntuación³⁸. A todo este proceso los lexicónomas han dado el nombre de *scrubbing*³⁹. Posteriormente, el programa divide el texto en palabras (*tokens*) y calcula las frecuencias absolutas y relativas para cada uno de los textos (o fragmentos de texto) y, por último, aplica la fórmula elegida para el cálculo de las distancias y dibuja el dendrograma⁴⁰. *Lexos* puede producir otros elementos de fantasía como nubes de palabras o diagramas de burbujas, elementos innecesarios para extraer conclusiones y tablas estadísticas con informaciones variadas. Por su parte, *stylo* también ofrece tablas estadísticas.

El Libro de la caza de las aves

El *Libro de la caza de las aves*, escrito por Pero López de Ayala⁴¹ mientras estuvo preso en Óvidos tras la derrota castellana de Aljubarrota (14/8/1385) es un texto que tiene dos fuentes principales: una versión al castellano del *Livro de cetraria* de Pero Menino⁴² y lo escrito por el propio Ayala. Está bastante claro cuáles son los capítulos que son traducción de la obra de Pero Menino, los capítulos 11–37 y 39, mientras que se consideran originales de Ayala los capítulos 1–10, 36 y 40–47. La pregunta que nos hacemos es la siguiente: ¿puede el análisis de grupos detectar qué capítulos son de uno y otro autor?

³⁸ También se pueden eliminar espacios en blanco (*whitespace*) e incluso pueden limpiar etiquetas XML, HTML, etc., por lo que no habría problemas para procesar textos codificados en TEI. E incluso se pueden introducir una serie de palabras, bien manualmente bien por medio de un fichero, que no se quiera procesar.

³⁹ Michael D. C. Drout, *Tradition and Influence*, 52 y Drout et al., *Beowulf Unlocked*, 17.

⁴⁰ Durante todo el proceso, el sitio web permite descargar los ficheros intermedios que se van generando: el fichero *limpiado*; una vez segmentado; la tabla con todos los datos tras *tokenizar* (palabras y frecuencias –filas– en cada fragmento de texto –columnas–) que se puede descargar en formato CSV o TSV; y se pueden generar otras estadísticas (número de palabras-tipo; hápax).

⁴¹ Pero López de Ayala, *Libro de la caza de las aves*, ed. José Fradejas Lebrero y José Manuel Fradejas Rueda (Barcelona: Castalia, 2016).

⁴² Pero Menino, *Livro de falcoaria*, ed. Manuel Rodrigues Lapa (Coimbra: Imprensa da Universidade, 1931).

Para llevar a cabo este análisis, hemos partido de la edición normativizada del *Libro de la caza de las aves* que se ofrece en el Archivo Iberoamericano de Cetrería⁴³. Al tratarse de una página web, se ha de limpiar de los códigos del HTML y de todo el material propio de una página web que no tiene interés alguno para nuestro análisis. Una vez limpio, se ha dividido por capítulos. Esta primera división nos ha proporcionado unas primeras estadísticas básicas, y curiosas (tabla 3)⁴⁴. Así el número de palabras-token va desde el máximo de 6967 (cap. 9) al mínimo de 110 (cap. 45), con una media de 806.

Tab. 3: Estadísticas básicas del *Libro de la caza de las aves*

Capítulo	Tokens	Tipos	Hápx	TTR	Tokens/Obra	Tipos/Obra	Hápx/Obra
AIC-01	1750	602	399	34.4	4.52231439129648	14.2148760330579	18.936877076412
AIC-02	1125	399	253	35.4666666666667	2.9072021086906	9.42148760330578	12.0075937351685
AIC-03	2093	579	346	27.663640707119	5.40868801199059	13.6717827626919	16.421452301851
AIC-04	706	250	158	35.4107648725212	1.82443083443161	5.90318772136954	7.49881347887992
AIC-05	854	347	238	40.632318501171	2.20688942295268	8.19362455726092	11.2956810631229
AIC-06	517	220	144	42.5531914893617	1.33602088017159	5.19480519480519	6.8343616516374
AIC-07	761	313	199	41.1300919842313	1.96656071530093	7.39079102715466	9.44470811580446
AIC-08	398	198	143	49.748743718593	1.02850350156343	4.67532467532468	6.78690080683436
AIC-09	6967	1397	801	20.0516721687958	18.00399796366643	32.987012987013	38.016136687233
AIC-10	446	202	137	45.2914798206278	1.15254412486756	4.76977567886659	6.50213573801614
AIC-11	384	190	134	49.4791666666667	0.992324986433057	4.48642266824085	6.35975320360702
AIC-12	1676	506	294	30.1903907875895	4.33108509703595	11.9480519480519	13.953488372093
AIC-13	533	257	181	48.2176360225141	1.3773677546063	6.06847697756789	8.59041290934979
AIC-14	424	188	119	44.3396226415094	1.09569217251983	4.43919716646989	5.64784053156146
AIC-15	537	243	168	45.2513966480447	1.38770447321498	5.73789846517119	7.9734219269103
AIC-16	531	219	144	41.2429378531073	1.37219939530196	5.17119244391972	6.8343616516374
AIC-17	478	213	141	44.5606694560669	1.23523787373698	5.02951593860685	6.69197911722829
AIC-18	244	117	79	47.9508196721311	0.630539835129338	2.76269185360094	3.74940673943996
AIC-19	703	280	184	39.8293029871977	1.8166782954751	6.61157024793388	8.7327954437589
AIC-20	360	175	115	48.6111111111111	0.930304674780991	4.13223140495868	5.45799715234931
AIC-21	581	248	159	42.6850258175559	1.50140837791043	5.85596221959858	7.54627432368296

⁴³ <http://www.aic.uva.es/clasicos/ayala/ayala-texto.html>.

⁴⁴ El texto del *Libro de la caza de las aves* según esta versión consta de 38697 palabras-token y 4235 palabras-tipo, con lo que obtenemos una TTR de 10.944. Hay 2107 casos de hápx, lo que supone el 5.446% del texto, y tan solo cinco de ellas –*e*, *que*, *de*, *en* y *el*– son las palabras que ocurren más de 1000 veces suponen el 22,201%. La primera palabra semántica aparece en la décima posición con una frecuencia absoluta de 478 (*falcón*) y si tuviéramos en cuenta la forma plural (de haberlo lematizado; *falcones* con 203 casos) ocuparía la séptima posición y hubiera supuesto el 1.760% del texto.

Capítulo	Tokens	Tipos	Hápax	TTR	Tokens/Obra	Tipos/Obra	Hápax/Obra
AIC-22	437	205	142	46.9107551487414	1.12928650799804	4.84061393152302	6.73943996203132
AIC-23	460	191	118	41.5217391304348	1.18872263999793	4.51003541912633	5.60037968675842
AIC-24	274	145	98	52.9197080291971	0.708065224694421	3.42384887839433	4.65116279069767
AIC-25	493	193	128	39.1480730223124	1.27400056851952	4.55726092089728	6.0749881347888
AIC-26	317	155	103	48.8958990536278	0.819184949737706	3.65997638724911	4.88846701471286
AIC-27	924	326	207	35.2813852813853	2.38778199860454	7.69775678866588	9.82439487422876
AIC-28	478	210	139	43.9330543933054	1.23523787373698	4.95867768595041	6.59705742762221
AIC-29	1020	382	245	37.4509803921569	2.63586324521281	9.02007083825266	11.6279069767442
AIC-30	733	308	211	42.0190995907231	1.89420368504018	7.27272727272727	10.0142382534409
AIC-31	462	190	124	41.1255411255411	1.19389099930227	4.48642266824085	5.88514475557665
AIC-32	942	342	218	36.3057324840764	2.43429723234359	8.07556080283353	10.3464641670622
AIC-33	241	116	76	48.1327800829876	0.62278729617283	2.73907910271547	3.60702420503085
AIC-34	721	298	199	41.3314840499307	1.86319352921415	7.03659976387249	9.44470811580446
AIC-35	558	250	171	44.8028673835125	1.44197224591054	5.90318772136954	8.11580446131941
AIC-36	613	259	165	42.2512234910277	1.58410212677985	6.11570247933884	7.8310393250119
AIC-37	440	184	117	41.8181818181818	1.13703904695454	4.34474616292798	5.55291884195539
AIC-38	411	163	96	39.6593673965937	1.06209783704163	3.84887839433294	4.5562411010916
AIC-39	292	163	119	55.8219178082192	0.75458045843347	3.84887839433294	5.64784053156146
AIC-40	686	263	168	38.3381924198251	1.77274724138822	6.21015348288076	7.9734219269103
AIC-41	273	152	109	55.6776556776557	0.705481045042251	3.58913813459268	5.17323208353109
AIC-42	918	337	231	36.7102396514161	2.37227692069153	7.95749704840614	10.9634551495017
AIC-43	775	324	235	41.8064516129032	2.0027392304313	7.65053128689492	11.1532985287138
AIC-44	185	107	82	57.8378378378378	0.478073235651343	2.52656434474616	3.89178927384907
AIC-45	110	80	62	72.7272727272727	0.284259761738636	1.88902007083825	2.94257273778832
AIC-46	1234	463	311	37.5202593192869	3.18887769077706	10.9327036599764	14.7603227337447
AIC-47	1623	506	318	31.1768330252619	4.19412357547097	11.9480519480519	15.0925486473659
AIC-48	1009	427	301	42.3191278493558	2.60743726903894	10.0826446280992	14.2857142857143

Posteriormente, se ha analizado el texto dividido en capítulos con *stylo*, y se le ha pedido que establezca el análisis de grupos de acuerdo con la fórmula Classic Delta con un máximo de 100 palabras. El resultado es el dendrograma de la figura 9 en el que se observa que hay dos grandes grupos de tamaño distinto: el de la izquierda, agrupa los capítulos 1–8 y 42–44, 46 y 48, es decir, la mayoría de los capítulos atribuidos normalmente a Ayala. Esto confirma, *grosso modo*, lo que sabíamos: que hay dos estilos diferentes y que lo podemos atribuir, gracias al conocimiento filológico previo, a una cuestión de fuentes. Sin embargo, hay otros problemas *menores*: el capítulo VIII (AIC-09) se ha asentado dentro de la rama de Menino.

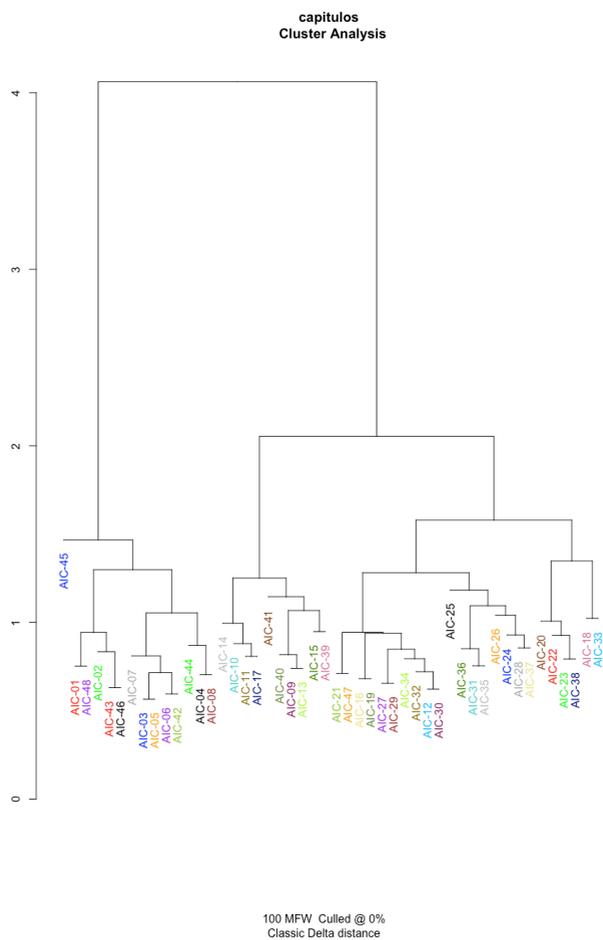


Fig. 9: Los capítulos del *Libro de la caza de las aves* (100 MFW)

Esto nos ha llevado a comprobar que ocurre si aumentamos el número de palabras que se han de tener en cuenta. Al aumentar a las 500 palabras más frecuentes, el resultado ha cambiado y el capítulo VIII (AIC-09) se ha situado cómodamente entre los que creemos de Ayala (fig. 10). Un nuevo aumento a las 1000 más frecuentes vuelve a ofrecer una geometría semejante como se puede observar en la figura 11.

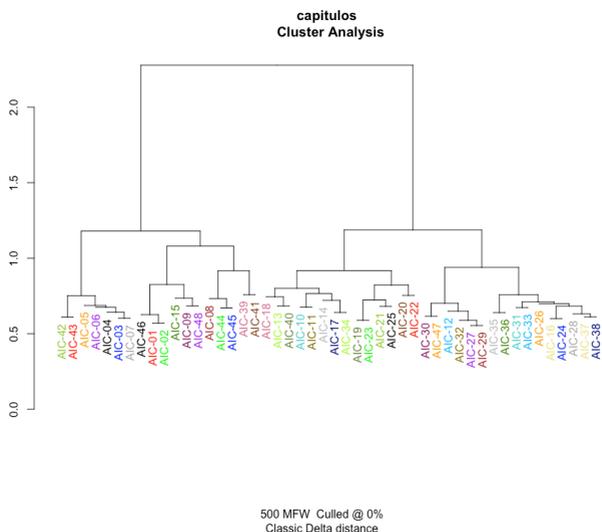


Fig. 10: Los capítulos del *Libro de la caza de las aves* (500 MFW)

Quizá el mayor problema es que es, sin lugar a dudas, el capítulo más extenso (él solo constituye el 18 % del texto) por lo que se procedió a analizar el texto dividido en segmentos de igual longitud. En un primer momento se dividió en pasajes de 1000 palabras para que *stylo* que considerara todas las palabras en cada fragmento. El resultado es una geometría semejante (fig. 12)), con variaciones en el orden de las hojas más bajas y con el capítulo VIII a caballo entre los dos grandes grupos (el capítulo VIII se halla entre los fragmentos 8 y 14), pero delimitado en un subgrupo de este segundo bloque.

Posteriormente, el tamaño de los fragmentos se redujo a 800 palabras, ya que esa era la media de palabras-token⁴⁵ para ejecutar la misma prueba con *stylo*. De nuevo, la geometría es semejante (fig. 13). El capítulo VIII, que se

⁴⁵ El último fragmento (AIC_800-48) es un poco más grande que los demás con 1090 palabras.

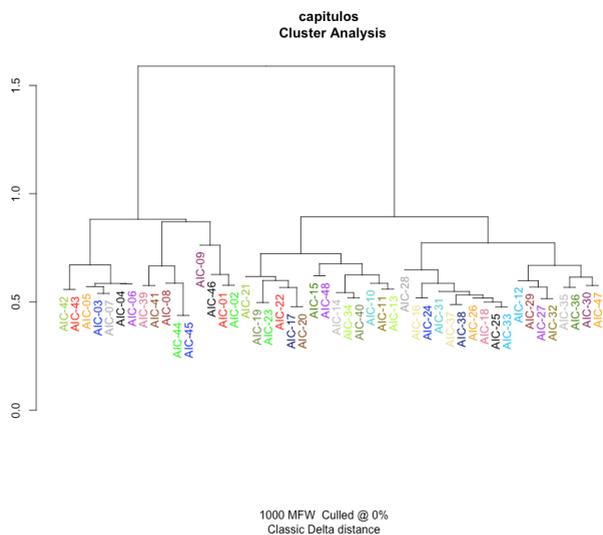


Fig. 11: Los capítulos del *Libro de la caza de las aves* (1000 MFW)

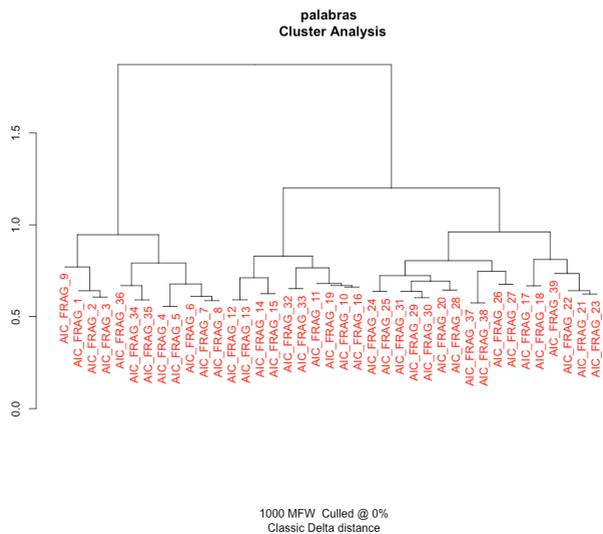


Fig. 12: El *Libro de la caza de las aves* dividido en fragmentos de 1000 palabras

encuentra repartido entre los fragmentos 11 y 20, se encuentra a caballo entre los dos grandes grupos, pero los pasajes del capítulo VIII se han agrupado dentro de un subgrupo homogéneo.

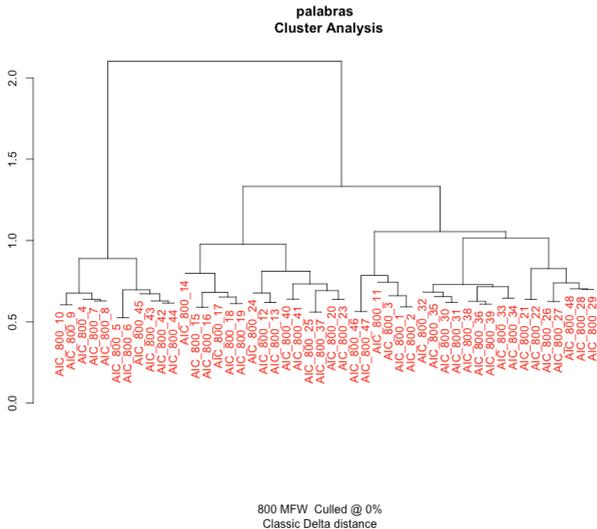


Fig. 13: El *Libro de la caza de las aves* dividido en fragmentos de 800 palabras

Por todo ello, es necesario revisar con mayor detenimiento este capítulo VIII y su atribución a López de Ayala, aunque no existe ninguna prueba textual de que este capítulo esté basado en la obra de Menino, ya que no tiene ningún lusismo ni falso amigo propio de la traducción ni, por otra parte, ningún testimonio de la obra de Menino habla, ni remotamente, de que el halconero portugués escribiera algo semejante a lo que se expone en este capítulo VIII.

El *Libro de la montería*

Otro texto medieval de interés es el *Libro de la montería*. Este tiene una complicada historia de atribuciones desde los que defienden a ultranza que el autor fue Alfonso X a los que creen que fue su biznieto Alfonso XI⁴⁶. Es un problema que no podrá resolverse porque no existe ningún escrito de Alfonso XI y

⁴⁶ José Manuel Fradejas Rueda, “El autor del Libro de la Montería: historia y comentario de seis siglos de controversia”, en *Actas II Congreso Internacional de la Asociación Hispánica de Literatura Medieval: Segovia, del 5 al 9 de octubre de 1987* (Alcalá de Henares: Universidad de Alcalá, 1992), 285–312.

de los presuntos libros de caza del rey Sabio, no se sabe nada, salvo que fue el promotor de la traducción al castellano del *Libro de los animales que cazan*⁴⁷, texto que, como se ha visto cómo anteriormente, se asienta con toda naturalidad entre las obras de origen árabe del *Scriptorium* alfonsí.

Esta obra es un amplio tratado de caza mayor, *chasse courrant* o *par force*, dividido en tres libros: el primero, con 46 capítulos, trata de la teoría de caza; el segundo, que a su vez se subdivide en dos partes, se ocupa de las heridas y enfermedades de los perros de caza y, el tercero, con 30 capítulos, describe los cazaderos de los dominios del rey de Castilla y León.

Aunque hay muchas ediciones en el mercado, unas mejores que otras, y para mantenernos dentro de la línea de las características editoriales utilizadas en los casos de don Juan Manuel y el corpus alfonsí, para este análisis se parte de la transcripción semipaleográfica de acuerdo con el sistema del Hispanic Seminary of Medieval Studies que realizó Dennis Seniff en los años 1980 para su tesis doctoral⁴⁸ y que corresponde al ms. escorialense Y.II.19, que se estima como el mejor testimonio de esta obra, aunque tiene varios problemas. Uno de ellos es que debió de caer en manos de un encuadernador torpe que descolocó algunos folios, lo cual añade un aliciente a la exploración del texto, puesto que no se pudo segmentar por capítulos, al margen de que habrían sido muy desiguales de tamaño, desde 27 palabras (2.2.13) a los varios cientos de palabras de cualquier capítulo del libro tercero, por lo que se decidió segmentarlo en fragmentos de 1000 palabras, de manera que se consiguieron 117 ficheros de los que los 116 primeros tienen 1000 palabras, mientras que el último tiene 918 (una auténtica suerte)⁴⁹.

Tras analizarlo con *stylo* y aplicando la distancia Classic Delta a las 1000 palabras más frecuentes, se obtuvo el dendrograma de la figura 14, ciertamente ilegible. Pero en él se observan dos grandes grupos muy separados entre sí. ¿A qué responde esta nítida división? Sencillamente, a que el grupo de la parte superior (fig. 15), el más pequeño, corresponde a una fuente específica del *Libro de la montería*: todos esos segmentos, aunque pueda haber solapamientos debido al corte cada 1000 palabras, se remontan a los libros cuarto y quinto del *Libro de los animales que cazan*, ese texto traducido del árabe que se asentaba con comodidad en medio del grupo árabe del *Scriptorium* alfonsí.

⁴⁷ Al-Bāzyār, *Libro de los animales que cazan* = *Kitāb al-Īawāriḥ*.

⁴⁸ Alfonso, *Libro de la montería: based on Escorial MS Y.II.19*, ed. Dennis P. Seniff (Madison: Hispanic Seminary of Medieval Studies, 1983).

⁴⁹ El texto tiene 115912 palabras-token y 7470 palabras-tipo, de las que 3650 son hápax.

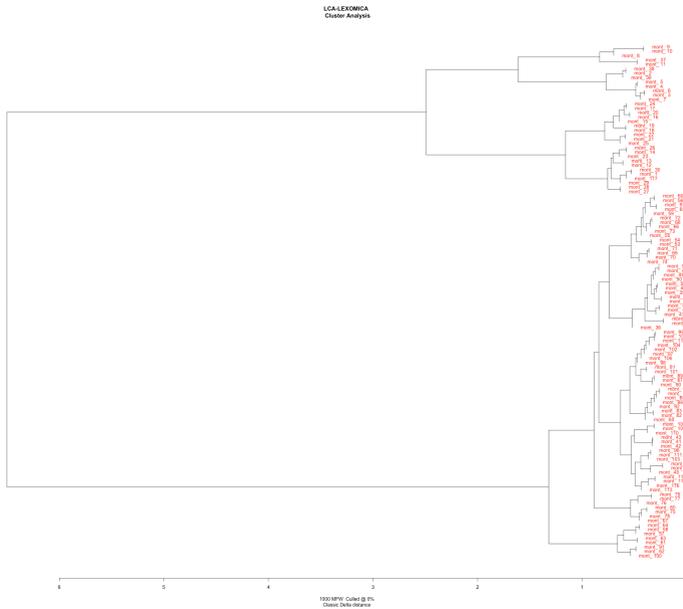


Fig. 14: Dendrograma del *Libro de la montería* dividido en fragmentos de 1000 palabras

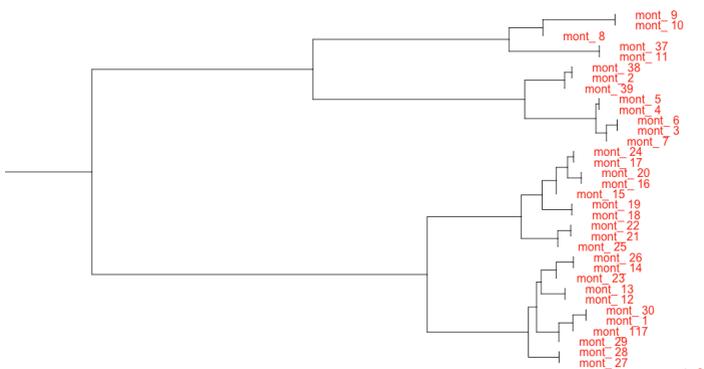


Fig. 15: Detalle de la figura 14

Conclusión

Convencidos de que los métodos estilométricos son válidos y de que pueden ayudarnos a decidir si un texto es de un autor o no (o, en el peor de los casos, para descartar atribuciones tradicionales), el gran problema es si es necesario normalizar, estandarizar, regularizar los textos que se pretenden analizar cuando se trata de textos de la Edad Media. Se ha visto que los investigadores del Wheaton College abogan por el uso de ediciones críticas, pero estos experimentos con los textos castellanos, tanto si están normalizados, como es el caso de las obras de don Juan Manuel o el *Libro de la caza de las aves* de Ayala, como si se trata de una edición semipaleográfica (Juan Manuel –Ayerbe-Chaux–, *Scriptorium* alfonsí, *Libro de la montería*, parece que no cambian sustancialmente los resultados. Somos conscientes de que estas son unas pequeñas pruebas sobre textos que sabía cómo debían de funcionar: el *Libro de la montería*, el *Libro de la caza de las aves*, el *Scriptorium* alfonsí; ahora bien, queda una pregunta final que aún no sé responder: ¿Qué es lo que hace que un sistema automatizado, que solo ha tenido en cuenta 100 palabras, haya detectado, sin error alguno, que un grupo de obras son traducción del árabe y el otro gran grupo son obras con fuentes, no diré traducción, latinas? Cuestión aparte es el errático comportamiento del capítulo VIII del *Libro de la caza de las aves*.

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.