

Cohesión estructural y capas redaccionales en el *Cantar de mio Cid* y en sus prosificaciones cronísticas: un análisis desde la estilometría estadística

José Luis Montiel Domínguez
(Universidad de Huelva)

1. Introducción teórica

La vieja polémica en torno a la doble autoría del *Cantar de mio Cid* (Menéndez Pidal, 1963) ya se abordó desde la informática en el trabajo pionero de Garci-Gómez (1993), pero, hasta donde llega mi conocimiento, no se ha vuelto a tratar con mecanismos más modernos. Montaner (2016), en plena convergencia con Deyermond (1987), deja firmemente establecida la coherencia orgánica del *Cantar*, aunque reconoce diferencias internas entre las tres partes tradicionales (*Destierro*, *Bodas* y *Corpes*) y posibles capas de redacción, que no implican doble autoría, sino que responden a una progresión narrativa bien construida. La crítica moderna del *Cantar* (Zaderenko y Montaner, 2018) tiene ya perfectamente asumida la estructura ascensional del texto desde la caída inicial a la recuperación del honor y la posterior culminación jurídica y simbólica. El diseño tripartito, que no ahoga la concepción unitaria del poema, desempeña en este proceso un papel no meramente editorial, sino funcional. La restauración de la honra, que articula todo el *Cantar*, parece describir una especie de curva en *W*, como se ha argumentado sólidamente en alguna ocasión, o bien una contraposición entre destierro y antidestierro, dentro de lo que Boix (2012) llama “epischen Doppelpunkt”. No obstante, es cierto que la conservación de símbolos comunes y la sensación de una progresión estructural uniformemente acelerada y optimista (Montaner, 2007) no invalidan la intuición subjetiva de que existe una división estilística apreciable entre el primer cantar y los dos restantes, lo cual no necesariamente constituye una prueba definitiva de doble autoría, aunque podría ser un indicio. En este trabajo pretendemos, entre otras cosas, explorar con herramientas estadísticas y estilométricas esa intuición, al menos, en lo que se refiere a las capas de redacción y, con más cautela, en cuanto a la doble autoría, sin excluir la posibilidad de que la configuración unitaria refleje una fase redaccional posterior, como ya afirmaron críticos de la talla de Horrent (1973).

Con respecto a las crónicas que prosifican el texto cidiano, vamos a analizar la *Versión amplificada o sanchina* de la *Estoria de España*, que contiene *Destierro*, y la *Interpolación cidiana*, que se da tras la laguna del capítulo 896 (PCG), donde se prosifica una versión novelesca de *Bodas* y *Corpes*. Esta es aproximadamente la misma tradición textual que refleja la *Crónica particular del Cid*, un retoño de la *Crónica de Castilla* (h. 1300), aunque incorpora escenas burgalesas en el *Destierro* y regulariza aspectos (parentescos e inconsecuencias) presentes en la *Interpolación cidiana*. Por último, vamos a estudiar el tratamiento de las secciones cidianas en la *Versión crítica* de la *Estoria de España*, donde el texto se somete a una abreviación consistente, aunque fiel.

Varios especialistas han abordado la constitución de las prosificaciones desde el importante trabajo de Rochwert-Zuili (1998), una tesis modélica sobre la perspectiva ideológica, que han continuado estudios sobre el simbolismo presente en las crónicas (Blašković, 2025). La crítica actual (Bautista, 2018) sostiene que esencialmente las crónicas comparten una prosificación común en *Destierro* y *Bodas*, y que la de *Corpes* estaría mucho más novelizada y entroncaría con los materiales hagiográficos de la *Leyenda de Cardeña*. Sin embargo, los estudios estilométricos más bien apuntan a una fuerte cohesión entre *Bodas* y *Corpes* en la PCG, y a que las capas de redacción en la

Versión crítica habrían nivelado ligeramente las diferencias subyacentes entre *Destierro* y *Bodas*.

Con respecto al carácter novelizado de la prosificación de *Corpes*, han incidido Bautista (2023) y más recientemente Hijano y Lacomba (2025), a la zaga de los estudios seminales de Catalán (2002). En este trabajo, valiéndonos de herramientas estadísticas y estilométricas, así como de un pormenorizado estudio de la “laguna cidiana”, que no se hacía desde Menéndez Pidal (1955), se defiende la hipótesis de que la prosificación empleada por la *Versión crítica* puede tener puntos en común con la que manejan los redactores de la *Interpolación cidiana*¹. Sin embargo, es indudable que ha sido sometida a un proceso de reforma a partir de *Bodas*, que puede identificarse como una segunda capa de redacción. Además, en consonancia con la opinión de Menéndez Pidal (1980: LXVII), tal redacción puede responder a un “gusto arcaizante” y acaso reactivo frente a la *Interpolación cidiana*.

A continuación, explicaremos en la sección 2 los métodos estadísticos propuestos para el análisis estilométrico, que abordaremos en la sección 3 para el *Cantar* y las crónicas, añadiendo un apartado sobre regresión logarítmica que nos va a permitir identificar el elevado índice de nivelización en la *Versión crítica*, y, secundariamente, en el propio *Cantar de mio Cid*. Concluimos con la sección 4 de discusión y conclusiones, donde realizamos un estudio concreto de la relación entre la *Interpolación cidiana* y la *Versión crítica* durante la historia valenciana del Cid, sección que hace de puente entre la prosificación del *Destierro* y de *Bodas*.

2. Métodos propuestos

2.1. Delta de Burrows

La estilometría clásica opera a partir de las palabras más frecuentes de un texto (MFW ‘Most frequent words’) para comparar los patrones de uso léxico en una escala común y detectar diferencias en los hábitos característicos de un autor, que suelen estar escasamente controlados de manera consciente, como ocurre con las muy extendidas palabras funcionales o *stopwords*. La medida Delta ha demostrado ser especialmente eficaz en estudios de atribución de autoría y análisis estilométricos (Burrows, 2002).

La distancia estilística entre las muestras se mide calculando, en primer lugar, los *z-scores*:

$$z_{i,A} = \frac{f_{i,A} - \mu_i}{\sigma_i}$$

donde la frecuencia relativa de la palabra *i* en el texto *A* ($f_{i,A}$) se resta a la media de esa palabra en el corpus de referencia μ_i , y se divide por la desviación estándar de la palabra en el corpus σ_i .

¹ Bautista (2024: 48) sitúa el ensamblado de la *Interpolación cidiana* hacia el final del reinado de Sancho IV, coincidiendo con el espíritu cruzado de la *Gran conquista de Ultramar*. Un dato precioso a tener en cuenta es la mención de Benito Pérez, cocinero de Sancho IV, como repostero del rey Alfonso VI: “no de otro modo se explica que el repostero del rey Alfonso VI, al que se le ordena adobar manjares para agasajar al de Vivar en las cortes de Toledo, se llame Benito Pérez (PCG, 615b₂₇), igual que uno de los cocineros del rey Sancho IV (Gaibrois de Ballesteros, (2019 [1922-1928: 41 n. 1])” (Montiel, 2009: 520 n. 9). Las siglas PCG corresponden a *Primera Crónica General*, editada por Menéndez Pidal en 1906, y reeditada en facsímil por el BOE (2022). Para Lacomba e Hijano (2025: 163) la *Interpolación cidiana* constituye “la verdadera fundación del género de historia novelesca”, que se proyecta teatralmente en una “historia dramatizada” con abundancia de diálogos y exageraciones.

En segundo lugar, se calcula la distancia Manhattan, definida como la media de la suma de las diferencias absolutas entre los vectores de *z-scores* correspondientes a las palabras más frecuentes de dos de los textos comparados:

$$\Delta(A, B) = \sum_{i=1}^M |z_{i,A} - z_{i,B}|$$

donde Δ es la distancia, A y B los dos textos, i la palabra frecuente en uno y otro texto, M el número total de palabras frecuentes consideradas, por el que se divide la suma para obtener la media. En definitiva, Delta es la media de las diferencias absolutas entre los valores estandarizados (*z-scores*) de cada palabra frecuente en los dos textos. Para calcularla, se puede recurrir al paquete *Stylo* de R (Eder, Rybicki y Kestemont, 2016), que no solo ofrece el estadístico, sino también figuras de dendrogramas de relación entre textos.

2.2. Distancia dinámica de Mahalanobis

Para analizar cómo evoluciona la separación estilística a medida que se incorporan palabras frecuentes, puede emplearse otra nueva métrica. La distancia dinámica de Mahalanobis (Montiel, en prensa) incorpora información sobre la variación incremental del contraste léxico. Partiendo de los *z-scores* de Burrows, y de k (las palabras más frecuentes), obtenemos:

$$D_{final,k}^2 = \frac{D_k^2 \overline{\Delta R}_k}{\overline{\Delta D}_k}, \quad \overline{\Delta D}_k > 0$$

Por un lado, la distancia de Mahalanobis D_k^2 (Mahalanobis, 1936) mide la diferencia entre dos textos ponderando por la matriz inversa de covarianzas del corpus de referencia, es decir, considerando las relaciones internas entre las palabras más frecuentes. En este sentido, la inversión de la matriz redistribuye la contribución efectiva de cada palabra en función de su varianza y de su correlación con las demás, lo que permite que afloren los pesos subyacentes del sistema léxico, al atenuar las redundancias y reforzar las dimensiones que aportan información independiente.

Pero la versión dinámica reescala este estadístico multiplicándolo por el promedio de la variación cuadrática discreta de inspiración riemanniana $\overline{\Delta R}_k$ –la media de los cuadrados de los incrementos sucesivos del contraste léxico, entendido como la distancia en la frecuencia estandarizada de cada palabra entre los dos textos– y lo divide por la tasa media acumulada de crecimiento de la propia distancia $\overline{\Delta D}_k$. Este procedimiento permite ponderar la magnitud global de la diferencia por su evolución estructural, y, al dividirla por el crecimiento medio de la propia distancia, evita que la adición de nuevas palabras distorsione la comparación o sobredimensione el estadístico.

Secundariamente, el valor obtenido puede calibrarse de manera aproximada mediante referencia a la distribución chi-cuadrado con k grados de libertad, lo que proporciona un p-valor orientativo (cosa de que no disponíamos con Delta de Burrows). Dado que el estadístico ha sido reescalado dinámicamente, esta referencia tiene un carácter heurístico y no debe interpretarse como una ley asintótica estricta.

3. Análisis estilométrico y estadístico

3.1. Distancia Delta en el Cantar y en las crónicas

Como se ha señalado, la distancia Delta se calcula a partir de un corpus de referencia que permite estandarizar las frecuencias mediante *z-scores*. De este modo, la comparación no se limita a los textos analizados de forma aislada, sino que se inscribe en un espacio estilométrico definido por un conjunto mayor, lo que proporciona una calibración más robusta de sus relaciones relativas. Ahora bien, el corpus de referencia no puede ser arbitrario ni indefinido, sino que debe constar de textos estrechamente relacionados entre sí, porque, de lo contrario, las diferencias internas del propio corpus pueden contribuir a inflar artificialmente las semejanzas entre las secciones de un mismo texto. Por tanto, el corpus de referencia manejado en este trabajo se compone de 12 textos muy vinculados: las tres secciones del *Cantar de mio Cid* (*Destierro*, *Bodas*, *Corpes*) en CMC (Montaner, 2016), PCG (Menéndez Pidal, 2022[1906]), Vc (De la Campa, 2009) y CrPart (Izquierdo, 2024). Hemos tenido en cuenta también las transcripciones de los manuscritos en el proyecto *Estoria de Espanna Digital* (Ward, 2026). Todos los textos presentan una extensión aproximada entre 5.000 y 10.000 palabras.

Como se aprecia en la siguiente Tabla 1, los resultados de Delta² apuntan a que en el *Cantar de mio Cid* las secciones de *Bodas* y *Corpes* forman un núcleo cohesionado, y *Destierro* presenta una diferenciación sistemática, especialmente frente a *Corpes*:

MFW	Destierro vs Bodas	Destierro vs Corpes	Bodas vs Corpes	Rango
10	0.837	0.601	0.706	0.236
20	0.776	0.608	0.567	0.209
50	0.647	0.691	0.562	0.129
100	0.664	0.705	0.623	0.082
200	0.703	0.807	0.661	0.145
500	0.674	0.793	0.680	0.120
1000	0.757	0.843	0.743	0.100
media_delta	0.722	0.721	0.649	0.145
pct_cercano	14.286 %	14.286 %	71.429 %	57.143
pct_intermedio	57.143 %	14.286 %	28.571 %	42.857
pct_lejano	28.571 %	71.429 %	0.000 %	71.429

Tabla 1. Distancias de Burrows (Δ) entre las secciones *Destierro*, *Bodas* y *Corpes* en función del número de palabras más frecuentes (MFW = 10–1000). Se presentan los valores de Δ para cada comparación por pares y el rango MFW (diferencia entre el valor máximo y el mínimo en cada fila). La fila *media_delta* recoge la media de Δ a lo largo de todos los MFW considerados, mientras que *pct_cercano*, *pct_intermedio* y *pct_lejano* indican el porcentaje de veces que cada comparación ocupa la posición relativa correspondiente dentro de cada nivel de MFW

El escenario es similar en el texto editado en PCG, que corresponde a dos secciones distintas. El *Destierro* está en la *Versión amplificada*; en cambio, *Bodas* y *Corpes* se encuentran en la llamada *Interpolación cidiana*, que contiene una versión novelizada del texto épico. La Tabla 2 muestra en PCG valores más extremos que en el CMC: mayor cohesión entre *Bodas* y *Corpes*, correlación intermedia entre *Destierro* y *Bodas*, y mayor distancia entre *Destierro* y *Corpes*.

² El cálculo se ha realizado con el programa RStudio. La distancia se establece entre dos textos, por lo que es necesario realizar tres operaciones para comparar tres textos entre sí.

MFW	Destierro vs Bodas	Destierro vs Corpes	Bodas vs Corpes	Rango
10	0.784	1.074	0.643	0.431
20	0.723	0.854	0.667	0.1
50	0.742	0.753	0.699	86
100	0.896	0.826	0.782	0.054
200	0.869	0.873	0.754	0.114
500	0.860	0.924	0.783	0.119
1000	0.847	0.920	0.783	0.141
media_delta	0.817	0.889	0.730	0.137
pct_cercano	0.000	0.000	100.000	0.168
pct_intermedio	85.714	14.286	0.000	100.000
pct_lejano	14.286	85.714	0.000	85.714

Tabla 2. Distancias de Burrows (Δ) entre *Destierro*, *Bodas* y *Corpés* de la PCG según el número de palabras más frecuentes (MFW = 10–1000). Se muestran los valores de Δ por comparación y el rango intra-MFW, junto con la media global y el porcentaje de posiciones relativas (cercano, intermedio, lejano) en el conjunto de niveles considerados

Por su parte, la *Crónica particular* presenta una estructura compatible con la de PCG, pues reelabora el texto a partir de la misma versión “novelizada” del *Cantar*, pero incorpora además escenas en los primeros compases del *Destierro*, y ofrece claras señales de regularización, como apuntábamos en la sección 1. La Tabla 3 refleja estas propiedades: los valores muestran una mayor moderación y una reducción progresiva del rango a medida que aumenta el número de MFW.

MFW	Destierro vs Bodas	Destierro vs Corpes	Bodas vs Corpes	Rango
10	0.439	0.712	0.569	0.273
20	0.724	0.641	0.672	0.083
50	0.679	0.729	0.682	0.050
100	0.781	0.735	0.695	0.085
200	0.771	0.810	0.717	0.092
500	0.833	0.848	0.807	0.041
1000	0.797	0.869	0.782	0.087
media_delta	0.718	0.763	0.704	0.101
pct_cercano	28.571 %	14.286 %	57.143 %	42.857
pct_intermedio	42.857 %	14.286 %	42.857 %	28.571
pct_lejano	28.571 %	71.429 %	0.000 %	71.429

Tabla 3. Distancias de Burrows (Δ) entre las secciones *Destierro*, *Bodas* y *Corpés* en la *Crónica particular del Cid* según distintos niveles de MFW (10–1000). Se presentan los valores de Δ por comparación, el rango intra-MFW, la media global y la distribución porcentual de posiciones relativas (cercano, intermedio, lejano)

No obstante, en la *Versión crítica*, aunque *Destierro vs Corpés* mantiene en promedio la mayor distancia (media $\Delta = 0.903$), las diferencias tienden a moderarse en los niveles altos de MFW, indicando una configuración interna relativamente equilibrada

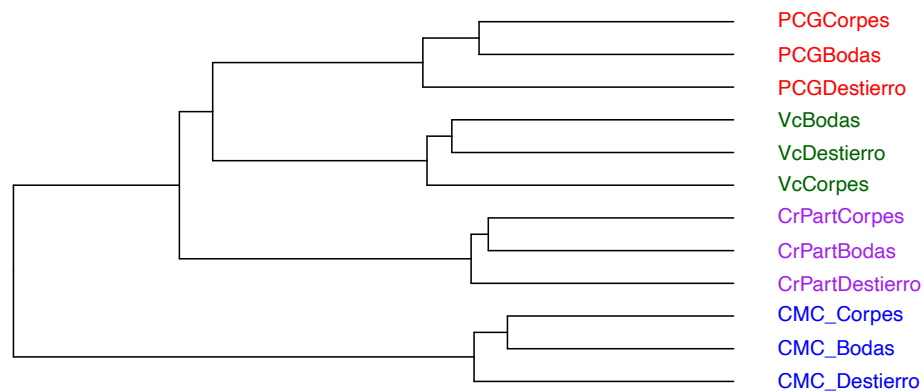
del conjunto textual, motivada tal vez por un efecto de nivelación estilística, como se desprende de la Tabla 4.

MFW	Destierro vs Bodas	Destierro vs Corpes	Bodas vs Corpes	Rango
10	0.808	1.030	0.821	0.222
20	0.709	0.949	0.788	0.240
50	0.838	1.044	0.919	0.205
100	0.898	0.862	0.805	0.093
200	0.813	0.829	0.795	0.035
500	0.803	0.802	0.829	0.027
1000	0.787	0.804	0.780	0.024
media_delta	0.808	0.903	0.819	0.120
pct_cercano	42.857 %	14.286 %	42.857 %	28.571
pct_intermedio	42.857 %	14.286 %	42.857 %	28.571
pct_lejano	14.286 %	71.429 %	14.286 %	57.143

Tabla 4. Distancias de Burrows (Δ) entre *Destierro*, *Bodas* y *Corpes* en la **Versión crítica**, según distintos niveles de MFW (10–1000). Se presentan los valores de Δ por comparación, el rango intra-MFW, la media global y la distribución porcentual de posiciones relativas

La Figura 1 muestra la vinculación entre los textos a partir de la media de Delta:

Dendrograma horizontal (Delta medio)
MFW: 10,20,50,100,200,500,1000



Distancia media (Burrows' Delta)

Figura 1. Dendrograma con las relaciones entre textos a partir de la distancia media de Delta en MFW

Todos los bloques textuales distinguen dos nodos, uno correspondiente al *Destierro* y otro al conjunto formado por *Bodas* y *Corpes*. No obstante, en la *Versión crítica* se singulariza la prosificación de *Corpes*, que aparece escindida, mientras que *Destierro* y *Bodas* quedan integrados en un mismo nodo.

3.2. Regresión logarítmica y nivelación textual

Ahora bien, este resultado no parece ser un efecto directamente atribuible al original del que parten los cronistas de la *Versión crítica*, sino a que precisamente los datos muestran una tendencia clara a la reducción del rango conforme aumenta el número de MFW, especialmente a partir de las 100 palabras más frecuentes, lo que sugiere un efecto de estabilización o nivelación estilométrica. Con objeto de ponderar ese grado de nivelación, que ciertamente (y en honor a los viejos maestros que ya lo intuyeron y argumentaron filológicamente) podría envolver una segunda capa de redacción, se puede aplicar una regresión lineal logarítmica a la cadena de rangos de cada obra, mediante la siguiente ecuación:

$$Rango_i = \beta_0 + \beta_1 \ln(MFW_i) + \varepsilon_i$$

donde β_0 es el intercepto (el valor estimado del rango cuando $MFW = 1$), β_1 la pendiente logarítmica, y ε_i el error residual (la parte de la variación que el modelo no logra explicar), y que se estima por mínimos cuadrados ordinarios. El p-valor asociado a la pendiente se obtiene mediante un estadístico t , dividiendo la estimación de la pendiente β_1 por su error estándar. Como puede apreciarse en la Tabla 5, el efecto de la regresión logarítmica resulta estadísticamente significativo en CMC y especialmente en la *Versión crítica* (Vc), lo que sugiere un proceso de nivelación estilométrica más marcado en estas tradiciones.

Modelo	Pendiente	p	R ²
CMC	-0.026	0.037	0.62
PCG	-0.042	0.177	0.33
CrPart	-0.027	0.158	0.36
Vc	-0.054	0.002	0.87

Tabla 5. Resultados de la regresión logarítmica del rango de las distancias Δ entre secciones en función del número de MFW para cada obra. Se presenta la pendiente del modelo, su nivel de significación (p) y el coeficiente de determinación (R²). Las pendientes negativas indican una tendencia a la reducción de la dispersión conforme aumenta el vocabulario considerado.

Este comportamiento es compatible con la coexistencia de dos niveles redaccionales: uno más divergente en los estratos léxicos más marcados y otro más homogéneo que se impone en los niveles altos de MFW, probablemente asociado a procesos de regularización estilística. Por otra parte, los altos coeficientes de determinación (R²) de Vc (0.87) y del CMC (0.62) sugieren que la tendencia estructural de la regresión logarítmica es estable.

La aplicación de una regresión lineal logarítmica sobre el rango de las distancias Δ no tiene por objeto modelizar un proceso lingüístico en sentido causal, sino detectar tendencias sistemáticas en la estabilización o dispersión de la diferencia estilométrica conforme aumenta el número de palabras más frecuentes consideradas. La elección de una transformación logarítmica responde al comportamiento típicamente decreciente y no lineal de la varianza léxica al ampliar el vocabulario funcional, fenómeno ampliamente documentado en análisis de frecuencia textual. El rango, más que la varianza absoluta, se adopta aquí como indicador sintético de dispersión entre secciones, permitiendo comparar obras con tamaños y estructuras similares bajo un mismo criterio relativo. La significación estadística de la pendiente se interpreta, por tanto, como evidencia de tendencia interna consistente, no como prueba de estratificación en sí misma; su valor

reside en señalar patrones de nivelación compatibles —y solo compatibles— con hipótesis filológicas independientes.

3.3. Distancia dinámica en el Cantar y en las prosificaciones

La incorporación de la distancia de Mahalanobis no pretende sustituir la métrica Delta de Burrows, sino complementarla en un aspecto que esta no controla: la correlación interna entre las palabras más frecuentes. Mientras que Delta opera sobre diferencias absolutas estandarizadas, Mahalanobis pondera dichas diferencias por la matriz inversa de covarianzas del corpus de referencia, evitando que la redundancia léxica o la colinealidad entre términos inflen artificialmente la distancia global. La versión dinámica aquí propuesta introduce, además, un factor de reescalado basado en la evolución incremental del contraste léxico, con el fin de evitar el crecimiento mecánico del estadístico al aumentar el número de MFW. Debe subrayarse que la calibración mediante la distribución χ^2 tiene un carácter estrictamente heurístico, dado que el reescalado altera las condiciones asintóticas clásicas; por tanto, los valores p se interpretan como indicadores comparativos internos y no como pruebas de significación en sentido fuerte. El interés principal del procedimiento radica en ofrecer una medida estructuralmente sensible a la evolución del contraste, no en establecer umbrales absolutos de decisión inferencial.

En la siguiente Tabla 6 se ofrecen los resultados para MFW 10 de la aplicación de la distancia dinámica de Mahalanobis a la comparación textual entre secciones³.

Comparación	Mahalanobis	DeltaR	Tasa acumulada	Mah_Final	p_valor
CMC_Destierro vs CMC_Bodas	21.523	1.283	2.389	11.558	0.316
CMC_Destierro vs CMC_Corpes	21.953	1.155	2.439	10.398	0.406
CMC_Bodas vs CMC_Corpes	21.774	0.659	2.415	5.939	0.820
PCGDestierro vs PCGBodas	15.565	4.157	1.729	37.415	0.000
PCGDestierro vs PCGCorpes	21.302	5.729	2.367	51.562	0.000
PCGBodas vs PCGCorpes	19.087	1.683	2.121	15.147	0.127
VcDestierro vs VcBodas	21.953	2.065	2.438	18.589	0.046
VcDestierro vs VcCorpes	21.988	3.603	2.426	32.655	0.000
VcBodas vs VcCorpes	21.897	1.930	2.422	17.443	0.065
CrPartDestierro vs CrPartBodas	8.223	0.787	0.911	7.106	0.715

³ El número de variables p (MFW) debe ser menor al número de observaciones n (textos) para garantizar la invertibilidad de la matriz de covarianzas de Mahalanobis, ya que en caso contrario sería singular (su rango máximo es $n - 1$).

CrPartDestierro vs CrPartCorpes	18.322	2.221	2.028	20.063	0.029
CrPartBodas vs CrPartCorpes	18.748	1.524	2.082	13.728	0.186

Tabla 6. Resultados del Mahalanobis dinámico con 10 MFW para las comparaciones por pares entre secciones (Destierro, Bodas y Corpes) en los cuatro textos (CMC, PCG, Vc y CrPart). Se reportan la distancia de Mahalanobis acumulada (D_k^2), la variación cuadrática media ($\overline{\Delta R_k}$), la tasa acumulada de crecimiento del estadístico ($\overline{\Delta R_k}$) y la distancia ajustada final ($D_{final,k}^2$), junto con su calibración χ^2 (p-valor).

Como puede observarse, la sección *Destierro* de PCG (en la *Versión ampliada*) se separa significativamente (p-valor: 0.000) de *Bodas* y *Corpes* (en la *Interpolación cidiana*), pero en Vc la diferencia es más acusada entre *Destierro* y *Corpes*, y marginal aunque apreciable entre *Destierro* y *Bodas* (p-valor: 0.04). En CMC, pese a valores altos de Mahalanobis acumulado, la corrección dinámica reduce el estadístico final y no se alcanzan niveles de significación, lo que apunta a mayor homogeneidad estructural entre secciones. Por su parte, la CrPart muestra un patrón intermedio; solo el contraste *Destierro vs Corpes* resulta significativo (p-valor: 0.029), mientras que el resto permanece dentro de la variación esperable.

Discusión y Conclusiones

El estudio estilométrico y estadístico apunta, pues, a una segunda capa de redacción en la *Versión crítica*, y probablemente también en el *Cantar de mio Cid*, no necesariamente debida a doble autoría o realizada en épocas diversas, aunque son hipótesis que no parecen completamente descartables. Me baso para tal afirmación, además de en los datos estilométricos comentados arriba, en el estudio pormenorizado de la historia valenciana del Cid –posterior a la laguna del capítulo 896 de PCG–, que hace de puente de unión entre la prosificación del *Destierro* y de *Bodas*.

La *Versión crítica* sigue muy de cerca el orden de la PCG (*Interpolación cidiana*), pero forma los capítulos combinando varios de la PCG. Se añade solo la cronología del año 28.º en el capítulo CCCXXX. Como es habitual, se producen importantes y frecuentes omisiones⁴, a veces pasajes extensos, tanto en el interior de los capítulos de la PCG como al principio (caps. 902, 904, 905) y al final (caps. 901, 903, 905, 906, 907), pero lo verdaderamente reseñable es que el texto que se conserva en la *Versión crítica* no se reformula, sino que coincide palabra por palabra con el que aparece en la *Interpolación cidiana*. Valga como ejemplo, entre los muchos que se pudieran aducir, este pasaje sobre el sartal de Zobayda:

⁴ En ningún caso, este sistema extremo de abreviación se conoce en secciones anteriores a la laguna. Las supresiones afectan particularmente a Juballa y al “almoxerifadgo” valenciano (por ejemplo, en la toma de Alcuía; PCG, cap. 904). También se resume, por ejemplo, el contenido de la carta de Abenial al Cid (PCG, cap. 900, 568a25-42 / Vc, CCCXXVIII77), o bien se omiten otras misivas en el avance de los almorávides (PCG, cap. 907). Pero lo más significativo es que las adiciones son muy contadas y se circunscriben a explicaciones o interpretaciones deducibles del contexto (PCG, cap. 897). Las adiciones informativas se reducen a tres pasajes: toma de “Çebolla” (Vc, CCCXXVIII78-84), los del Cid no se hacen con el dinero enviado a los almorávides (Vc, CCCXXX8-11), fecha de la salida de los almorávides de Valencia (Vc, CCCXXXI54-56), que pueden corresponder a deducciones del compilador, pasajes lagunosos o glosas marginales.

*Et quando saliera este rey del alcaçar sacara consigo de sus thesoros del maspreciado et mas noble aljófar que podrie ser quel non podien fallar en ningun logar tal nin meior; et otrossi de piedras preciosas, muchos çafiles et de rubis et de esmeraldas et de otras piedras muy preciosas; et saquara vna arqueta que era toda de oro lenna et de todas estas cosas; et tenie en su çinta vn sarta de piedras preciosas et daliofar tal qual nunca rey ouiera nin cosa tan preciada commo aquel sartal era, et diz que fue de Zobayda, mujer de Haron Arraxit el que fue señor de *Baldac, et passo despues a los reyes que dizen Benhumaya que fueron sennores del Andaluzia, despues fue este sartal de Almemon el que fue señor de Toledo, et ouieralo Yahia rey de Valencia de su madre que fuera fija de Almemon. Et quando sopo Abeniaf daquel sartal et de las otras cosas preciadas que tenia este que fuera rey de Valencia, creçiol grant cobdicia dello, et cuydo en su coraçon commo lo auria que gelo non sopiesse ninguno. (PCG, cap. 898, 567a19-42; en cursiva lo conservado literalmente por la Vc, CCCXXVII59-61)*

O este otro sobre el respeto a los labradores en el cerco valenciano, impuesto por el Cid:

Et enbiaua el Çid sus algaras que corriesen a Valencia dos uezes al dia, los vnos yuan a la mañana et los otros contra la noche, et robauan los ganados et catiuaban a quantos que fallauan, fueras a los labradores que labrauan por pan; ca el Çid tomara pleito et omenage a los caualleros et a los adalides et a los almocadenes que non fiziessen mal a los labradores, mas que los falagassen et les dixiessen que labrassen et fiziessen algo; et quando fuese al tiempo de coger pan, si algun acorro les viniessen que aurien que comer, “et si non nos viniere ajuda auremos otrossi que comer et uos”. (PCG, cap. 901, 568b37-569a4; subrayado lo conservado literalmente en la Vc, CCCXXVIII91-92)

Son muy destacables las omisiones de carácter ideológico⁵, siendo contadas y escasísimas las adiciones⁶. Con respecto al empalme entre esta sección y la del comienzo

⁵ La supresión ideológica de capítulos enteros intenta paliar la crueldad del asedio: elegía de Alhuacaxí (PCG, caps. 909-910), penalidades de los valencianos y escenas dantescas (PCG, caps. 913-915, y cap. 916 salvo algunas líneas conservadas) y evacuación de Valencia (PCG, cap. 921). La Vc omite también la supuesta autoría del relato: “Et diz Abenalfarax en su arauigo, onde esta estoria fue sacada» (PCG, cap. 911, 578b30-31), y la referencia proléptica a la *Leyenda de Cardeña* en cuanto a la conversión de Alhuacaxí (PCG, 591a5-12). Se suprime por doquier el dramatismo del alza de precios en Valencia, o se sustituye por una frase de este tenor: “En todo esto era la carestía muy grande en Valencia” (Vc, CCCXXXIII121, CCCXXXIV144; PCG, 574b45-575a7, 575b17-29). Sin embargo, los precios bajan después de la entrada del Cid en la ciudad (PCG, 588a21-33). En la *Versión crítica* se omiten también las promesas incumplidas del Cid entre el primer discurso a los valencianos (PCG, cap. 918) y el segundo (PCG, cap. 919). Se lleva a cabo un arreglo resumiendo enormemente el primer discurso y enlazándolo con el segundo, donde se les exige a los valencianos la entrega de Abeniaf. Así, las palabras del segundo discurso no suenan a rectificación, como parece en la PCG, ni se conservan comentarios que pudieran interpretarse como reproches a la actuación del Cid (PCG, 590a19-23). No aparece tampoco en la *Versión crítica* la justificación de la crueldad del asedio como consecuencia de la obcecación de los valencianos, ni la disculpa del Cid por las penalidades sufridas (PCG, 589b12-18). Los cronistas de la *Versión crítica* retocan también los pasajes sobre la acusación y ejecución de Abeniaf para evitar dudas sobre su culpabilidad (Vc, CCCXXXVI154-58); por ejemplo, suprimiendo la consulta con el almojarife cidiano. En estos últimos capítulos de la sección la *Crónica de 1344* y la *Crónica particular del Cid* ofrecen un texto más completo.

⁶ La más reseñable (Vc, CCCXXXIV129-144) es una adición en la que se combinan informaciones tomadas de la *Historia Roderici* (§ 57-58; § 47 para lo del castillo de Borja) y el *Cantar* (vv. 1181-1205). El capítulo acaba con la misma fórmula de abreviación sobre la carestía que el precedente, donde también se habían prosificado versos del *Cantar* sobre el cerco de Murviedro.

de la prosificación de *Bodas* en la *Versión crítica*, como es sabido se prosifica el episodio del cerco de Murviedro (Vc, CCCXXXIII82-107 / CMC, 1095-1171), que presenta un grado de conservación⁷ de alrededor del 40 % de los versos del episodio, mucho menor que el del último episodio del *Cantar del destierro*.

En resumidas cuentas, se dan diferencias notables en la compilación de la *Versión crítica* antes y después de la laguna del capítulo 896. Tales diferencias podrían explicarse por la existencia de un arreglo en la *Versión crítica* que cubriese las secciones posteriores a la laguna, incluyéndose la decisión de prosificar el *Mío Cid* del siglo XIII a partir del *Cantar de las bodas* en contra del resto de la tradición cronística, que asume una versión refundida o novelizada. La utilización del *Cantar* original se explicaría por la marcada tendencia a la abreviación, impuesta en la *Versión crítica* después de la laguna. El arreglo que estamos considerando sería realizado posiblemente por otro equipo de trabajo y con posterioridad al tratamiento de las secciones anteriores a la laguna –lo que encuentra además apoyo en el estudio estilométrico y estadístico realizado arriba–, y en virtud del rechazo que debían causar los materiales de fuente árabe (Ibn ‘Alqama), muy reducidos después de la laguna en comparación con lo que ocurría antes. Los nuevos redactores los sustituyeron en algunos casos por informaciones procedentes de la *Historia Roderici* y del *Cantar de mio Cid*.

Así pues, las crónicas realizan cambios notables conducentes a la verosimilitud del relato⁸ para facilitar su coherencia y justificar el halo cruzado y, por ende, hagiográfico, con el que ha permanecido el Cid en el inconsciente colectivo hispánico. Con estos datos se puede afirmar que a nivel estilométrico la prosificación de la *Versión crítica* del segundo y del tercer cantar difiere sustancialmente de la del primero, por lo que sufrió una reelaboración sustancial tras operarse la llamada “laguna cidiana”. Esa hipótesis la defendió Montiel (2015, I: 521-525), pero encuentra ahora respaldo adicional en los indicios estadísticos y estilométricos. Además, la reelaboración y abreviación de la primitiva prosificación alfonsí en la *Versión crítica* se hizo teniendo muy presente la de la *Interpolación cidiana*; es, pues, en su capa superficial un texto posterior a la versión novelizada del *Cantar*, y posiblemente condicionado por su existencia previa.

No obstante, los resultados estilométricos obtenidos no deben interpretarse como demostración concluyente de una segunda autoría o de una estratificación cronológica diferenciada, sino como la detección de patrones cuantitativos de cohesión y nivelación interna que resultan compatibles con tales hipótesis. La estilometría, en este sentido, no sustituye el análisis filológico tradicional, sino que lo complementa, ofreciendo un criterio independiente para evaluar intuiciones previamente formuladas por la crítica. La convergencia entre los indicios estadísticos –especialmente la tendencia significativa a la nivelación en la *Versión crítica*– y los datos textuales relativos a omisiones, reordenaciones y regularizaciones posteriores a la laguna cidiana refuerza la plausibilidad de una intervención redaccional diferenciada. Tales resultados deben entenderse como evidencia probabilística dentro de un marco interpretativo más amplio y no como prueba determinante en sentido estricto.

⁷ También se conservan en la Vc los pregones para asediar Valencia (CMC, 1181-1205). La Vc y la PCG comparten la prosificación de los versos sobre la entrada en la ciudad (CMC, 1209-1220).

⁸ Montaner (2014: § 24) equipara el concepto de verosimilitud entre los cronistas al de autenticidad, de tal manera que los cambios que operan en sus fuentes –no siempre consistentes, en mi opinión–, como el de la racionalización cronística del episodio de Alcoer, buscan ofrecer una imagen más confiable del héroe, acorde con los valores del *volksgeist* medieval. Sin embargo, este enfoque objetivista en la historia, al modo hegeliano (Hegel, 2004), ha sido puesto reiteradamente en cuestión por otros filósofos, que solo ven en la historia la mano subjetiva del historiador (Collingwood, 1993).

Obras citadas

- Bautista Pérez, Francisco. "The *Poema de mio Cid* in 13th- and 14th- Century Romance historiography". En Alberto Montaner Frutos e Irene Zaderenko eds. *A Companion to the "Poema de Mio Cid"*. Leiden: Brill, 2018. 412-461.
- . Interpolación cidiana". *Cuadernos de Historia de España* 90 (2023): 177-200.
- . "Un Cid para la modernidad: contexto y tradición de la *Crónica particular del Cid*". En Vicenç Beltrán coord. *Crónica del famoso cavallero Cid Ruy Díez Campeador. Crónica particular del Cid*. México: Frente de Afirmación Hispanista, A. C. (2024): 11-222.
- Blašković, Marija. "Material performances: Heroic threads in the *Crónica de Castilla*". *Studia Historica. Historia Medieval* 43/1 (2025): 199-218. <https://doi.org/10.14201/shhme.31404>
- Boix Jovaní, Alfonso. *El "Cantar de mio Cid": adscripción genérica y estructura tripartita*. Vigo: Academia del Hispanismo, 2012.
- Burrows, John. "Delta: a measure of stylistic difference and a guide to likely authorship". *Literary and Linguistic Computing* 17 (2002): 267-287. <https://doi.org/10.1093/lc/17.3.267>
- Catalán, Diego. *El Cid en la historia y sus inventores*. Madrid: Fundación Ramón Menéndez Pidal, 2002.
- Collingwood, Robin George. *Idea de la historia*. México: Fondo de Cultura Económica, 1993.
- De la Campa, Mariano. *La "Estoria de España" de Alfonso X. Estudio y edición de la "Versión Crítica" desde Fruela II hasta la muerte de Fernando II*. Málaga: Universidad de Málaga, 2009.
- Deyermond, Alan D. *El "Cantar de mio Cid" y la épica medieval española*. Barcelona: Sirmio, 1987.
- Eder, Maciej, Rybicki, Jan, y Kestemont, Mike. "Stylometry with R: a package for computational text analysis". *The R Journal* 8 (2016): 107-121.
- Gaibrois de Ballesteros, Mercedes. *Historia del reinado de Sancho IV de Castilla*. Madrid: Agencia Estatal del Boletín del Estado, Real Academia de la Historia, 2019 [1922-1928], 3 vols.
- Garci-Gómez, Miguel. *Dos autores en el "Cantar de Mio Cid": Aplicación de la informática*. Cáceres: Universidad de Extremadura, 1993.
- Hegel, Georg Wilhelm Friedrich. *Lecciones sobre la filosofía de la historia universal*. Madrid: Alianza, 2004.
- Horrent, Jules. *Historia y poesía en torno al "Cantar del Cid"*. Barcelona: Ariel, 1973.
- Izquierdo Andreu, Almudena ed. *Crónica de famoso cavallero Cid Ruy Díez Campeador. Crónica particular del Cid*. México: Frente de Afirmación Hispanista, A. C., 2024
- Lacomba, Marta, e Hijano Villegas, Manuel. "La *Crónica de Castilla*, paradigma de la historia novelesca". *Revista de Poética Medieval* 39 (2025): 153-185. <https://doi.org/10.37536/RPM.2025.39.1.113147>
- Mahalanobis, P. C. "On the generalised distance in statistics". *Proceedings of the National Institute of Sciences of India* 2/1 (1936): 49-55.
- Menéndez Pidal, Ramón. "Tradicionalidad de las Crónicas Generales de España. A propósito de los trabajos de L. F. Lindley Cintra". *Boletín de la Real Academia de la Historia*, 136 (1955): 131-197.
- . "Dos poetas en el *Cantar de mio Cid*". *Romania* 84 (1963): 145-200.
- . *Reliquias de la poesía épica española; acompañadas de Epopeya y Romancero*. Madrid: Gredos, 1980.

- ed. *Primera Crónica General. Estoria de España que mandó componer Alfonso el Sabio y se continuaba bajo Sancho IV en 1289*. Madrid: Agencia Estatal Boletín Oficial del Estado, 2022 [1906].
- Montaner Frutos, Alberto. “Tal es la su auze: El héroe afortunado del *Cantar de Mio Cid*”. *Olivar* 8/10 (2007): 89-105.
- . “Historicidad medieval y protomoderna: lo auténtico sobre lo verídico”. *e-Spania. Revue interdisciplinaire d’études hispaniques médiévales et modernes* 19 (2014): 1-32.
<https://doi.org/10.4000/e-spania.24054>
- ed. *Cantar de mio Cid*. Madrid: Real Academia Española, 2016.
- Montiel Domínguez, José Luis. “Modelos textuales y cognitivos en la obra alfonsí”. En Teresa Bastardín Candón y Manuel Rivas Zancarrón eds. *Estudios de historiografía lingüística*. Cádiz: Universidad de Cádiz, 2009. 513-529.
- Montiel Domínguez, José Luis. “Estudio comparativo de las prosificaciones del *Cantar de mio Cid* en la historiografía de Alfonso X el Sabio”. Tesis doctoral. Pons Rodríguez, Lola dir. Universidad de Sevilla, 2015, 2 vols.
- (en prensa). “A Dynamically Regularized Mahalanobis Distance for Correlated Multivariate Profiles (with an Application to Burrows’ Delta)”. *Journal of Classification*.
- Rochwert-Zuili, Patricia. “Du Poème a l’Histoire. La Geste cidienne dans l’historiographie alphonsine et neo-alphonsine (XIII^e-XIV^e siècles)”. Thèse sous la direction de M. le Professeur Georges Martin, Université Paris XIII, 1998.
- Ward, Aengus et al. ed. *Estoria de Espanna Digital*, University of Birmingham, 2026
<https://blog.bham.ac.uk/estoriadigital/>
- Zaderenko, Irene, y Alberto Montaner eds. *A Companion to the Poema de mio Cid*. Leiden: Brill, 2018.