



Diferencias internas en *Don Quijote*. Cambios en proporciones y cambios estructurales

Freddy López Quintero
Departamento de Matemática, IVIC

RESUMEN:

En este trabajo se investigan los cambios estilísticos que ocurren dentro de *Don Quijote* utilizándose primeramente pruebas simples de diferencia de proporciones y análisis de varianza multivariante entre las dos partes de *Don Quijote* y luego se buscan los puntos precisos donde ellos han ocurrido. Las variables utilizadas han sido: los *tokens* más frecuentes y una variable de riqueza verbal. Una vez ha sido mostrada la diferencia entre los grupos de variables, se consideran los datos como series temporales y se realizan varias pruebas de cambio estructural sobre ellas. Una de las posibles causas para los cambios encontrados podría ser atribuible a la colaboración de otros autores en el trabajo de Cervantes.

ABSTRACT:

In this paper we investigate the stylistic changes that occur within the Spanish masterpiece *Don Quixote* performing, firstly, simple proportion differences and Multivariate Analysis of Variance between its two parts and then looking for the precise epochs (chapters) where they happen. Variables used were: the most frequent *tokens* and one variable of verbal richness. Once it is found that there are univariate differences in a large group of variables, we consider data as observations of a time series and then we perform several tests to find structural changes in these measures. One of the possible causes for these changes could be attributable to the contribution of other authors in this work by Cervantes.

1. Introducción

Sobre el libro de *Don Quijote de la Mancha*, publicado en dos partes: una en 1605 y otra en 1615, tituladas respectivamente *El ingenioso hidalgo don Quixote de la Mancha* y *Segunda parte del ingenioso caballero don Quixote de la Mancha*, mucho se ha escrito a lo largo de la Historia de la Literatura. Sin embargo, sobre el texto y sobre el autor pesan varias irregularidades.

A continuación se mencionan algunas:

- El giro completo que se le da al protagonista en los dos libros. En el primero, Don Quijote es un personaje demente que alucina situaciones mientras que en la se-

gunda son los demás personajes los que suelen jugarle y representarle algunos espectáculos (de Riquer, 1969). Con su escudero y amigo, Sancho Panza, sucede algo similar. En la primera parte es tratado casi con desprecio por Cervantes ('poca sal en la mollera') mientras que en la segunda, Sancho llega a dar grandes muestras de discreción y sabiduría. Estas y otras variantes han sido resaltadas por Criado de Val (2007).

- No queda claro tampoco por qué Cervantes muestra tantas dudas al apellidar a Don Quijote: ¿Quijada, Quesada, Quijana? (ver Stagg, 1964). Existen también muchas vacilaciones relacionadas con los títulos en la Primera parte y los cambios de palabras en las primeras ediciones, que no son fáciles de aclarar si fueron realizadas por el mismo autor o por los impresores/libreros. Por ejemplo, está documentado que hubo una edición previa del *Quijote* publicada en Valladolid en 1604 que constaba de cuatro partes (lo que conocemos hoy como Primera Parte y que se hace palpable en ediciones modernas como las de Rico, 1998, 2008; o Montgomery, 2009). Eisenberg (2006) señala que no es posible que haya habido solo una primera parte del *Quijote*, sino varias.
- El caso, canónico ya, del robo del rucio de Sancho Panza. Eisenberg (1990) apunta que los acontecimientos del rucio posibilitan que debiera estudiarse el primer libro, o estos pasajes, con tanto escepticismo como otras atribuciones cervantinas. Algunos estudios recientes de atribuciones cervantinas han sido hechos, además de Eisenberg (1990), por Madrigal (2003, 2004, 2006).
- El paralelismo desconcertante entre los primeros seis capítulos del *Quijote* y el llamado 'Entremés de los Romances' no debe pasarse por alto (una edición moderna del entremés, se puede consultar en Eisenberg y Stagg, 2002). El 'Entremés de los romances' es una pequeña obra que utiliza en forma irónica algunos versos de romances de la época para describir las peripecias de un señor que, de tanto leer romances, enloquece y cree que está viviéndolos. La fecha de composición y autor, o autores, del 'Entremés' es desconocida. Al igual que en el *Quijote*, el protagonista abandona su familia y parte por un ideal. Aún más, el mismo Don Quijote y Bartolo (nombre del protagonista del entremés) utilizan los mismos versos en algunos pasajes de la primera parte (ver Capítulo V, Primera Parte y versos 292-3 y 324-5 del *Entremés*). Además, son varios los lugares en los que se reseña que Don Quijote está siguiendo un romance y no un libro de caballerías, como era de esperar de un personaje que ha enloquecido por leer libros de este tipo (Stagg, 2002). Mucho se ha escrito sobre cuál se escribió primero, si *Don Quijote* o el *Entremés*, y si alguno de ellos precede al otro (para conocer quiénes se han propuesto como autores del *Entremés* y un poco de su historia, consultar Rey Hazas, 2007).

Por otro lado, existe la sospecha que circuló, bien en forma impresa o manuscrita, una especie de 'novela ejemplar' que trataba el tema de *Don Quijote de la Mancha* antes de la publicación de lo que se conoce como *Primera Parte*. Esta novelita llevaría por título 'Hidalgo de la Mancha', justificando así la unidad cerrada de los primeros capítulos y explicaría la mención que se hace de este material por algunas personas antes de 1605. Esta propuesta la han intentado demostrar varios críticos, razonando que ciertos párrafos y

oraciones fueron intercalados para dar armonía a un libro de la dimensión completa del *Quijote*. Esta tesis ha sido defendida por Stagg y descartada tajantemente por Menéndez Pidal (Stagg, 1964).

Sobre Miguel de Cervantes, Eisenberg (2006) indica que: a) no consta que haya existido, puesto que su segundo apellido era Cortinas y no Saavedra; tampoco consta que haya nacido en Alcalá de Henares (sino solo que fue bautizado en esta ciudad) y b) cuando firmaba no usaba la letra v sino la letra b (Cerbantes). Se ha llegado incluso a especular que Cervantes y Shakespeare eran la misma persona y que Thomas Shelton, primer traductor al inglés del libro, nunca vivió y fue un ardid para que Sir Francis Bacon, bajo el pseudónimo de Cervantes, publicara la obra (una revisión de estas últimas teorías puede consultarse en Carr, 2004). Además, sobre Shelton existe la teoría que no tradujo todo el *Quijote*, pese a que su nombre esté en esas primeras versiones al inglés. Ver Montgomery (2008).

En esta investigación se buscarán diferencias de estilo dentro de las dos partes del *Quijote* atendiendo a las variables estilométricas tradicionales, a saber: palabras, signos de puntuación, riqueza verbal y frases de dos y tres palabras más frecuentes. En caso de existir estas diferencias de estilo interno, se buscarán los puntos (capítulos) donde ocurren estos cambios y se hará una breve revisión de las posibles causas para que estos hayan ocurrido.

2. Metodología

En esta Sección presento brevemente la metodología utilizada. La exposición no se hace técnicamente en profundidad y exhorto a los lectores a consultar las fuentes indicadas para ahondar en las técnicas.

2.1. Datos

Las pruebas y métodos utilizados descansan sobre el supuesto de que el libro de *Don Quijote* fue escrito secuencialmente, esto es, el capítulo $(j+1)$, $j=1, \dots, n-1$, siempre fue escrito posteriormente al capítulo j , para los n capítulos del libro¹. Por 'libro' se entenderá la *Primera y Segunda Parte* en un solo volumen, contando así, inicialmente, con $n=52+74=126$ capítulos.

Para todo el texto de *Don Quijote* se han contabilizado las 30 palabras más frecuentes y se han medido en cada uno de los capítulos. Adicionalmente, se han contado los 9 signos de puntuación frecuentes, las tres frases de dos palabras más usadas y la frase de tres palabras más utilizadas. Cada una de estas contabilizaciones fue dividida por la cantidad de *tokens* que tiene cada capítulo para así tener registros estandarizados.

También se ha construido una medida de riqueza verbal (RT), cuyo cálculo es $RT = (\#Palabras\ diferentes) / (\#Total\ de\ palabras)$ para cada uno de los $(n-1)$ capítulos². Es-

1.- Sin embargo, como ya se ha señalado, algunos estudiosos creen que ciertos pasajes fueron añadidos a determinados capítulos para darle unidad al texto. No obstante, creemos que estas añadiduras no han influido notablemente en las medidas.

2.- El listado de las variables investigadas es: -, !, ,, , , : , ; , ? , j , ¿ , a, al, como, con, de, de_la, de_lo_que, del, dijo, don, el, en, es, la, las, le, lo, lo_que, los, más, me, mi, no, por, que, que_no, quijote, RT, sancho, se, si, su, un, y, yo.

tas serán las variables de estudio. Luego de este conteo, cada una de esas variables es medida en los 126 capítulos, con lo que se cuenta con una matriz de datos de dimensiones 126×44 .

Los cálculos de este trabajo fueron realizados con la ayuda del software estadístico R (R Development Core Team, 2011) y el paquete *strucchange* (Zeileis *et al.*, 2002).

2.2. Cambio en las proporciones

La primera aproximación a las diferencias internas de *Don Quijote* necesariamente está relacionada con la disimilitud de proporciones.

Esto respondería a la pregunta: ¿ha cambiado el libro de *Don Quijote* en términos de las variables estilométricas? Para ejecutarla, se debe fijar un punto en el cual haya sospecha de cambio (un *antes* y un *después*).

Un primer punto que parece natural está alrededor del fin de la primera parte y el inicio de la segunda; más concretamente: ¿son distintas, en promedio, las variables estilométricas en la primera y la segunda parte de *Don Quijote*?

Esta pregunta puede ser respondida haciendo uso de pruebas *t* tradicionales. Aproximadamente la mitad de las 44 variables (21), incluyendo la riqueza verbal, presentan un cambio en relación a la primera y segunda parte de *Don Quijote* ($p \leq 0.1$). A continuación en la Tabla 1 se muestran las 10 variables con más diferencias entre la primera y la segunda parte de *Don Quijote*.

Variable	Parte I	Parte II	Valor p
no	0,01291395	0,01742504	0,00
con	0,01420961	0,01228126	0,00
los	0,01135672	0,01392935	0,00
las	0,00828184	0,00989052	0,00
su	0,01044162	0,00806862	0,00
le	0,01055663	0,00744563	0,00
,	0,10449256	0,10017246	0,00
.	0,02010702	0,03008195	0,00
-	0,01569964	0,00997378	0,00
!	0,00240486	0,00398444	0,00

Tabla 1: Pruebas *t* sobre las variables utilizadas. La negación 'no', por ejemplo, representa en promedio el 1.29% del total de palabras de cada capítulo de la primera parte y el 1.74% en la segunda.

Para evaluar la diferencia conjunta entre el vector que compone los promedios de las variables en estudio para la primera parte y las mismas variables en la segunda parte se utilizarán pruebas de análisis de varianza multivariante (MANOVA, por sus siglas en inglés) en su versión clásica (Johnson y Wichern, 2002). Más formalmente, se evaluarán las hipótesis:

$$H_0: \mu_{\text{1ra. Parte}} = \mu_{\text{2da. Parte}}$$

$$H_1: \mu_{\text{1ra. Parte}} \neq \mu_{\text{2da. Parte}}$$

donde μ es el promedio multivariado de todas las variables bajo examen. Los resultados se muestran a continuación en la Tabla 2:

Test	gl	Estadístico	Aprox. F	gl num.	gl den.	Valor p
Pillai	1	0,68572	40,167	44	81	0,00
Wilks	1	0,31428	40,167	44	81	0,00
Hotelling-Lawley	1	2,1819	40,167	44	81	0,00
Roy	1	2,1819	40,167	44	81	0,00

Tabla 2: Pruebas MANOVA tradicionales

Notar que los resultados corroboran que existen diferencias multivariantes estadísticamente significativas, en base a estas variables, entre la primera y la segunda parte de *Don Quijote*.

2.3. Cambio estructural

Los contrastes de cambio estructural (Green, 2003) son utilizados para saber si los parámetros asociados a un modelo de regresión o de series de tiempo han cambiado dentro de un período de estudio. Cuando se conoce el momento exacto en que este cambio debió ocurrir, la prueba, en el contexto de análisis de regresión, es denominada *contraste de Chow* (Chow, 1960) y se realiza con un contraste corriente F .

Este tipo de pruebas suele hacerse cuando se implementa una nueva política y se desea saber si efectivamente tuvo algún resultado que se reflejase en los coeficientes de un modelo de regresión. Cuando los datos se están modelando por medio de series de tiempo, el problema puede abordarse por medio del *análisis de intervención* (Enders, 2004).

No obstante, muchas veces, no se puede tener certeza de cuándo ha ocurrido este cambio y una manera de abordar el problema fue sugerido inicialmente por Brown *et al.* (1975) basado en *residuos recursivos*. Para esta prueba, la hipótesis nula es que el vector de parámetros, $(\beta_0, \beta_1, \dots, \beta_p)$, permanece constante para cada período. La hipótesis alternativa es que en algún período al menos un parámetro cambia su valor.

Como la estilometría se ha fundado en que las características medidas son propias e invariables para cada escritor en un plazo razonable de tiempo, se espera que ellas no cambien significativamente de unos capítulos a otros (Bagavandas y Manimannan, 2008). Cuando la estabilidad se rompe, se puede especular que es debido a la colaboración de un autor adicional. Véase por ejemplo el trabajo de Riba y Ginebra (2006) en el cual se expone la posibilidad que Joanot Martorell y Martí Joan de Galba no hayan sido los únicos autores de *Tirant lo Blanc*. Los autores llegan a esta conclusión analizando la 'diversidad de vocabulario' que se muestra a lo largo del libro.

Entonces, considerando el modelo sencillo

$$y_t = \beta_0 + \varepsilon_t \quad (1)$$

donde y_t es una de las 44 variables anteriormente señaladas; ε_t es un término de error con media cero y varianza constante y $t=1, \dots, 126$; se espera que β_0 sea el mismo para todos los

t . Note el lector que cada uno de estos t se corresponde con un capítulo de las dos partes de *Don Quijote*. En caso que se hallara que β_0 no es constante, se estaría demostrando, al menos, que la manera de escribir de Cervantes sufrió un cambio estadísticamente cuantificado durante el tiempo que escribió el libro.

Suponiendo que se encuentre algún cambio estructural, la ecuación (1) adopta la forma

$$y = \beta_0^{(j)} + \varepsilon_t \text{ with } t = n_{j-1}+1, \dots, n_j \text{ and } j=1, \dots, m+1 \quad (2)$$

donde $j=1, \dots, m$ son los períodos donde existen m interceptos distintos. El conjunto índice n_1, \dots, n_m son los *puntos de cambio desconocidos*. Cada uno de ellos da el lugar preciso dónde se da un cambio en la forma de escribir. El objetivo es averiguar cuándo o en qué momento ocurrieron estos cambios. Este problema es también conocido como *fechar* el cambio estructural y su solución ha sido abordada utilizando un algoritmo de programación dinámica que permite el cálculo de los puntos de cambio estimados como el minimizador global de la suma de cuadrados de los residuos (para mayores detalles, ver Bai y Perron, 2003).

La primera aproximación para hallar los puntos de quiebre, los lugares donde cambia el estilo de Cervantes, se hará siguiendo la metodología propuesta por Brown *et al.* (1975) y se pueden ver en la Figura 1. En ella se dispone el trabajo de 4 de las variables en estudio en términos de los residuos recursivos de cada serie original. Sobre ellas, se observan dos bandas límites: una por encima y otra por debajo. Cuando la serie evaluada *se escapa* de estos límites, se puede sospechar que allí ha habido un cambio de estilo.

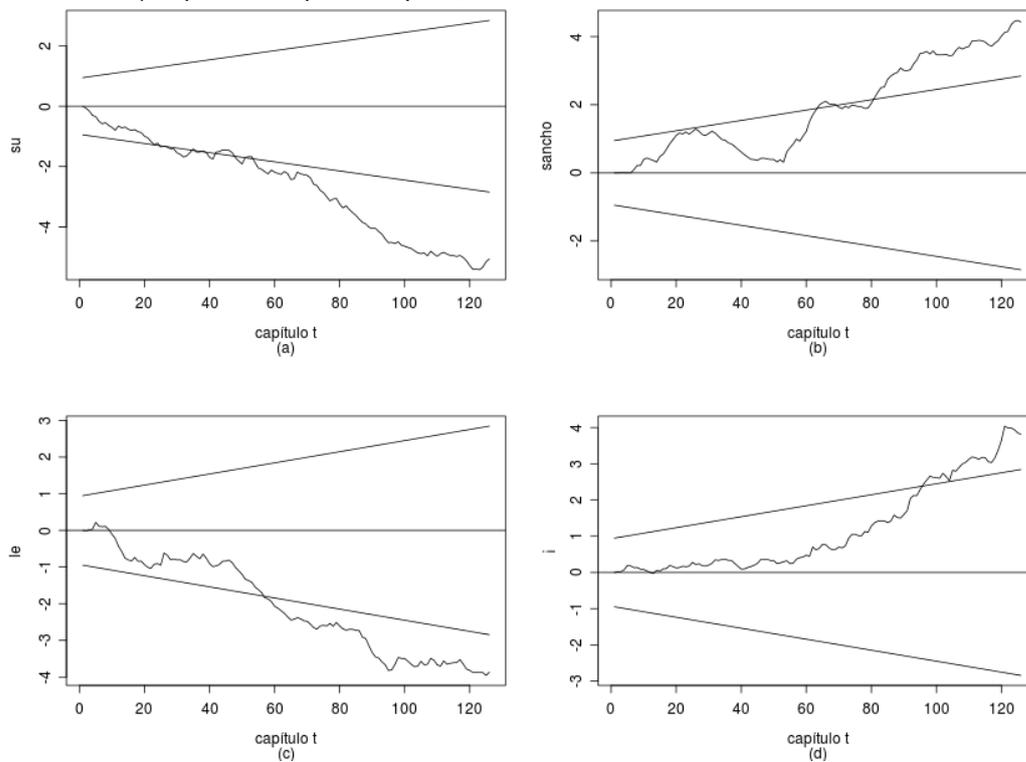


Figura 1: Pruebas en base a los residuos recursivos CUMSUM

De esta figura se desprende:

- Figura 1a: se puede observar que a partir del capítulo 20 los residuos recursivos se escapan de la banda inferior, si bien no se hace expresamente hasta, aproximadamente, el capítulo 60.
- Algo similar ocurre en la Figura 1b, con los residuos dirigidos hacia arriba. Es importante observar que para esta palabra, «sancho», es razonable esperar que tenga un cambio cerca del capítulo 7 dado que el personaje de Sancho Panza hace su aparición en este capítulo.
- En la Figura 1c los residuos solo se escapan a partir, aproximadamente, del punto $t=60$.
- La Figura 1d presenta un patrón distinto, saliéndose de las bandas límites luego del capítulo 100, aproximadamente.

Para estar más seguros de dónde ocurren estos puntos de corte, se procederá a seguir la estrategia propuesta por Bai y Perron (2003) mostrada en la Figura 2 y 3. En la primera de ellas se presentan algunas medidas que reportaron solo un punto de cambio y en la segunda se presentan algunas medidas que presentaron dos puntos de cambio. En estas Figuras se puede ver con precisión en qué capítulos han ocurrido los cambios.

Se ha procurado utilizar medidas diferentes en cada uno de los gráficos (su, sancho, le, ;, las, que, y, la, lo que) para que el lector aprecie que los cambios se dan aproximadamente hacia los mismos lugares. El lugar preciso de cada punto de cambio, con un intervalo de confianza del 95%, puede hallarse en el Apéndice, en las Tablas 3 y 4.

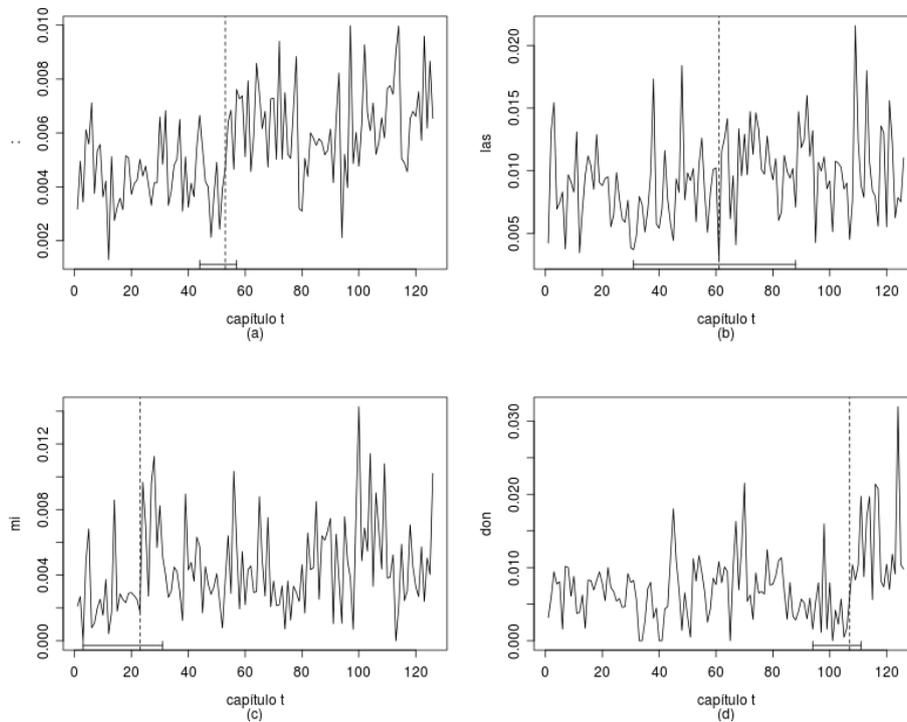


Figura 2: Algunos ejemplos de cambio estructural con un punto de cambio

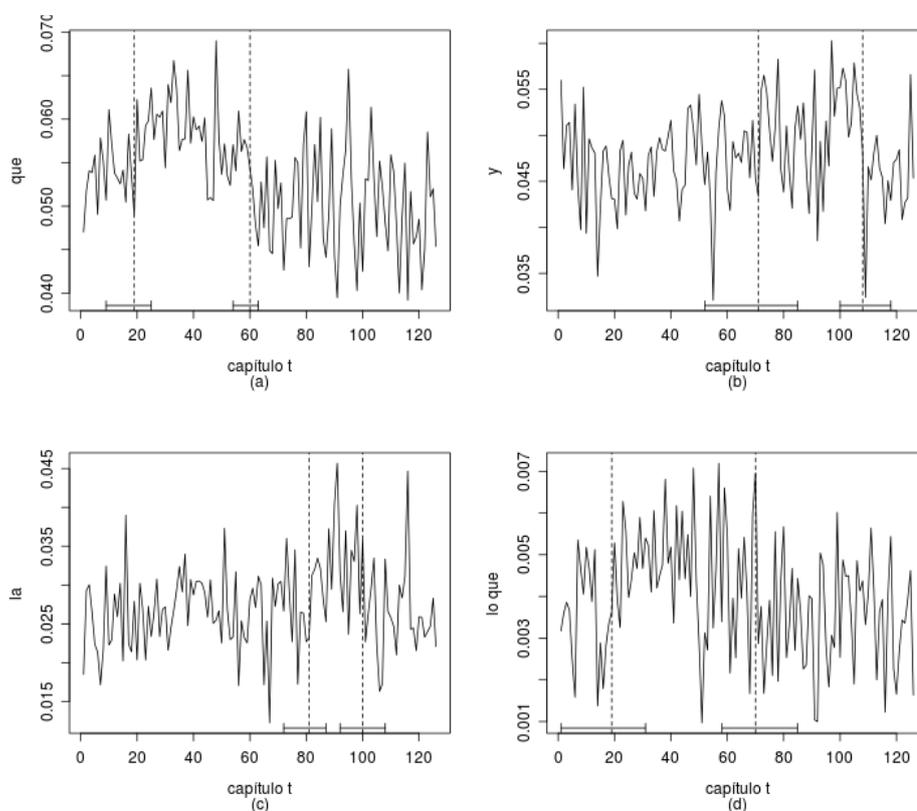


Figura 3: Algunos ejemplos de cambio estructural con dos puntos de cambio

3. Discusión

Según las figuras del presente estudio, se observa que hubo varios cambios en algunas medidas. El primero de ellos está en torno al $t=20$. Esto concuerda con la teoría que Cervantes escribió, de corrido, los primeros 19 capítulos sin divisiones y sin epígrafes abandonando el proceso creativo para dedicarse a otros proyectos (Anderson y Pontón Gijón, 1998). Este punto de cambio estructural también hace patente que Cervantes mantuvo el mismo estilo en estos 19 capítulos, incluyendo la posible influencia del ‘Entremés de los romances.’

Otro cambio que se presenta en repetidas oportunidades sucede alrededor de $t=60$, es decir, en torno al capítulo 8 de la *Segunda Parte*. Eisenberg (1976) cree que la *Segunda Parte* no fue compuesta de corrido sino que hubo largos lapsos en el proceso creador, uno de ellos estimulado por la novela *Segundo tomo del ingenioso hidalgo Don Quijote de la Mancha, que contiene su tercera salida, y es la quinta parte de sus aventuras*, del desconocido y célebre autor Alonso Fernández de Avellaneda (una edición del llamado *Quijote apócrifo* en García Salinero, 1972).

Según Eisenberg, el estilo de la *Primera Parte* continúa los primeros 4 o 5 capítulos de la *Segunda Parte* y el que haya un cambio en torno a $t=60$ justifica parcialmente esta apre-

ciación. Coufal y Joula (2010) sugieren que no fue Cervantes quien terminó *Don Quijote* por un cambio sistemático que ocurre, especialmente, de los capítulos 6 a 64 de la *Segunda Parte* respecto a la primera.

Note el lector que es naturalmente posible que un autor desplace su estilo de escribir a lo largo del tiempo. Por ejemplo, Can y Patton (2004) lograron obtener más del 90% de clasificación exitosa en un estudio de atribución para los escritos ‘viejos’ y ‘nuevos’ de dos autores turcos. No obstante, el tiempo que separaba los trabajos ‘viejos’ y ‘nuevos’ oscilaba entre 40 años para un autor y 27 para otro: lapsos considerablemente mayores a la composición, al menos, de la segunda parte de *Don Quijote*.

Se dan dos cambios más: uno cercano a $t=80$ y otro en torno a $t=100$ (capítulos 28 y 48 de la *Segunda Parte*, respectivamente). El primer cambio ha sido notado por Eisenberg (1976), y, en sus propias palabras, la historia del capítulo 29 «no encaja bien con los capítulos que la rodean». No obstante, no tenemos documentación en relación al cambio ocurrido luego del capítulo 48.

4. Conclusiones

He realizado un estudio tomando prestadas herramientas utilizadas comúnmente en la econometría para determinar los puntos donde *El Quijote* cambia de estilo en función de algunas medidas como palabras, frases y signos más frecuentes y se encontraron 4 puntos bien diferenciados, al menos 3 de ellos documentados con anterioridad en trabajos literarios.

Entre las hipótesis que podemos manejar para explicar estos cambios:

- El tiempo que invirtió Cervantes en componer la obra y en los cambios propios que supone ir madurando e ir encontrando su propio estilo.
- La influencia que pudo tener Cervantes al escribir algunos capítulos (como los que se comparan con el ‘Entremés de los Romances’).
- La existencia de uno o varios autores que hayan contribuido a componer la obra.
- Por medio del análisis mostrado no es posible establecer con seguridad ninguna de esas hipótesis pero sí se aseguran los cambios internos de estilo.

Tampoco se puede excluir la posibilidad de la interacción de las tres hipótesis: un Cervantes que con el paso del tiempo fue cambiando su manera de escribir, que en un tímido comienzo, como la mayoría de los artistas, imitó un texto que le llamó la atención y que recibió, en algunos momentos, la ayuda de otros escritores.

5. Apéndice

	Límite inferior	Capítulo	Límite superior
Fig. 2a	44	53	57
Fig. 2b	31	61	88
Fig. 2c	3	23	31
Fig. 2d	94	107	111

Tabla 3: Intervalos de confianza sobre el punto de cambio de la Figura 2.

	Primer punto de cambio			Segundo punto de cambio		
	Límite inferior	Capítulo	Límite superior	Límite inferior	Capítulo	Límite superior
Fig. 3a	9	19	25	54	60	63
Fig. 3b	52	71	84	98	107	118
Fig. 3c	72	81	87	91	100	108
Fig. 3d	1	19	31	58	70	85

Tabla 4: Intervalos de confianza sobre el punto de cambio de la Figura 3.

Bibliografía

- Anderson, Ellen M., y Gonzalo Pontón Gijón. «La composición del 'Quijote'». *Don Quijote de la Mancha*. Editor Francisco Rico. Crítica, Barcelona, 1998.
- Bagavandas, M., y G. Manimannan. «Style Consistency and Authorship Attribution: A Statistical Investigation.» *Journal of Quantitative Linguistics*, 15(1): 100–110, 2008.
- Bai, Jushan, y Pierre Perron. «Computation Analysis of Multiple Structural Change Models.» *Journal of Applied Econometrics*, 18:1–22, 2003.
- Brown, R.L., Durbin, J., y J.M. Evans. «Techniques for testing constancy of regression relationships over time.» *Journal of the Royal Statistical Society*, B, 37: 149–63, 1975.
- Can, Fazli. y Jon M. Patton. «Change of writing style with time.» *Computers and the Humanities*, 38:61–82, 2004.
- Carr, Francis. *Who Wrote Don Quixote?* Xlibris Press, London, 2004.
- Cervantes, Miguel de. *Don Quixote*. Edición a cargo de James H. Montgomery. Hackett Publishing Company, Inc., 2009.
- Cervantes, Miguel de. *Don Quijote de la Mancha*. Edición a cargo de Francisco Rico. Crítica, Barcelona, 1998.
- Cervantes, Miguel de. *Don Quijote de la Mancha*. Edición a cargo de Francisco Rico. Santillana Ediciones Generales, 2008.
- Chow, Gregory. «Test of Equality Between Sets of Coefficients in Two Linear Regressions.» *Econometrica*, 28:79–99, 1960.
- Coufal, Christopher y Patrick Joula. Authorship Discontinuities of El Ingenioso Hidalgo don Quijote de la Mancha as detected by Mixture-of-Experts. Agosto 2010 <<http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-755.html>>
- Criado de Val, Manuel. «Las figuras y la iconografía del Quijote (1987).» *Revista de Estudios Cervantinos*, 3 (octubre-noviembre):53–75, 2007.
- de Riquer, Martin. *Aproximación al Quijote*. Biblioteca Basica Salvat Editores, 1969.
- Eisenberg, Daniel. «El rucio de Sancho y la fecha de composición de la Segunda Parte de 'Don Quijote'.» *Nueva Revista de Filología Hispánica*, 25:94–102, 1976.
- Eisenberg, Daniel. «Repaso crítico de las atribuciones cervantinas.» *Nueva Revista de Filología Hispánica*, 38:477–92, 1990.
- Eisenberg, Daniel. «No hay una primera parte del Quijote.» *El Quijote desde América*. Editores Gustavo Illades y James Iffland. BUAP, 57–79, 2006.
- Eisenberg, Daniel y Geoffrey Stagg. «Entremés de los romances.» *Bulletin of Cervantes Society of America*, 22 (2):151–74, 2002.
- Enders, Walter. *Applied Econometric Time Series*. Wiley, NJ, 2 edition, 2004.
- García Salinero, Fernando. «Introducción crítica sobre la obra y su autor.» *Segundo tomo del ingenioso hidalgo Don Quijote de la Mancha, que contiene su tercera salida, y es la quinta parte de sus aventuras*. Edición a cargo de Fernando García Salinero. Castalia, 1972.
- Green, William. *Econometric Analysis*. Prentice Hall, NJ, 5 edition, 2003.
- Johnson, R.A., y D.W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, International, 5 edition, 2002.
- Madrigal, José Luis. «De cómo y por qué La tía fingida es de Cervantes.» *Artifara*, 2, 2003.
- Madrigal, José Luis. «Algunas reflexiones en torno a la atribución cervantina del 'Diálogo Entre Cilenia y Selanio sobre la vida del campo'.» *Bulletin of Cervantes Society of America*, 24 (1):217–52, 2004.
- Madrigal, José Luis. «Miguel de Cervantes y la 'Relación de las fiestas' de 1605.» *Bulletin of Cervantes Society of America*, 25 (2):271–302, 2006.

- Montgomery, James H. «Was Thomas Shelton the Translator of the 'Second Part' (1620) of Don Quixote?» *Bulletin of Cervantes Society of America*, 26 (1):209–17, 2008.
- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Rey Hazas, Antonio. «Estudio del Entremés de los romances.» *Revista de Estudios Cervantinos*, 1(junio-julio), 2007.
- Riba, Alex, y Josep Ginebra. «Diversity of Vocabulary and Homogeneity of Literary Style.» *Journal of Applied Statistics*, 33(7): 729–41, 2006.
- Stagg, Geoffrey. «Sobre el plan primitivo del 'Quijote.'» *Actas del primer congreso de la Asociación Internacional de Hispanistas*. Editores F. Pierce and C. A. Jones. Oxford: The Dolphin Book, 463–71, 1964.
- Stagg, Geoffrey. «Don Quijote and the 'Entremés de los romances.' A Retrospective.» *Bulletin of Cervantes Society of America*, 22 (2):129–50, 2002.
- Zeileis, A., F. Leisch, K. Hornik, y C. Kleiber. strucchange: «An R package for testing for structural change in linear regression models». *Journal of Statistical Software*, 7(2):1-38, 2002. URL <http://www.jstatsoft.org/v07/i02/>.