

Philosophische Fakultät

Institut für Romanistik

Betreuerin: PD Dr. Reißler-Pipka, Nanette

**«Atribución de autoría y Humanidades Digitales: Métodos de
estilometría y aplicación al Siglo de Oro»**

Abschlussarbeit im Fach TransRomania-Studien: Romanische Sprachen, Literaturen
und Kulturen

Master-Studiengang zur Erlangung des Akademischen

Grades "Master of Arts" (M.A.) und "M.Education Spanisch/Französisch"

der Georg-August-Universität Göttingen.

Vorgelegt am: 20.01.22

von Alejandro Cienfuegos Pérez

aus Xixón, Asturias, Spanien.

Matrikelnummer: 11732051

E-mail: a.cienfuegosperez@stud.uni-goettingen.de

Alejandro Cienfuegos Pérez

Erstprüferin: PD Dr. Reißler-Pipka, Nanette

Zweitprüfer: Prof. Dr. Mensching, Guido

Índice

1	Introducción	1
2	Aspectos teóricos	4
2.1	La estilística y el estilo	4
2.2	Estilo para la estilometría.....	8
2.3	La importancia del autor: estudios de la teoría literaria en torno a la figura autorial.....	10
3	El análisis cuantitativo de textos	11
3.1	Orígenes y desarrollo de los métodos cuantitativos para la atribución de autoría	11
3.2	Atribución de Autoría.....	14
3.3	Bases para la atribución de autoría.....	16
4	La estilometría	17
4.1	Estilometría desde las H.D. en español.....	17
4.2	Estilometría en la obra de Cervantes	19
5	Corpus	22
5.1	Problemática de los textos del Siglo de Oro	22
5.2	Fuentes para la recopilación de textos del corpus.....	26
5.3	Cosecha de los textos y configuración del corpus	28
5.4	Relevancia de «La Tía Fingida»	32
6	Stylo de R: presentación de la herramienta	35
7	Pruebas estilométricas	37
7.1	Análisis de clúster para ajustar el corpus	38
7.2	Precisión a través de la validación cruzada y clasificación.....	48
7.3	Verificación de autoría	51
7.4	Stylo.network () y gephi	57
8	Análisis literario: la novela en el Renacimiento y Barroco	60
8.1	Cervantes: sus novelas	61
8.2	La novela corta en Cervantes: generalidades técnicas.....	63
9	Análisis literario: <i>La tía fingida</i> vs. <i>El celoso extremeño</i>	63
9.1	Análisis de los contenidos	65
9.1.1	<i>La tía fingida</i>	65
9.1.2	<i>El celoso extremeño</i>	68
9.2	Análisis de las técnicas literarias	71
9.2.1	<i>La tía fingida</i>	71

9.2.2	<i>El celoso extremeño</i>	79
9.3	Análisis de la expresión.....	83
9.3.1	<i>La tía fingida</i>	84
9.3.2	<i>El celoso extremeño</i>	94
10	Síntesis del análisis literario	104
11	Conclusiones	106
12	Bibliografía	113
13	Anexos	120
13.1	Anexo: Código R.....	120
13.2	Anexo Corpus.....	132

Índice de tablas y figuras

Tabla 1: obras atribuidas a Cervantes. BVMC.....	20
Tabla 2: autor, título, fecha, edición y versión digital de <i>La tía fingida</i>	31
Tabla 3: resultados de la validación cruzada, 500MFW, <i>Cosine Delta</i>	49
Tabla 4: resultados <i>classify</i> , corpus + La TF, 500MFW, <i>Cosine Delta</i>	50
Tabla 5: resultados de la función imposters con <i>Delta de Burrows</i>	52
Tabla 6: resultados <i>imposters</i> en las pruebas individuales para cada autor.....	53
Tabla 7: resultados imposters con <i>Cosine Delta</i> , <i>Eder's Delta</i> y <i>Burrow's Delta</i>	54
Tabla 8: resultados de imposters.optimice, rangos de fiabilidad de los datos.....	55
Tabla 9: <i>Voyant Tools</i> , vocabulario de la TF y <i>El celoso</i> . Textos fragmentados.....	77
Tabla 10: <i>Voyant Tools</i> , vocabulario de la TF y <i>El celoso</i> . Textos completos.....	77

Tablas del anexo:

Tabla 1 corpus: prosa narrativa breve.....	132-133
Tabla 2 corpus: prosa narrativa extensa.....	134-137
Tabla 3 corpus: otros textos en prosa.....	138-141

Imágenes

Imagen 1: prosa narrativa breve, análisis de clúster.....	39
Imagen 2: prosa narrativa extensa, a. clúster, 500 MFW, <i>Cosine Delta</i>	40
Imagen 3: detalle del dendograma (Imagen 4), partes del texto problemáticas.....	41
Imagen 4: prosa narrativa breve y extensa, a. clúster, 500 MFW, <i>Cosine Delta</i>	42
Imagen 5: sección 1 del dendograma (Imagen 4).....	43
Imagen 6: sección 2 del dendograma (Imagen 4).....	43
Imagen 7: sección 3 del dendograma (Imagen 4).....	43
Imagen 8: sección 4 del dendograma (Imagen 4).....	43
Imagen 9: prosa narrative breve y extensa, <i>BCT</i> , 500 MFW, <i>Cosine Delta</i>	44
Imagen 10: otros textos en prosa, análisis de clúster, 500MFW, <i>Cosine Delta</i>	45
Imagen 11: corpus completo, análisis de clúster, 500MFW, <i>Cosine Delta</i>	47
Imagen 12: detalle 1 del dendograma (img. 11).....	47
Imagen 13: detalle 2 del dendograma (img. 11).....	47
Imagen 14: detalle 3 del dendograma (img. 11).....	47
Imagen 15: detalle 4 del dendograma (img. 11).....	48

Imagen 16: detalle clúster corpus «completo» + TF, 500MFW, <i>Cosine Delta</i>	57
Imagen 17: <i>Gephi</i> del BCT de 4 autores <i>Cosine Wurzburg</i> (500MFW).....	58
Imagen 18: detalle de la imagen 17, <i>La Tía Fingida</i> y sus nexos.....	59
Imagen 19: <i>Word Cloud</i> (Cirrus) de las MFW, <i>La tía fingida</i>	77
Imagen 20: <i>Word Cloud</i> (Cirrus) de las MFW, <i>El celoso extremeño</i>	78

Agradecimientos

A mi madre: por su apoyo incondicional y su ayuda que permitió iniciar esta aventura alemana. Sin tu respaldo no hubiera llegado hasta aquí.

A Johanna: por ser la mejor compañera, darme tranquilidad, inspirarme y soportar los vaivenes emocionales que acompañaron la redacción de esta tesis.

Al sistema educativo alemán y a la universidad de Göttingen: por ser universal, gratuita, inclusiva y democrática permitiéndome crecer académicamente.

A Nanette Reißler-Pipka: por introducirme al interesante ámbito de las H.D., acompañarme durante la redacción de esta tesis y ocuparse de que no tuviera preocupaciones laborales durante el proceso.

A José Calvo Tello: por tu trabajo y paciencia para ayudarme a entender las herramientas de la estilometría empleadas en este trabajo.

A mis compañeros y amigos de Detmold y Göttingen: que hicieron posible que llegara hasta aquí, acompañándome en mis estudios y en mi vida en Alemania, haciendo de ambos una experiencia inolvidable.

A todos vosotros, os doy mis más sinceras gracias.

1 Introducción

Este trabajo tiene como labor un doble objetivo, en primer lugar, emplear y comprobar la utilidad de la estilometría aplicada al Siglo de Oro en un caso concreto de autoría dubitada y, en segundo lugar, llevar a cabo un análisis estilístico literario para unir ambas vertientes de la estilística: la literaria y la lingüística. Partimos de la hipótesis de que los nuevos métodos de la estilometría, y en particular el empleo de la herramienta *stylo* de R (Eder, 2018; Eder y Rybicki, 2012; Rybicki y Eder, 2012) puede si no dar una solución definitiva al menos aportar nuevas perspectivas a una discusión literaria largamente debatida que es la de la autoría de *La tía fingida*¹ atribuida por muchos a Cervantes. Así mismo, también consideramos relevante no obviar la perspectiva literaria cuando la determinación de autoría se ocupa del estilo literario de un autor^{2 3}.

Para empezar, se presentarán de forma breve los conceptos de estilística y estilo con el fin de acotar las definiciones que como veremos son abundantes pues dependen del ámbito que las aborde. Particularmente relevante es dejar claro de antemano cuál es el concepto de estilo que utiliza la estilometría. También será necesario comentar en estos prolegómenos el debate en torno a la figura autorial en la literatura que se intensificó con la publicación de Barthes *La muerte del autor* (1968). Aquí se expresarán de forma escueta nuestras consideraciones al respecto para justificar la importancia de conocer la autoría de un texto.

Con estas informaciones, podremos pasar al plano del análisis cuantitativo de textos, ¿dónde veremos de forma concentrada la evolución de los métodos en el marco de la atribución de autoría hasta llegar a los procedimientos que se emplean hoy en día para lo cual nos serviremos del trabajo de Holmes y

¹ Nos referiremos a ella como *La tía fingida* y la TF indistintamente.

² Aunque la frase pueda sonar redundante no son pocos los trabajos que desde diferentes perspectivas afrontan la atribución de autoría utilizando métodos de diversa índole que obvian el contexto literario. Por ejemplo, López, F. (2011)

³ A pesar de que consideramos la importancia del uso del lenguaje inclusivo y la distinción autor/autora y similares, evitaremos su empleo a fin de garantizar una mejor legibilidad del texto. En cualquier caso, entiéndase que utilizamos un masculino genérico para referirnos a ambos sexos.

Kardos *Who was the autor?* (2003). En este apartado, enunciaremos también las bases teóricas y requisitos de la atribución de autoría⁴ de la mano de Juola (2008) Holmes (1994) y Bailey (1979).

Continuaremos después con un capítulo dedicado a la estilometría en el que nos centraremos en exponer los trabajos que se han venido llevando a cabo en las Humanidades Digitales en español y continuaremos con los estudios que en este entorno concreto han tratado la obra de Cervantes. Además, veremos aquí el estado de la cuestión de las investigaciones en torno a algunos de los textos atribuidos a Cervantes para poder comenzar a entender otro de los motivos que sustentan la elección de *La tía fingida* como objeto de análisis. De esta manera pretendemos también justificar el método utilizado que no surge de manera espontánea y aislada si no que parte del conocimiento y lectura de estos trabajos en los que además se basa.

Proseguiremos con una de las partes que consideramos más complejas en el entorno del Siglo de Oro que es llevar a cabo la compilación de un corpus suficientemente representativo y fiable de textos digitales. El camino desde las fuentes, pasando por la problemática y limitaciones hasta la cosecha de textos y configuración del corpus será expuesto aquí explayadamente. También afecta y es por tanto ineludible la relación que todo esto tiene con el texto dubitado. De él se comentará su relevancia subrayando los avatares desde la aparición del manuscrito original hasta llegar a las múltiples versiones que existen hoy en día del mismo y se comentarán lacónicamente algunas de las posiciones que nutren el agitado debate en torno a su autoría.

Llegados a este punto se presentará de forma sucinta la herramienta *stylo*, se justificará su elección en base a las advertencias de Juola (2008) y se procederán a llevar a cabo las pruebas estilométricas. Para ellas, se utilizarán diferentes funciones a parte de la ya citada *stylo: classify, imposters, stylo.network*. La primera, *stylo*, servirá para organizar el corpus en un dendograma que mostrará una serie de *clusters* según el autor y también para comprobar que la organización de los textos no dubitados es correcta. Esto nos ayudará a configurar el corpus antes de llevar a cabo cualquier prueba que

⁴ Nos referiremos a ella también como AA

implique a la TF. A partir de esos resultados y una vez verificada la fiabilidad de la herramienta por medio de la verificación cruzada llevaremos a cabo dos experimentos: uno de atribución y otro de verificación de autoría. El primero de ellos (función *classify*) parte de la base de que el autor del texto dubitado es anónimo y buscará posibilidades entre los candidatos, el segundo (método *imposters*) sirve para comprobar estos resultados pues se fundamenta en que sabemos quién es el autor (Cervantes, p.ej.) pero existiendo dudas es preciso verificar esta afirmación cuyo resultado es dado por un coeficiente asignado al candidato y los denominados «impostores». Esta metodología parte de la base objetiva de que no conocemos al autor pues la atribución a Cervantes no ha sido a nuestro parecer fehacientemente justificada ni refutada. Así pues, pretendemos aplicar un método de atribución para después verificar dicha atribución con uno de verificación.

En base a los datos obtenidos y utilizando la función *stylo.network* y el programa de visualización de datos *Gephi*, crearemos un gráfico que muestre las obras más cercanas, es decir, que exponga los nexos entre las obras del corpus. De esta manera podremos pasar a la siguiente sección en la que analizaremos la TF y aquella/s novela/s que presenten una mayor similitud estilométrica, pero desde una perspectiva literaria.

En esta última parte del trabajo dedicada a la lectura cercana y el análisis del estilo desde la literatura nos basaremos en la metodología que propone Castagnino (1987) en su libro «*El análisis literario: introducción metodológica a una estilística integral*». En consonancia con su propuesta, comenzaremos por un análisis de los contenidos para seguir con las técnicas literarias y terminar con el análisis de la expresión. En este último, seleccionaremos algunas figuras retórico-literarias que en base a Castagnino se dividen en relaciones de la expresión intencionalidad y figuras de economía expresiva. También se expondrán algunos fenómenos morfológicos que resulten llamativos en la lectura cercana de las novelas sin intención de ser exhaustivos en este contexto. Una vez acometida la tarea se efectuará una síntesis comparativa del análisis literario de ambas obras que servirá de preámbulo a las conclusiones finales de esta tesis.

2 Aspectos teóricos

2.1 La estilística y el estilo

A la hora de abordar un trabajo de atribución de autoría nos vemos confrontados con la cuestión de tratar de descifrar el estilo de un autor. En nuestro caso este autor es Miguel de Cervantes. Sin embargo, el concepto de estilo es muy heterogéneo y su definición varía dependiendo del área que lo trate. En nuestro caso vamos a realizarlo desde la estilometría computacional que como veremos tiene una visión particular y restringida de lo que es el estilo y también desde la literatura. No obstante, haremos un repaso y aclaración del concepto para entender mejor la perspectiva que nos ocupa.

De forma muy simple se puede definir la estilística como el estudio del estilo. Sin embargo, esta definición etimológica es evidentemente imprecisa como apunta Fernández Retamar (1963: 4) quien comenta la existencia de una «diversidad de estilísticas». Para acercarnos a una aclaración del concepto Fernández Retamar (1963) ve necesario comenzar hablando de las funciones del lenguaje que distinguió Bühler (1879) que son la representativa (la palabra alude a una cosa), la apelativa (la palabra solicita del oyente una actitud determinada) y la expresiva (que nos permite percibir en base a ciertas características -pronunciación, tono, etc.- ciertas emociones). Son estas dos últimas funciones, la apelativa y expresiva las que manifiestan el aspecto extra lógico del lenguaje y las que estudia la estilística (cfr. Fernández Retamar, 1963: 4).

La estilística al considerar la obra literaria toma contacto con una serie de disciplinas como son la filología, la retórica, la crítica literaria, ciencia de la literatura, teoría de la literatura e historia de la literatura. Así es como ambos conceptos (estilo y estilística) son definidos de diversas formas en base a diferentes enfoques que se deben a la influencia que han tenido las ramas de la crítica literaria y de la lingüística. Mukherjee (2005) aporta una descripción general de lo que se entiende por estilística y estilo que nos puede resultar aclaratoria para empezar:

«La estilística es la descripción y el análisis de la variabilidad de las formas lingüísticas en el uso real de la lengua. Los conceptos de "estilo" y "variación estilística" en la lengua se basan en el supuesto general de que, dentro del sistema lingüístico, el mismo contenido puede codificarse en más de una forma lingüística. Al operar en todos los niveles lingüísticos (por ejemplo, lexicología, sintaxis, lingüística textual y entonación), los estilistas analizan tanto el estilo de textos específicos como la variación estilística entre textos. Estos textos pueden ser de naturaleza literaria o no literaria. En general, el estilo puede considerarse como una elección de medios lingüísticos; como una desviación de una norma; como una recurrencia de formas lingüísticas; y como comparación.»⁵

Vemos pues en esta definición nuevamente que en **la estilística** no existe una concepción unívoca de «estilo». Alicia Yllera (1979: 150) hace un recorrido por las diferentes corrientes que han existido y en términos similares subraya que dentro de este ámbito han habido siempre vacilaciones en torno al concepto de estilo señalando tres oposiciones:

- 1) Una **individualista** (focalizando al autor o a la obra) y una **social o genérica** (estilo como «formas de época» o características de un tipo de habla)
- 2) Estilo como **plano de la expresión** o estilo como **plano de la enunciación** (expresión y contenido)
- 3) Estilo como **desvío**, resultado de una elección o como **coherencia** de la estructura totalizante de una obra.

No obstante, esta misma autora comenta que en los primeros estudios sobre estilística moderna se reconocen dos concepciones: el estilo como expresividad/rasgos expresivos y el estilo como desvío de una norma. Esta dicotomía sigue la hipótesis saussureana que considera que un texto es parte del dominio del habla por lo que la creación individual, el estilo, se define por referencia a una norma como un *desvío* (cfr. Dubois et al., 1979: 246). Sin embargo, Yllera (1979) señala que ambos criterios (rasgos expresivos y desvío) aparecen unidos en un cierto número de teóricos.

La dicotomía se traduce en el trabajo que nos ocupa en dos términos pues abordamos el estilo desde la lingüística y la literatura. Así, en este contexto el estilo se define como **idiolecto** para la lingüística forense y como **estilo de autor**

⁵ Traducción propia del inglés (original).

(autor/es, género o movimiento literario) desde la perspectiva literaria. Para la justificación de la importancia de ambas vertientes es necesario recordar las palabras de R. Jakobson (1969) cuando dice que: «Un lingüista sordo a la función poética, así como un especialista de la literatura indiferente a los problemas e ignorante de los métodos de la lingüística es hoy un anacronismo» (Jakobson, 1969: 248)

Por otra parte, desde la literatura es relativamente fácil caer en generalizaciones subjetivas en relación al estilo si nos basamos solamente en la recepción del texto. Cuando leemos a un determinado autor/a podemos emitir juicios sobre diversos aspectos como la riqueza de vocabulario, el ritmo de la narración, etc., pero emitiendo una interpretación propia, individual y subjetiva, que a nuestro parecer puede ser limitada para la investigación académica. Este es además uno de los problemas de la estilística como estudio científico que Dubois et al. (1979) pone de relieve en su *Diccionario de lingüística* cuando comenta que «en [...] las estilísticas modernas se siguen [...] empleando como criterio de pertinencia el juicio o el gusto del estilólogo» (Dubois, 1979: 246). Dubois et al. no pone en duda el carácter específico y el análisis de la estilística empírica, pero si considera necesaria una base científica.

En la estilística literaria se han desarrollado diferentes perspectivas desde las cuales se tratan de analizar los elementos que contenidos en el texto influyen la recepción del mismo, pero sin focalizar los esfuerzos exclusivamente en la figura del autor, pues se tienen en cuenta autor, texto y lector. Es este el motivo por el cual, los teóricos de la recepción de línea **psicologista o biográfica** y también de línea **estructuralista y formalista** encontraron en la estilística un interesante ámbito de estudio (*cfr.* Reyes, 2018)

Así, **la estilística en la literatura** se entiende como parte de los estudios literarios o la historia de la literatura, pero también como disciplina lingüística. Pérez (2009) comenta que, aunque esos dos polos se hayan presentado en ocasiones como opuestos tienen de hecho grandes coincidencias pues como parte de la literatura que es también incluye el análisis de las formas y su interpretación, tratando de describir el sentido último de la obra literaria para lo cual es imprescindible la lingüística. En relación con esto es reseñable la

afirmación que hace Fernández Retamar en torno a las distintas perspectivas al afirmar que:

«sea cual fuere el punto de vista que finalmente se adopte, el campo de estudio de la estilística es el mismo: la totalidad de la lengua [...], aunque buscando en ella fines distintos a los que persigue la lingüística general; buscando en ella, en vez del estudio de las estructuras lingüísticas por sí mismas, el estudio de aquéllas que permiten la expresión del hombre» (Fernández Retamar, 1963: 37).

Este planteamiento, tiene en cuenta la advertencia de muchos estudios que siguiendo la visión de Todorov (1970: 171)⁶ subrayan la necesidad de tener en cuenta la interpretación del conjunto. En este sentido, vemos en la estilometría un puente entre la lingüística y la literatura. Es este el motivo por el cual realizaremos un análisis estilométrico para tratar de determinar la autoría de *La tía fingida* en base a métodos cuantitativos que son principalmente lingüísticos y continuaremos con un análisis literario comparativo de las obras que presenten en base al primer análisis similitudes estilísticas para de esta manera tratar de profundizar en estas similitudes y ver por qué se producen o lo que es lo mismo, si entre los textos similares existen relaciones estilísticas literarias más allá del uso de unas palabras frecuentes determinadas. Al plantearlo así, estaremos teniendo no solo en cuenta el desvío como recurrencia o elección de unas formas lingüísticas y literarias sino también como comparación y desviación del lenguaje estándar.

En resumidas cuentas, en este trabajo se emplea el concepto de estilo de la estilometría que profundizaremos en el próximo apartado y también el de la literatura, donde nos centraremos en aspectos retóricos y morfológicos principalmente. En esencia, los fundamentos teóricos de los estudios de atribución en este ámbito son en palabras de Pascual y Urbón (2009: 28) que «un autor, al escribir, produce un discurso que indefectiblemente salvo manipulaciones específicamente concebidas para ocultar ciertas huellas presenta una serie de marcas verbales, patrones de escritura, que lo caracterizan e identifican frente a otros discursos salidos de diferente mano». Es

⁶ Todorov considera que existe un nexo entre ambas disciplinas pues según su criterio para conocer la literatura se lleva a cabo una descripción formal (usando herramientas lingüísticas) de obras concretas y de la interpretación del conjunto (1970: 171)

precisamente «la convicción de que cada hablante/escritor posee un idiolecto propio, [...] es decir unos hábitos personales que se escapan a lo consciente, lo que ha propiciado el establecimiento de unos algoritmos de atribución que, [...] ante un texto anónimo posibiliten proponer una autoría con cierta seguridad.» (Pascual y Urbón, 2009: 28)

2.2 Estilo para la estilometría.

En el apartado anterior hemos delimitado en cierto modo los conceptos de estilo y estilística. Como hemos visto, la estilística trata desde diferentes puntos de vista de estudiar el estilo. Este podría definirse en base a Dubois et al. (1979) como la marca de la individualidad del sujeto en el discurso (Dubois et al., 1979: 250). En términos generales, se pueden reconocer dos vertientes en la estilística dentro de la lingüística que desde sus orígenes se diferencian por tener diferentes objetivos siendo en ambos casos el producto lingüístico lo que se analiza, ya sea como manifestación **de la lengua o del habla**.

En todo caso, lo visto hasta ahora permite diferenciar a grandes rasgos las líneas teóricas que en el ámbito de la estilística han seguido las ciencias de la literatura y la lingüística del de la estilometría. El concepto de estilo autoral que utiliza la estilometría es una interpretación basada en una definición muy específica y limitada del mismo. Reißler-Pipka (2018) al respecto, apunta que: «el estilo es una compleja reunión de elementos contables determinados por la interpretación»⁷ y cita en este punto también al trabajo de Hermann et al. (2015) quien propone la siguiente definición de estilo para esta rama de la investigación: «el estilo es una propiedad de los textos constituida por un conjunto de características formales que pueden ser observadas cuantitativa y cualitativamente»⁸ (Hermann et al., 2015: 44). Por su parte Holmes (1994: 87) también hace una breve definición de este concepto: «podríamos definir el estilo

⁷ Trad. propia. Original: «*Stil ist eine komplexe Zusammenkunft von zählbaren und durch Interpretation bestimmten Elementen*» (Reißler-Pipka, 2018) En línea en: https://zfdg.de/2018_004 (recuperado el 05.19.21)

⁸ Trad. propia. Original: «*Style is a property of texts constituted by an ensemble of formal features which can be observed quantitatively or qualitatively*» (Hermann et al., 2015: 14)

como un conjunto de patrones medibles que podrían ser únicos de un autor»⁹. Como se aprecia, estas acepciones consideran que en los textos existen una serie de elementos contables tanto cuantitativa como cualitativamente que podrían ser particulares y característicos de un determinado autor o autora. En base a ellos podríamos reconocer ciertos patrones que servirían para identificar el estilo. Esto puede aportar a la literatura múltiples posibilidades de análisis y también con ello podríamos determinar la autoría de un texto mediante la comparación de los elementos comunes o las divergencias.

Esta definición se sustenta pues en la medición de las diferencias entre textos reconocidas cuantitativamente hoy en día mediante herramientas digitales utilizando determinados algoritmos y que son medibles en una gradación que se traduce en valores estadísticos. Los textos a analizar deben coincidir en cuanto a extensión (número de palabras), época y género para poder obtener resultados válidos (cfr. Reißler-Pipka, 2019: 7). A su vez, es necesario subrayar que los procedimientos de la estilometría no se limitan a la atribución de autoría si no que pueden incluirse en las ciencias de la literatura para diversas investigaciones (cfr. Reißler-Pipka, 2019: 7) que van desde el análisis de sentimientos (Zahonero, 2020), clasificación tipológica y estilística de textos literarios (Boto, 2017), género del autor y época (Bia, 2016), estudios lingüísticos históricos y huellas de los traductores (Rueda, 2019) entre otros¹⁰.

La estilometría aporta así valiosos datos pudiendo llevar a nuevas perspectivas que sea cual sea su ámbito permiten seguir investigando entorno a la búsqueda de los factores que son característicos y determinantes para tratar de definir el «singular o exclusivo» estilo autorial. De esta manera se sitúa nuevamente en el debate que veremos a continuación en torno a la figura del autor, sus funciones, etc. (cfr. Reißler Pipka, 2019: 8).

⁹ Trad. propia. Original: «we may define "style" as a set of measurable patterns which may be unique to an author» (Holmes, 1994: 87)

¹⁰ Juola (2018: 299-306) comenta de forma resumida varias posibles aplicaciones de la atribución de autoría (que aquí llamamos estilometría entendiendo que la AA es una aplicación más en este campo). Juola (2018) nombra: género, datación de documentos y cambios del lenguaje, variables socioeconómicas (nacionalidad, dialecto, religión, nativo de la lengua o no, nivel educativo), personalidad y salud mental.

2.3 La importancia del autor: estudios de la teoría literaria en torno a la figura autorial.

Cuando hablamos de la figura autorial dentro del debate autor-obra, la cuestión de estilo toma otros derroteros dentro de la teoría literaria contemporánea al plantear la duda de si es el texto (discurso escrito) y el lector (recepción) o es el autor lo más relevante a la hora de llevar a cabo una lectura, comprensión y análisis del texto literario.

Si observamos la evolución de la figura autorial a lo largo de la historia, vemos que su relevancia va cambiando a lo largo de las épocas y en paralelo la obra se va colocando en primer plano o en una posición subordinada con respecto al mismo. Para Barthes (1968) situar al autor sobre la obra supone limitar la comprensión del texto. Así, es con este teórico con el que la discusión llega a un punto de inflexión en su célebre *La muerte del autor* (1968). Desde su punto de vista lo relevante es la obra y no aquel que la haya creado pues no es posible explicar la obra literaria en base a los acontecimientos y vivencias que rodean la vida del autor. Las consideraciones de Barthes (1968) invitaron a repensar muchos aspectos teóricos al proponer los postulados entorno a la problemática del autor y el concepto de su desaparición para la Nueva Crítica Literaria. Esta publicación tuvo una rápida respuesta en la obra *¿Qué es un autor?*, de Michel Foucault (1969) que redefine la figura del autor entendiendo a este como «función del texto» y en otros trabajos como los de Genette, Eco, etc., para cuyo resumen recomendamos el trabajo de Fumaz (2012).

La relación entre la muerte y la redefinición o resurrección del autor (en términos de Zapata, 2011) corresponde según Rißler-Pipka «por un lado, con la necesidad que tiene la literatura de llevar a cabo una clasificación histórica de autores, estilos y épocas y, por otro lado, el objetivo de realizar una descripción lingüística de los estilos del lenguaje». (2019: 5)

En nuestra opinión, dentro de este debate, consideramos que la importancia del autor sigue estando presente. En el caso que nos ocupa, Cervantes, es entendido en base al conjunto de su obra. Son muchas las relaciones intertextuales que se establecen en su corpus y que nos ayudan a entender determinados textos. Su obra más célebre, *El Quijote*, es quizás el

ejemplo más obvio: ¿cómo entender, interpretar o analizar esta novela sin tener en cuenta sus trabajos dramáticos o las novelas breves? Dar relevancia al autor no nos parece centrarnos en la mera figura autorial, si no que nos permite analizar múltiples aspectos comunes en un corpus que constituye y da significación a un todo. Si analizamos, por ejemplo, los recursos metadieгéticos cervantinos es preciso saber que Pedro de Urdemalas y el entremés del Retablo de las Maravillas son de Cervantes de modo que entendemos también el hábil empleo de este recurso y lo mismo ocurre con otros aspectos como por ejemplo el tratamiento de las figuras femeninas, el sentido didáctico de los textos etc. También es importante para la historia de la literatura si consideramos que los diferentes géneros, movimientos, etc. no surgen espontáneamente, sino que son producto de una herencia literaria de la que se nutren y en la que se desarrollan. El conocimiento del nexo de unión entre las obras, representado por el nombre del autor nos permite un análisis mucho más global del texto entendido no como unidad aislada si no como conjunto que expresa y completa toda una serie de significados e ideas trascendentales para la crítica literaria e historia literaria. Así, en base a estas consideraciones, estimamos relevante llevar a cabo un trabajo de atribución de autoría para tratar de añadir o desechar de ese conjunto los contenidos de la obra dubitada.

3 El análisis cuantitativo de textos

3.1 Orígenes y desarrollo de los métodos cuantitativos para la atribución de autoría

En este apartado, como apunta el título, hablaremos de los orígenes de los métodos cuantitativos para la atribución de autoría y su evolución que nos lleva a la estilometría computacional. Esto nos servirá para hacer una aproximación sincrónica a la metodología que va a ser utilizada permitiendo a su vez entender el recorrido del ámbito de investigación. Para ello, en todo este apartado nos basaremos en la excelente síntesis que realizan Holmes y Kardos (2003) en su trabajo *Who was the autor?*.

Los orígenes de la estilometría se remontan a 1851 con el lógico inglés August de Morgan que comentó por carta a un amigo que la autoría podría

determinarse si, comparando dos textos, uno de ellos presentara diferencias en la longitud de las palabras. Esta hipótesis llamó la atención del físico americano Thomas Mendenhall que en los años 80 del siglo XIX publicó un trabajo en el que mostraba un espectro de vocabulario («*word spectra*») con la frecuencia de las palabras de diferentes longitudes para comprobar si Francis Bacon era en realidad el autor de los trabajos de Shakespeare. Mendenhall comprobó que los espectros de palabras no podían servir como argumento exclusivo para determinar la autoría, pero su trabajo también le permitió observar que los textos de Christopher Marlowe (otro autor contemporáneo de cuyo puño y letra se suponen algunos textos de Shakespeare) presentaba características similares (la preponderancia de palabras de cuatro letras) lo que demostraría una fuerte influencia de Marlowe en el trabajo del Bardo.

Casi 30 años más tarde, el trabajo de Mendenhall fue seguido por las investigaciones del estadístico inglés G.Udny Yule y el lingüista americano George Zipf que profundizaron en características alternativas del estilo. Zipf encontró que existía una relación entre el número de ítems de vocabulario que aparecían en un número determinado de veces («*r times*») en un texto ($r = 1, 2, 3, \text{etc.}$) y «*r*» en sí. De esta manera, utilizando escalas logarítmicas descubrió una relación lineal que denominó la «Zipf's law» e ideó una medida objetiva de esta característica de las frecuencias de palabras que se conoce como «Yule's characteristic K». Sin embargo, la investigación demostró más tarde que la Yule's K no es suficientemente fiable para establecer la autoría.

Es ya en el siglo XX a comienzo de los 60 cuando el misterio de la autoría de los *Federalist Papers* permite a dos célebres estadísticos americanos, Frederik Mosteller y David Wallace, aplicar sus conocimientos y demostrar de forma convincente el potencial de la estilometría. Los *Federalist Papers* son un conjunto de 85 ensayos políticos publicados en los periódicos de Nueva York en 1787 y 1788 que trataban de persuadir a los votantes para la ratificación de la constitución de los Estados Unidos. Aunque estos ensayos fueron todos firmados como «Publius», se sabe que fueron escritos por Alexander Hamilton, James Madison y John Jay como ellos mismos afirmaron con posterioridad. Sin embargo, la autoría de doce de los ochenta y cinco papeles, los denominados «*disputed papers*» fue reclamada tanto por Hamilton como por Madison. Es en

este punto en el que Mosteller y Wallace, basándose en el trabajo de Ellegard (1962)¹¹ emplearon las palabras funcionales (preposiciones, conjunciones y artículos) como discriminadores. Esto les permitió obtener los porcentajes de aparición de una determinada palabra por grupo de palabras (p.ej. «*upon*» que aparece 3,24 veces por cada 1000 palabras en los textos de Hamilton y solamente 0,23 veces en los textos de Madison). En base a estos datos, emplearon probabilidades numéricas para expresar el grado de creencia de autoría para cada uno de los dos autores y utilizaron el teorema de Bayes para ajustar las probabilidades. La investigación concluyó con la atribución de los doce textos a Madison, pero también se subrayó que fue un caso de atribución particularmente difícil debido al contexto político y las similitudes de estilo entre ambos autores.

El paso del uso de un método estadístico a el uso de varios combinados lo encontramos a finales de los 80 y principios de los 90 del siglo XX con la publicación por parte de John Burrows de una serie de documentos de referencia que para Holmes y Kardos (2003) son tan relevantes como el trabajo de Mosteller y Wallace. Burrows utilizó amplios sets de palabras funcionales con una alta frecuencia, computó sus porcentajes de ocurrencia por cada mil palabras en los textos del candidato y empleó el análisis de componentes principales en el «multivariate data array» resultante. Burrows utilizó su método con autores y géneros variados (Austen, Bronte, y los diferentes géneros de Scott y Byron) y obtuvo resultados importantes pues pudo demostrar que existe una forma diferente y distintiva de los autores en la utilización de palabras funcionales comunes. De este modo, la aplicación del «método Burrows» es un paso fundamental en los problemas de atribución de autoría y el uso de técnicas estadísticas multivariantes es casi la regla en la estilometría.

Así, es basándose en la metodología llevada a cabo en otros trabajos de atribución de autoría utilizando una versión actualizada y mejorada del algoritmo de Burrows (Cosine Delta) de lo que nos serviremos para tratar de encontrar respuesta a las preguntas que nos planteamos en este trabajo.

¹¹ Ellegard, A. (1962), "A Statistical Method for Determining Authorship: The Junius Letters 1769-1772," *Gothenburg Studies in English*, 13, Sweden: University of Gotenborg. En Holmes y Kardos (2003)

3.2 Atribución de Autoría

La atribución de autoría para Patrick Juola (2008) es en una definición amplia «la ciencia de inferir características del autor a partir de las características de los documentos escritos por ese autor»¹² (Juola, 2008: 233) y también «cualquier intento de inferir las características del creador de un archivo lingüístico»¹³ (Juola, 2008: 238). Estas definiciones tan amplias pretenden incluir por lo tanto las investigaciones hechas en otras áreas como el reconocimiento del habla (speech recognition).

En cuanto a la selección del corpus, Juola diferencia dos tipos: cerrado y abierto. El primero consiste en un texto que se sabe que es de uno de un conjunto de autores por lo que el objetivo es determinar cuál es. El segundo consiste en el mismo procedimiento, pero en este caso se trata de determinar quién es el autor si es que estuviera entre los autores seleccionados. Por lo tanto, la diferencia principal en este primer problema de la atribución de autoría es ¿sabemos a ciencia cierta que el/la autor/a está entre los que hemos seleccionado?

Juola, 2008 comenta algunos de los problemas a los que se enfrentan los trabajos de atribución de autoría entre los que destacan la precisión de las técnicas y las cuestiones implícitas y explícitas de las características de los textos y de su selección. Sobre la precisión de la atribución de autoría, Juola (2008: 246) comenta que hay tres aspectos interrelacionados en este aspecto. El primero es que existen prácticamente un ilimitado número de técnicas que han sido propuestas para este fin y que por lo tanto cabe preguntarse cuál de ellas es más fiable y en qué circunstancia particular. El segundo son las cuestiones de género, representatividad y tamaño del corpus ya que son aspectos que hacen variar los resultados estadísticos. El tercero es la facilidad para crear programas que calculen aspectos lingüísticos como longitud de las palabras, número de palabras, etc. lo cual puede llevar, según Juola, a un abuso en la aplicación de herramientas fuera de su área de validez, es decir que hay que

¹² Trad. Propia. Original: «Authorship attribution, the science of inferring characteristics of the author from the characteristics of documents written by that author» (Juola, 2018: 233)

¹³ Trad. Propia. Original: «any attempt to infer the characteristics of the creator of a piece of linguistic data» (Juola, 2008: 238)

tener en cuenta aspectos como por ejemplo la lengua original de un texto o un corpus y otros aspectos para poder utilizar una herramienta, pues incluso si las técnicas son precisas, estas podrían no ser adecuadas en una investigación determinada.

Por otra parte, los textos en sí, deben ser lógicamente también considerados. Juola (2008: 247) subraya la importancia de unas muestras que no estén corruptas. Si bien también indica que determinar un texto limpio para el análisis de autoría puede ser una labor muy complicada o incluso imposible. Esto se debe a que en teoría solamente se deberían utilizar para el análisis de autoría textos cuyas características fueran responsabilidad directa del autor lo cual en la práctica es difícil de determinar. Además, otro problema que existe en relación con el corpus que contiene los textos de control consiste en responder a las cuestiones de qué autores, qué documentos y qué partes de documentos deben ser incluidos en el análisis. En resumidas cuentas, en el corpus de control solo deben ser incluidos textos de los que sepamos con total seguridad quienes son los autores.

Una vez mostrados los problemas fundamentales de un trabajo de atribución de autoría, Juola (2008: 249) propone un posible sistema metodológico ideal. La primera consideración es la precisión de resultados es decir que el sistema debe utilizar métodos que hayan dado resultados satisfactorios sin incurrir en errores con frecuencia. Dichos resultados deben poseer ya un historial en diferentes áreas que permitan tanto definir los niveles previstos de precisión como los límites donde la precisión se reduciría. A parte de esto, el sistema debería disponer de explicaciones que detallen las características que permiten determinar la autoría lo cual hace que los resultados del programa puedan ser más útiles para un amplio público. Por último, el sistema ideal debería poder ser modular y extensivo para posibilitar la adición, eliminación o combinación de nuevas propuestas que permitan acercarse al conjunto de características ideal y estar teóricamente bien fundamentado de forma que se pueda explicar la efectividad de las funciones en términos de características lingüísticas fisiológicas y/o neurológicas del proceso de escritura.

Advertiremos más adelante que la herramienta *stylo* cumple con todos estos requisitos: ya ha dado resultados cuya precisión es medible, hay

numerosos trabajos que han utilizado este método, posee explicaciones de su funcionamiento que muestran de forma transparente como se lleva a cabo la atribución de autoría, podemos seleccionar la/s característica/s a analizar y es extensivo pues es factible modificar los parámetros pudiendo además añadir o eliminar textos de las carpetas para incluir nuevas propuestas. A continuación, vamos a resumir los requisitos, en parte ya vistos, que son necesarios para atribuir la autoría de un texto.

3.3 Bases para la atribución de autoría

Como vimos, Holmes (1994) ya comenta las virtudes de la atribución de autoría desde la estilometría y afirma que un estudio estadístico de autoría disputada debe incluir la comparación del texto dubitado con las obras de cada uno de los posibles autores candidatos mediante pruebas estadísticas apropiadas sobre las características cuantificables de los textos, características que reflejan el «estilo» en base a la definición del mismo que ya comentamos. (cfr. Holmes, 1994)

El mismo autor cita a Bailey (1979)¹⁴ quien en un contexto de lingüística forense propone tres reglas básicas para poder proceder a establecer hipótesis de autoría y que se pueden aplicar perfectamente al ámbito literario (Bailey, 1979; como se citó en Holmes, 1994: 87-88):

- i) *the number of putative authors should constitute a well-defined set;*
- ii) *the lengths of the writings should be sufficient to reflect the linguistic habits of the author of the disputed text and also those of each of the candidates;*
- iii) *the texts used for comparison should be commensurate with the disputed writing.*

Por lo tanto, debemos de tener en cuenta tres aspectos básicos: un número definido de posibles autores, longitud de los textos suficiente, textos similares (aquí debemos tomar en consideración la longitud y el tipo de texto).

¹⁴ Bailey, R. W. (1979). Authorship Attribution in a Forensic setting, *Advances in Computer-aided Literary and Linguistic Research*. Eds. D.E. Ager, F.E. Knowles and J. Smith. Birmingham: AMLC.

Por su parte Holmes (1994) expone en su trabajo un listado de criterios que pueden servir de base para cuantificar características de los textos que sirvan para la atribución de autoría y también ejemplifica con diferentes trabajos que se han realizado tomando dichos criterios como base del análisis. El listado no es totalmente exhaustivo, pero sirve como referencia para hacerse una idea de las múltiples posibilidades disponibles: sílabas; longitud de las palabras; longitud de las frases; Part of Speech distribution; Function Words; The Type-Token Ratio; Simpson's Index (D); Yule's Characteristic (K); Entropía; distribución de vocabulario: modelos para las curvas de vocabulario; comparar vocabulario en textos de diferentes longitudes; frecuencias de palabras; hápax legomena; hápax dislegomena; estudios multivariados (cfr. Holmes, 1994). Como vemos existen criterios estadísticos variados a tener en cuenta. Juola (2008) en su trabajo *Autorship attribution* teoriza sobre estos y otros aspectos problemáticos.

Como se muestra, hay una importante variedad de criterios lo cual lleva a preguntarse cuál de estas es más efectiva para un objetivo determinado. En este trabajo nos basamos para la elección de los mismos en otras investigaciones realizadas en el mismo contexto (atribución de autoría en la literatura y más específicamente la del Siglo de Oro). Adelantamos que se trata de un análisis multivariado y que como ya se ha dicho, tendrá en cuenta principalmente la frecuencia de las palabras funcionales. A continuación, veremos algunos de estos trabajos exponiendo previamente el estado de las investigaciones en el ámbito del español.

4 La estilometría

4.1 Estilometría desde las H.D. en español.

Hace ahora cinco años, en el 2016, José Calvo Tello comentaba en su trabajo *Entendiendo Delta desde las humanidades digitales* que a pesar del desarrollo de la investigación en Humanidades Digitales del ámbito de la estilometría pocos estudios habían sido realizados con textos en español o desarrollados en países hispanohablantes. Desde entonces, se puede observar que se han multiplicado los trabajos y que estos abarcan diferentes aspectos de

la estilometría como por ejemplo clasificación por géneros, atribución de autoría, análisis de sentimientos, etc. Uno de los escollos que añadía dificultad a la investigación según Calvo Tello eran las escasas aportaciones que existían en español sobre la estadística, la lingüística computacional o la informática, ámbitos necesarios para entender la estilometría. Afortunadamente, en la actualidad se encuentran trabajos que exponen estas informaciones de forma clara facilitando así el uso de las herramientas que aportan estos campos a investigadores/as del ámbito de la lingüística y la literatura. En este apartado, expondremos algunos de los trabajos como muestra del punto en el que nos hallamos, sin pretender ser completamente exhaustivos para evitar una compilación innecesaria de trabajos y datos.

Para empezar, subrayaremos el ya citado trabajo de Tello (2016) pues en él, este autor explica de forma sencilla el funcionamiento del algoritmo que utilizaremos más tarde para nuestro análisis, es decir, la medida de distancia textual de Burrows llamada Delta que como el autor indica es uno de los métodos principales en la estilometría.

Otro trabajo destacable es el que realiza Bravo (2017) pues en él utiliza la herramienta *stylo* para llevar a cabo una agrupación tipológica y estilística de la obra de Eduardo Mendoza con buenos resultados. También destaca en sus conclusiones la necesidad de tener en cuenta tanto el macroanálisis basado en la estadística como el estudio estilístico tradicional.

En el ámbito de la atribución de autoría que nos ocupa aquí, existen numerosas referencias. Podemos destacar los aportes de Fradejas Rueda (2016) donde muestra la efectividad de la herramienta *stylo* para la AA. También del mismo Fradejas Rueda (2020) es el libro *Cuentapalabras*, disponible en la red y que es un manual imprescindible para que los filólogos puedan aplicar paso a paso y de forma relativamente sencilla algunas de las funciones que ofrece *stylo*. Otra publicación interesante del mismo autor muestra la capacidad de la herramienta *stylo* para el reconocimiento del origen de las fuentes utilizadas para las traducciones de textos medievales (vid. Fradejas Rueda, 2019).

Siguiendo con la atribución de autoría, Rissler-Pipka (2016) aborda la problemática de Avellaneda y el otro Don Quijote desde las nuevas herramientas

digitales. A partir de su análisis subraya algunas limitaciones que tiene *stylo* para la AA, ya que como herramienta sirve para establecer hipótesis y no conclusiones definitivas. No obstante, el experimento logra hallar semejanzas estilísticas¹⁵ entre Avellaneda y Cervantes.

Podríamos continuar con otros trabajos igualmente importantes, sin embargo, en la próxima sección vamos a centrarnos en algunos de los que se han llevado a cabo para analizar la obra de Cervantes desde la estilometría computacional. Por este motivo daremos aquí este por concluido este apartado, subrayando que existen otras muchas publicaciones al respecto. Sirvan las obras aquí citadas como muestra ejemplar de algunas de las múltiples posibilidades de la aplicación de las herramientas digitales y también de demostración de la validez y el incremento de la importancia de esta rama de la investigación en español en los últimos años cuyos resultados además justifican el planteamiento de nuestro trabajo.

4.2 Estilometría en la obra de Cervantes

La obra de Cervantes es particularmente llamativa para las Humanidades Digitales debido a la existencia de numerosos textos que, aunque se le hayan inscrito no se sabe su paternidad con total seguridad. El detonante de las atribuciones viene de largo pues la existencia de obras sin nombre es sugerida ya por el propio autor en el Prólogo de las *Novelas Ejemplares*, donde en la descripción de su autorretrato comenta:

«éste digo que es el rostro del autor de La Galatea y de Don Quijote de la Mancha, y del que hizo el Viaje del Parnaso, a imitación del de César Caporal Perusino, y otras obras que andan por ahí descarriadas y, quizá, sin el nombre de su dueño. Llámase comúnmente Miguel de Cervantes Saavedra.»
(Cervantes, *Novelas Ejemplares*)

Sembrada la duda fueron surgiendo atribuciones de toda clase (prosa, verso, epistolarios...) y siendo catalogadas como ciertas o falsas en base a una

¹⁵ Donde por «estilo» se entiende el uso de las palabras más frecuentes de 100 a 5000 (Rissler-Pipka, 2016: 50)

gran variedad de criterios. Una recopilación de las mismas puede encontrarse en los trabajos de Ford y Lansing (1931), de Marín (1948) o de Eisenberg (1990). Actualmente, una simple búsqueda en la página web de la Biblioteca Virtual Miguel de Cervantes¹⁶ nos permite tener acceso a un listado de (algunas) de las obras que presentan dudas en cuanto a su atribución al manco de Lepanto junto a ediciones digitales de las mismas:

<ul style="list-style-type: none"> ▪ <i>El Buscapié</i>
<ul style="list-style-type: none"> ▪ Ed. Sevilla Arroyo y Rodríguez Rodríguez
<ul style="list-style-type: none"> ▪ <i>La conquista de Jerusalén</i>
<ul style="list-style-type: none"> ▪ Ed. Sevilla Arroyo ▪ Manuscrito II/460 (13) 1058632 Biblioteca Real
<ul style="list-style-type: none"> ▪ <i>Los habladores</i>
<ul style="list-style-type: none"> ▪ Ed. Manuel de Foronda
<ul style="list-style-type: none"> ▪ <i>Las semanas del jardín</i>
<ul style="list-style-type: none"> ▪ Ed. Daniel Eisenberg
<ul style="list-style-type: none"> ▪ <i>La tía fingida</i>
<p><u>Versión del Códice A2-141-4 Biblioteca Colombina</u></p> <ul style="list-style-type: none"> ▪ Ed. Schevill y Bonilla
<p><u>Versión del Manuscrito 56-4-34 Biblioteca Colombina</u></p> <ul style="list-style-type: none"> ▪ Ed. Sevilla Arroyo
<p><u>Versión Porras de la Cámara</u></p> <ul style="list-style-type: none"> ▪ Según edición de García Arrieta <ul style="list-style-type: none"> ➤ Por Imprenta de la viuda de Vallín, 1814 ▪ Según edición de Franceson-Wolf <ul style="list-style-type: none"> ➤ Ed. Schevill y Bonilla ➤ Ed. Sevilla Arroyo ➤ Por G. C. Nauck, 1818

Tabla 1: obras atribuidas a Cervantes. BVMC

Insistimos en que esta lista no es completa, sin embargo, al contrario que otras obras atribuidas a Cervantes que no aparecen en la BVMC estas están disponibles en formato digital, motivo por el cual la tomaremos como referencia para nuestro experimento. Entre estas obras de la BVMC ya han sido varios los trabajos que se han ocupado de utilizar las herramientas digitales en general y

¹⁶ http://www.cervantesvirtual.com/portales/miguel_de_cervantes/estados_atribuciones/

la función *stylo* en particular para tratar de encontrar hipótesis que puedan ayudar a dar una respuesta a la duda:

- Del *Buscapié* ha demostrado ser una falsificación de Adolfo de Castro (vid. Eisenberg, 1990: 483; Romero-Ferrer, 2000: 180-181) su inclusión en la lista de la BVMC no aparenta tener fundamento a nuestro parecer.
- *La Conquista de Jerusalén* ocupó a Soler y Tello (2019) en un trabajo que utiliza también la herramienta *stylo* en el que en base a los resultados demostraron que esta obra ocupaba casi sin lugar a dudas un lugar en el corpus cervantino.
- *Las Semanas del Jardín* está incompleta y solo se conserva el *Diálogo entre Cilenia y Selanio sobre la vida del campo*. El fragmento se puede encontrar en la BVMC y es parte del trabajo de Eisenberg (1988) *Las Semanas del jardín de Miguel de Cervantes: estudio, edición y facsímil del manuscrito* (Vol. 3). Consideramos que por sus características y al existir solo una parte de un texto más amplio no es adecuada para nuestro análisis.
- De *La Tía Fingida* hablaremos amplio y tendido en el próximo capítulo, ya que como ya adelantamos, utilizaremos esta obra para nuestro análisis. Podemos decir de antemano que no existen trabajos que utilicen *stylo* para la atribución de autoría lo cual la hace ya un objeto de estudio interesante.
- El entremés de *Los Mirones* que no aparece en la lista de la BVMC fue investigado por Blasco (2019) utilizando *stylo*. Blasco da muestra de las dificultades con las que se encuentra el análisis estadístico al trabajar este subgénero debido a la inestabilidad de los textos inherente a las circunstancias de transmisión del entremés.
- De *Los habladores* comenta Arellano-Torres (2018) que podría ser de Cervantes debido a ciertas coincidencias lexicales, sin embargo, este autor no pretende abordar la polémica de la autoría sino realizar una edición crítica del texto. El propio título de su trabajo implica el reconocimiento de la obra como dubitada pues versa «El entremés de Los habladores, atribuido a Cervantes». No existe a nuestro saber ningún trabajo hasta la fecha que haya afrontado este texto específicamente

utilizando los recursos de las H.D. No obstante, Blasco (2019) al analizar el entremés de *Los Mirones*, incluye *Los Habladores* en su corpus y este resulta asociado a Quevedo, pero en una subrama (del dendograma) muy próxima al resto de entremeses cervantinos.

Vemos pues que existen entorno a la obra cervantina algunos trabajos que ya han aplicado la estilometría computacional utilizando también *stylo* para tratar de encontrar al autor tras los textos. En todos los casos sirvió si no para determinar la autoría de manera fehaciente al menos para plantear hipótesis que permitan seguir investigando. La obra que elegimos para nuestro análisis es la *Tía Fingida*. En los próximos capítulos justificaremos esta elección exhaustivamente, pero adelantamos que el hecho de que no exista ningún trabajo desde las H.D. utilizando la herramienta *stylo* y sí muchos desde otras perspectivas que no han conseguido determinar la autoría convincentemente es ya a nuestro parecer un motivo de peso. Llegados a este punto, es necesario sumergirse en el siguiente paso previo al experimento que es la recopilación y justificación de un corpus con el que contrastar la TF.

5 Corpus

5.1 Problemática de los textos del Siglo de Oro

En este capítulo vamos a comprender por qué la literatura del Siglo de Oro español es particularmente interesante para el análisis cuantitativo de textos y en particular para los trabajos de atribución de autoría debido a diversos motivos que tienen que ver sobre todo con las circunstancias de transmisión de los mismos. A pesar de ello, las circunstancias de transmisión también dan lugar numerosos problemas que también expondremos.

Actualmente están a disposición de la investigación numerosos textos y ediciones digitales variadas. Sin embargo, como comenta (Rißler-Pipka, 2018) para el análisis estilométrico es necesario disponer de los textos digitales, pero también los metadatos deben estar completos y ser correctos. Estos metadatos son algunas de las informaciones extralingüísticas de los textos: autor, fecha y lugar de publicación, etc. Su importancia se debe al hecho de que la calidad de las ediciones de los textos de este periodo es difícilmente comprobable incluso

en ediciones contemporáneas. Juola (2008: 247) apunta en este sentido a otro inconveniente que puede existir en las ediciones digitales donde pone como ejemplo los textos existentes en Google Scholar, JSTOR, o el Project Gutenberg que podrían haber sido corrompidos debido al proceso de escaneado o reescritura.

Otros de los aspectos en los que se presentan dificultades para la atribución de autoría es la determinación del corpus como abierto o cerrado (open/closed set) que vimos en Juola (2008). En este sentido, la primera pregunta que habría que responder es ¿sabemos a ciencia cierta que el posible autor se encuentra entre los seleccionados? En determinados casos, como *el Lazarillo*, esto supone un gran reto al ser esta una obra anónima, por lo que se pueden plantear muchos candidatos (vid. De la Rosa Pérez, 2016: 9). En nuestro caso, tenemos también una obra anónima, pero esta ha sido atribuida ya a Cervantes. Partimos pues de la verificación de autoría. En todo caso un análisis exhaustivo conllevaría reunir todo el corpus de obras de los posibles autores, que aun suponiendo que pudiera ser Cervantes y existiendo determinadas hipótesis podría ser cualquiera que, por ejemplo, hubiera realizado trabajos del mismo género (novela breve) en torno a la fecha de publicación de la obra cuya autoría queremos determinar. Como veremos más adelante, existen limitaciones para poder reunir este corpus ideal.

Además de este aspecto, en muchos casos no existen ediciones modernas de los manuscritos de la Edad de Oro española lo cual ya es un impedimento para su digitalización como comenta De la Rosa Pérez (2016: 9). En este sentido, Fradejas Rueda en una conferencia realizada en el marco de los cursos de verano de la UNED¹⁷¹⁸ en 2021 comenta el procedimiento utilizado en la investigación de atribución de autoría realizado por Madroñal en colaboración con el proyecto ETSO¹⁹ sobre una comedia de un tal Miguel de Bermúdez que resultó ser de Lope de Vega²⁰. En este caso, se escanearon los manuscritos y

¹⁷ Universidad Nacional de Educación a Distancia.

¹⁸ BibliotecasUNED. (23.09.21) Qué es la estilometría, usos y aplicaciones [Video]. YouTube. Obtenido de <https://www.youtube.com/watch?v=tVZ5tFDfPrA>

¹⁹ ETSO. *Estilometría aplicada al Teatro del Siglo de Oro*. <http://etso.es/>

²⁰ El resultado de este trabajo se puede consultar en: Madroñal, A. (2021) *El renacer del Fénix: Yo he hecho lo que he podido, Fortuna lo que ha querido. Una nueva comedia de Lope de Vega* (Colección «Olmedo Clásico», Número 17), Valladolid-Olmedo, Universidad de Valladolid-Ayuntamiento de Olmedo.

después se pasaron por un programa llamado *Transkribus*²¹ que como su nombre delata transcribe automáticamente textos. La efectividad de este programa muestra un acierto del 99,3% y aunque exige una revisión para solucionar ese 0,7% de error, aporta una inestimable ayuda a la hora de disponer de materiales fiables para el análisis digital. A pesar de ello la disponibilidad de los manuscritos no digitales es en muchas ocasiones también limitada o inexistente y en cualquier caso en nuestro trabajo no pretendemos asumir la digitalización de los textos.

Siguiendo con los manuscritos es necesario reseñar un importante escollo añadido que es que los textos que mejores resultados estadísticos dan en cuanto a fiabilidad son aquellos cuyo contenido es el expresamente creado por el autor. Pues bien, incluso tomando como punto de partida los manuscritos del Siglo de Oro, estos presentan una serie de problemas dado que los procesos de impresión no eran tal y como los conocemos hoy en día. Francisco Rico (2000) publicó un trabajo coordinado al respecto de esta problemática con el nombre *Imprenta y crítica textual en el Siglo de Oro*. En este libro, podemos no solamente hacernos una idea de las ediciones que se realizaban por el propio impresor de los textos originales sino también del origen del problema que nos ocupa: la atribución de autoría.

Por otra parte, Arellano (1997: 41-42) comenta las vicisitudes de la transmisión literaria en el Siglo de Oro subrayando la necesidad de preguntarse sobre la trayectoria de la difusión de los textos desde el manuscrito pasando por la edición y la posterior venta del libro. Distingue dos vías posibles para la adquisición del texto por parte de un el lector del Siglo de Oro: una colectiva (la representación teatral, p.ej.) y otra individual (el libro, la lectura). En base a estas circunstancias Arellano (1997: 42) se cuestiona sobre la vía individual y en particular sobre la cantidad y la facilidad para acceder a la lectura de los textos en el Siglo de Oro. El autor divide su artículo distinguiendo y explicando los avatares de los textos según su tipo: textos impresos poéticos/ manuscritos poéticos; textos impresos teatrales/ los manuscritos; textos impresos en prosa/ manuscritos. Nos interesa particularmente la información que aporta en cuanto

²¹ <https://readcoop.eu/transkribus/>

a los textos impresos en prosa. Dentro de estos, Arellano (1997: 51) da muestra de algunos ejemplos de los procesos de impresión exponiendo ciertos aspectos y problemas que dificultaron el paso del manuscrito al libro. Así, en un apartado que dedica a continuación a las reescrituras por razones varias de motivo externo incluye las *Novelas ejemplares* de Cervantes que se publican en 1613. Algunas novelas eran anteriores, pues en el *Quijote* de 1605 se cita al *Rinconete y Cortadillo*. Entre 1604-1606 fue compilada una colección miscelánea de obras para el cardenal Fernando Niño de Guevara por parte de Francisco Porrás de la Cámara, racionero de la catedral de Sevilla. En ella incluye sendas versiones con variantes de *Rinconete y Cortadillo* y del *Celoso extremeño*, además de *La tía fingida*. El manuscrito de Porrás de la Cámara, desapareció de la biblioteca de los Reales Estudios de San Isidro entre 1810-1823²², cuando estaba en posesión del erudito liberal Bartolomé José Gallardo, que al menos había tenido tiempo de sacar algunas copias. Hoy tenemos la copia manuscrita («basada» en el códice Porrás de 1604-1606) y la versión impresa de la novela *El celoso extremeño*. (1613). Lo curioso de estas dos versiones es que el final difiere pues en la primera versión que recoge el manuscrito de Porrás de la Cámara el acto sexual adúltero se consumaba (Arellano, 1997: 54). Esto ha llevado a discrepancias entre los estudiosos: Maregalli (1992) achaca el final de 1613 a la censura, Canavaggio (1992) sin embargo considera que el mismo Cervantes habría llevado a cabo la reescritura para dar a la trama y los personajes mayor complejidad (cfr. Arellano, 1997: 55). En cualquiera de los casos, nuestra intención no es en este punto entrar en este debate, si no exponer algunos ejemplos de las vicisitudes y los cambios realizados en los textos que son otro de los variados problemas que son inherentes en la literatura del Siglo de Oro.

Una vez expuestos los aspectos limitantes más generales, pasaremos al siguiente apartado donde expondremos las fuentes de las que disponemos para poder compilar un corpus de trabajo.

²² Arellano (1997:54) comenta que «terminó en las aguas del Guadalquivir en 1923». Lucía Megías, J. M. (2018) indica que fue «con posterioridad a 1810». Aprovechamos para recomendar el trabajo de Lucía Megías (2018) para conocer mucho más sobre el códice Porrás. En su trabajo también podemos observar las distintas manos que de una forma u otra pudieron intervenir en los textos originales y que sirven para reforzar la problemática sobre la fiabilidad de los textos que pretendemos aclarar en este apartado.

5.2 Fuentes para la recopilación de textos del corpus

En otro orden de aspectos y dejando a un lado todos los obstáculos anteriormente mostrados, es indispensable para un trabajo de estilometría reunir un amplio corpus que sea fiable y representativo. Para este fin, existen en la actualidad algunos recursos que son de gran ayuda pero que sin embargo presentan en muchos casos ciertas limitaciones. José Calvo Tello en Github²³ pone a disposición una lista de portales que aportan recursos digitales en español. Entre ellos podemos destacar la base de datos de la **Biblioteca Virtual Miguel Cervantes (BVMC)** que da acceso a un inmenso catálogo de libros entre los que se hayan una gran cantidad del Siglo de Oro. Sin embargo, como comenta Rißler-Pipka (2018) muchos de estos libros son ediciones antiguas por cuestiones de derechos de autor lo cual es un perjuicio a la hora de realizar un escaneado además de que su lenguaje puede no corresponder con el español moderno. Estos aspectos se sumarían al que comentaba Juola (2008) con respecto a la corrupción de los textos en este proceso, ya que como apunta Rissler-Pipka (2018) en base al trabajo de Fradejas-Rueda (2016: 199) también pueden existir errores en los metadatos, en este caso particular, se trata del año de publicación de una edición disponible en el BVMC del *Libro de la caza de las aves* (S.XIV) de Pedro López de Ayala. Otras bases de datos reseñables que pueden servir de fuente para la recopilación de textos son: **AHCT** (Association for Hispanic Classical Theater); **CORDE** (Corpus diacrónico del español) que no pone a disposición ningún texto completo permitiendo solamente la búsqueda de títulos y autores en un periodo determinado. Es destacable también proyecto: Canon 60. *La colección esencial del TC/12* realizada por Joan Oleza como coordinador en el año 2014²⁴. Actualmente este corpus está disponible en la BVMC no obstante, esta es una colección de obras teatrales, por lo que no nos resulta de gran utilidad. Existen otras numerosas fuentes que, aunque ofrecen los textos no presentan metadatos o estos no son fiables. Vemos pues que existen también trabas a la hora de elaborar un corpus para el Siglo de Oro.

²³ <https://github.com/morethanbooks/Atlas-de-Datos/blob/master/atlas%20de%20datos.csv>

²⁴ Disponible en: <http://www.cervantesvirtual.com/nd/ark:/59851/bmckw791>

Todos los aspectos problemáticos y limitantes expuestos en este apartado podrían hacer pensar que un trabajo como el que nos ocupa podría no presentar la suficiente fiabilidad y demasiados escollos para poder ser conclusivo. Pues bien, tampoco lo pretende, recuperemos en este punto las palabras de Holmes y Kardos (2003) que afirmaban que «la estilometría [...] no pretende anular la escolástica tradicional de los expertos en literatura e historia, más bien trata de complementar su trabajo proveyendo significados alternativos a los trabajos de investigación sobre proveniencia dudosa»²⁵ (Holmes y Kardos, 2003: 1) y adjuntaban una insistencia final donde enuncian que «la evidencia estilométrica tiene que ser contrastada con la provista por estudios más convencionales hechos por los estudios literarios»²⁶ (Ibid., 2003: 5). Por ende, el objetivo de aportar nuevas perspectivas sigue siendo legítimo y posible pese a las limitaciones que la época impone al corpus. Además, aunque debemos de tener en cuenta todos los aspectos condicionantes aquí señalados, estos no han sido óbice para obtener resultados satisfactorios en otros trabajos en los cuales nos basamos para aplicar sus métodos.

Por último, consideramos que la utilización de las MFW como marcador de estilo puede en parte hacer frente a algunos de las restricciones como la alteración que pueden producir en la estadística las modificaciones y ediciones del texto dado que pensamos que siendo estas a menudo las denominadas «palabras funcionales» son menos susceptibles de ser editadas pues no varían el significado o el mensaje de un determinado texto. Por otra parte, en algunos casos pueden faltar partes del texto, para lo cual no existe a nuestro saber ninguna solución. No obstante, esta consideración es una mera hipótesis. Una respuesta a esta conllevaría un estudio y análisis profundo de las vicisitudes y los cambios más frecuentes en los textos del Siglo de Oro que estableciera exactamente cómo se modificaban los textos.

²⁵ Traducción propia. Original: «Stylometry – the statistical analysis of literary style – does not seek to overturn traditional scholarship by literary experts and historians, rather it seeks to complement their work by providing an alternative means of investigating works of doubtful provenance» (Holmes y Kardos, 2003: 1)

²⁶ Traducción propia. Original: «stylometric evidence must always be weighed in the balance along with that provided by more conventional studies made by literary scholars» (Ibid, 2003:5)

Vistas las fuentes, pasaremos pues en el próximo apartado al proceso de compilación de nuestro corpus y la explicación de su configuración.

5.3 Cosecha de los textos y configuración del corpus

Para compilar el corpus con el que vamos a trabajar llevamos a cabo en primer lugar una consulta en el CORDE²⁷. Dentro de este, seleccionamos todos los autores entre 1547 y 1620, periodo que coincide prácticamente con la vida de Miguel de Cervantes (1547-1616). Es una selección amplia que pretende abarcar todas las hipótesis posibles por lejanas que parezcan. Dentro de este periodo nos interesa saber que autores realizaron **prosa narrativa breve** en España, tanto culta como tradicional. El buscador arroja un total de 63 documentos con 677554 palabras. De estos resultados eliminaremos los numerosos anónimos que aparecen, pues no nos interesa incluir más textos dubitados. También se excluirán cartas, memoriales, premáticas y otros textos que por su extensión excesivamente breve o su tipología pueden ser problemáticas al compararlas con la TF.

De esta manera nos quedamos con los textos que aparecen en las tablas expuestas en el anexo «corpus» (Tablas 1, 2 y 3). Si bien, ahí no termina el cribado. Como se puede observar en las tablas del anexo, los textos de las celdas marcadas en rojo serán excluidos por diferentes motivos ya sean cuestiones literarias o limitaciones que impone el estado de la digitalización de los textos. Estas obras marcadas en rojo, son en muchos casos textos que en un corpus ideal para el objetivo que nos ocupa estarían incluidos pero dado que no disponemos de ellos en formato digital o implican la necesidad de modificarlos profundamente para ser leídos por los programas deben ser dejados a parte en este trabajo (ver anexo: Corpus 1 – prosa narrativa breve. Tabla 1)

En vista de que tenemos tan solo 4 autores es necesario ampliar nuestro corpus incluyendo al vecino más próximo que es en este caso la **prosa narrativa extensa** y dentro de ella la novela y formas similares. En un análisis ideal dispondríamos de suficientes textos de prosa breve similares a las *Novelas Ejemplares* de diferentes autores de los que además tendríamos numerosos

²⁷ REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE) [en línea]. Corpus diacrónico del español. <<http://www.rae.es>> [30.10.21]

ejemplos, no obstante, no es el caso y no hay nada que podamos hacer para remediarlo. Una nueva búsqueda incluyendo la narrativa extensa da en el CORDE como resultado 39 documentos con un total de 4061193 palabras. Las *Novelas Ejemplares* tienen alrededor de 5000 hasta ca. 23000 palabras y como vimos, para el análisis estilométrico, necesitamos textos que sean lo más similares en cuanto a su extensión, sin embargo, dados los límites del corpus de la prosa breve no estamos en condiciones de rechazar muestras. Por este motivo trataremos de incorporar algunos de los textos que se pueden ver en el anexo (Corpus 2 – prosa narrativa extensa. Tabla 2).

A este corpus debemos adjuntar **otros textos** que tienen algunas similitudes con las *Novelas Ejemplares* pero que pertenecen a una época posterior a Cervantes. La autoría es en este caso más hipotética pues como bien comenta Madrigal (2003): «la colección de Porras de la Cámara no puede ser posterior a 1609, pues el cardenal Niño de Guevara, a quien va dirigida la colección, muere en enero de ese año, con lo cual la posibilidad de un imitador de Cervantes *avant la lettre* resultaría poco menos que imposible». Incluiremos estos textos como parte del experimento pues también servirán para añadir más fuentes de autores que ya están en los corpus anteriores a fin de reforzar las fuentes para identificar su estilo, son pues una herramienta más para cotejar el estilo y la autoría. Consideramos de antemano que cualquier similitud con autores posteriores que se pudiera observar en el análisis cuantitativo podría darse por influencias cervantinas en los artífices de estas obras y no al revés. Esto se debe a que creemos que es poco probable que alguno de los presentes de este subcorpus hayan sido los creadores de la TF, pues esta obra precede a sus primeras publicaciones como escritores, no obstante, debemos dejar la puerta abierta a cualquier posibilidad remota. Además, no tenemos otra opción si queremos reunir suficientes textos que se asemejen a las características más generales (extensión y género literario) de la TF. Estos textos tienen también un argumento a favor para su inclusión que es su disponibilidad en formato txt. a través del *Github* de Fradejas Rueda²⁸. Por otra parte, existe un problema y es que no disponemos de todos los metadatos acerca del origen de la edición de

²⁸ <https://github.com/7PartidasDigital>

los textos más allá de la explicación de su propietario²⁹ por este motivo trataremos de contrastar los textos y completar estos datos marcando en la tabla (nuevamente con fondo rojo) aquellas partes en las que no podamos encontrar fuentes que nos permitan comparar los textos. (ver anexo: Corpus 3 – Otras novelas del Siglo de Oro. Tabla 3)

Es necesario señalar que los autores que están representados por una sola obra son problemáticos para la clasificación utilizando *stylo* pues el algoritmo va a tratar de encontrar los textos que presentan una mayor similitud. Al no existir más ejemplos del mismo autor, es probable que no se puedan observar las características estadísticas de forma clara y el texto sea situado en algún lugar erróneo dando prioridad a otras cuestiones como el género, la extensión o la temática y genere de este modo irregularidades. Por este motivo y para aplacar el obstáculo que supone la diferente extensión de los textos, vamos a fragmentar los documentos del corpus igualando su longitud entorno a las 6000-7000 palabras, lo cual también nos posibilita multiplicar las muestras de un determinado autor e incorporarlo incluso disponiendo de un solo texto de base.

Estas muestras que al igual que algunos de las anteriores presentan diferencias evidentes con la TF, podrían ser útiles a la hora de configurar los parámetros de la herramienta donde servirán quizás para tratar de anular algunas de las señales que se desvelan tras las pruebas estilométricas que como resumen Tello y Soler (2019: 237) citando algunos ejemplos de trabajos al respecto, son: el género literario (Kestemont, 2011), la época de composición (Jockers, 2014) y otras.

Tras este trabajo de compilación, hemos realizado varios (sub)corpus: un primer corpus de la prosa narrativa breve contemporánea a Cervantes, muy reducido; un segundo corpus con la prosa narrativa extensa también contemporánea a Cervantes que presenta un número de palabras por texto muy superior al de las *Novelas Ejemplares*; y un tercer corpus de obras del Siglo de

²⁹ «Los textos relacionados con M.^a de Zayas y Alonso Castillo Solórzano proceden de los ficheros de Alejandro García-Reydi (USAL), los del entorno del Quijote de Avellaneda los he cosechado en la Cervantes Virtual y algunos me los ha pasado Javier Blasco Pascual (UVA)». Fradejas Rueda en: <https://github.com/7PartidasDigital/NovelaBarroca>

Oro de autores que suponemos menos susceptibles de guardar relación con la TF.

Al corpus añadiríamos por último la versión de la obra dubitada. Encontramos diferentes ediciones en la BVMC por lo que surge la cuestión en torno a cuál de ellas seleccionar. Jose Manuel Lucía Megías (2018) hace un comentario que nos parece relevante con respecto a las ediciones:

«La tía fingida, [...] a pesar de los esfuerzos editoriales de los últimos años, aún carece de la edición crítica que dé cuenta de la complejidad de sus materiales textuales y el hecho de contar con testimonios de muy diversa naturaleza — copias antiguas y modernas—que dan cuenta de dos redacciones de la obra.» (Lucía Megías, 2018: 346)

De entre los testimonios que disponemos seleccionaremos la versión de Porras de la Cámara realizada por Franceson y Wolff (Berlín, 1818) en la edición realizada por Florencio Sevilla Arrollo, pues la edición berlinesa es en palabras de Lucía Megías (2018: 344-345) *«muy correcta y fiel a su testimonio base, en que solo se aleja de los tres grandes grupos de cambios lingüísticos y ortográficos»* que como indica y presenta profusamente en el apéndice son: cambio de grafías, cambio de mayúsculas y minúsculas y cambios en la puntuación. Para nuestro análisis la modernización de las grafías es un aspecto positivo que hemos buscado también en el resto de textos para poder disponer de cierta homogeneidad. Los cambios en la puntuación no son relevantes pues no son un elemento fiable en los textos de esta época por lo que no serán tenidos en cuenta en el análisis cuantitativo.

Autor	Título	Fecha de publicación	Edición	Versión digital
Desconocido ¿Cervantes?	<i>La Tía fingida</i>	16??	Novela de la tía fingida [versión Porras de la Cámara por Francenson-Wolf] / obra atribuida a Miguel de Cervantes; edición de Florencio Sevilla Arroyo (en formato HTML)	http://www.cervantesvirtual.com/nd/ark:/59851/bmc1z4h1

Tabla 2: autor, título, fecha, edición y versión digital de La tía fingida.

No se nos escapa el hecho de que el corpus que hemos recopilado para este trabajo parte de los datos que nos da el CORDE. Como ya comentamos, nos basamos en aspectos sincrónicos seleccionando todas aquellas obras

disponibles que correspondieran con la categoría de prosa narrativa. Sin embargo, el CORDE tiene en cuenta aspectos lingüísticos no literarios y por lo tanto da como resultado una serie de textos que corresponden con diferentes géneros literarios. Es por eso que es necesario aclarar ciertos aspectos pues el corpus está formado por novelas del Siglo de Oro y aunque en base a los resultados del buscador del CORDE hay una cierta homogeneidad en realidad desde un punto de vista literario nos encontramos con un corpus más bien heterogéneo. Profundizaremos en las diferencias dentro del apartado dedicado a la literatura tras haber llevado a cabo los experimentos estilométricos con *stylo*.

5.4 Relevancia de «La Tía Fingida»

La Tía Fingida es sin lugar a dudas un texto polémico y particularmente problemático. Prueba de ello son los numerosos trabajos que han tratado de descifrar su autoría o que han plasmado sus opiniones al respecto. En las siguientes líneas mostraremos el problema desde su origen, veremos resumidamente algunos de los estudios que se han realizado y justificaremos el interés que tiene esta obra para la estilometría computacional.

Uno de los primeros escollos que presenta la TF es el hecho de que se trate de una obra excepcional si la comparamos dentro del corpus cervantino. Se trata de un relato corto comparable quizás a las Novelas Ejemplares como inducen a pensar las ediciones que la incluyen junto con éstas, pero diferente por su longitud.

El origen del dilema sobre su autoría tiene su primer episodio ya en el contexto de su aparición. De la TF existen dos versiones: una en un códice de la Biblioteca Colombina en Sevilla y por otra parte varias copias de un manuscrito que ya presentamos en el apartado anterior, denominado «Porras» por haber sido Francisco Porrás de la Cámara en el S.XVII (ca.1606 según Aylward, 1999: 40) el compilador de los textos que lo componen. Este manuscrito desapareció en su totalidad en torno a 1810-1823 por lo que no tenemos acceso a tan valioso documento. La primera copia del manuscrito que conservamos fue realizada en

1788 por Isidoro Bosarte. La versión de la Biblioteca Colombina³⁰ es completamente diferente manteniendo la misma historia. Vemos pues que existen motivos externos e internos relativos a la obra que dificultan su atribución y que a la vez han sido la causa de un debate al respecto con posturas a favor y en contra de la autoridad cervantina. Aylward (1999) resume este debate en su artículo «*Significant Disparities in the Text of La tía fingida vis-à-vis Cervantes El casamiento engañoso*». Por su parte Eisenberg (1990: 481-483) le parecía sorprendente que tras dos siglos la discusión no haya alcanzado un quorum al respecto. Este autor considera que se ha dado gran importancia a las publicaciones en contra de la legitimidad de esta obra por parte del cervantismo oficial (p.ej. Manuel Criado de Val) mientras que se ignoró a los que, como Bonilla, Gallardo y Astrana Marín defendían la atribución a Cervantes. Eisenberg (1990) concluye que la causa del enfrentamiento podría ser el hecho de que la TF trate el tema sexual de forma franca.

Para hacernos una idea de los posicionamientos en cuanto al debate que van desde el rechazo, pasando por la duda hasta la certeza y también de sus fundamentos que son de variada índole resumiremos aquí algunos de ellos sin ser exhaustivos dado que esto conllevaría un trabajo específico al respecto:

A favor de la autoría cervantina

- Márquez Villanueva (1991) → Define la obra como «llena de pinceladas narrativas perfectamente ajustadas a [...] un firme encuadre semiológico» (124). Pero también admite los argumentos de Icaza en cuanto a su semejanza con los *Ragionamenti* de Aretino. Considera posible la autoría cervantina pero también afirma que podría ser de «un temprano imitador consciente y sistemático» que podría ser Luis Vélez de Guevara o Juan de Sali.
- Gallardo (1832) → Ve en la TF y en el corpus cervantino semejanzas estilísticas y temáticas.
- Astrana Marín (1944 y 1953) → Reconoce paralelismos entre «algunas analogías de lenguaje, modismos y estilo que ofrece La tía fingida con ciertas producciones cervantinas» además de afirmar que es «novela tan excelente, de un estilo tan semejante al suyo o que no desdice de él». (1944: 40)

³⁰ Biblioteca Colombina de Sevilla, codex A²-141-4, folios 77-^a a 83

- Madrigal (2003) → Apoya la autoría cervantina contradiciendo la tesis de Icaza. Utiliza para ello el método de Bosarte para llevar a cabo la comparación de ciertas expresiones y pasajes que Icaza conectaba con los *Ragionamenti* pero que también se encuentran en otros textos cervantinos.
- López-Vázquez (2013) → Utiliza métodos estadísticos para buscar coincidencias en los textos que aporta el CORDE para los textos publicados entre 1600 y 1620 de cinco autores (Cervantes, Quevedo, Espinel, López Úbeda y Salas Barbadillo). Encuentra ciertas expresiones coincidentes. El mismo autor en otro trabajo (2018) compara 43 índices utilizando la estilometría y llega a la conclusión de que la TF es de Cervantes
- Corral (1981) → Encuentra semejanzas en torno al concepto que tiene el autor de 'honestidad' en las heroínas de las *Novelas Ejemplares* y en el personaje de Esperanza en la TF.
- Apraiz (1906) y Bonilla (1917) → Observan coincidencias lingüísticas entre la obra de Cervantes y la TF.

En contra de la autoría cervantina:

- Francisco A. Icaza (1916) → Investigación léxica de arcaísmos en palabras y expresiones, empleo de determinados regionalismos o coloquialismos que no son clarificados por el autor de la TF. Semejanzas con los *Ragionamenti* de Aretino (1535)
- Criado de Val (1953) → Compara la frecuencia en el uso de ciertas formas del subjuntivo y el condicional en el conjunto de las *Novelas Ejemplares* y en la TF
- Inamoto (2004) → Ve diferencias en el acento rítmico del soneto intercalado en la TF considerando que la TF o el soneto insertado no es de Cervantes.
- Avalor Arce (1973) → Opina subjetivamente que Cervantes no la escribió.
- Foulché-Delbosc (1899) → Considera que la TF no es de Cervantes en base a su estilo, la configuración de las frases y el desarrollo del relato que considera lejano al complutense.

No son estos todos los trabajos al respecto, ni probablemente lo serán en un futuro si incluyéramos todos los que hoy se han publicado. Lo que si evidencia este listado es la existencia de un auténtico muestrario de argumentos con más o menos fundamento. Es precisamente el hecho de que la polémica acerca de la autoría de la TF no haya concluido ni hallado aún una respuesta satisfactoria para todos/as los/las investigadores/as nos lleva a ver en ella un particular interés

como objeto de análisis de la estilometría computacional. Recordemos aquí que Holmes (1994) consideraba necesario tener en cuenta la escolástica tradicional antes de abordar un trabajo de atribución desde la estilometría. En el caso de la TF podemos observar que existen suficientes intentos de dar respuesta a la autoría de esta obra que nos permiten poder justificar el aporte de una nueva perspectiva de análisis. Si bien, entendemos los límites que las circunstancias imponen de antemano al no disponer de las premisas que en cuanto a la fiabilidad del texto vimos en Juola (2008) quien recordemos, afirmaba que solamente se deberían utilizar textos cuyas características fueran responsabilidad directa del autor. No obstante, consideramos que es quizás posible aportar nuevas hipótesis, dado que no existen trabajos que analicen la obra desde una perspectiva estilométrica digital³¹ ³², pero sí se dan por válidos ciertos análisis pseudoestadísticos, hermenéuticos, etc. que se sirven en definitiva de los mismos textos que utilizaremos.

6 *Stylo* de R: presentación de la herramienta.

En este punto ya tenemos listo un corpus considerable de textos en prosa del Siglo de Oro. Así pues, el siguiente paso antes de poner en marcha el análisis es presentar la herramienta que vamos a utilizar para llevarlo a cabo para así poder también comprender los resultados que arrojará y su índice de fiabilidad.

R es un entorno y lenguaje de programación que está enfocado en el análisis estadístico pero que como apunta Fradejas-Rueda (2019) y muestran los

³¹ Existe de hecho un trabajo que podría considerarse similar, pues aborda este problema desde la estadística: López, F. (2011). Donde se muestran algunos resultados de atribución de autor en torno a la obra cervantina. *Revista Colombiana de Estadística*, 34(1), 15-37. Este trabajo afronta la autoría desde la estadística utilizando una gran muestra de métodos que requieren amplios conocimientos matemáticos. Nuestro método, difiere al utilizar *stylo* de R como vía de acceso a la estadística desde el ámbito de la lingüística y la literatura. Además, utilizamos un corpus más extenso, se incluyen más autores, se utiliza el Cosine Delta como método de clasificación y sobre todo, no obviamos la importancia del análisis literario para este fin.

³² Otro trabajo que aborda la TF desde la estadística es el ya citado López-Vázquez (2013). Este autor utiliza palabras como discriminantes para acotar la posible autoría. La metodología es evidentemente diferente y compartimos la opinión de Rissler-Pipka (2016: 50) de que la estilometría (con *stylo*) parece «más independiente de presuposiciones que la comparación de un vocabulario específico [...]». No consideramos pues que comparar algunas palabras raras de un autor sirva para establecer su distintivo de manera convincente.

numerosos estudios que se han servido de este entorno tiene un gran potencial para su uso en los estudios filológicos. Como ejemplo de los trabajos realizados utilizando R, Fradejas Rueda (2019) hace referencia a estudios del ámbito de la lingüística (Gries 2013, Gries 2017, Desagulier 2017, Levshina 2015, Winter 2019) y de la literatura (Jockers 2014). Esta lista, se puede ampliar inmensamente, prueba de otros trabajos son los ya comentados en los apartados sobre estilometría en español y en torno a la obra de Cervantes.

Dentro de este entorno, para llevar a cabo el análisis estilométrico existen múltiples herramientas. Una de las más recientes y que ha dado buenos resultados en varias investigaciones es el paquete *Stylo* basado en R (Eder et al., 2016) mantenido y desarrollado por el grupo de estilística computacional³³. Este paquete aporta diferentes funciones para el análisis estilométrico además de una interfaz de uso sencillo y diagramas de gran calidad.

Recordemos en este punto que Juola (2018) comentaba respecto a las herramientas a utilizar para el análisis estilométrico que lo que es más importante para el usuario casual es la capacidad de seleccionar los algoritmos y las características a utilizar dinámicamente, en función del idioma, el género, el tamaño, de los documentos disponibles (Juola, 2018: 250). En este sentido *stylo*, nos parece una herramienta que cumple con todas estas necesidades. Juola (2018) también añade que un «buen sistema» debería respaldar su precisión con un historial; un conjunto bien validado de resultados en una variedad de áreas y problemas para ayudar a definir tanto los niveles esperados de precisión como demarcar las áreas donde se podría esperar que los niveles bajen. Con respecto a la precisión nos remitiremos al trabajo de Blasco Pascual y Ruiz Urbón (2009) que realizan una evaluación y cuantificación de algunas técnicas de atribución de autoría en textos españoles. En su trabajo, muestran que el *perfil de palabra simple*³⁴ es el más efectivo para los textos del Siglo de Oro con un 70% de precisión en un corpus de 10 autores que se eleva al 90% entre dos autores. Además, dan cuenta de que el perfil más efectivo es el de *marca de puntuación*

³³ <https://computationalstylistics.github.io/resources/>

³⁴ El *perfil de palabra simple* contrasta la frecuencia relativa las palabras más frecuentes en los textos dubitados (que resultan ser las denominadas palabras funcionales: de, la, que, el, en, y, a, los...) con las frecuencias que en cada uno de los autores tienen esas palabras

*simple*³⁵ que presenta un 85% de fiabilidad en un corpus de 10 autores sin embargo este método no es aplicable al Siglo de Oro dado que como ya comentamos en el apartado sobre la problemática de los textos de esta época, los textos, y en particular su puntuación no es suficientemente fiable³⁶. De este modo, de los diferentes métodos posibles para el análisis solo queda uno efectivo: «Desgraciadamente del *top ten* obtenido para el español sólo resulta operativo el *perfil de palabra simple*» (Blasco Pascual y Ruiz Urbón, 2009: 44).

Así pues, llegados a este punto queda explicado el motivo del uso de *stylo* su objetivo: medir una de las realidades textuales que es la frecuencia de palabras y dentro de estas el *perfil de palabra simple* que parece ser teóricamente el más adecuado y efectivo para este corpus. Con el corpus compilado y la herramienta presentada es hora de comenzar el experimento.

7 Pruebas estilométricas

El primer paso que vamos a realizar son una serie de experimentos con el fin de determinar si nuestro corpus es organizado de manera lógica (por autores) por parte de la herramienta. Por el momento dejaremos la obra dubitada a un lado. Vamos a realizar un *Cluster Analysis* de cada uno de los subcorpus para crear un corpus final con el que trabajaremos.

En el corpus podríamos aplicar las consideraciones de Maciej Eder (Eder, 2015) quien realizó diversos análisis con diferentes configuraciones en un corpus

³⁵ Este perfil es calculado mediante la división de la frecuencia de una serie de signos de puntuación entre el número total de caracteres contenidos en el texto.

³⁶Pascual y Urbón (2009) comentan en sus conclusiones que hay dos razones que suponen un problema para la atribución de autoría y que tienen que ver una con la fase de creación y otra con la fase de edición de los textos. Son las siguientes:

a) *Los hábitos de escritura en los Siglos de Oro son poco conocidos, pero pensamos que la existencia de varias manos tras la conformación de una obra literaria era más frecuente de lo esperado (bien por verdadera colaboración entre autores, bien por actuar uno como corrector de otro).*

b) *Los correctores, componedores y cajistas de imprenta podrían ser en alta medida los responsables de muchas de las marcas que nuestros procedimientos de medición actuales contemplan. Las lecciones accidentales (ortografía y puntuación) conservadas en los diversos testimonios en que nos han llegado no son las originarias de los escritores, y, por tanto, imposibilitan que las variables basadas en mediciones relacionadas con los grafemas o las marcas de puntuación puedan ser aplicadas o, al menos, que deba concedérseles el mismo grado de confianza.* (Blasco Pascual y Ruiz Urbón, 2009:44)

de textos de diferentes lenguas con el fin de tratar de determinar a partir de qué extensión de texto la señal autorial presenta mayor fiabilidad garantizándose así los resultados del análisis estilométrico. Sus experimentos mostraron que la extensión varía según la lengua y el género literario pero este límite se puede establecer en las 5000 palabras. Lorenzo (2019: 192) añade que tiene sentido que al analizar la narrativa el estilo tienda a estar más diseminado a lo largo del texto que por ejemplo en la poesía. En nuestro caso analizamos novela breve, es por lo tanto narrativa, pero es un género que entendemos que concentra el estilo. También utilizamos algunas novelas (extensas), pero las hemos dividido en partes similares. Por este motivo y porque hemos fragmentado todo el corpus vamos a utilizar tan solo el rango de las 500MFW.

7.1 Análisis de clúster para ajustar el corpus

Corpus 1: Prosa narrativa breve

Tras varias pruebas con la función `stylo` (), encontramos que la organización de los textos es eficiente en la siguiente configuración: análisis de las 100 a las 500 MFW utilizando *Cosine Delta*. También se observó en los primeros experimentos que existían algunas irregularidades que podrían ser debidas a la distinta extensión de los textos. Esta heterogeneidad de la de los textos está presente con mayor o menor medida en todo el corpus como se puede observar en el anexo (corpus 1, 2, 3). Es por ese motivo por el cual hemos decidido fragmentar las obras en partes de aproximadamente 6000 palabras con el fin de homogeneizar lo más posible nuestros experimentos. La organización de los textos con 500 MFW se puede ver en la siguiente página. La herramienta reconoce sin problemas la obra de los autores y los organiza como se esperaba.

Si observamos el eje vertical, llama la atención que la herramienta reconoce dos Cervantes diferentes y que algunos de los textos no están organizados con las partes que les corresponde. Esto no es ningún impedimento, simplemente muestra una cierta heterogeneidad en este corpus cervantino y pone quizás de manifiesto la relación entre ciertas partes de diferentes obras (por ejemplo, la cuarta parte de *La gitanilla* y *La fuerza de la sangre*). Vemos

pues que en comparación a otras pruebas hechas utilizando los textos completos, la división parece ayudar a que los textos se organicen mejor. Así mismo esto nos permitirá más adelante incluir algunos autores de los que solo disponemos de un texto que de otra manera sería imposible por no ser suficientemente representativo del «estilo» ya que el sistema trataría de asociarlo a un texto cercano y al no disponer de él aparecería junto a la obra de algún otro autor. El corpus siguiente va a incluir por lo tanto todos los textos de los corpus 1 y del corpus 2.

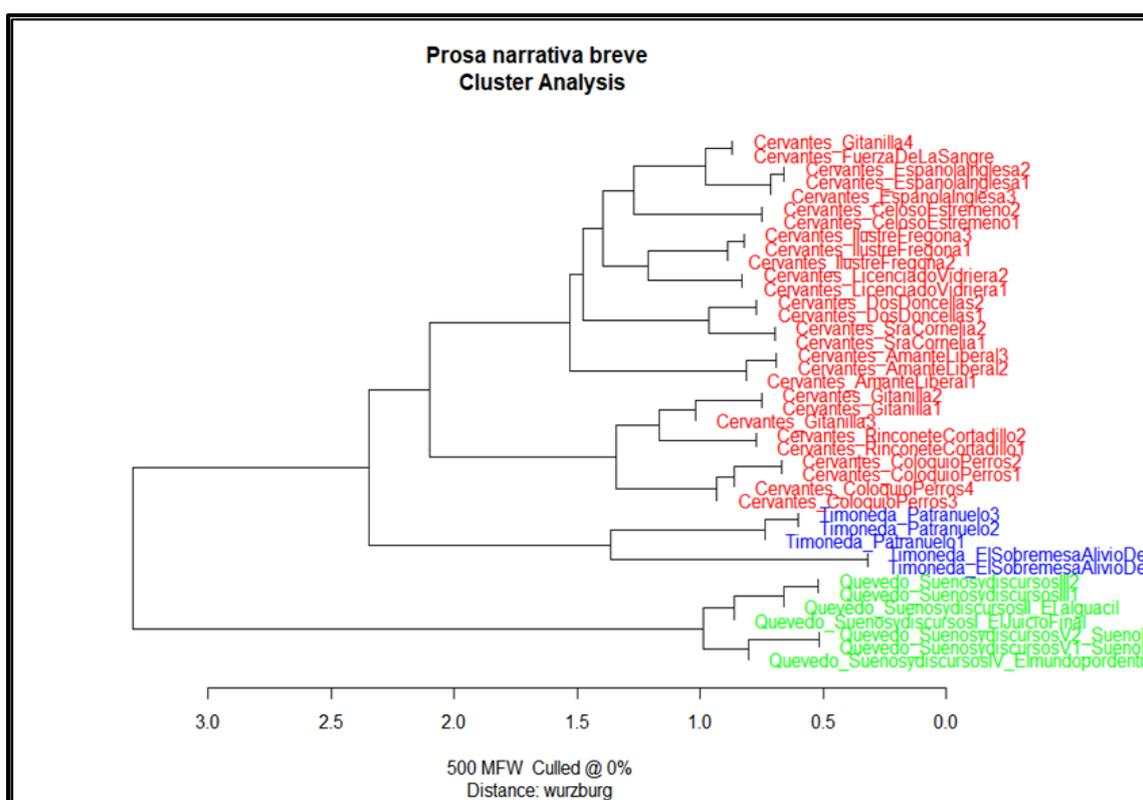


Imagen 1: prosa narrativa breve, análisis de clúster.

Corpus 2: Prosa narrativa extensa

Continuamos el experimento manteniendo la misma configuración de 500MFW y Cosine Delta. Esta vez con nuestro segundo subcorpus que contiene las obras que corresponden a la prosa narrativa extensa. Vemos que en los resultados la organización también es la esperada. Solamente sorprende la parte diecisiete de *La Galatea* que aparece junto a la séptima parte *Los siete libros de Diana* de Montemayor. Igualmente, las tres primeras partes del *Guzmán de Alfarache* de Alemán aparecen separadas del resto de la obra en un subcluster

en el que encontramos los textos de Salas Barbadillo. Aunque este dato podría ser interesante para el análisis literario, no es el tema que nos ocupa. Nuestro objetivo es conseguir un entorno libre de interferencias en el que podamos incorporar un texto dubitado. Por este motivo y considerando que las obras están suficientemente representadas, vamos a prescindir de esas partes de los textos. Al repetir la prueba sin estas partes, la organización de las obras por autores es satisfactoria. En la imagen se puede ver el dendograma con los primeros resultados:

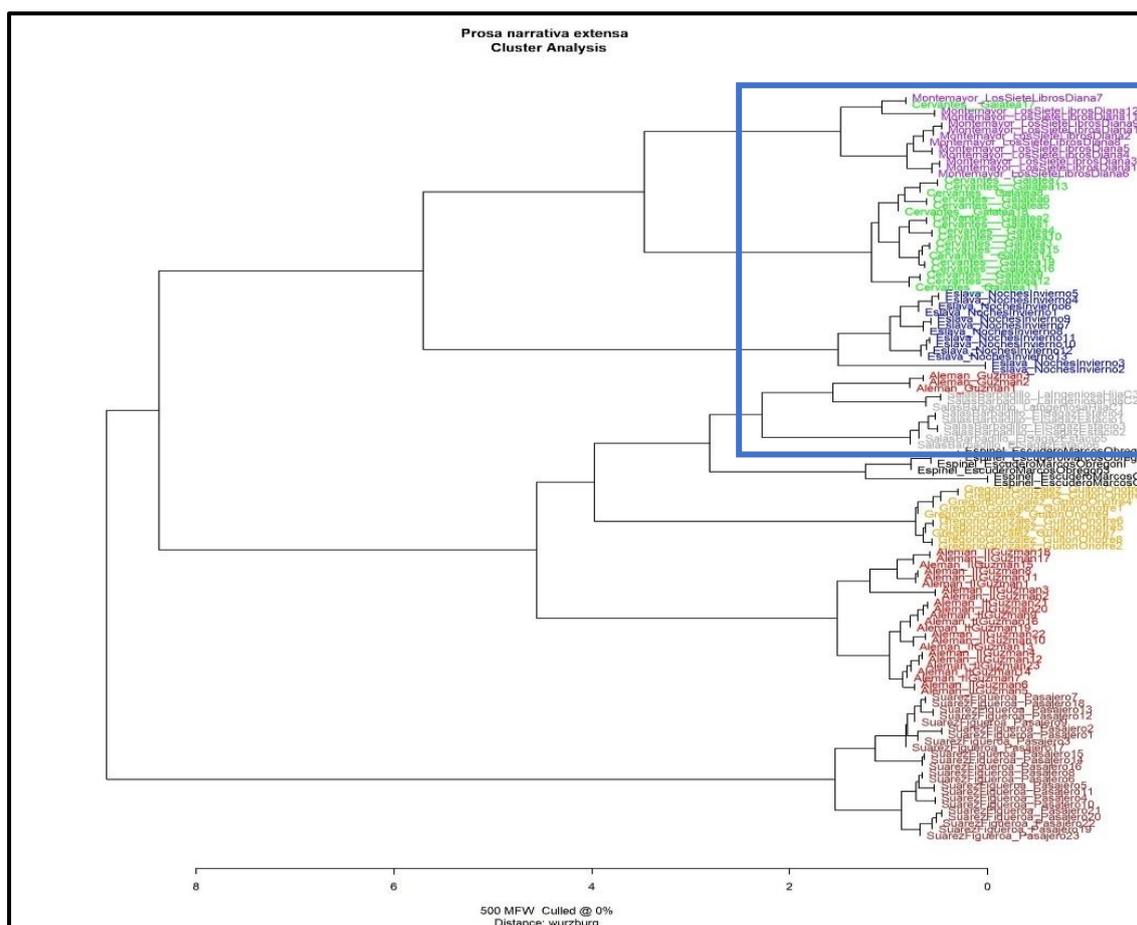


Imagen 2: prosa narrativa extensa, análisis de clúster, 500 MFW, *Cosine Delta*.

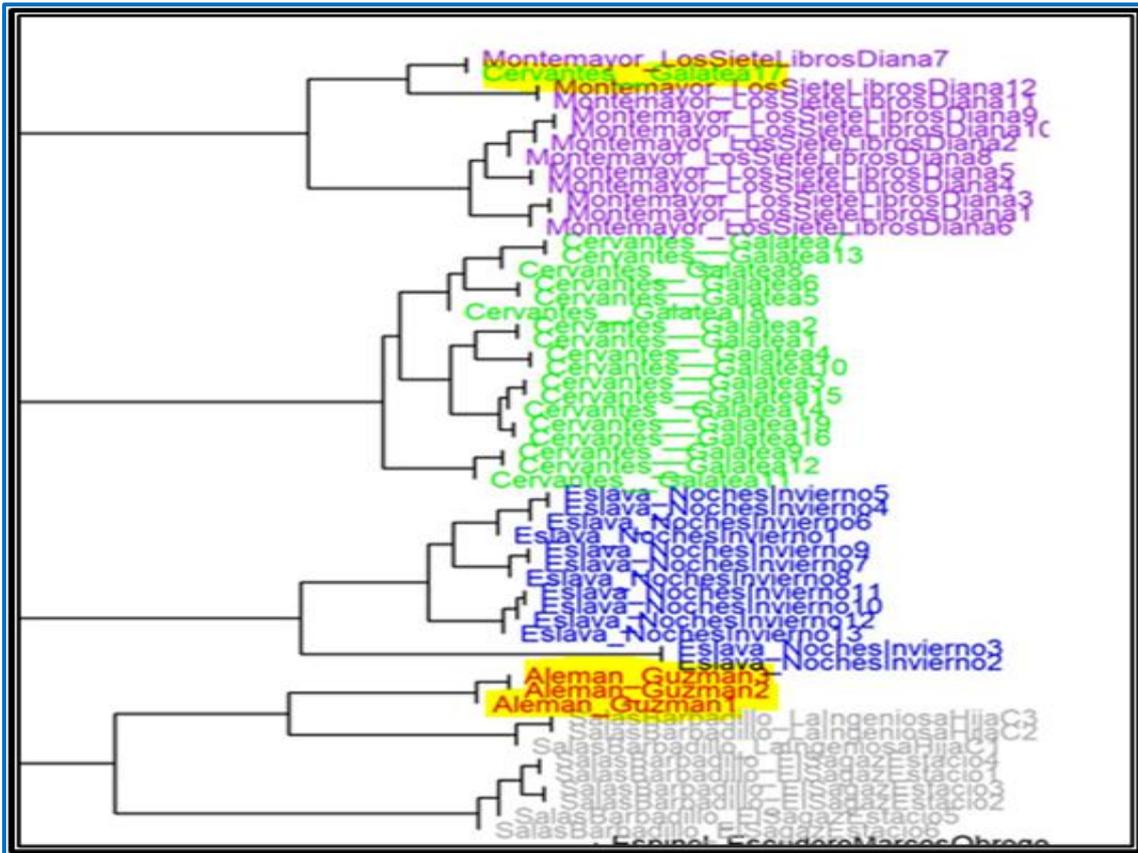


Imagen 3: detalle del dendrograma (Imagen 4), partes del texto problemáticas.

Corpus 1 y 2: Prosa narrativa breve + extensa

Llegados a este punto, vemos que una vez excluidos los fragmentos de los textos que causaban irregularidades, la herramienta es capaz de organizar las obras por autores en los subcorpus por separado por lo que ya podemos proceder a juntar ambos subcorpus para ver como se comportan en conjunto. La imagen de la página siguiente muestra que la organización de los autores y sus obras en *clusters* es correcta utilizando las 500MFW y Cosine Delta. El amplio número de textos hace que la imagen del dendrograma se vea con dificultad por lo que incluimos imágenes con los detalles de las diferentes partes:

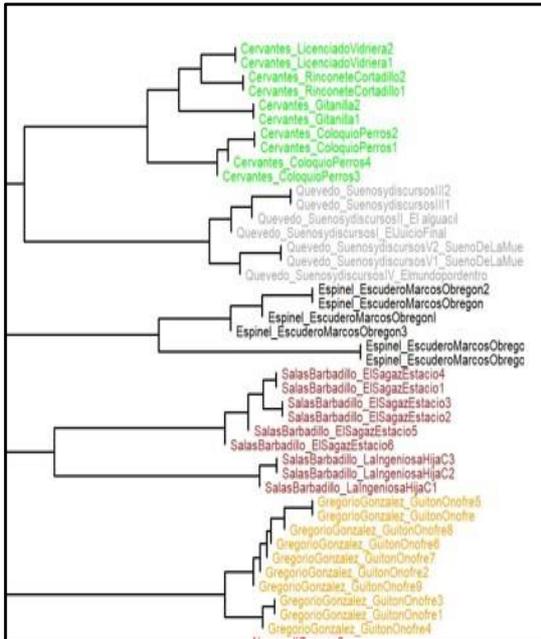


Imagen 5: sección 1 del dendograma (Imagen 4)

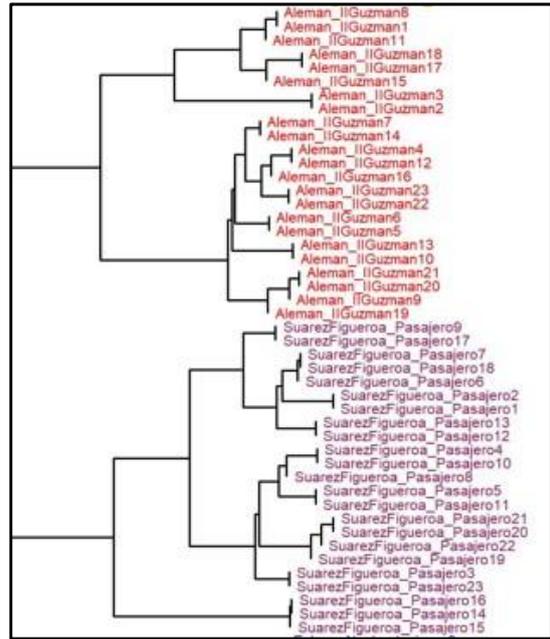


Imagen 6: sección 2 del dendograma (Imagen 4)

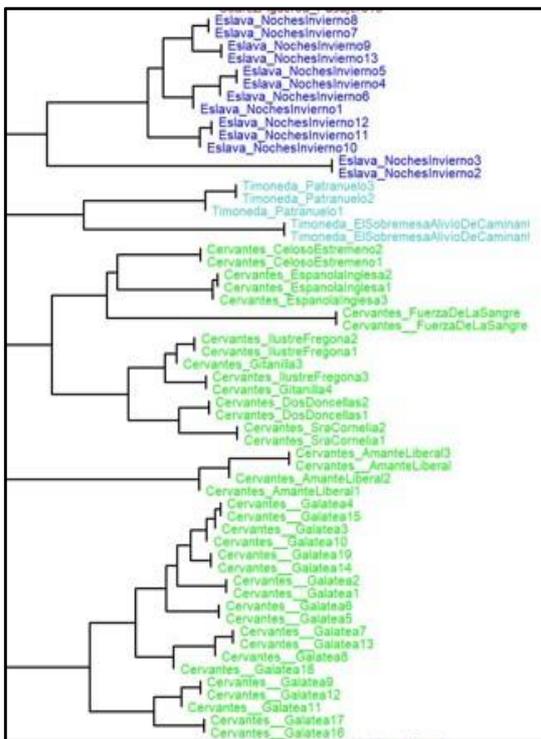


Imagen 7: sección 3 del dendograma (Imagen 4)

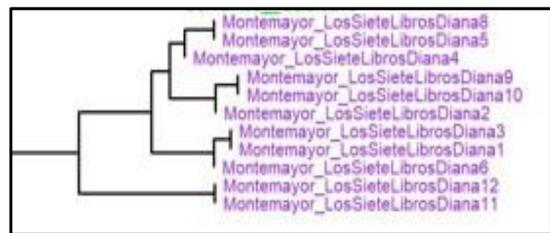


Imagen 8: sección 4 del dendograma (Imagen 4)

Algunos resultados dignos de mención son la clasificación de dos de las partes de *El escudero Marcos Obregón* de Espinel (imagen 1), en *Noches de Invierno de Eslava* (imagen 3) o de *La fuerza de la sangre* de Cervantes (imagen 3). En los dos primeros casos entendemos que son partes de los textos que se

diferencian del resto por múltiples posibles motivos. Podrían ser una muestra de la señal del género literario o temática, por ejemplo. Esto no obstante no va a ocupar nuestro trabajo, aunque resulta interesante y podría analizarse en un trabajo específico. Lo que es para nosotros fundamental es que dichas partes fueron agrupadas correctamente en base a sus autores.

Ahora vamos a utilizar el árbol de consenso (BCT) que reúne los resultados de los dendogramas que van desde las 100 MFW hasta las 500 MFW con el incremento especificado de 100MFW vemos que la clasificación de autores y obras es consistente en el sentido de que no se mezclan que como ya comentamos es lo que nos interesa en este punto. Sin embargo, también se puede observar que las ramificaciones correspondientes a Cervantes (verde claro) que salen del eje central corresponden a cada uno de los textos lo cual sugiere una gran heterogeneidad en su obra. Se podría decir que el programa reconoce a doce Cervantes. Por otra parte, la aparición de *La Galatea* a la izquierda separada del resto de su obra responde a la diferencia del género y su cercanía a la novela *Los siete libros de la Diana* de Montemayor lo respalda, pues esta novela pertenece también al género pastoril.

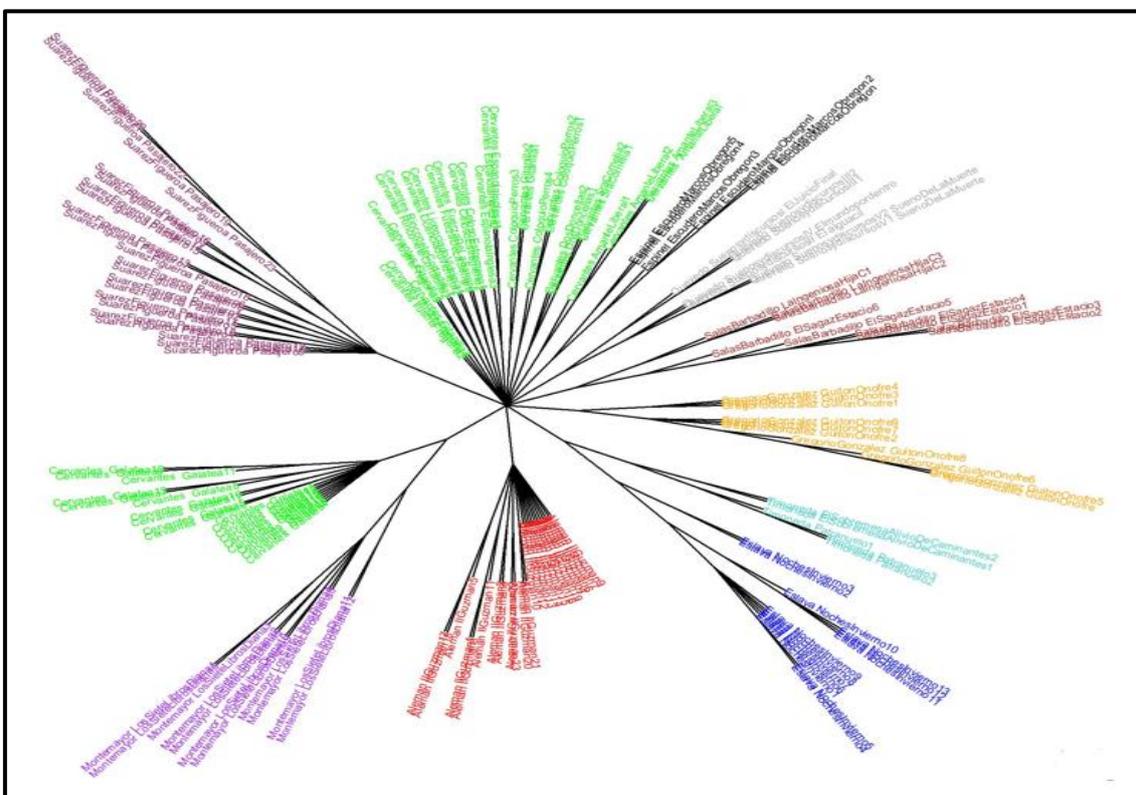


Imagen 9: prosa narrativa breve y extensa, *Bootstrap Consensus Tree*, 500 MFW, Cosine Delta.

Corpus 3: Otros textos en prosa

El último subcorpus que queda por organizar es el correspondiente a aquellos textos que incluyen a los autores que en un principio consideramos menos susceptibles de ser los creadores de la TF por una cuestión de cronología. Aquí realizamos nuevamente un análisis de clúster. Los resultados son los esperados utilizando nuevamente las 500 MFW con *Cosine Delta*:

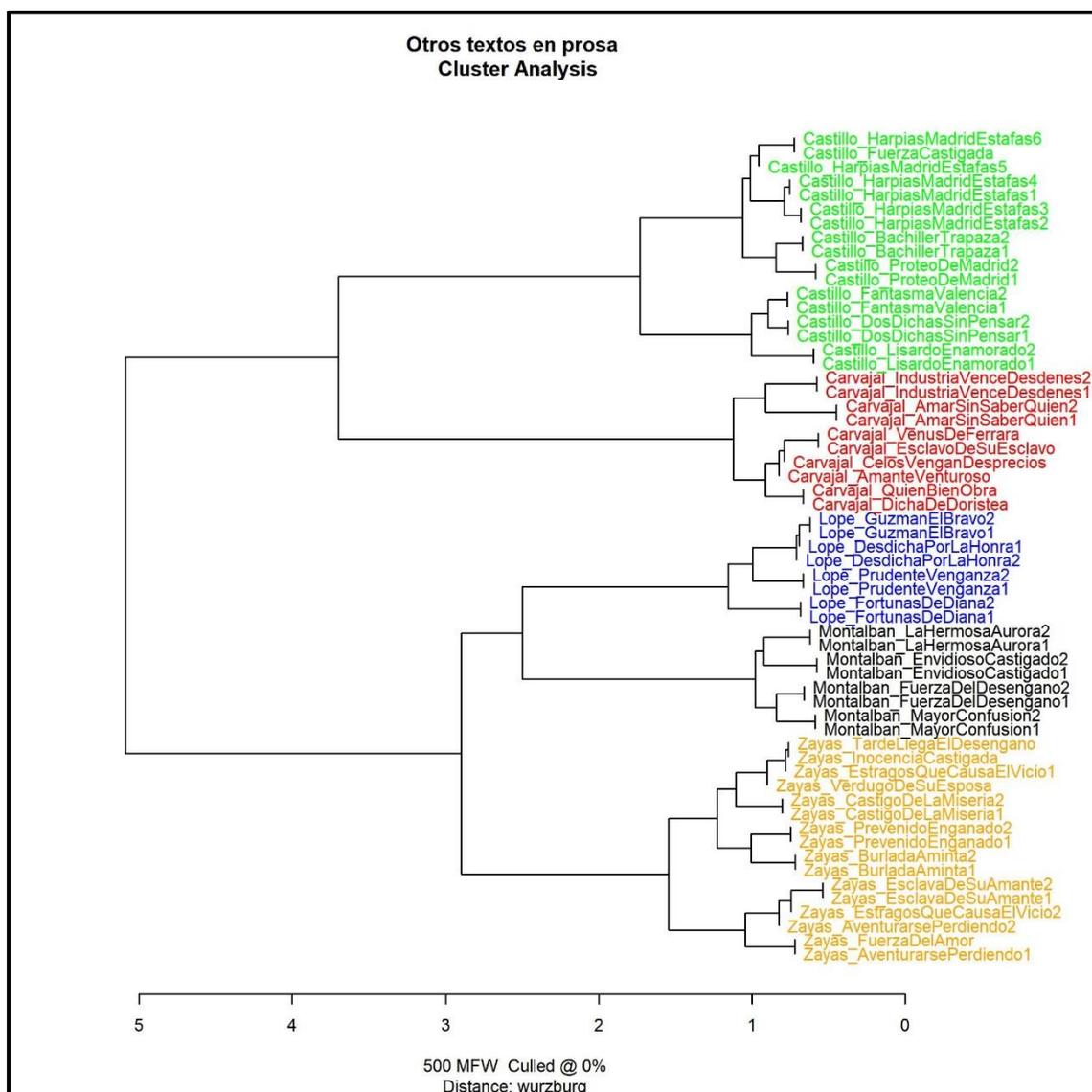


Imagen 10: otros textos en prosa, análisis de clúster, 500MFW, Cosine Delta.

En este caso, no hay grandes sorpresas pues todos los textos son organizados en clústeres en los que se incluye un solo autor.

Corpus Completo

Tras esta progresiva comprobación de la organización de nuestro subcorpus por parte de *stylo* llegamos al que hemos llamado «corpus completo». En él incluimos todos los textos y autores de las pruebas anteriores para tratar de comprobar si la organización sigue siendo la esperada. Los resultados se vuelven más difícilmente visibles (imagen 11). Sí se puede apreciar por los colores que la organización de los textos por autores es correcta. Si bien, sigue existiendo como es lógico bastante heterogeneidad en el corpus cervantino. Esto da lugar a que se cuelen dentro de un mismo *cluster* aunque separado en los subclusters *El patrañuelo* y *El sobremesa o alivio de caminantes* de Timoneda junto a las *Novelas Ejemplares*. También cercano encontramos el texto de Eslava, *Noches de Invierno*, pero este está algo más alejado como se puede observar en el eje horizontal donde se muestra que la relación aparece con respecto al *cluster* en el que están las *Novelas Ejemplares* y el texto de Eslava. Por otra parte, *La Galatea* se desmarca también y aparece en un *cluster* en el que también está la obra de Montemayor. Ya comentamos a que se puede deber esto: se trata de dos novelas (prosa extensa) que pertenecen al género pastoril.

En este punto podemos quizás plantearnos retirar la obra de Timoneda. *El sobremesa y alivio de caminantes* es una compilación de breves cuentos algunos de los cuales llegan a tal brevedad que se asemejan a proverbios. La temática es muy variada al igual que la forma y el origen (algunos de ellos son traducciones del Decamerón). Por otra parte, *El patrañuelo* también es una colección de 22 novelas que el autor denomina «patrañas». Estas son de origen diverso no siendo ninguna creación propia del autor. La fuente principal son ciertas narraciones cortas italianas de las cuales se hace una adaptación (de Boccaccio, Bandello, Ariosto, etc.). Por lo tanto, vemos que tanto las primeras como las segundas pueden ser problemáticas para determinar el estilo de un autor también desde un punto de vista literario, sobre todo, porque las fuentes de las que proceden no son originales. Dicho esto, consideramos justificado eliminar estas obras del corpus por los motivos enunciados y porque generan inconsistencias que consideramos problemáticas cuando nuestro fin es introducir además un texto dubitado en la ecuación. Al hacer esto los resultados son idénticos, pero desaparece Timoneda del dendograma:

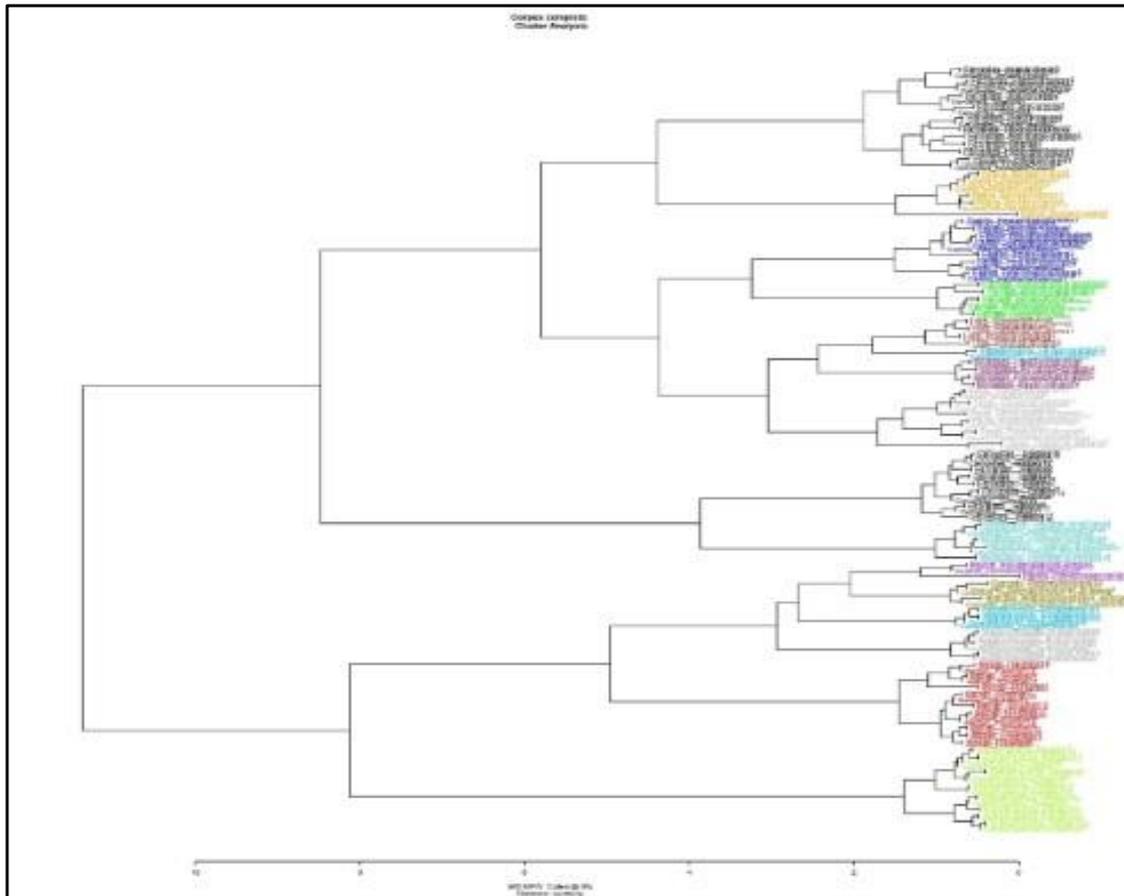


Imagen 11: corpus completo, análisis de clúster, 500MFW, *Cosine Delta*

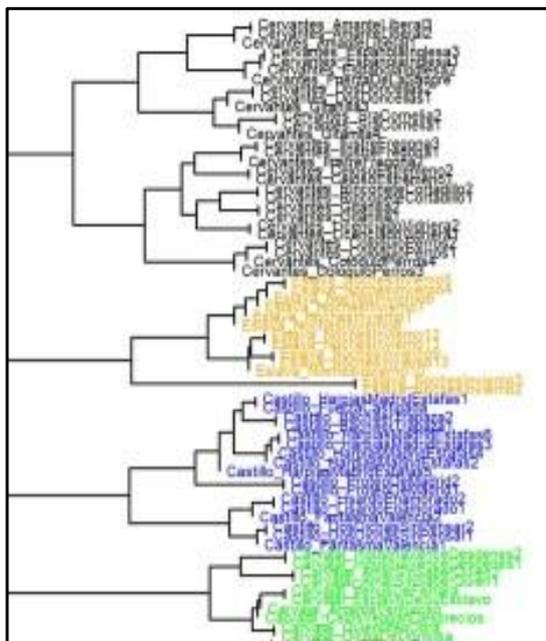


Imagen 12: detalle 1 del dendograma (img. 11)

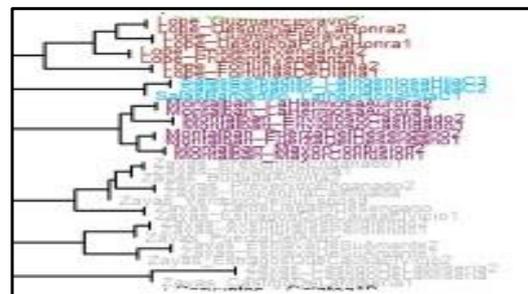


Imagen 13: detalle 2 del dendograma (img. 11)

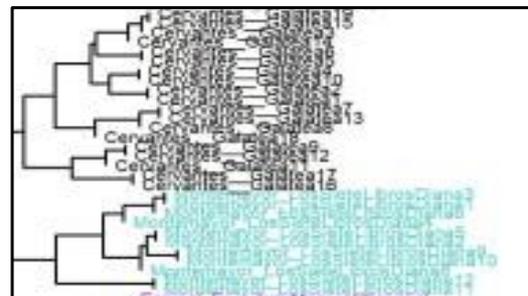


Imagen 14: detalle 3 del dendograma (img. 11)

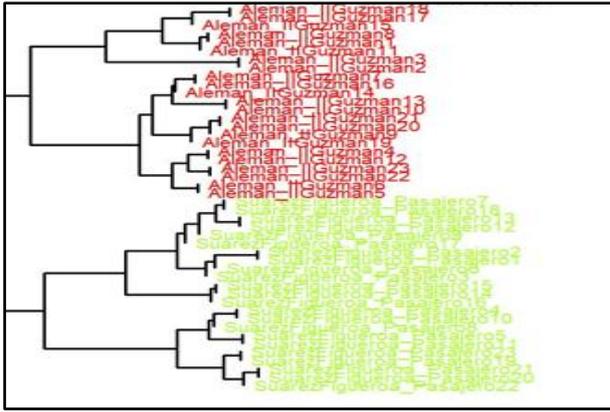


Imagen 15: detalle 4 del dendrograma (img. 11)

En este punto ya hemos puesto a prueba la herramienta en su función más básica y hemos comprobado que la organización de los textos en base a la figura autorial funciona correctamente con los diferentes subcorpus y también en el conjunto de todos ellos eliminando algunas partes y algunos textos completos. En el siguiente apartado vamos a proceder a la muy importante tarea de comprobar la efectividad de estos resultados antes de plantearnos introducir nuestro texto dubitado.

7.2 Precisión a través de la validación cruzada y clasificación

Para llevar a cabo la organización de los textos según su autor hemos utilizado la configuración de las 500MFW y la distancia *Cosine Delta*. Sin embargo, aún nos queda responder a una pregunta de gran relevancia ¿con qué precisión se están organizando los textos? ¿tenemos la misma con todos los autores y textos contenidos en el último corpus que hemos compilado? En este sentido, Eder añadió una revisión a sus afirmaciones en un nuevo estudio realizado en 2017 en el que observó que el límite fiable de las MFW puede variar según el contexto que establece el corpus utilizado en el análisis. En base a esto, Eder (2017b) llegó a la conclusión de que el factor más determinante de un texto es la fuerza de la señal autorial: a algunos les basta un número reducido de palabras para poder alcanzar una gran fiabilidad mientras que otros continúan siendo dudosos pese a una mayor extensión. Así, con una fuerte señal autorial nos encontramos que las 2000 MFW serían suficientes. En nuestro caso como ya apuntamos utilizamos aún menos (500MFW) adaptándonos a las

circunstancias que impone la brevedad y diferentes longitudes de los textos del corpus. Las observaciones llevadas a cabo por Eder tienen como consecuencia que para poder tener constancia de la fiabilidad de atribución de autoría para un autor determinado es preciso llevar a cabo un análisis supervisado cuyo fin es comprobar el grado de fiabilidad que puede esperarse de la organización de los diferentes textos por parte de *stylo* antes de utilizarlo sobre el texto dubitado (cfr. Lorenzo, 2019: 193-194). El código de este experimento puede consultarse en el anexo «código» (1).

Los resultados muestran que se reconocen quince clases en base a quince textos que se han tomado como referencia. Estos representan a los 14 autores que hemos seleccionado (Cervantes está repetido). El porcentaje general de acierto es del 100% con una desviación estándar del 0% utilizando desde las 100 hasta las 500MFW con un incremento de 100MFW. Si accedemos a los datos sobre el acierto atendiendo al número de palabras frecuentes utilizadas vemos que ya con las 100MFW este porcentaje es del 100% y se mantiene así hasta las 500MFW.

Sabiendo estos datos, utilizamos la función *classify ()* para la validación cruzada cuyo código y explicación se encuentra nuevamente en el anexo «código» (2). Los resultados de esta prueba son otra vez correctos. Los autores predichos para los textos dados son los siguientes:

Aleman	Carvajal	Castillo	Cervantes	Cervantes
1	1	1	1	1
Eslava	Espinel	GregorioGonzalez	Lope	Montalban
1	1	1	1	1
Montemayor	Quevedo	SalasBarbadillo	SuarezFigueroa	Zayas
1	1	1	1	1

Tabla 3: resultados de la validación cruzada, 500MFW, *Cosine Delta*.

Así pues, vemos que utilizando tan solo las 500 MFW el resultado de acierto general es del 100%. El sistema ha conseguido también reconocer la existencia de 15 clases correspondientes a los 14 autores tal y como era esperable. Para poder llegar a estos resultados hemos tenido que dividir los textos de aquellos autores de los cuales disponíamos de un solo ejemplo y homogeneizar las longitudes de todo el corpus. También tuvimos que prescindir

de algunos textos que causaban irregularidades e imprecisión por lo que consideramos que no nos permitirían introducir un texto dubitado en el próximo paso. Esto se debe en muchos casos a la carencia de suficientes muestras representativas que sirvan en el corpus de prueba (`primary_set`) para que el algoritmo pueda aprender las características de los autores. Allí donde solo disponemos de un texto nos vemos en la coyuntura de tener que utilizar tan solo una parte de la misma obra en cada uno de los subcorpus lo cual es una gran diferencia frente a otros autores de los que disponemos de al menos cuatro textos.

Por último, ya solo nos queda introducir el texto dubitado en este proceso. Así pues, repetimos la última prueba, pero esta vez introduciendo el texto dubitado en la carpeta `secondary_set` que es la que contiene las muestras de diferentes autores que van a ser comparadas con el `primary_set` conteniendo numerosos textos de ejemplo que le sirven al programa para «aprender» el estilo. Los resultados de las clases esperadas y predichas se pueden ver en la siguiente tabla:

Esperado («expected»)	Predicho («predicted»)
[1] "Aleman"	[1] "Aleman"
[2] "Carvajal"	[2] "Carvajal"
[3] "Castillo"	[3] "Castillo"
[4] "Cervantes"	[4] "Cervantes"
[5] "Cervantes"	[5] "Cervantes"
[6] "Desconocido"	[6] "Cervantes"
[7] "Eslava"	[7] "Eslava"
[8] "Espinel"	[8] "Espinel"
[9] "GregorioGonzalez"	[9] "GregorioGonzalez"
[10] "Lope"	[10] "Lope"
[11] "Montalban"	[11] "Montalban"
[12] "Montemayor"	[12] "Montemayor"
[13] "Quevedo"	[13] "Quevedo"
[14] "SalasBarbadillo"	[14] "SalasBarbadillo"
[15] "SuarezFiguroa"	[15] "SuarezFiguroa"
[16] "Zayas"	[16] "Zayas"

Tabla 4: resultados de la función *classify*, corpus + La TF, 500MFW, Cosine Delta

Como se puede ver en la tabla, el texto *Desconocido* que no es otro que *La tía fingida*, ha sido clasificado como perteneciente a Cervantes. Por lo tanto, el resultado que arroja la prueba de atribución de autoría con la función *classify* es que entre los autores seleccionados el texto dubitado presenta mayor similitud estilística con Cervantes y por lo tanto parece pertenecer a este autor.

7.3 Verificación de autoría

Llegados a este punto en el que hemos comprobado la efectividad y hemos visto que la TF es atribuida a Cervantes podemos plantearnos continuar con las pruebas en este caso de verificación de autoría dado que suponemos que el autor es Cervantes, pero queremos ponerlo a prueba. Esto sirve también como validación cruzada de los resultados anteriores. Para realizar este experimento vamos a utilizar otra característica de *stylo* denominada *General Imposters* (GI) también conocida como segundo sistema de verificación (o2) que fue introducido por Koppel y Winterr (2014) y aplicado en el estudio de los escritos disputados de Julio Cesar (Kestemont et al., 2016). Lo que hace esta característica es en palabras de Kestemont et al. (2016: 88):

«La [...] GI no consiste en evaluar si dos documentos son simplemente similares en cuanto al estilo de escritura, dado un vocabulario de características estático, sino que pretende evaluar si dos documentos son significativamente más similares entre sí que otros documentos, a través de una variedad de espacios de características estocásticas y en comparación con selecciones aleatorias de los llamados autores distractores (Juola, 2015), también llamados ‘impostores’.»³⁷

Para ello seguimos en todo momento el script que propone Eder (2018) adaptándolo a nuestro corpus. Todo el código puede encontrarse en el anexo código (3) en el que se puede ver como *tokenizamos* los textos obteniendo también una indexación de las 49 obras. La primera de las pruebas consiste en comparar las frecuencias relativas de las MFW de la TF con todos los demás

³⁷ Eder, M. (2018). Authorship verification with the package *stylo*. Retrieved on, 8, 2019. En línea en: <https://computationalstylistics.github.io/docs/imposters>

textos a través de la función `imposters()`³⁸. Utilizamos la medida de distancia por defecto que es Delta de Burrows obteniendo los siguientes datos:

Alemán	Carvajal	Gregorio Gonzalez	Lope
0.00	0.00	0.00	0.00
Castillo	Cervantes	Montalban	Montemayor
0.02	0.98	0.00	0.00
Eslava	Espinel	Quevedo	Salas Barbadillo
0.00	0.02	0.00	0.01
Suarez Figueroa	Zayas		
0.02	0.05		

Tabla 5: resultados de la función *imposters* con Delta de Burrows.

Para entender los coeficientes solo hace falta saber que la atribución se da en los valores cercanos a 1 y se rechaza en el 0. Los resultados de esta tabla son los que nos da el algoritmo al tratar de comparar un texto anónimo determinado con un conjunto de textos de candidatos entre los que se encuentra también el probable autor. Dado que no indicamos otra cosa, el método trata de comprobar uno tras otro todos los autores disponibles que considera como candidatos potenciales. Sin embargo, también podemos focalizarnos en un solo autor haciendo que el método compare el texto dubitado con todos los textos de un candidato. Así pues, le indicamos al programa que queremos hacer esto. Lo llevamos a cabo de manera iterativa y con todos los autores siguiendo el consejo de Eder (2018) de forma que los resultados nos sirvan también a modo de validación cruzada. Tras llevar a cabo este laborioso proceso, vemos que los valores de las ocho pruebas varían levemente con respecto a la anterior destacando Zayas que obtiene un 0,15 en vez de 0,5, el resto de resultados de esta validación corresponden con los datos en el primer experimento con una desviación de 0,01/ 0,02 (ver siguiente página, tabla 6).

³⁸ El script con el código tras esta función puede consultarse en el Github del Computational Stylistics Group: <https://github.com/computationalstylistics/stylo/blob/master/R/imposters.optimize.R>

Alemán	0	Eslava	0.01		
Carvajal	0	Espinel	0.01	Montemayor	0
Castillo	0.02	Gregorio González	0.01	Quevedo	0
Cervantes	0.97	Lope	0	Salas Barbadillo	0.03
Suarez F.	0.01	Zayas	0,15	Montalbán	0

Tabla 6: resultados de la función *imposters* en las pruebas individuales para cada autor.

Como dijimos, los valores más cercanos a 1.00 corresponderían con la atribución a un determinado autor. En este caso solo hay uno que se acerque a ese rango y es Cervantes. El único candidato (candidata en este caso) que parece mostrar cierta similitud es Zayas. Sigamos pues con el experimento para ver si estos resultados son estables. Realizaremos de nuevo la primera prueba utilizando la función *imposters*, pero esta vez vamos a emplear *Wurzburg Delta* que es la medida de distancia que hemos aplicado desde los primeros experimentos pues según Eder (2018) muestra una mejoría notable con respecto a las otras (Burrow's Delta, Eder's Delta)³⁹. También incluiremos Eder Delta para poder contrastar los resultados. Tras realizar este proceso obtenemos los siguientes datos:

Autores	Delta Wurzburg (Cosine)	Eder	Delta Burrows
Alemán	0.03	0.02	0.00
Carvajal	0.00	0.00	0.00
Castillo	0.00	0.02	0.02
Cervantes	0.96	0.95	0.98
Eslava	0.00	0.04	0.01
Espinel	0.00	0.02	0.02
Gregorio González	0.00	0.01	0.00
Lope	0.00	0.01	0.00

³⁹ La forma de introducir este parámetro puede consultarse también en el anexo «código» (3)

Montalbán	0.00	0.00	0.00
Montemayor	0.00	0.00	0.00
Quevedo	0.05	0.00	0.00
Salas Barbadillo	0.00	0.00	0.01
Suarez Figueroa	0.00	0.01	0.02
Zayas	0.04	0.13	0.05

Tabla 7: resultados de imposters con diferentes distancias, *Cosine Delta*, *Eder's Delta* y *Burrow's Delta*.

Para justificar y explicar los resultados, volvemos a la documentación de Eder (2018) quien afirma que en términos teóricos cualquier resultado por encima de 0,5 puede ser indicativo de que la verificación de autoría para un candidato dado fue exitosa. Entonces, con los datos que vemos en la tabla podemos concluir que Cervantes es siempre reconocido como el autor estilísticamente más cercano a la obra dubitada en todas las distancias estadísticas utilizadas. No obstante, existen algunos autores que presentan un cierto grado de similitud. En un principio podríamos obviar este hecho dada la notable diferencia que presentan con respecto al alcalaíno, pero podemos implementar una última comprobación de los resultados.

Ahora disponemos de una tabla de datos, pero nos gustaría establecer en qué rangos los resultados son fiables. Valores como los de Zayas pueden resultar llamativos a pesar de que la diferencia con Cervantes es abismal. No obstante, podríamos dudar sobre si 0.13 es mucho o poco. Para dar una respuesta a esta cuestión existe la función `imposters.optimize()` que utilizando un enfoque de búsqueda en cuadrícula, intenta definir una zona gris en la que las puntuaciones de atribución no son fiables (cfr. Eder, M., 2018). También nos permite introducir diferentes parámetros para p.ej. poder hacer la prueba con distintas distancias.

Los resultados del empleo de esta función se pueden ver en la tabla inferior (tabla 8). Muestran el rango inferior por debajo del cual la atribución sería negativa y el superior por encima del cual la atribución sería positiva. Los valores que estuvieran entre ese rango corresponderían a una zona de grises en la que no se puede descartar con absoluta certeza. Sin embargo, esta medición se

utiliza cuando los datos están en torno al 0,5 donde estaríamos a medio camino entre el 1 (atribución) y el 0 (no atribución) y vimos que en nuestros resultados no existía ningún autor que se aproximara a ese punto medio:

Medidas de distancia	Delta Wurzburg	Eder	Delta Burrows
Resultados	0.00 0.89	0.00 - 0.95	0.00 0.78

Tabla 8: resultados de *imposters.optimize*, rangos de fiabilidad de los datos .

- Con la distancia **Delta de Wurzburg** los resultados muestran que por encima de 0.89 el resultado sería una atribución clara. El único autor que da valores por encima de ese rango es Cervantes con 0,96. Alemán, Quevedo y Zayas tenían 0.03, 0.05 y 0.04 respectivamente. Se encuentran por lo tanto en la zona de grises.
- La **distancia Eder** otorgaba a Zayas 0.07 y a Cervantes 0.95. Los resultados de la función `imposters.optimize()` muestran al igual que la anterior que en ese rango solo hay una atribución segura que es nuevamente Cervantes. Por su parte Zayas con 0,13 y otros autores no son completamente descartados, pues se quedan en el rango de la zona de grises.
- **Delta de Burrows** da unos resultados muy parecidos a Delta de Eder y encontramos el mismo problema con Zayas que tenía 0,05 y se encuentra en la zona de grises mientras que Cervantes con 0,98 es la única atribución.

Podríamos pensar en base a esta última prueba que hay dudas con respecto a los autores que se hayan en la zona de grises. Sin embargo, debemos de volver a considerar que esta herramienta se utiliza en los valores que se encuentran en torno al 0,5 (por ejemplo 0,40) donde la interpretación del resultado (entre 0 y 1) puede dar lugar a dudas. En nuestro caso algunos autores han dado resultados del 0,01 al 0,13 muy alejados de esa zona en la cual se

plantean incógnitas. No obstante, hemos querido probar esta función para ver su utilidad y para comprobar los resultados que nos da para Cervantes. Gracias a ella, ahora sabemos que utilizando diferentes distancias (Eder, Delta de Burrows, Delta de Würzburg) hay una sola atribución con una alta probabilidad pues la función `imposters()` apunta a que *La tía fingida* es de Cervantes. También vimos que la función `classify()` reconocía que el texto dubitado es del complutense. En definitiva, tanto la atribución como la verificación de autoría dan como resultado que el autor de esta novela es el Manco, aunque hay que insistir que en este corpus no están todos los posibles autores y textos que se podrían tener en consideración debido en parte a las limitaciones ya comentadas.

Por lo tanto, la única posible pega que vimos es que el programa parece tener algún problema para descartar de forma definitiva a varios de los autores destacando entre ellos Zayas que se queda siempre en la zona «de grises»⁴⁰. Sin embargo, la atribución cervantina es tan constante y clara que no consideramos esto un impedimento. Además, en este punto conviene recordar el apunte que hace Eder (2018) con respecto a los resultados «cuando realizas el test nuevamente, el resultado final puede diferir levemente debido a la naturaleza estocástica del test». Efectivamente, al repetir la prueba, algunos autores dan como resultado 0,00 y son descartados, si bien Zayas permanece. Por este motivo consideramos que debemos continuar indagando un poco. Además, nos interesa descubrir con qué obras se establecen los nexos entre la *Tía Fingida* y *Las Novelas Ejemplares* para después poder llevar a cabo el análisis literario que nos permita comparar el estilo desde esa perspectiva.

En primer lugar, vamos a realizar un dendograma tal y como hicimos al principio para determinar la configuración de nuestro corpus. Utilizamos por lo tanto el «corpus completo» y empleamos nuevamente la función `stylo()` con las 500MFW y la distancia *Cosine Delta*. Habíamos dejado para el final la comprobación de los resultados incluyendo a *La tía fingida*. La imagen que muestra el dendograma es idéntica a la que se mostraba en el «corpus completo». La diferencia estriba en Cervantes, donde nos encontramos con *La*

⁴⁰ Estas mismas pruebas fueron realizadas utilizando los textos completos (sin fragmentar) y en un rango de MFW más alto (3200). Esto nos obligaba a dejar fuera algunos autores de los cuales solo tenemos una obra, pero en ese caso la posible atribución a Zayas quedaba en todos los casos por debajo del rango inferior que daba la función `imposters.optimize()` y por lo tanto descartada.

tía Fingida que, aunque está sola en una ramificación, aparece asociada en un mismo subcluster con *El celoso extremeño*:

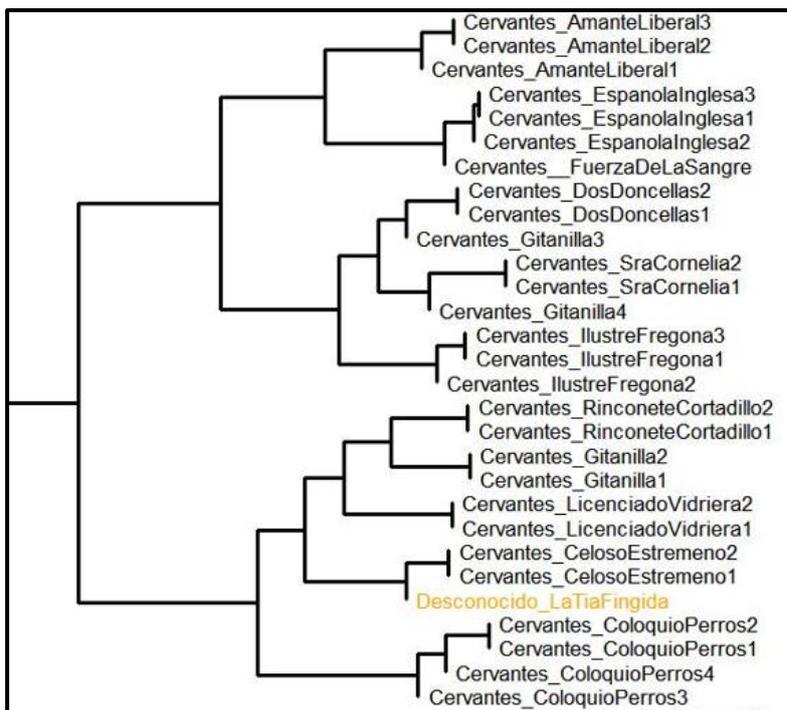


Imagen 16: detalle (obras de Cervantes) del análisis de clúster con el corpus «completo», 500MFW, *Cosine Delta*.

Para seguir con la comprobación de los resultados podemos echar un vistazo a las relaciones entre todos los autores y el texto dubitado. Esto lo vamos a realizar en el siguiente apartado donde llevaremos a cabo un experimento que nos permitirá visualizar los parentescos entre textos.

7.4 Stylo.network () y gephi

En la última prueba se puso de manifiesto que parece existir una relación entre la obra cervantina y la TF. Además, sabemos que Zayas, Quevedo y Alemán parecen tener un cierto grado mínimo de similitud estilística con Cervantes. El problema es que no conocemos donde residen estas relaciones ¿cuáles son exactamente las obras de Cervantes y de los otros autores que presentan vínculos con la TF? Para obtener esta información ya hemos utilizado el dendrograma y este nos muestra como resultado un parentesco con *El celoso extremeño*. Ahora nos queda por comprobar un segundo método de visualización: el Bootstrap Consensus Tree (BCT) a partir de la frecuencia

relativa de las 100 a las 500MFW utilizando aquellos autores que la función *imposters()* no acababa de descartar del todo. La función *stylo()* nos va a servir para crear el gráfico. En particular, vamos a utilizar la función *stylo.network()* con la que vamos a crear un BCT cuyos resultados exportaremos a *Gephi*⁴¹. Se trata de dos tablas de Excel en formato .csv conteniendo los datos de los nodos y las aristas. Con ello empleamos el *layout* de Fruchterman-Reingold (1991) con el que, tras ajustar los parámetros de visualización, obtenemos las siguientes imágenes:

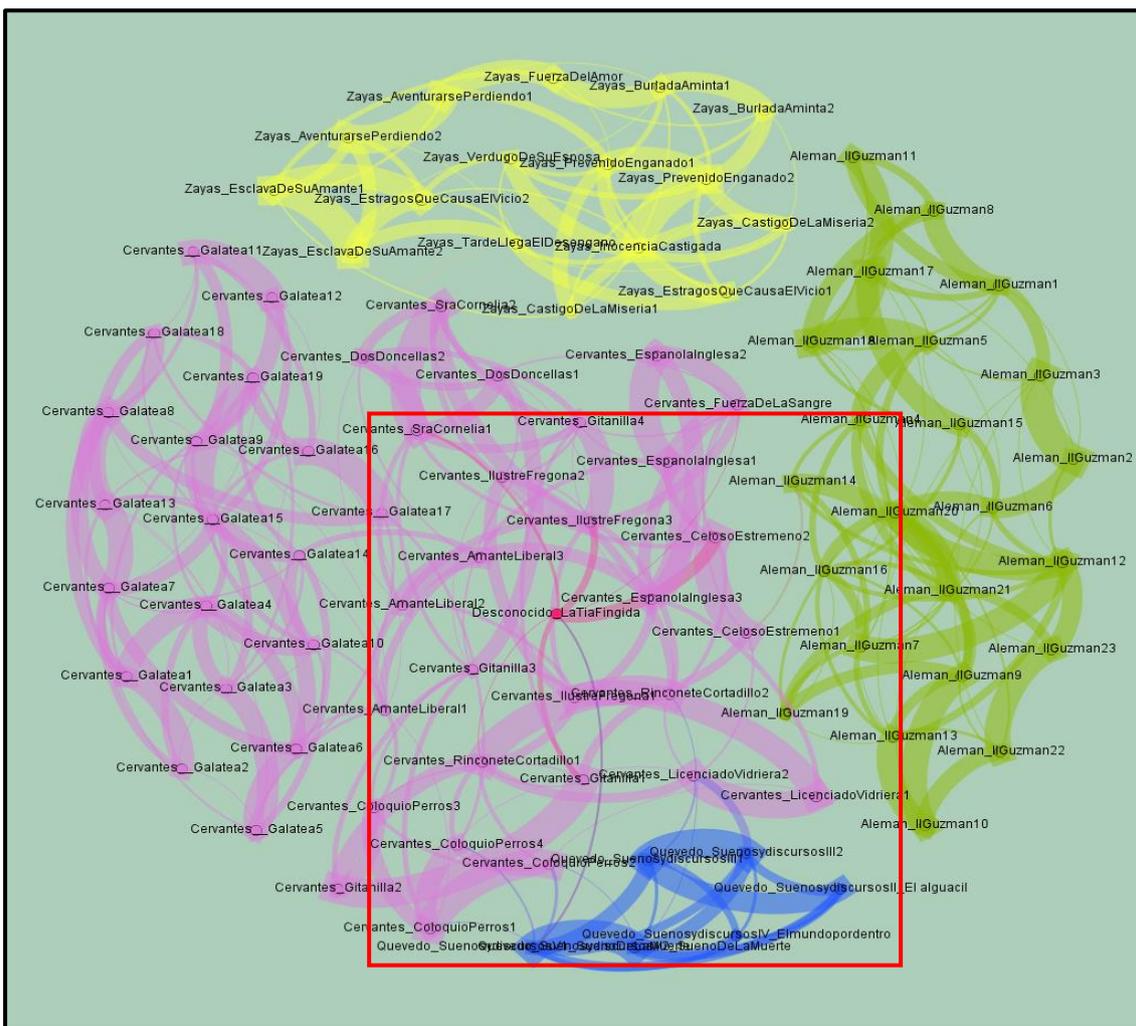


Imagen 17: gráfico Gephi del BCT de cuatro autores (Zayas, Cervantes, Quevedo, Alemán) con distancia *Cosine Wurzburg* (500MFW)

⁴¹ Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open-source software for exploring and manipulating networks. International AAAI Conference on Weblogs and social media. En <https://gephi.org/users/publications/> (consultado el 10.12.21)

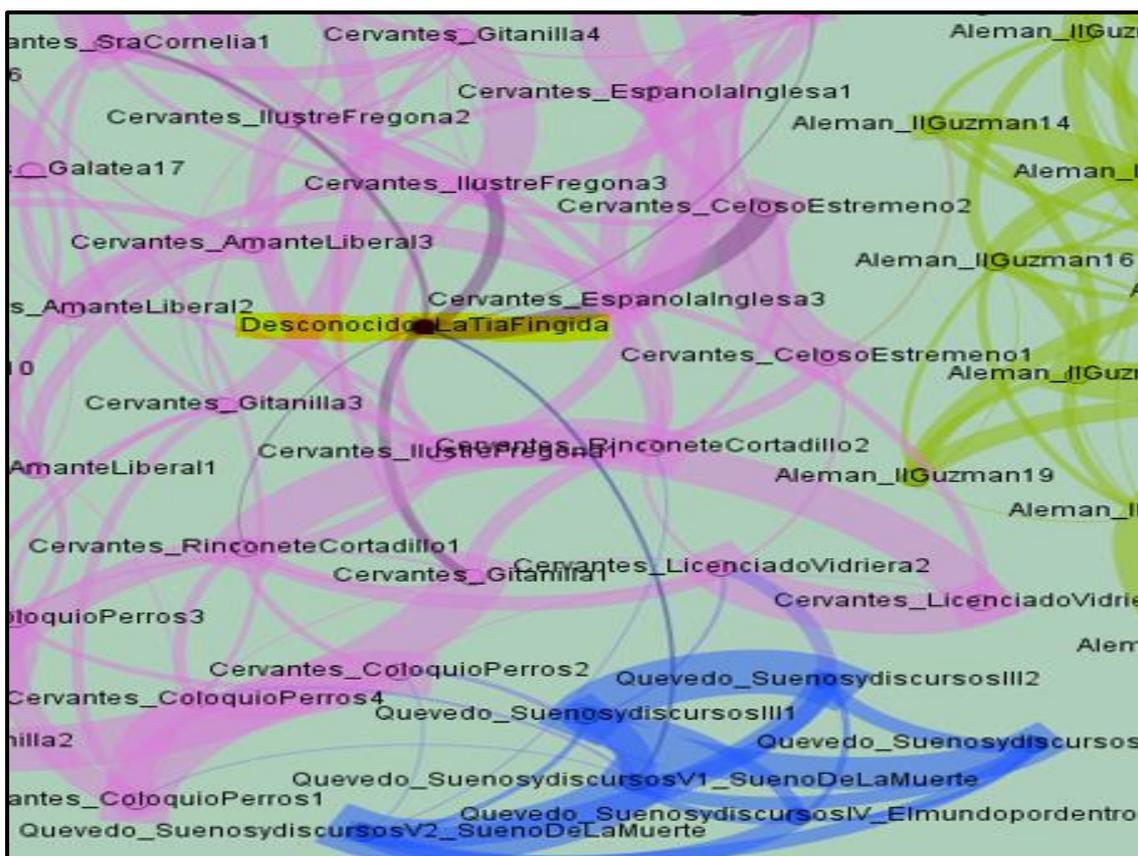


Imagen 18: detalle de la imagen 17, *La Tía Fingida* y sus nexos.

A través de este gráfico podemos ver con mayor claridad las relaciones que tiene la TF con otros textos en base a los parámetros que seleccionamos con la función *stylo()*. Vemos efectivamente que no tiene ningún nexo con Zayas, pero sí alguna conexión con el primer fragmento de la quinta parte de los *Sueños y discursos*, *el sueño de la muerte* de Quevedo. Sin embargo, las relaciones con las *Novelas Ejemplares* de Cervantes son mucho más notorias y variadas: el tercer fragmento de *La ilustre fregona*, el primer fragmento de *La Gitanilla*, el primer fragmento de *La señora Cornelia* y en la parte superior (fuera de la imagen) está *La fuerza de la sangre*. Por último, destaca una relación que sobresale por encima del resto. Se trata de la segunda parte de *El celoso extremeño*.

En este punto y en vista de estos resultados podríamos hacer un análisis literario de todas y cada una de las obras que presentan algún tipo de relación para tratar de dilucidar dónde se encuentran las semejanzas y diferencias entre ellas. Sería interesante quizás analizar la conexión con Quevedo. No obstante,

por una cuestión de economía de texto y dada la evidente relación que aparenta tener *La Tía Fingida* con el de *El celoso extremeño* procederemos al análisis literario de ambas obras y las confrontaremos. De esta manera esperamos poder descifrar algunos de los elementos que se esconden tras los datos que hemos visto hasta ahora.

8 Análisis literario: la novela en el Renacimiento y Barroco

Para comenzar con la contextualización más general de la novela es necesario tener en cuenta que lo que hoy entendemos como novelas no se definía así en aquella época. El nombre se reservó tal como comentan Díez-Echarri y Roca Franquesa (1972: 222) para «un reducido número de obras de carácter amatorio, conforme al camino que trazó Cervantes en sus *Doce novelas ejemplares de honestísimo entretenimiento*». Así, los mismos autores consideran que la denominación de novela se puede aplicar a todas las formas redactadas en prosa, de carácter imaginario, que tienen como objetivo «el placer estético, la distracción de los lectores» (Ibidem)

Otro aspecto relevante es que precisamente es en el Siglo de Oro cuando la novela que en los siglos anteriores no había destacado alcanza un hito en cuanto a calidad y cantidad. Además, también debe recordarse que es Cervantes quien desde la historia de la literatura es considerado «el verdadero creador de la novela moderna» (Díez-Echarri y Roca Franquesa, 1972: 222).

Así, la riqueza de la novela de los siglos XVI y XVII da lugar a un extenso campo en el que nos encontramos diferentes formas de novelas. Díez-Echarri y Roca Franquesa (1972: 222) resumen esquemáticamente las siguientes: Novela de caballerías, pastoril, picaresca, histórico-morisca, bizantina, cortesana y formas no definidas. Evitaremos aquí hacer un resumen de cada una de ellas por una cuestión de economía de texto. En todo caso Díez-Echarri y Roca Franquesa (1972) llevaron a cabo una valiosa labor sintetizando las particularidades de cada uno de estas formas que puede servir de referencia⁴².

⁴² Las características de las formas novelísticas del Siglo de Oro pueden ser consultadas en Díez-Echarri, E., y Roca Franquesa, J.M. (1972: 222-227)

La categorización no termina ahí pues dentro de estas formas también podemos encontrar determinados tipos, por ejemplo, en la novela picaresca se reconocen: la idealístico-satírica, la realístico-optimista y la novelesco-descriptiva (vid. Díez-Echarri y Roca Franquesa, 1972: 252). Así pues, nuestro corpus incluye novelas que se circunscriben a varias de estas formas y tipos lo que le da la ya señalada heterogeneidad siendo este un aspecto relevante y que como comentamos no se puede observar en la búsqueda hecha en el CORDE para recopilar los autores y las obras. No obstante, su uso lo justifica el hecho de facilitar una necesaria visión general de los autores que fueron coetáneos a Cervantes y las obras que se escribieron en prosa en ese periodo, pero sabiendo que en esta búsqueda hay una gran variedad de formas y tipos.

8.1 Cervantes: sus novelas

Cervantes destacó principalmente como novelista y como ya señalamos en palabras de Díez-Echarri y Roca Franquesa (1972) puede ser considerado el maestro de la novela moderna. Estos mismos autores afirman que Cervantes es el creador de alguno de los tipos de novela: la novela corta (Díez-Echarri y Roca Franquesa, 1972: 356)

En nuestro corpus utilizamos *La Galatea* que es su primera novela y que está dividida en seis libros. Corresponde con el género de la novela pastoril presentando las características propias del género: convencionalismo en el lenguaje, en las ideas y en los sentimientos. Por otra parte, el núcleo principal de nuestro corpus lo constituyen las *Novelas Ejemplares*, porque estas son las más parecidas en cuanto a extensión y forma siendo además a las que se circunscribe *La Tía Fingida* que es incluso editada en ocasiones como una *Novela Ejemplar* más justificándolo con el contexto de su aparición junto con una copia manuscrita de *Rinconete y Cortadillo* y *El celoso extremeño* o en base a argumentos para su atribución a Cervantes.

La colección apareció en Madrid en 1613 suponiendo una etapa intermedia entre la Primera y la Segunda parte del *Quijote* aunque se considera que debió de haber empezado con su redacción tiempo atrás dado que en la Primera parte del *Quijote* se alude a *Rinconete y Cortadillo*. Las *Novelas Ejemplares* destacan según Díez-Echarri y Roca Franquesa (1972: 358) por dos hechos: el primero,

es la primera vez que en español se denomina a un texto «novela» con esta acepción que con anterioridad significaba más bien invención imaginaria no histórica o cuento; el segundo, es que introdujo el género de la novela corta que era desconocida antes de Cervantes.

El término *Ejemplares* se refiere al carácter principalmente moralizador que el autor quiso darles, aunque como comentan Díaz-Echarri y Roca Francesa (1972: 358) «la realidad quedara algunas veces bastante lejos de su intención [ya que] abundan en pasajes resbaladizos y alguna vez tocan en lo escabroso»

En cuanto a los aspectos técnicos, que nos interesan para tratar de determinar el estilo literario, las *Novelas Ejemplares* son narraciones breves donde se toma un tema que suele ser de la vida real y sin demorarse en detalles superfluos se desarrolla de forma sucinta. Las descripciones son breves, el retrato de los personajes se limita a los rasgos más destacables o a generalidades. Con respecto a la temática esta es variada y corresponde a distintos géneros. La clasificación que proponen Díaz-Echarri y Roca Francesa (1972: 359) es la siguiente:

- Género italianizante: *El amante liberal*, *Las dos doncellas*, *La española inglesa*, *La señora Cornelia*.
- Asuntos de aventuras: *La gitanilla*
- Género picaresco: *Rinconete y Cortadillo*
- Género morisco: *El amante liberal*.
- Género realista: *La ilustre fregona*, *El celoso extremeño*, *El casamiento engañoso*.
- De sátira lucianesca: *El licenciado Vidriera*, *El coloquio de los perros*.

Sin embargo, esta clasificación no es definitiva pues siempre hay que tener en cuenta que existen interferencias entre las características de los géneros de forma que un mismo texto puede clasificarse en dos o incluso tres géneros simultáneamente. Quizás por este motivo existen otras propuestas de clasificación por parte de la crítica. Es extendida la división de las *Novelas Ejemplares* en dos grupos: novelas idealistas (abunda la imaginación) y novelas realistas (abunda la observación de la realidad). En las realistas se sitúan *El coloquio de los perros*, *El casamiento engañoso*, *Rinconete y Cortadillo*, y *El celoso extremeño*; en las idealistas, *La señora Cornelia*, *Las dos doncellas*, *La española inglesa*, *El amante liberal*, *La fuerza de la sangre* y *La ilustre fregona*.

Por su parte *La gitanilla* y *El licenciado vidriera* son consideradas en ocasiones «ideorrealistas» por encontrarse entre realidad e idealismo.

8.2 La novela corta en Cervantes: generalidades técnicas

Las novelas cortas de Cervantes destacan según Díez-Echarri y Roca Franquesa (1972: 362) por su originalidad que resumen citando los argumentos de Pfandl (1933)⁴³:

- a) Finalidad ética. Aunque haya pasajes escabrosos estos son impuestos por el tema o la situación. La intención es siempre honesta.
- b) Aplicación del principio de *deleitar aprovechando*; motivo por el cual comenta en el prólogo «Si bien lo miras, no hay ninguna de quien no se pueda sacar algún ejemplo provechoso»
- c) Uso frecuente del estilo directo y del diálogo. Numerosas descripciones directas y diálogos que aproximan las escenas evidenciando el parentesco de la novela con el drama.
- d) Mezcla armoniosa de lo cómico y alegre con lo serio que estimula sentimientos mixtos, lo cual logra un gran efecto en las obras literarias.
- e) Desenlace feliz: todas las Novelas Ejemplares terminan con una boda o reconciliación, sugieren una perspectiva optimista.
- f) Concepto exacto de lo que es la novela corta: ésta actúa sobre un suceso fragmentario de cierta relevancia y dificultad y se dirige sin grandes ornamentos accesorios hacia el desenlace.

9 **Análisis literario: *La tía fingida* vs. *El celoso extremeño***

Como hemos visto, estilométricamente *La tía fingida* guarda particular relación con una de las *Novelas Ejemplares*: *El celoso extremeño*. En vista de estos resultados que tienen en cuenta las 500MFW de ambos textos vamos a realizar un análisis literario-estilístico de las mismas para tratar de dilucidar desde este punto de vista si existen similitudes o diferencias notorias.

⁴³ Pfandl, L. (1933). Historia de la Literatura nacional española en la Edad de Oro, trad. del alemán por Balaguer Rubio, J., Barcelona.

Para llevar a cabo el análisis literario del estilo tendremos en cuenta la propuesta de la estilística integral de Castagnino (1987) que considera que no hay estilísticas sino tan solo una estilística. Esta es una disciplina integral que reúne lo puramente lingüístico y lo literario. Además, considera que confluyen en la disciplina de forma colateral algunos aspectos históricos, críticos, fórmulas de la preceptiva y otras disciplinas en apariencia extraliterarias (cfr. Castagnino, 1987: 13). De este modo Castagnino propone un análisis estilístico integral en el que se pretende limpiar de accesorios para introducirse en el texto llegando hasta «los secretos de los efectos que produce» (Op. Cit., pág 14) que sería su técnica estilística.

Castagnino considera también que el estilo es una desviación, así se entiende cuando el teórico afirma:

«En el análisis estilístico interesa la elección hecha por el autor con la palabra, en cuanto atañe a las cualidades sensoriales de esta, [...] es símbolo de idea. Por lo que respecta a la oración, cabe atender a su construcción; a la unidad de pensamiento que transmite y al reflejo psíquico que esa transmisión comporta en cuanto es clara, oscura, hermética, original, clisé, etc. En cuanto va cargada de sentimientos, afectos y emociones: en cuanto supone una jerarquización de las ideas, etc.» (Castagnino, 1987: 179)

La propuesta metodológica de Castagnino nos parece pues adecuada para nuestro trabajo pues el teórico propone una perspectiva en base a la cual el objetivo es «integrar» los aspectos lingüísticos y los literarios incluyendo además como ya dijimos otros aspectos extratextuales. Ya hemos incluido a lo largo de nuestros experimentos una parte que corresponde a los aspectos más lingüísticos como es la relación de las palabras funcionales. Así pues, comenzaremos aplicando una parte del esquema de análisis que propone Castagnino (1987) con *La tía fingida* para después tratar de contrastar en el siguiente texto las semejanzas y/o diferencias que puedan reforzar o contradecir los datos que hemos visto hasta ahora en este trabajo. Así mismo, para el análisis de las técnicas narrativas utilizaremos la conocida terminología de Gerard Genette (1989, 1998).

9.1 Análisis de los contenidos

9.1.1 *La tía fingida*

El título de la obra hace referencia a uno de los protagonistas como sucede a menudo en el resto de novelas ejemplares (por ejemplo: El licenciado vidriera, La señora Cornelia, etc.) en este caso se trata de doña Claudia de Astudillo y Quiñones que finge ser la tía de una bellísima joven llamada Esperanza.

El argumento es relativamente sencillo. En Salamanca dos estudiantes de origen manchego se encuentran con la casa donde vive la anciana Doña Claudia de Astudillo y Quiñones que en realidad es una alcahueta y que está acompañada por dos dueñas, un escudero y la bella joven doña Esperanza de la que dice ser su tía. Los jóvenes tratan de cortejar a la joven y le dan serenata, pero son espantados por una de las dueñas. Ante esta negativa se dirigen a la casa de un caballero llamado Don Félix a quien le informan sobre la belleza de la joven Esperanza. Interesado por la información, Don Félix habla con una de las dueñas y acuerda ir esa noche a la casa de Esperanza. Al llegar a la vivienda espera oculto lo cual da lugar a que escuche la conversación entre la anciana y la joven en la cual sale a relucir el engaño pues la vieja ya había explotado los encantos de la moza hasta en tres ocasiones, prostituyéndola como presunta virgen. Don Félix es descubierto, pero intenta cerrar el trato. En ese momento se produce una algarabía entre la dueña Grijalba y Doña Claudia que llama la atención de las autoridades que oportunamente ya estaban informadas del negocio que tenía lugar allí. El corregidor envía a las mujeres a la cárcel, pero por el camino Esperanza es liberada y raptada por los dos estudiantes. Uno de ellos trata de abusar de ella, pero el otro lo evita por lo que el primero decide pedir su mano para poder «gozarla» legítimamente y se escapa con ella al pueblo. Finalmente, el estudiante se casa con Esperanza y Doña Claudia es castigada públicamente.

El asunto es en palabras de según Díez-Echarri y Roca Franquesa «más descarnado y libre que los de las otras *Novelas Ejemplares* de Cervantes» (1972: 363) lo cual ha sido uno de los motivos para que existan dudas sobre su autor.

La temática gira en torno a una proxeneta de tintes celestinescos que en la acción narrada es descubierta y termina pagando por su inmoralidad.

En cuanto a la relación de **personajes** los participantes en el relato son:

- Personajes principales: estudiante 1, estudiante 2, Doña Claudia, Esperanza y Don Félix.
- Personajes secundarios: vecino, músicos, poeta, extremeño, dueña2, Grijalba (dueña 1), escudero, corregidor y el padre de uno de los estudiantes.

Tenemos pues una constelación limitada de figuras como es de esperar en una novela corta donde algunos de los personajes secundarios son accesorios pero necesarios. La mayoría de ellos, aunque forman parte de la narración no participan en los diálogos (músicos, poeta, dueña2, escudero, padre de uno de los estudiantes). La participación del resto de figuras secundarias (vecino, extremeño, corregidor) son intervenciones muy breves. Las descripciones de los personajes son así mismo reducidas con la excepción de las dos mujeres que protagonizan la novela. De Esperanza se dice:

«moza, al parecer, de diez y ocho años, de rostro mesurado y grave, más aguileño que redondo; los ojos negros, rasgados y al descuido adormecidos; cejas tiradas y bien compuestas, pestañas negras y encarnada la color del rostro; los cabellos plateados y crespos por artificio, según se descubrían por las sienes; saya de burriel fino, ropa justa de contray o frisado; los chapines de terciopelo negro, con sus clavetes y rapacejos de plata bruñida; guantes olorosos, y no de polvillo sino de ámbar. El ademán era grave, el mirar honesto, el paso airoso y de garza. Mirada en partes, parecía muy bien, y en el todo, mucho mejor;» (Cervantes, La Tía Fingida)⁴⁴

Por su parte, Doña Claudia es descrita así al comienzo de la novela:

«una reverenda matrona, con unas tocas blancas como la nieve, más largas que una sobrepelliz de un canónigo portugués, plegadas sobre la frente con su ventosa y con un gran rosario al cuello de cuentas sonadoras, tan gordas como las de Santenuflo, que a la cintura la llegaba; manto de seda y lana, guantes blancos y nuevos sin vuelta, y un báculo o junco de las Indias con su remate de plata en la mano derecha, y de la izquierda

⁴⁴ La versión que utilizamos de la misma es: Novela de la tía fingida [versión Porras de la Cámara por Francenson-Wolf], edición de Florencio Sevilla Arroyo. Esta se encuentra disponible en línea en la BVMC en su versión digital, lo que permite una rápida búsqueda de los términos y pasajes. Por ende, no incluiremos las páginas en las citas a partir de ahora.

Enlace: <http://www.cervantesvirtual.com/nd/ark:/59851/bmc1z4h1>

la traía un escudero de los del tiempo del conde Fernán González, con su sayo de velludo, ya sin vello, su martingala de escarlata, sus borceguíes bejaranos, capa de fajas, gorra de Milán con su bonete de ahuja, porque era enfermo de váguidos, y sus guantes peludos, con su tahalí y espada navarrisca.» (Cervantes, La Tía Fingida)

Es notorio que la descripción que se hace de Esperanza refleja principalmente sus cualidades físicas, aunque también la forma de vestir. Con esto, el autor parece querer destacar la belleza de la protagonista para dar sustento al empeño de los estudiantes por seducirla y por lo tanto sentido al relato. De Doña Claudia por otra parte se destaca solamente su manera de vestir queriendo quizás hacer una referencia al contenido mismo de la novela pues es tan solo la apariencia lo que esta figura tiene de verdadero fingiendo ser una señora «*de mucha autoridad*». Además, como contraste a estas profusas descripciones se pueden observar las síntesis que se llevan a cabo con otros personajes. En algunos casos (dueñas, corregidor, padre del estudiante, músicos) tan solo se nombran sin añadir ninguna descripción, los que sí son descritos tienen un carácter general:

- Los dos estudiantes: «*dos estudiantes mancebos y manchegos, más amigos del baldeo y rodancho que de Bártulo y Baldo*» (Cervantes, La Tía Fingida)
- El vecino: «*un oficial vecino*» (Cervantes, La Tía Fingida)
- El poeta: «*un poeta de los que sobran en aquella ciudad*» (Cervantes, La Tía Fingida)
- Un extremeño: «*un bellacón de los circundantes*» (Cervantes, La Tía Fingida)
- Don Félix: «*cierto caballero amigo suyo, de los que llaman generosos en Salamanca y se asientan en cabeza de banco: el cual era mozo, rico, gastador, músico, enamorado, y sobre todo amigo de valientes*» (Cervantes, La Tía Fingida)

Los espacios geográficos de la novela se sitúan en Salamanca y son: la casa de Doña Claudia, las calles de Salamanca, la casa de Don Félix, el pueblo del estudiante y la plaza. Hay pocas referencias a los espacios que se dividen en la calle y la casa. Las descripciones son vagas «*cierta calle de Salamanca*»,

«una casa», «la calle», «en cierta calle tras cantón» «llegó a casa de su padre» «en medio de la plaza». Por lo tanto, no se especifica en ningún momento el espacio en el que transcurre la historia más allá de que es en diferentes lugares de Salamanca. Podemos considerar que estos son elementos accesorios y que el autor prescinde de ellos para abreviar. Como vimos, una de las generalidades técnicas de las Novelas Ejemplares era que presentan un concepto exacto de lo que es la novela corta y prescinden para ello de ornamentos dirigiéndose siempre hacia el desenlace.

Por último, **el tiempo** en el que transcurre la historia son dos días (tres si contamos la partida de Salamanca de Esperanza y el estudiante) sin anacronías y siendo por lo tanto un relato «*Ad Ovo*», es decir lineal, con un orden cronológico que sigue el principio de causa-efecto. Sí existe al final una muy breve prolepsis externa en la cual se avanza hasta las consecuencias que tuvo la historia para los protagonistas y el estudiante manchego. Esta sirve como epílogo conduciendo a su término la línea de acción (cfr. Genette, 1980: 11) y se introduce por medio de las expresiones «*Se le averiguó*» y «*Súpose luego*».

Vistos pues los contenidos, continuaremos nuestro análisis con los mismos aspectos de *El celoso extremeño* para poder contrastar las semejanzas y diferencias entre ambas novelas. Después continuaremos con las técnicas literarias de ambas obras por separado para ver a conciencia qué estrategias se utilizan para la presentación de los contenidos. Terminaremos, siguiendo el mismo sistema, con el análisis de la expresión en el que podremos contrastar los recursos retóricos y algunos aspectos morfológicos destacables.

9.1.2 *El celoso extremeño*

El título de la novela hace nuevamente referencia al protagonista, en este caso se trata de un anciano de 68 años llamado Filipo de Carrizales cuyo mayor defecto son sus excesivos e incontrolables celos.

El tema de *El celoso extremeño* tiene un largo antecedente: se basa en el matrimonio entre un viejo y una joven por motivos de desigualdad social.

El argumento se centra en la figura del rico y viejo indiano Carrizales que se casa con una joven bellísima y cándida llamada Leonora. El marido, que sufre de unos celos extremos la incomunica con el mundo exterior para evitar cualquier atisbo de seducción. No obstante, el galán Loaysa encuentra la manera de poder tener acceso a ella. Finalmente, el marido los encuentra durmiendo juntos. Existen dos finales alternativos que corresponden con la del manuscrito Porras de Cámara de 1606 en la cual se consuma el adulterio, y con la versión retocada en la cual no hay adulterio porque los amantes caen en un sueño profundo. En ambas versiones Carrizales muere de pena, insta a Leonora a que despose a Loaysa y le deja como herencia toda su riqueza. Sin embargo, Leonora, llena de arrepentimiento, decide entrar en un convento.

El elenco de **personajes** al igual que en la TF es reducido siguiendo las generalidades del género. Está compuesto por las siguientes figuras:

- Personajes principales: Filipo de Carrizales, Leonora y Loaysa
- Personajes secundarios: Marialonso (dueña de la mansión), Luis (esclavo eunuco de Carrizales), esclavas, doncellas.

Las descripciones de los personajes son breves al igual que en la mayoría de figuras de la TF siguiendo las generalidades técnicas del género. En este caso no tenemos una descripción profusa de los protagonistas. Sobre Leonora se dice: «*doncella, al parecer de edad de trece a catorce años, de tan agradable rostro y tan hermosa que, [...] Carrizales rindió la flaqueza de sus muchos años a los pocos de Leonora*». A lo largo de la novela se utilizan adjetivos que insisten en su apariencia física y poca experiencia vital: *la simple Leonora, la ignorante Leonora, la simple e incauta Leonora, sus hermosas manos*. En el personaje de Carrizales se remarca el vicio de su personalidad: los celos. Se comenta que «*era el más celoso hombre del mundo*» (hipérbole), se repite el adjetivo antepuesto 'celoso' («*celoso hombre*», «*celoso marido*», «*celoso anciano*») y se reitera en su avanzada edad con el adjetivo 'viejo' y 'anciano'. De Loaysa tenemos dos descripciones enfrentadas. La que lleva a cabo el narrador sobre su apariencia: «*Era mozo y de gentil disposición y buen parecer*». Junto al comentario más general que hace también el narrador sobre la gente de su tipo:

«género de gente ociosa y holgazana, [...] gente baldía, atildada y meliflua [...]»

Y la que es producto de su disfraz de pobre músico tullido:

«se puso unos calzones de lienzo limpio y camisa limpia; pero encima se puso unos vestidos tan rotos y remendados, que ningún pobre en toda la ciudad los traía tan astrosos. Quitóse un poco de barba que tenía, cubrióse un ojo con un parche, vendóse una pierna estrechamente, y, arrimándose a dos muletas, se convirtió en un pobre tullido»

El resto de personajes es descrito brevemente y a menudo por medio de epítetos. De la dueña Marialonso se dice que es «*de mucha prudencia y gravedad*». Más adelante ante su actuación los adjetivos tienden a ser peyorativos «*astuta y reportada*», «*falsa dueña*», «*la maldita dueña*», «*la malvada de la dueña*», «*la falsa de Marialonso*». Luis (el negro) es definido como «*un negro viejo y eunuco*» y «*pobre negro*». Los personajes que completan la constelación son tan solo citados o se expresan algunos rasgos básicos sobre su personalidad o apariencia: «*negras bozales*», «*Guiomar, la negra, que no era muy ladina*», «*Guiomar, la negra, que por ser portuguesa*».

En cuanto a **los espacios de la novela** al igual que en la TF el principal es la casa que en este caso es denominada también en una ocasión «*fortaleza*» y está situada en Sevilla. Dentro de ella, se distinguen diferentes partes: el portal con su torno, varias estancias, el patio y el pajar. También de forma similar a la TF tiene relevancia la calle, espacio opuesto a la casa y de cuyos peligros guarda Carrizales su tesoro con celo. Otros espacios de menor relevancia son Italia, Flandes y sobre todo las Indias de donde viene Carrizales y a donde escapa al final Loaysa. Al igual que estos, el resto de espacios son tan solo citados: una calle, la estancia, el pajar, el portal, las Indias, Flandes. La única descripción relevante en este aspecto es la de la casa. Particularmente cuando al comienzo es transformada en fortaleza:

«compró una [casa] en doce mil ducados, en un barrio principal de la ciudad, que tenía agua de pie y jardín con muchos naranjos; cerró todas las ventanas que miraban a la calle y dioles vista al cielo, y lo mismo hizo de todas las otras de casa. En el portal de la calle, que en Sevilla llaman casapuerta, hizo una caballeriza para una mula, y encima della un pajar y apartamiento [...]; levantó las paredes de las azuteas de tal manera, que el que entraba en la casa había

de mirar al cielo por línea recta, sin que pudiesen ver otra cosa; hizo torno que de la casapuerta respondía al patio.» (Cervantes, El celoso extremeño)

Por último, **el tiempo** en el que transcurre la historia son varios años, aunque el nudo se desarrolla en varios días. La narración comienza con un momento indeterminado de finales del siglo XVI o comienzos del XVII, «*no ha muchos años*». Carrizales pasa varios años en las Indias «*veinte años que en ellas estuvo*». Cuando regresa sabemos que tiene 68 primaveras, y se entiende que hay una elipsis implícita en la que debieron pasar varios años hasta que comienza el nudo en el que ya vive con Leonora pues al final de la novela dice «*los casi ochenta [años] míos*». Al comienzo de la segunda escena se dice «*pasaron un año y medio de noviciado*» y durante todo este nudo y hasta el desenlace tenemos indicaciones del paso del día y las noches: «*al cabo de algunos días*», «*de aquí a dos días*» etc. Al final el anciano Carrizales nos da el dato exacto, indicio irónico de su obsesión, de la duración del matrimonio «*hace hoy un año, un mes, cinco días y nueve horas que me entregastes a vuestra querida hija*». El relato es lineal, aunque incluye saltos temporales a medida que avanza la historia. Es al igual que la TF «Ad Ovo» y se diferencia de esta tan solo en que el tiempo de la historia y los saltos temporales son más dilatados.

9.2 Análisis de las técnicas literarias

9.2.1 *La tía fingida*

Técnicas narrativas

a) Modo narrativo: focalización, distancia, relato de palabras

Para hablar del punto de vista utilizaremos la terminología de Genette (1989, 1998). Cuando hablamos del ángulo de visión del narrador nos referimos **a la focalización**. En este relato el narrador sabe más que todos los personajes y accede a sus sentimientos más íntimos de modo que responde a una focalización cero. Aunque el carácter breve de la obra tenga como consecuencia que haya escasas referencias, se puede observar que el narrador conoce los pensamientos y sentimientos de los personajes. Algunos ejemplos son:

- «*don Félix [...] satisfecho de todo cuanto saber quería*»

- «él [don Félix] quedó pensando en su ida y aguardando la noche, que le parecía se tardaba mil años»

También se observa lo mismo con más claridad en el siguiente pasaje en el cual el narrador conoce y describe las intenciones de una de las dueñas y de Esperanza para quedarse con los beneficios que pudiera aportar don Félix:

«muy deseosa que sus conciertos viniesen a efecto, según su señora la moza y ella lo tenían ordenado; cuales eran que, sin que la Claudia lo supiese, todo aquello cuanto con que don Félix cayese y pechase fuese para ellas solas, sin que la vieja tuviese que ver ni haber de ello»

En cuanto a **la distancia**, vemos que el relato de los acontecimientos tiene lugar por medio de la diégesis: el narrador habla por si mismo sin que se delegue en los personajes en ningún momento. Esto tiene como consecuencia que abunde la narración y la descripción más que el diálogo.

Con respecto a los diálogos, hablamos en términos de Genette (1998) de «**relato de palabras**» que se define de forma más concisa como «*Modos de (re)producción del discurso y el pensamiento de los personajes en el relato literario escrito*» (Genette, 1998: 36) y de los cuales encontramos en la TF bastante variedad. Por una parte, hay **discurso narrativizado** en el cual el narrador solo indica que ha habido acto del habla y añade puntualmente el tópico:

«quedáronse los señores en la calle, pensativos y medio enamorados, dando y tomando brevemente en qué hacer debían»

«Informóse del paje la astuta Claudia de la calidad y condiciones de su señor, de su renta, de su inclinación y de sus entretenimientos y ejercicio»

También encontramos **discurso transpuesto** en el que se narran las palabras de los personajes en oraciones subordinadas. De este tipo hay ejemplos en **estilo indirecto libre** en el que se introduce el contenido de una conversación sin darle entrada con un verbo del habla (comentar, decir, expresar, etc.). Por ejemplo:

«El caballero, pues, que era de los del campo través, no tardó mucho en ofrecerles que él la conquistaría para ellos, costase lo que costase»

Pero también hay muestras del **estilo indirecto marcado** en el que si encontramos los verbos del habla precediendo a una oración subordinada:

«Respondióla el caballero que todo cuanto le había dicho del merecimiento, valor y hermosura, honestidad, recogimiento y principalidad -por hablar a su modo- de su ama lo creía»

En términos generales se puede reconocer una mayor presencia de la narración y de diferentes técnicas de empleo de los diálogos en estilo indirecto en las primeras partes de la novela (A, B, C. Véase: estructura narrativa). A partir del monólogo de Doña Claudia al que sigue un diálogo con Esperanza y que aparece titulado como «Consejo de Estado y Hacienda» (parte D y D') hay una presencia más notoria de **discurso restituído o citado** que es el discurso del estilo directo en el que el narrador simplemente da la palabra a los personajes.

b) Voz narrativa: niveles narrativos, tiempo del relato

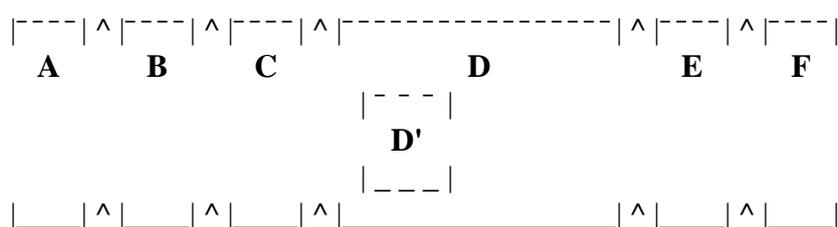
Siguiendo con la terminología de Genette, en los **niveles narrativos** se distinguen tres: el extradiegético, el intradiegético y el metadiegético. La metadiégesis en Cervantes es un recurso conocido, basta echar un vistazo a los trabajos sobre el entremés del *Retablo de las Maravillas*, a la comedia *Pedro de Urdemalas* o a la forma en que el autor introdujo algunos de sus textos en el *Quijote*. No obstante, en esta novela el uso de estos recursos podría resultar contraproducente pues alejaría quizás al autor de su objetivo de ser breve en su narración. Así en *La tía fingida* la voz narrativa se mantiene en el nivel **extradiegético** pues es la voz del narrador la que lleva adelante la narración. Además, en cuanto a su relación con la historia se observa a través del uso de las formas verbales que narra en tercera persona (Ej. Al comienzo: *vieron, pareciéndoles*, etc.) y no es un personaje de ella por lo que es **heterodiegético**.

Por último, queda por comentar brevemente el **tiempo del relato** para poder ubicar a la instancia narradora es decir para saber en que tiempo se encuentra la historia. Como se puede observar en el uso de los tiempos verbales pretéritos, la **narración es ulterior** pues está contada en tiempo pasado en su totalidad. El uso de otros tiempos verbales se reserva a los diálogos en estilo

directo en los que hay más variedad, aunque domina el uso de las formas del presente.

c) Estructura narrativa

La estructura narrativa de la TF es resumida de forma sucinta por E.T.Aylward (1999) por lo que tomaremos como referencia su esquema de la división de las escenas. (Aylward, 1999: 49):



En *La tía Fingida* encontramos cinco escenas principales que como se comentó siguen una cronología lineal. No tenemos grandes saltos al pasado, ni al futuro, ni existen tramas accesorias por lo que estamos ante una estructura que puede considerarse sencilla. La última escena (F) representa el breve epílogo en el que reconocíamos una prolepsis externa que resume el destino de los protagonistas. Desde la escena A a la escena E tienen lugar el planteamiento (A y B), nudo (C y D) y el desenlace (E + F). Se sigue así la estructura más básica de la narración. Vamos a verlas por partes:

- A) Escena 1: En Salamanca, dos estudiantes manchegos que se encuentran en la calle descubren la casa en la que habita Doña Claudia y comienzan a indagar sobre sus habitantes. Un vecino les informa y ven llegar a Esperanza, Doña Claudia, las dos dueñas y el escudero. Los estudiantes quieren tratar de seducir a Esperanza y contratan a un poeta y músicos para darle serenata.
- B) Escena 2: Por la noche, los dos estudiantes junto con su comitiva de músicos acuden a los exteriores de la casa. Se genera una algarabía a causa del espectáculo de manera que los estudiantes terminan huyendo.

- C) Escena 3: Los estudiantes se dirigen a la casa de Don Félix para que les ayude en el cortejo. Don Félix acepta y envía una propuesta a Doña Claudia. Después Don Félix y Grijalba se encuentran y acuerdan que el pretendiente visite a Esperanza por la noche de forma secreta para llevar a cabo un encuentro nocturno.
- D) Escena 4: Al atardecer, cuando Don Félix se cuela en la casa de Doña Claudia asiste escondido a una conversación que incrimina a la vieja y la joven que aparece titulada como «Consejo de Estado y Hacienda»

D') «Consejo de Estado y Hacienda»

Se trata de un diálogo entre Claudia y Esperanza en el que la primera da muestra de su conocimiento acerca de las características de los varones según su procedencia en relación con su negocio. Sale a relucir que ambas se dedican a obtener dinero de los pretendientes que quieren gozar de la belleza y supuesta virginidad de Esperanza.

Don Félix es descubierto a causa de sus estornudos y sale de su escondite. En ese momento tiene lugar una airada discusión entre Grijalba y Claudia. Inmediatamente después entran en la sala el Corregidor y su séquito que también han escuchado la conversación incriminatoria por lo que envía a las mujeres a la cárcel a pesar de los intentos de Don Félix por persuadirle de lo contrario.

- E) Escena 5: En la calle, los dos estudiantes acompañados por seis de sus amigos tratan de liberar a Claudia y Esperanza de camino a la prisión. Consiguen liberar solamente a Esperanza. Uno de los estudiantes trata de aprovecharse de Esperanza, pero el segundo se lo impide por lo que el estudiante 1 decide pedirle matrimonio para poder tener sexo con ella de forma legítima. Mas tarde, se lleva a Esperanza a su pueblo donde se la presenta a su padre el cual se muestra satisfecho con la pareja de su hijo.
- F) Epílogo: Cuenta que durante las interrogaciones se descubre que Claudia no es la tía de Esperanza si no una niña a quien había tomado de la puerta de la iglesia y que se había aprovechado de otras tres más vendiéndolas

por doncellas a diferentes personas. Por este motivo es sentenciada a recibir cuatrocientos latigazos y a estar en una escalera con una jaula y coraza en medio de la plaza. Por su parte Esperanza vive apaciblemente su matrimonio con el estudiante manchego.

Para concluir este apartado, cabe señalar que las escenas están divididas en la obra por medio de pequeñas elipsis. Cuando hablamos del tiempo de la historia dijimos que este es lineal y efectivamente estas elipsis no suponen una brecha en la historia, pero la hacen avanzar y sirven como elemento de separación de las escenas:

- **A a B** → *Llegóse en esto la noche*
- **B a C** → *cuasi al alba sería*
- **C a D (D')** → *serían las nueve de la noche cuando...*
- **D a E** → *de allí a poco*
- **Epílogo** → aquí no aparece una elipsis al comienzo donde se relata el sino de Claudia, pero si al hablarnos de Esperanza cuando se emplea «*súpose luego*».

d) Vocabulario: *La tía fingida* vs. *El celoso extremeño*

Con respecto al vocabulario nos interesa particularmente examinar la riqueza del mismo. Para este fin utilizaremos otra vez una herramienta digital: *Voyant Tools*⁴⁵. Así mismo, aquí adelantaremos el posterior análisis de *El celoso extremeño* para poder contrastar mejor los datos. Vamos a incluir las dos partes del texto que ya utilizamos en anteriores experimentos (resultado de la fragmentación en longitudes similares en torno a las 6000-7000 palabras). Al utilizar los textos fragmentados seguimos los consejos que nos da la propia herramienta⁴⁶. Los datos que muestra *Voyant Tools* sobre estos textos son los siguientes:

⁴⁵ <https://voyant-tools.org/>

⁴⁶ «The type/token Ratio value can be a useful way of expressing vocabulary richness, but the value is somewhat sensitive to document length and should be considered with circumspection. A more reliable way of measuring vocabulary richness is to average the type/token ratios from equally long segments in a text (e.g., the mean of type/token ratios for each 1,000-word segment in the text). » En <https://voyant-tools.org/docs/#!/guide/documents> (consultado el 29.12.21)

	Título	Palabras	Tipos	Proporción ↑	Palabras/Oración
1	Cervantes_CelosoExtremeno1	6,950	1,749	25%	40.6
2	Cervantes_CelosoExtremeno2	6,747	1,752	26%	39.2
3	Desconocido_LaTiaFingid.	6,992	1,967	28%	46.0

Tabla 9: *Voyant Tools*, datos sobre el vocabulario de la TF y *El celoso extremeño*. Textos fragmentados.

De estos resultados se puede interpretar que la el ratio tipo/token es muy similar en ambos textos. Está claro que esto puede cambiar si empleamos los textos en su totalidad. Al hacerlo obtenemos el siguiente resultado:

	Título	Palabras	Tipos	Proporción	Palabras/Oración
1	Cervantes_CelosoExtremeno	13,697	2,824	21%	39.9
2	Desconocido_LaTiaFingida	6,992	1,967	28%	46.0

Tabla 10: *Voyant Tools*, datos sobre el vocabulario de la TF y *El celoso extremeño*. Textos completos.

Obsérvese que, pese a que *El celoso extremeño* contiene el doble de palabras, la riqueza del vocabulario de *La tía fingida* es algo superior. Esto demuestra que es un texto que tiene una riqueza de vocabulario considerable pese a su brevedad. No obstante, para la comparación es más fiable tener en cuenta los resultados de la tabla superior. El dato de las palabras por oración no es relevante si tenemos en cuenta los comentarios que hicimos sobre la poca fiabilidad de los signos de puntuación de las ediciones del Siglo de Oro por lo cual obviaremos esa ratio. Si ahora utilizamos la herramienta *cirrus* que ofrece *Voyant Tools* para ver las 55 palabras más frecuentes nos da el resultado que se ve en la imagen:



Imagen 19: *Word Cloud* (Cirrus) de las palabras más frecuentes, *La tía fingida*.

Este simple análisis de la riqueza del vocabulario de ambos textos nos lleva nuevamente a la estilometría computacional a través de la herramienta *Voyant Tools*. Los resultados parecen reforzar de forma bastante evidente que existe un parentesco entre ambas obras que con una densidad de vocabulario similar presentan así mismo una relación tipo/token cercano. Nos hemos alejado solo brevemente de los aspectos más puramente literarios (son difíciles de aislar) pero volvemos a ellos al observar las 10 palabras más comunes de ambos textos ya que se reconocen ciertos aspectos que pueden explicar los resultados que vimos en nuestros experimentos con *stylo* pues son manifestaciones que bien podrían ser los marcadores propios de la obra (personajes, espacios), el género (verbos en pasado: narración, relato ulterior) y del autor (palabras funcionales).

No obstante, al tomar como referencia otro texto de características similares, vemos que las palabras más frecuentes responden al mismo patrón. Así, *La inocencia castigada* de María de Zayas tiene una densidad de vocabulario similar de 0.223 y las palabras más frecuentes aportan los mismos rasgos que ya vimos en los anteriores: a (248); no (178); le (101); Inés (84); doña (83); más (65); ella (65); don (64); había (56); diego (53); casa (47). Por este motivo, aunque hay divergencias interesantes (como el diferente uso de los tiempos verbales) no podemos darle gran relevancia a este aspecto. De todos modos, esto podría servir para mostrar que la atribución llevada a cabo por *stylo*, *classify e imposters* a través de los algoritmos tiene un fundamento mucho más global que el hecho de tomar solo algunas palabras frecuentes y la densidad del vocabulario para comparar dos textos.

9.2.2 *El celoso extremeño*

Técnicas narrativas

a) Modo narrativo: focalización, distancia, relato de palabras

La focalización en esta novela también responde a una focalización cero, pues el narrador sabe más que todos los personajes y accede a sus sentimientos más íntimos. En este caso, hay una referencia clara del conocimiento de los pensamientos en el siguiente ejemplo:

«[Carrizales] **hablando consigo mismo**, decía:

Esta muchacha es hermosa, y a lo que muestra la presencia desta casa, no debe de ser rica; ella es niña, sus pocos años pueden asegurar mis sospechas; casarme he con ella; [...] Alto, pues: echada está la suerte, y ésta es la que el cielo quiere que yo tenga.»

Más adelante tenemos otro ejemplo en el que el narrador conoce los pensamientos de Marialonso:

*«No quiso la mala dueña esperar a las reprehensiones **que pensó** le darían los padres de su señora»*

En cuanto a **la distancia**, vemos que hay algunas diferencias con respecto a la TF. Aunque en el relato de los acontecimientos es mayormente el narrador el que habla por sí mismo, hay largas partes en las que se delega en los personajes mediante el diálogo. Tenemos por lo tanto momentos de mayor y menor distancia. Se observa por lo tanto abundancia de narración, pero a diferencia de la TF tenemos mucho más diálogo. En este sentido, a nivel estructural se diferencia más narración en el planteamiento, más diálogo en el nudo y una combinación de ambos, más o menos equilibrada, en el desenlace.

Los diálogos es decir el «relato de palabras» según el término de Genette presentan al igual que en la TF una gran variedad. Hay tanto **discurso narrativizado** en el cual el narrador solo indica que ha habido acto del habla y añade puntualmente el tópico:

«habló con los padres de Leonora, y supo como, aunque pobres, eran nobles; y, dándoles cuenta de su intención y de la calidad de su persona y hacienda, les rogó le diesen por mujer a su hija»

También encontramos a menudo el uso del **discurso transpuesto**. De este tipo se utiliza sobre todo el **estilo indirecto marcado** en el que la oración subordinada se introduce con un verbo del habla:

«pidieron al negro les dijese quién era tan milagroso músico»

En términos generales podemos resumir que se aprecia una mayor presencia del diálogo en el nudo que es a su vez la parte más larga de la novela. Los diálogos están en estilo directo (discurso restituido o citado) e indirecto

siendo el paso de uno a otro es en ocasiones muy fluido y natural. Una muestra es el siguiente pasaje:

«A esto contradijo su señora con muchas veras, diciendo que no se hiciese la tal cosa ni la tal entrada, porque le pesaría en el alma, pues desde allí le podían ver y oír a su salvo y sin peligro de su honra. [Diálogo indirecto, discurso transpuesto en estilo indirecto marcado] ¿Qué honra? dijo la dueña. ¡El Rey tiene harta!» [Diálogo directo, discurso restituído o citado]

Esto contrasta con la TF en la que a pesar de la variedad había un paso más brusco entre las técnicas como por ejemplo en el comienzo de el «Consejo de Hacienda y Estado».

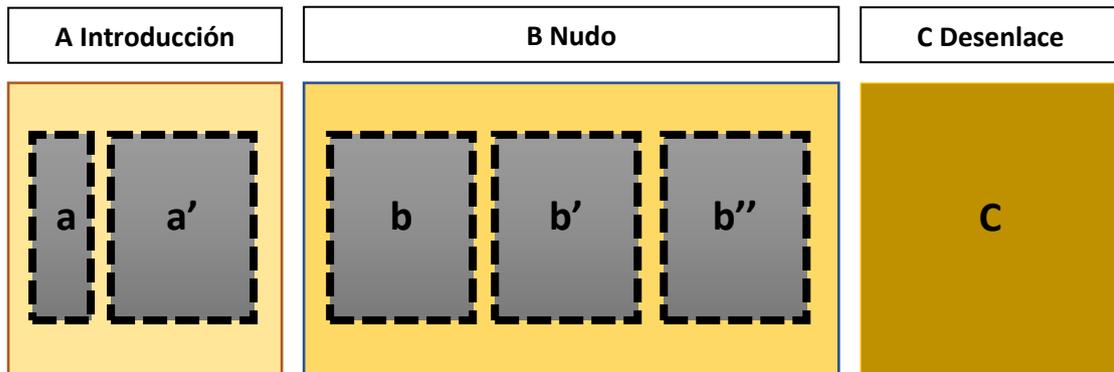
b) Voz narrativa: niveles narrativos, tiempo del relato

En relación con los **niveles narrativos** al igual que sucedía en *La tía fingida* la voz narrativa se mantiene en el nivel **extradiegético**: es la voz del narrador la que lleva adelante la narración. Además, de la misma manera que en la TF se observa a través del uso de las formas verbales que narra en tercera persona (Ej. Al comienzo: *salió, anduvo*, etc.) y no es un personaje de ella por lo que es **heterodiegético**.

Por último, queda por comentar el **tiempo del relato**. La instancia narradora utiliza los tiempos verbales pretéritos, la narración es por consiguiente **ulterior** pues está contada en pasado. El uso de otros tiempos verbales se reserva nuevamente a los diálogos en estilo directo en los que hay más variedad, aunque domina el uso de las formas del presente.

c) Estructura narrativa

En *El celoso extremeño* se diferencian las tres partes básicas que corresponden con el planteamiento, nudo y desenlace de la obra. Dentro de el planteamiento podemos distinguir a su vez dos partes. La parte mas extensa la constituye el núcleo de la novela dentro del cual se pueden distinguir el comienzo, desarrollo y conclusión del ardid. Esquemáticamente podríamos dividir la obra de la manera que se ve en la siguiente página:



A) Introducción

Escena a → Hay una breve introducción en la que se sintetiza la agitada vida de Filipo de Carrizales desde su juventud, con su paso por América, hasta el momento en que, ya anciano, se encuentra casualmente con Leonora.

Escena a' → El anciano habla con los padres de Leonora y se casan. Los celos invaden a Carrizales motivo por el cual convierte su casa en un lugar inaccesible y aislado del exterior. Designa al negro eunuco y viejo como encargado del portal, y a cuatro esclavas blancas y dos negras. También encarga la guarda de la joven muchacha a una dueña llamada Marinesca y como compañía dos doncellas de su edad.

Elipsis: *De esta manera pasaron un año de noviciado*

B) Núcleo

Escena b → Loaysa descubre la casa y tras informarse de sus habitantes decide tratar de expugnarla. Para ello se disfraza de músico pobre y tullido que interpreta junto a la entrada de la casa. Es a través de la excusa de darle clases de guitarra a Luis (el negro) que guarda el portal como consigue tras múltiples triquiñuelas colarse en la vivienda.

Elipsis: *Durmieron lo poco que de la noche les quedaba, y a obra de las seis de la mañana...*

Escena b' → La música despierta el interés de las damas que no pueden ver quien se encuentra en el aislado portal y también por conocer a Loaysa. Por mediación del joven galán planean darle un fuerte somnífero a Carrizales para que este se quede dormido y puedan verse en persona.

Elipsis: *Vino, pues, la noche y la hora acostumbrada de acudir al torno*

Escena b'' →. Seguidamente Loaysa y las mujeres llevan a cabo su plan que en un primer momento funciona. Hay bailes y diversión. Sin embargo, Marinesca trata de aprovechar la situación y promete a Loaysa interceder para que pueda gozar a Leonora. A pesar de esto, cuando llega el momento, la joven evita el adulterio, pero galán y muchacha se quedan dormidos juntos. Es en este momento en el que Carrizales despierta y descubre a la pareja. Ante la sorpresa y el disgusto le da un ataque al corazón.

Elipsis: *Llegóse en esto el día*

C) Conclusión

Escena C → El anciano, ya en su lecho de muerte, decide no vengar su deshonra y deja una buena herencia a Leonora con la única condición de que esta se case con Loaysa. La joven trata de explicar que no hubo ofensa, pero no es escuchada. Finalmente, la muchacha decide tomar los hábitos y Loaysa, despechado, se embarca a América,

Vistas estas partes del análisis reconocemos ya numerosos parecidos de tipo formal y algunas diferencias entre ambas obras. A continuación, vamos a ver el último aspecto de este análisis que se centra en la expresión, y en particular los recursos retóricos y algunos rasgos morfológicos destacable. Como venimos haciendo, comenzaremos con *La tía fingida*.

9.3 Análisis de la expresión

Aunque Castagnino distingue aquí bastantes aspectos más estos son para ser aplicados en cualquier análisis del estilo literario. Para las obras que utilizamos aquí nos vamos a centrar solo en dos: las relaciones de la expresión-intencionalidad, y la morfología y estilo. La primera es una clasificación de las muchas que hay para afrontar los recursos que utiliza la función poética del lenguaje. Recordemos que la función poética del lenguaje es también conocida como función estética y como lo explicaba Jakobson en su publicación *Linguistics and Poetics* (1960) se ocupa de la forma del mensaje, de los estilismos y recursos literarios que este utiliza pues su objetivo es dar énfasis o embellecer la información que se transmite. Esta función está por lo tanto muy

ligada a la literatura y al estilo como elección de los recursos literarios por parte del autor. Para todas las definiciones de las figuras retóricas que vamos a utilizar nos hemos servido del *Diccionario práctico de figuras retóricas y términos afines* de Fernández, V.H. (2007). Advertimos que no se persigue llevar a cabo un análisis completamente exhaustivo, aunque sí presentar aquellos elementos que tienen una presencia y relevancia destacada en las obras.

9.3.1 La tía fingida

a) Las relaciones de la expresión-intencionalidad

Aquí nos centramos en el análisis de aquellas figuras retóricas que ayudan al autor a reforzar la intención estilística. En el método de análisis que propone Castagnino se diferencian dos tipos: de intensificación expresiva y de economía expresiva.

Intensificación expresiva:

Sinonimia

Figura de lenguaje que consiste en usar palabras de significado semejante, para reforzar la expresión de una idea. Encontramos pocos ejemplos evidentes. Uno es el que utiliza Félix para insistir en su buena voluntad diciendo que no pretende la «*Deshonra y menoscabo*» de las damas. También es empleado por el narrador para hablar del plan de la dueña y resaltar que se trata de un: «*embuste o enredo*». Otra muestra es la que utiliza nuevamente el narrador para remarcar que los dos estudiantes manchegos estaban preparados «*prontos y listos*» para liberar a las damas de la justicia. Esta disposición se refuerza aún más en esta frase en la que se emplea la figura que vamos a ver a continuación: «*que si fuera para ir a algún solemne banquete*» (hipérbole).

Hipérbole

Tropo que consiste en aumentar o disminuir excesivamente la verdad de lo que se dice. Encontramos algunas como: «*con otras **mil circunstancias** con que quedó don Félix -que así se llamaba el caballero-, satisfecho*» «*le parecía se tardaba **mil años***». O como la siguiente, utilizada para describir las tocas de

doña Claudia y enfatizar exagerando que son blanquísimas mediante un símil y larguísimas con la hipérbole: «*Una reverenda matrona, con unas tocas **blancas como la nieve, más largas que una sobrepelliz de un canónigo portugués***». También se usan con frecuencia hipérboles que utilizan el concepto del globo terráqueo para la exageración de la totalidad o cantidad como: «*estás en Salamanca, que es llamada en todo el mundo madre de las ciencias*» (epíteto), «*sin que haya persona en el mundo que pueda decir otra cosa*», «*todo el mundo se engaña*», «*nadie sabrá en el mundo*», «*la más fea y abominable catadura del mundo*».

Paralelismo

Figura de lenguaje –sintáctica– que consiste en repetir estructuras dentro de un texto. El Paralelismo consiste en la repetición de la estructura gramatical y pertenece al grupo de figuras de dicción. Aquí el narrador utiliza este recurso⁴⁷ para enfatizar y advertir de las consecuencias que para todas las personas tiene la forma de actuar de las protagonistas mediante la reiteración conjunción + verbo tener + sustantivo + conjunción + sustantivo: «***tal fuerza tiene la discreción y hermosura, y tal fin y paradero tuvo la señora doña Claudia de Astudillo y Quiñones, y tal le tienen y tendrán todas cuantas su vivir y proceder tuvieren.***».

En el siguiente ejemplo encontramos la repetición de la estructura gramatical: Verbo + (preposición + artículo = contracción) + adjetivo. Sirve aquí para enfatizar que Esperanza es concedora de lo que se debe hacer en su negocio con cada persona en base a sus defectos:

«incitar al tibio, provocar al casto, negarse al carnal, animar al cobarde, alentar al corto, refrenar al presumido, despertar al dormido, convidar al descuidado, acordar al olvidado, requerir al... escribir al ausente, alabar al necio, celebrar al discreto, acariciar al rico y desengañar al pobre?»

Otro ejemplo, es el utilizado por una de las dueñas en un pasaje que enfatiza las cualidades de Esperanza mediante este recurso. En este caso la

⁴⁷⁴⁷ Si atendemos solamente a la repetición de la conjunción consecutiva ‘tal’ podríamos considerar que se trata de polisíndeton. Sin embargo, justificamos esta clasificación en base a la repetición de la estructura gramatical. En todo caso, podríamos afirmar que se utilizan las dos figuras retóricas. A nivel morfológico encontramos además el empleo del polípote, o derivación que es la repetición de una palabra utilizando diferentes morfemas flexivos.

repetición de la estructura gramatical consiste en la repetición del adverbio 'muy' + adjetivo: «*porque es mi señora muy principal, muy honesta, muy recogida, muy discreta, muy graciosa, muy música, y muy leída y escribida»*

Eufemismo

Figura de pensamiento que consiste en atenuar la expresión de una idea, para que no resulte demasiado dura, violenta o malsonante. Los eufemismos que hemos encontrado en la TF tratan la temática sexual. Ya al comienzo de la novela se emplean eufemismos para referirse a la profesión de las personas que podrían vivir en la casa de las protagonistas:

«hay casas, así en Salamanca como en otras ciudades, que llevan de suelo vivir siempre en ellas mujeres cortesanas, y por otro nombre trabajadoras o enamoradas.»

Más adelante, cuando don Félix quiere tener relaciones sexuales con Esperanza cueste lo que cueste, afirma: «*no saldré de la tienda sin comprar toda la pieza*». Sobre su virginidad añade que le hubiera gustado ser el que la desvirgara de la siguiente forma: «*ser el primero que esquilmará este majuelo o vendimiara esta viña». La virginidad es denominada por medio de un eufemismo comúnmente utilizado: «*la flor*». Esta ha sido tres veces vendida: «*Tres flores he dado y tantas ha vuesa merced vendido*». El corregidor al detener a las damas hacer referencia a esta virginidad tres veces vendida con una forma eufemística y muy original al denominar a Esperanza de la siguiente manera: «*colegial trilingüe en el disfrute de su heredad*». También el narrador cuenta que la dueña le dice a Félix que la joven «*estaba de tres mercados, o [...] de tres ventas*». Otro eufemismo de uso común que evita hablar explícitamente de sexo, acto sexual o cópula es el que utiliza el verbo gozar, es decir disfrutar: «*queriendo el que la hubo quitado a la justicia gozarla aquella noche*».*

Símil o comparación

El símil es una figura de pensamiento que consiste en comparar dos cosas entre sí. Aparece una comparación de las cuentas que lleva doña Claudia que se parecen a las de Santenuflo: «*cuentas sonadoras, tan gordas como las de Santenuflo*». Este es un personaje desconocido recurrente en algunas obras de diferentes autores al que Icaza le dedica todo un capítulo (Icaza, 1917: 139). El

siguiente símil hace referencia a la apariencia de los gusanos, imaginamos que las mujeres estarían dormidas y envueltas en sus sábanas: «*estaban de dos dormidas, como gusanos de seda*. Encontramos nuevamente un símil para recordar la pulcritud de la dama. Es una comparación con el estado más inmaculado del cuerpo que es el del nacimiento: «*pulcela como su madre la parió*». Más adelante se hace una comparación con las aves. Las golondrinas no suelen volar en grandes grupos. Aunque son abundantes en España, se entiende que la comparación se hace en relación con la percepción de cantidad que dan. De la plata se añade que la gastan como si fuera hierro un metal mucho menos valioso: «*los vizcaínos, aunque son pocos, como las golondrinas cuando vienen, así gastan en su servicio y sustento la plata, como si fuese hierro*». También encontramos un doble símil en el que se destaca la cualidad de tener de todo con el ejemplo de los boticarios y se hace lo que entendemos como una crítica a la alquimia para la que sus pretensiones de transformar metales como el plomo en oro nunca se cumplieron y por lo tanto se dice que hay lo que hay: «*Los extremeños tienen de todo, como boticarios y son como la alquimia: que, si llega a plata, lo es [...]*». Por último, para hablar sobre la reconstrucción del himen que lleva a cabo la vieja para aparentar la virginidad de Esperanza se utiliza un símil con función eufemística: «*Dar puntadas en ella, [mis carnes] como en ropa descosida o desgarrada*».

Metáfora

Esta figura ha sido tradicionalmente considerada una comparación abreviada; por ella se designa una cosa mediante el nombre de otra con la cual tiene una relación de semejanza. En la metáfora se produce un desplazamiento semántico; La designación no está usada en sentido literal sino figurado. Por usar las palabras en sentido traslaticio, la metáfora es un tropo. Son bastante abundantes en la novela, así ya al comienzo tenemos una de las primeras cuando se dice que los estudiantes son «*más amigos del baldeo y rodancho que de Bártulo y Baldo*», es decir quizás más aficionados a la bebida (baldear = pasar un líquido de un vaso a otro) y la acción (rodancho = tipo de escudo) que al estudio de la obra de los dos juristas italianos. Estos personajes son además «*deshollinadores de cuantas ventanas tenían albahacas con tocas*» esto es que son dado al cortejo, mujeriegos. La casa en la que viven las mujeres es «*de buen*

peaje» por lo tanto de elevado alquiler. Para decir que Grijalba quedó estupendamente se utiliza: «*quedó hecha una amapola*». Claudia por su parte emplea también una metáfora para expresar marcadamente que nunca había recibido unos golpes como los que le propina Grijalba cuando dice: «*desde que Dios me arrojó en este mundo*». También encontramos esta figura retórica en la descripción de Félix que era «*de los del campo través*» lo cual quiere expresar figuradamente que tendía a dejar el camino y atravesar el campo, es decir, salirse de las normas o de lo convencional. Después se comenta la facilidad con la que la dueña cuenta toda la verdad a Félix para lo que no fue necesario «*dar otra vuelta al cordel del ruego, ni atezarle los garrotes*». La misma dueña, la Grijalba, se refiere a Esperanza con una metáfora cuando le dice a Félix «*suya es la joya*». Después este se prepara para ir a la casa de Claudia y va «*hecho un San Jorge*» es decir como un heroico caballero con armadura, figura en la que se podría ver también una hipérbole.

Pregunta retórica: erotema

Es una figura de pensamiento. Consiste en la realización de una pregunta de la que no se espera respuesta por estar ya contenida en la interrogación o porque no hay posibilidad de darle contestación. Se utiliza para enfatizar ideas o sentimientos y la encontramos tan sólo en el siguiente pasaje en el que, ante la sorpresa del descubrimiento de Félix en la casa, Claudia dice: «*¿Qué gran desventura y desdicha es ésta? [...] ¿Hombres en mi casa, y en tal lugar y a tales horas? [...] ¿Y mi honra y recogimiento? ¿Qué dirá quien lo supiere?*»

Alegoría

Es un tropo que consiste en una serie de metáforas, cuyo conjunto ofrece dos sentidos: uno recto, literal, y otro figurado, de manera que expresa una cosa, pero dando a entender otra distinta. Sobre los jóvenes se comenta su actitud depredadora por su similitud a la de los cuervos: «*la condición e inclinación de los dos manchegos era la misma, que es la de los cuervos nuevos, que a cualquier carne se abaten, vista la de la nueva garza, se abatieron a ella*». Tenemos otro ejemplo de alegoría en el siguiente fragmento en el que se presentan una serie de metáforas marinas en las que doña Claudia advierte a

Esperanza sobre los peligros y le aconseja que se deje guiar por ella para que su negocio no se eche a perder.

*«Mira, pues, Esperanza, con qué variedad de gentes has de tratar, si será necesario, **habiéndote de engolfar en un mar de tantos bajíos e inconvenientes, te señale yo y enseñe un norte y estrella por donde te guíes y rijas, porque no dé al través el navío de nuestra intención y pretensa, que es pelallos y disfrutallos a todos; y echemos al agua la mercadería de mi nave»***

También Claudia añade otra pequeña alegoría marinera en su conversación con Teresa en relación a la actitud de los hombres. Entendemos que expresa la disposición de las posibles víctimas que no siempre buscan a una mujer pero que eso depende del estímulo: *«no todas veces lleva el marinero tendidas las velas de su navío, ni todas las lleva cogidas, porque según es el viento, tal el tiento»*.

Prosopografía

Podríamos considerar esta figura dentro de los estímulos sensoriales y en particular de la visión. Sin embargo, dado que no vamos a utilizar todo el esquema propuesto por Castagnino por las limitaciones que impone este trabajo incluiremos aquí esta figura que puede ser también vista por su forma de llevar a cabo la descripción como una forma de intensificación de la expresión. En efecto, la definición de la prosopografía es: figura de pensamiento pintoresca que consiste en la descripción del aspecto físico de una persona o de un animal. El único ejemplo que encontramos de la misma es el siguiente:

«moza, al parecer, de diez y ocho años, de rostro mesurado y grave, más aguileño que redondo; los ojos negros, rasgados y al descuido adormecidos; cejas tiradas y bien compuestas, pestañas negras y encarnada la color del rostro; los cabellos plateados y crespos por artificio, según se descubrían por las sienes; saya de buriel fino, ropa justa de contray o frisado; los chapines de terciopelo negro, con sus clavetes y rapacejos de plata bruñida; guantes olorosos, y no de polvillo sino de ámbar. El ademán era grave, el mirar honesto, el paso airoso y de garza. Mirada en partes, parecía muy bien, y en el todo, mucho mejor»

Apóstrofe

Figura de pensamiento patética que, mediante una exclamación, expresa un arrebató del ánimo. El primer ejemplo sirve para expresar la sorpresa de doña Claudia cuando descubre a don Félix escondido en su casa en el que también se empleaba la pregunta retórica: «- ¡Jesús, valme! [...] ¡Desdichada de mí! ¡Desventurada fui yo!». La anciana continúa su teatral actuación exclamando:

«- ¡Ay, sin ventura de mí, [...] y a cuántos peligros están puestas las mujeres que viven sin maridos y sin hombres que las defiendan y amparen! ¡Agora sí te echo menos, malogrado de ti, Juan de Bracamonte (no el arcediano de Jerez), mal desdichado consorte mío, que si tú fueras vivo, ni yo me viera en esta ciudad, ni en la confusión y afrenta en que me veo!»

El segundo representa la reacción/comentario del propio narrador ante la propuesta de matrimonio que va a ser realizada por parte de uno de los estudiantes:

«¡Oh, sucesos estraños del mundo! ¡Oh, cosas que es necesario contarlas con recato para ser creídas! ¡Oh, milagros del amor nunca vistos! ¡Oh, fuerzas poderosas del deseo, que a tan estraños casos nos precipitan!»

Figuras de economía expresiva

La Ironía

Figura de pensamiento que consiste en dar a entender, por las palabras o por el tono de voz, lo contrario de lo que se dice. Existen diferentes tipos: el sarcasmo, la antífrasis, el asteísmo, el carientismo, el clenasma y la mímesis. En la TF encontramos los siguientes ejemplos:

→ El sarcasmo

Figura de pensamiento que consiste en una burla feroz; es una ironía mordaz. En el ejemplo vemos una comparación en la cual se observa la ironía con cierto nivel de ofensa pues el oficio del cardador no guarda ninguna relación con el arte de la poesía si no todo lo contrario. En el trabajo que se lleva a cabo para procesar la lana, el cardado es el primer paso que consiste en lavar y desenredar (o escardar) los vellones para poder proceder al hilado y su posterior

uso textil. Se confronta por lo tanto el refinamiento de la poesía con el esforzado pero rudo trabajo del cardador. No obstante, según el narrador este poeta, compone como tal: *«mordiéndose los labios y las uñas, y rascándose las sienes y frente, forjó un soneto, como lo pudiera hacer un cardador o peraille»*

→ Carientismo

Figura de pensamiento que consiste en disimular la ironía o la burla mediante expresiones ingeniosas. Aquí el autor comenta, como vimos con anterioridad, que el poeta es algo tosco y también denomina al poema «descomulgado». Sin embargo, uno de los personajes que hemos denominado «Bellacón extremeño» hace su primera y única intervención para alabar las virtudes del soneto. El autor redobra su burla extendiéndola a todos los presentes debido a que lo ven a partir de su intervención como hombre docto y versado en el arte poética. Esta ironía se podría entender incluso como una crítica social, pues en una ciudad de estudiantes, ninguno de los presentes está capacitado para reconocer la poca calidad del soneto:

«- ¡Voto a tal, que no he oído mejor estrambote en todos los días de mi vida! ¿Ha visto vuesa merced aquel concordar de versos, y aquella invocación de Cupido, y aquel jugar del vocablo con el nombre de la dama, y aquel imploro tan bien encajado [...] y de allí adelante quedó en opinión de todos por hombre docto y versado en la arte poética, sólo por haberle oído desmenuzar tan en particular el cantado y encantado soneto.»

→ Mímesis

Imitación de una persona, tanto del modo de hablar como de los gestos, que habitualmente se hace con ánimo de burla. En este caso, el narrador introduce las interjecciones que utiliza la dueña para expresar asombro y su comentario sobre las perlas para denominarla burlonamente «dueña del huy y las perlas» añadiendo además el adjetivo «repulgada» para expresar la falta de naturalidad, lo forzado de la mujer:

«Huy, huy -dijo la dueña-, en eso, por cierto, está mi señora doña Esperanza de Torralba Meneses y Pacheco. Sepa, señor mío [...] no hará lo que vuesa merced le suplica aunque la cubriesen de perlas.»

«Estando en este deporte y conversación con la repulgada dueña del huy y las perlas»

b) La morfología y estilo

El adjetivo (adjetivo + adjetivo)

Entre las características más evidentes de la obra, destaca la presencia de múltiples adjetivos. Estos aparecen a menudo en grupos de dos unidos mediante la conjunción copulativa 'y' de forma reiterativa. No haremos una enumeración de todas, pero algunos ejemplos son: *mancebos y manchegos, hermosa y honesta, mesurado y grave, cejas tiradas y bien compuestas, benditos y corteses, pensativos y medio enamorados, etc.*

El sustantivo (sustantivo + sustantivo)

Se observa que lo mismo que pasa con los adjetivos se repite con los sustantivos. Esta combinación tiene menos frecuencia que los primeros, no obstante, es abundante: *Seda y lana, donaire y belleza, voz y guitarra, calle y casa, profesión y estado, utilidad y servicio, músicos y acompañados, refacción y deshecha, gravedad y fausto, etc.*

El Epíteto

Se llama así al adjetivo que expresa una cualidad que es propia o natural del nombre al que se aplica, al adjetivo que caracteriza al nombre. Puede no limitarse al adjetivo y consistir también en añadir un complemento a un sustantivo por lo general, antepuesto. No son numerosos: *buena señora, buen escudero, buen sufrimiento* (que es además un oxímoron), *vieja Claudia*. También encontramos una utilizada en el soneto para insistir en el gran tamaño del gigante: *Será quien la alcanzare un gran gigante.*

El adverbio

Si observamos los adverbios terminados en -mente: encontramos tan solo ocho adverbios que utilizan el sufijo -mente: *Principalmente, brevemente, juntamente, medianamente, cortésmente, extraordinariamente, juntamente, fácilmente.* Quizás el limitado empleo de estos tiene relación con la preferencia

del uso de adjetivos. El empleo del adjetivo 'principal', por ejemplo, aparece en tres ocasiones frente a una del adverbio.

Conjunciones: polisíndeton

Figura de construcción que consiste en repetir las conjunciones para dar mayor fuerza al discurso. Encontramos solo una muestra con la conjunción copulativa 'ni'. El énfasis en este caso lo hace Grijalba para reforzar la idea de que no hay persona que se iguale a don Félix: «*ni papa, ni emperador, ni Fúcar, ni embajador, ni cajero de mercader, ni perulero, ni aun canónigo*»

Aunque se podría seguir profundizando en el análisis de todos los elementos que constituyen la morfología y estilo de esta novela (artículos, otros adverbios, más combinaciones morfológicas, el pronombre, las interjecciones, locuciones, conjunciones) vamos a concluir en este punto para evitar alargar esta tesis *ad infinitum*. La profundización hasta llegar a un análisis integral completo tal y como lo comenta De Leal (1984: 196) daría lugar a un comentario excesivamente amplio. Para ello propone dos soluciones: no agotar el análisis seleccionando los fenómenos más destacados o la selección de un solo plano para el análisis (*Ibid.*,197). En este caso hemos optado por la primera de las propuestas: lo que hicimos fue presentar los contenidos de ambas obras y después analizar la forma en la que estos se emplean por medio de la selección de las características retóricas y morfológicas que nos resultaron más llamativas.

Con respecto a la sintaxis ya comentamos en el apartado sobre la problemática de los textos del Siglo de Oro los límites en cuanto a la fiabilidad. En este sentido dado que las obras eran a menudo producto de modificaciones por parte de copistas e impresores hemos preferido dejarla, al menos de momento, a un lado. Por lo tanto, continuaremos con el análisis de la expresión de *El celoso extremeño* para luego pasar a la síntesis del análisis literario de ambas obras.

9.3.2 *El celoso extremeño*

a) Las relaciones de la expresión-intencionalidad

En el análisis de esta segunda obra evitaremos repetir las definiciones de las figuras retóricas que ya vimos, mostrando tan solo los ejemplos y puntualmente su explicación.

Intensificación expresiva:

Sinonimia

Encontramos algunos ejemplos, no siendo estos numerosos. Algunos de ellos son: el que se aplica a Carrizales que cuando fue a Indias era «*pobre y menesteroso*»; el relativo al veneno que según dice Loaysa no hace «*mal ni daño*»; las intenciones de este personaje que son según él dar «*gusto y contento*» y el estado del anciano al final que está «*atónito y embelesado*»

Hipérbole

Tenemos una mayor presencia de este recurso que en la TF. Hay numerosos ejemplos que destacan determinados momentos o resaltan mediante la exageración conceptos relacionados con el argumento.

Una de las hipérbolas más evidentes es la que transmite Luis a Loaysa sobre el extremo defecto de su amo, que más allá de ser celoso: «*es el más celoso hombre del mundo*». De manera similar, Loaysa utiliza este recurso para adular a Luis. Animándole a recibir sus clases de guitarra apunta: «*que la mejor voz del mundo pierde de sus quilates cuando no se acompaña con el instrumento*». Esto se repite cuando Luis comenta a las mujeres quien es su maestro: «*el mejor músico que hay en el mundo*». Más tarde cuando la dueña recibe la llave de la casa que Leonora sustrae, lo hace: «*con la mayor alegría del mundo*».

Una de las esclavas en su conversación con Luis, expresa su desesperación por tener contacto con el mundo exterior, más allá de los muros de la casa, Luis les exige discreción para evitar que el señor Carrizales se entere, a lo que una esclava dice: «*Callaremos más que si fuésemos mudas*» y continúa «*porque te prometo, amigo, que me muero por oír una buena voz*»

Ante el «endemoniado» son de la Zarabanda que interpreta Loaysa, hay un comentario del narrador que expresa el efecto que este tiene sobre su público no acostumbrado a tales estímulos. Esta música forma parte del planeado cortejo de Loaysa y explica la posterior actuación de las damas de ahí la exageración: «*No quedó vieja por bailar, ni moza que no se hiciese pedazos*»

Leonor, se acerca al anciano para untarle con el somnífero ungüento con gran temor. Su miedo se reitera con la expresión: «*casi sin osar despedir el aliento de la boca*»

El ungüento es muy efectivo en primera instancia, esto es reforzado por medio de la siguiente figura que emplea Leonora: «*Dame albricias, hermana, que Carrizales duerme más que un muerto*».

Después, se vuelve a insistir en el efecto del veneno. Loaysa asegura la calidad del ungüento y apremia a las mujeres a hablar alto y sin temores mientras Carrizales duerme. Leonora está de acuerdo, y comenta que dada su obsesión que le quita el sueño: «*si así no fuera, ya él hubiera despertado veinte veces*».

Se utiliza también para resaltar la atracción que siente la dueña por Loaysa que explica en parte por qué realiza el malicioso trato con él: «*A trueco de cumplir el deseo que ya se le había apoderado del alma y de los huesos y médulas del cuerpo*»

Más tarde, en la conversación en la que Marialonso trata de convencer a Leonora de las virtudes del joven pretendiente, el narrador refuerza la idea de la capacidad de convicción de ésta cuando afirma que las palabras de la ladina dueña son capaces de mover no solo el corazón de «*la simple e incauta Leonora, sino el de un endurecido mármol*». Dentro de esta hipérbole se utiliza además el pleonismo «*endurecido mármol*» con el que se insiste en una cualidad que es evidente y propia de esta roca, lo que subraya aún más la carencia de sentimientos de la interesada dueña.

Cuando Carrizales se despierta y descubre a su mujer durmiendo con Loaysa, se consolidan todos sus miedos. El narrador nos describe aquí una reacción cargada de retórica que remarca la intensidad de los sentimientos:

«Con la amarga vista [sinestesia⁴⁸] de lo que miraba [pleonasma; vista de lo que miraba]; la voz se le pegó a la garganta [hipérbole], los brazos se le cayeron de desmayo, y quedó hecho una estatua de mármol frío» [hipérbole].

Se puede observar que de manera general la hipérbole se utiliza de forma no indiscriminada y bien meditada. Es utilizada en determinados pasajes en los que el autor trata de remarcar los efectos de objetos (el veneno), sentimientos (por ejemplo: la decepción o sorpresa del anciano, el temor de la joven, la atracción de la dueña) o acciones cuando estas son de relevancia para el argumento (dormir, despertar, respirar).

Paralelismo

Aparecen en esta obra también algunos ejemplos de esta figura. En particular en el momento en que Marialonso convence a Leonora para que acuda junto a Loaysa. En este párrafo hay, al igual que el encuentro de Carrizales con los jóvenes durmientes, varios recursos retóricos que veremos por separado. El paralelismo que encontramos aquí incide en el efecto que tuvieron los argumentos de la dueña sobre Leonora: «*Leonora se rindió, Leonora se engañó y Leonora se perdió*» (nombre propio + pronombre indirecto + verbo 3ª persona del indefinido)

Símil o comparación

El uso de las comparaciones retóricas lo encontramos nuevamente en pasajes en las que los acontecimientos de la historia son intensos. Ya hablamos sobre el veneno, que es un elemento importante del plan de Loaysa. Sobre el exceso en su aplicación, el narrador comenta: «*fue lo mismo que haberle embalsamado para la sepultura*». Más adelante Marialonso exige a Loaysa que jure no excederse con las damas de la casa, aprovecha para remarcar que pese a su edad ella es virgen al igual que el resto de las doncellas con excepción de la deseada Leonora. La virginidad es expresada aquí mediante un símil que ya habíamos visto en la TF «*desta casa somos doncellas como las madres que nos parieron, excepto mi señora*».

⁴⁸ Evitaremos fragmentar aún más el análisis incluyendo la sinestesia y el pleonasma que se produce en este pasaje. La sinestesia, según el modelo de Castagnino pertenecería a los estímulos sensoriales, un aspecto que no vamos a tratar aquí.

En el pasaje del descubrimiento de Carrizales que ya comentamos, el uso de figuras retóricas incluye el siguiente símil que compara el sueño de la pareja con el del anciano «[Carrizales vio a la pareja] *durmiendo tan a sueño suelto como si en ellos obrara la virtud del unguento*». Al final, Leonora está arrepentida, aunque esté ausente de culpa. Se resalta aquí mediante la comparación, la forma en la que cuida al enfermo Carrizales y se interesa por su estado: «*preguntándole qué era lo que sentía, con tan tiernas y amorosas palabras, como si fuera la cosa del mundo que más amaba*».

En su último parlamento, el anciano resume el devenir de los acontecimientos con un símil en el que se compara a sí mismo con un gusano de seda y a su casa con el capullo que los protege. «*Yo fui el que, como el gusano de seda, me fabriqué la casa donde muriese*». Puede entenderse también como una frase metafórica pues el gusano no muere, si no que se tras su etapa de crisálida se transforma en mariposa y emprende el vuelo (ascenso al cielo: la muerte). Este símil es similar a uno que vimos en la TF en la que Claudia y Esperanza estaban dormidas como «*gusanos de seda*»

Metáfora

El uso de la metáfora es recurrente y se encuentran numerosos ejemplos. Trataremos de abreviarlos señalando el pasaje o escena con el que se relacionan.

La visión que Leonora tiene de su esposo al comienzo utiliza dos metáforas consecutivas que forman un símil: «*La plata de las canas [Metáfora] del viejo, a los ojos de Leonora, parecían cabellos de oro puro [Metáfora][Símil]*».

El sueño del viejo y sus consiguientes ronquidos son de gran agrado para Leonora que espera poder recibir al músico esa noche en su patio: «*comenzó a dar el viejo tan grandes ronquidos, que se pudieran oír en la calle: música, a los oídos de su esposa, más acordada que la del maeso de su negro*»

Marialonso exige un juramento a Loaysa previa entrada a la casa. Con esta metáfora se remarca la división exterior-interior como si de dos diferentes reinos se tratara: «*vuesa merced nos ha de hacer [juramento], primero que entre en nuestro reino*»

La reacción de las damas al ver a Loaysa en persona está llena de metáforas que inciden en lo perfecto de su apariencia con mención a sus dientes y sus ojos:

«¡Ay, qué blancura de dientes! ¡Mal año para piñones mondados, que más blancos ni más lindos sean! [metáfora] " Otra: "¡Ay, qué ojos tan grandes y tan rasgados! Y, por el siglo de mi madre, que son verdes; que no parecen sino que son de esmeraldas»

El viejo descubre a los amantes durmiendo y el autor utiliza la siguiente metáfora para que el lector se haga una imagen mental de la manera: *«cogió a los nuevos adúlteros enlazados en la red de sus brazos»*. Después, para el anciano Carrizales, el cuidado y las caricias de Leonora son muy dolorosos. Este dolor se expresa afirmando que *«cada palabra o caricia que le hacía [era] una lanzada que le atravesaba el alma»*. El rencor que siente se expresa también en el siguiente pasaje que abunda en figuras retóricas:

«Oyó la voz de la dulce enemiga [sinestesia] suya el desdichado viejo, y, abriendo los ojos desencasadamente, como atónito y embelesado [sinonimia], los puso en ella, y con grande ahínco, sin mover pestaña [metáfora]»

En el discurso final del celoso marido, se utiliza una metáfora que también aparecía en la TF empleada de manera similar esto es para referirse a la protagonista, cuando se dice *«del mundo, quise guardar esta joya»*

El final de la novela tiene también un buen número de metáforas, primero cuando Leonora y sus padres escuchan el discurso de Carrizales con gran atención, información que el narrador nos dice a través de la intervención del anciano: *«veo la suspensión en que todos estáis, colgados de las palabras de mi boca»*

Después se suceden las reacciones de tristeza por parte de los asistentes que son exaltadas utilizando nuevas metáforas: *«a Leonora se le cubrió el corazón», «y a las gargantas de los padres de Leonora se les atravesó un nudo (metáfora) que no les dejaba hablar palabra»*.

El narrador termina preguntándose por qué Leonora no puso más empeño en explicar el malentendido, a lo que aduce: *«la turbación le ató la lengua, y la priesa que se dio a morir su marido no dio lugar a su disculpa»*

Alegoría

La más evidente de estas figuras retóricas en *El celoso extremeño* es la que se refiere al grupo de mujeres y su modo de actuar ante el galán. Así, el narrador habla del grupo de mujeres que acuden a escuchar la música la siguiente metáfora: «*la banda de las palomas acudió al reclamo de la guitarra.*». Aquí se oculta también el recurso de la animalización pues implícitamente se entiende que las mujeres son palomas. Más adelante cuando Carrizales hace amago de despertarse, se utiliza otra vez una alegoría siguiendo la misma lógica que resalta profusamente la pavorosa huida en una sucesión de metáforas: «*Quien ha visto banda de palomas [grupo de mujeres] estar comiendo en el campo, sin miedo [bailando, disfrutando], *lo que ajenas manos sembraron, que al furioso estrépito de disparada escopeta [ante el aviso de Guiomar] se azora y levanta, [se asusta] y, *olvidada del pasto, confusa y atónita, cruza por los aires [huye despavorida], tal se imagine que quedó la banda y corro de las bailadoras, pasmadas y temerosas, oyendo la no esperada nueva que Guiomar había traído.*».**

También podríamos hablar largo y tendido de las referencias mitológicas como «*él era la ronda y centinela de su casa y el Argos de lo que bien quería.*» Esta y otras referencias alegóricas tienen una función irónica que puede consultarse en Gómez Iñiguez (1990: 637) quien realiza un excelente trabajo sobre la ironía en *El celoso extremeño*.

Prosopografía, pregunta retórica y eufemismo.

En esta novela, al contrario que en la TF, así como encontramos más ejemplos de las figuras retóricas vistas hasta este punto, no hay muestras de prosopografía o pregunta retórica. El uso de los eufemismos tampoco es destacable, quizás porque el tema tratado no es tan escabroso como en el texto dubitado. La muestra de eufemismo más evidente tiene que ver también con la sexualidad, y más en particular con la virginidad pues es la denominación «*doncellas*» que utiliza Marialonso al hablar de las mujeres de la casa.

Apóstrofe

La exclamación o apóstrofe aparece utilizado de forma similar que en la TF por parte del narrador ante las convincentes palabras que Marialonso dirige a Leonora con el objetivo de que lleve a cabo su encuentro con Loaysa. Aquí aclama:

«¡Oh dueñas, nacidas y usadas en el mundo para perdición de mil recatadas y buenas intenciones! ¡Oh, luengas y repulgadas tocas, escogidas para autorizar las salas y los estrados de señoras principales, y cuán al revés de lo que debíades usáis de vuestro casi ya forzoso oficio!».

Inmediatamente después de este pasaje, el narrador repite este recurso después de enumerar las precauciones que tomó Carrizales y recordar lo poco que le sirvieron:

«¡Todo aqueso derribó por los fundamentos la astucia, a lo que yo creo, de un mozo holgazán y vicioso, y la malicia de una falsa dueña, con la inadvertencia de una muchacha rogada y persuadida!"»

De la misma manera, el anciano al día siguiente del descubrimiento se lamenta patéticamente transmitiendo sus emociones así: *«¡Desdichado de mí, y a qué tristes términos me ha traído mi fortuna!».*

Por último, el narrador remarca nuevamente el patetismo de la escena del discurso del moribundo anciano a los padres de Leonora exclamando: *«¡estraño y triste espectáculo para los padres, que a su querida hija y a su amado yerno miraban!»*

Figuras de economía expresiva

La ironía

En cuanto a la ironía se reconocen diferencias sustanciales entre las dos obras. En la TF encontramos apenas tres muestras utilizando diferentes tipos de ironía. Esta era en dos de los casos muy evidente (sarcasmo y mímesis). Por otra parte, en *El celoso extremeño* hay ironía en toda la obra presentado de forma evidente y también velada. Gómez Iñiguez (1990) de quien ya hemos hablado sobre su trabajo al respecto de la ironía, desgrana el uso de este elemento tan cervantino. En él, divide este recurso en tres tipos: la ironía verbal,

la dramática y la ironía del sino (Gómez Iñiguez, 1990: 634). Basándose en ellos, reconoce que el tipo dominante es diferente en cada una de las tres partes principales que componen la estructura de la obra: en la primera parte, abunda la ironía verbal (juegos de palabras, ambigüedades léxicas, etc.); en la segunda, con la aparición de Loaysa en la trama, la obra gana en teatralidad y es la ironía dramática (apoyada sobre la verbal) la que domina; en la tercera parte, aparece la ironía del sino con Carrizales. (cfr. Gómez Iñiguez, 1990: 634). Existen por lo tanto numerosos aspectos que se podrían analizar a este respecto, desde los excesos de Carrizales en la primera parte, pasando por el disfraz y la preparación teatral del engaño en la segunda (proceso del cual el lector es cómplice) o el paradójico final en el que se cristalizan todos los miedos del anciano en la tercera. Evitaremos aquí extendernos demasiado recomendando el artículo de Gómez Iñiguez (1990) como referencia y pasaremos a ver algunos ejemplos de sarcasmo y carientismo.

→ El sarcasmo

En la segunda parte, entre la ironía dramática encontramos una muestra de sarcasmo (ironía verbal) cuando la dueña espeta a Luis el siguiente comentario que evidencia la falta de talento musical del eunuco: *«Por cierto dijo la dueña que, si no es algún demonio el que te ha de enseñar, que yo no sé quién te pueda sacar músico con tanta brevedad.»*

→ Carientismo

Podríamos identificar como tal al uso de la silepsis como recurso de ironía verbal. El calificativo «liberal» como explica también Gómez Iñiguez (1990: 636) es empleado de forma polisémica (con sentido recto y figurado). Se puede encontrar a lo largo de la novela hasta en cuatro ocasiones para referirse al protagonista evidenciando el uso premeditado por parte del autor. La contradicción entre la actitud de Carrizales con su esposa y el adjetivo es bastante evidente y supone una burla solapada.

→ Mímesis

Ejemplo de este tipo de ironía es el comentario de Guiomar al quedar como guarda de Carrizales: *«¡yo, negra, quedo; blancas, van; Dios perdone a todas!»*.

Aquí Cervantes, imita de forma burlesca la manera de hablar de la portuguesa en un perfecto ejemplo de parodia que después explicita diciendo a través del narrador: «*Guiomar, la negra, que por ser portuguesa y no muy ladina, era extraña la gracia con que la vituperaba*».

b) La morfología y estilo

El adjetivo (adjetivo + adjetivo) y el sustantivo (sustantivo + sustantivo)

En la TF encontrábamos que se empleaba esta construcción con mucha frecuencia, se observa que aquí hay también diferencias con *El celoso extremeño* pues es mucho menos común. Algunos ejemplos son blando y próspero, anchas y espaciosas, diversas y buenas, etc. Lo mismo sucede también con los sustantivos en los que no encontramos esta construcción empleada con el mismo exceso del que hace gala la TF. Por ejemplo, industria y diligencia, quietud y sosiego, eficacia y vehemencia, recogimiento y recato.

El epíteto

Llama la atención al comienzo de la obra la enumeración de ingeniosos epítetos con los que se denomina a América:

«las Indias, refugio y amparo de los desesperados de España, iglesia de los alzados, salvoconduto de los homicidas, pala y cubierta de los jugadores [...] añagaza general de mujeres libres, engaño común de muchos y remedio particular de pocos.»

Abundan también los epítetos algunos de ellos irónicos principalmente se utilizan con los personajes principales resaltando sus virtudes o defectos:

- Carrizales: *El buen viejo Carrizales, el recatado Carrizales, el viejo Carrizales, el discreto Carrizales, el celoso viejo, el lastimado viejo, el celoso anciano*
- Leonora: *tierna Leonora, simple Leonora, ignorante Leonora, simple e incauta Leonora, desmayada Leonora*
- Loaysa: *el buen Loaysa*
- Marialonso: *la buena dueña, la maldita dueña, la pestífera dueña, la mala dueña*

Vemos que donde la TF utilizaba este recurso en contadas ocasiones, en esta obra es de uso frecuente para reforzar y acotar las características de los personajes o su estado físico/anímico. También se observa que algunos de ellos tienen un matiz claramente irónico: *el buen Loaysa, la buena dueña*; utilizados en pasajes en los que su actuación es reprochable.

El adverbio

En comparación con la TF el empleo de los adverbios con el sufijo -ente son mucho más abundantes. Si en la TF encontrábamos tan solo ocho muestras en *El celoso extremeño* tenemos veinticuatro⁴⁹, el triple, lo cual revela un uso más frecuente a pesar de la diferente longitud de los textos. Esta elección estilística es una divergencia bastante obvia dado que el empleo de estos adverbios revela una elección a nivel paradigmático entre un número limitado de opciones, por ejemplo: perpetuamente, de manera/forma perpetua, con perpetuidad. Los adverbios de este tipo no se repiten en ningún caso por lo que muestran bastante variedad.

Conjunción

Polisíndeton: encontramos este recurso cuando Loaysa en un alarde de cinismo se describe a sí mismo alabando sus virtudes. El autor utiliza aquí esta figura retórica para remarcar por medio de la conjunción consecutiva 'tan' lo sublime de las cualidades que se atribuye a la vez que ralentiza el ritmo del discurso del personaje en un guiño que quizás refiere a la charlatanería del mismo: «*tan sencillo, tan manso y de tan buena condición, y tan obediente*»

Asíndeton: aparece al principio cuando el narrador nos cuenta como fue el regreso a España del ahora rico Carrizales. Este recurso consigue sintetizar aún más y con gran agilidad el desarrollo de los acontecimientos:

«*Desembarcó en Sanlúcar; llegó a Sevilla, tan lleno de años como de riquezas; sacó sus partidas sin zozobras; buscó sus amigos: hallólos todos muertos; quiso*

⁴⁹ Son los siguientes: finalmente, comúnmente, estrechamente, juntamente, sospechosamente, suavemente, mansamente, ahincadamente, encarecidamente, especialmente, estremadamente, delicadamente, largamente, blandamente, estrechamente, desencasadamente, brevemente, dolorosamente fingidamente, perpetuamente, ordinariamente, fácilmente, honradamente, amargamente.

partirse a su tierra, aunque ya había tenido nuevas que ningún pariente le había dejado la muerte.»

Es utilizado por Carrizales en su discurso dándole agilidad e imprimiendo cierta pasión al pasaje:

«Alcé las murallas desta casa, quité la vista a las ventanas de la calle, doblé las cerraduras de las puertas, púsele torno como a monasterio; desterré perpetuamente della todo aquello que sombra o nombre de varón tuviese. Dile criadas y esclavas que la sirviesen, ni les negué a ellas ni a ella cuanto quisieron pedirme; hícela mi igual, comuniquéle mis más secretos pensamientos, entreguéla toda mi hacienda»

Llegados a este punto, terminaremos nuestro análisis en base a las razones que ya justificamos anteriormente. Vamos pues a continuar con una síntesis en la que expondremos de manera breve la comparación de los elementos analizados en ambos textos.

10 Síntesis del análisis literario

En este apartado hemos llevado a cabo un repaso por algunas características relevantes de ambas obras. Siguiendo el modelo de Castagnino (1987), analizamos los contenidos, las técnicas literarias y por último la expresión. En el nivel de los contenidos nos encontramos con dos obras pertenecientes a un mismo género, el de la novela breve y se reconoce un mismo tipo, el realista. La semejanza del argumento se encuentra en sus protagonistas (las jóvenes y bellas Esperanza y Leonora) que son cortejadas por un galán (Félix / Loaysa) y en las que aparece un personaje con tintes celestinescos (Claudia, Marialonso). También los espacios principales presentan semejanzas (la casa, la calle) y en cuanto al tiempo de la historia, esta sigue en ambos casos una cronología lineal.

El segundo aspecto, el de las técnicas literarias sigue el mismo camino, hallamos semejanzas en la focalización (cero) aunque esta sea más evidente en *El celoso extremeño*. El narrador es extradiegético, heterodiegético y la narración es ulterior en las dos obras por lo que el uso de determinadas formas verbales

propias de la narración en tercera persona y en pasado son elementos que se repiten en las dos novelas. También encontramos en sendos textos una combinación de diálogo y narración recayendo la diégesis siempre en el narrador. La mayor diferencia aquí, estriba en que la obra no dubitada da muestras de un mayor empleo del relato de palabras (diálogo) tanto transpuesto como marcado (indirecto) y restituido (directo). Además, las transiciones de una de las técnicas a otra son en la mayor parte de los casos muy fluidas mientras que en la TF resultan más repentinas y bruscas. La estructura narrativa por su parte, es básica en ambos casos y sigue las tres partes elementales de planteamiento, nudo y desenlace separadas por breves elipsis, aunque en *El celoso extremeño* el nudo es mucho más extenso y complejo que en la primera.

Con respecto al vocabulario, nos ocupamos de la riqueza y vimos que ambas obras presentan una densidad de vocabulario y una ratio tipo/token similar, tanto si dividimos la novela del celoso como si no. En este ámbito llama la atención que las diez palabras más frecuentes de ambas obras son muestras que podrían revelar el género (verbos del habla en pasado: narración), la obra (personajes y espacios) y lo más importante para nosotros, el autor (palabras funcionales). Sin embargo, también vimos que este es un patrón que se repite en otras novelas de nuestro corpus de características similares, como *La inocencia castigada* de Zayas.

El último análisis que corresponde con la expresión se ocupó de las figuras retórico-literarias (relaciones de la expresión-intencionalidad) y de algunos aspectos morfológicos. Ambas obras dan muestra del empleo de numerosos recursos para la intensificación expresiva. Destaca el uso de la hipérbole que es más abundante en el texto no dubitado. En los dos textos abundan las comparaciones a través de símiles y en metáforas. En estas segundas, llama la atención el uso compartido de metáforas sobre las aves (golondrinas y cuervos en la TF, palomas en *El celoso extremeño*). De la misma manera, las obras exponen el uso aunque solo puntual de otros recursos como la pregunta retórica y el apóstrofe. El eufemismo y la prosopografía la encontramos solo en la TF: el primero, relacionado siempre con la temática sexual; el segundo, para describir a Esperanza.

La gran divergencia entre ambas obras la encontramos en el empleo de la ironía. La TF presenta muestras de sarcasmo, carientismo y mimesis que sirven como contrapunto cómico y alegre al relato. Por su parte *El celoso extremeño*⁵⁰ está empapado de este recurso desde el comienzo afectando a la obra en su totalidad. Se encuentran diferentes tipos de ironía que basándonos en la tipología que expone Gómez Iñiguez, 1990: 634) son la verbal, la dramática y la del sino. También se emplean a nivel verbal el sarcasmo, el carientismo y la mimesis, pero este va mucho más allá y lo encontramos también en la morfología (uso de determinados epítetos).

A nivel morfológico hay varias diferencias sustanciales. La primera es el uso constante y repetido de las construcciones 'adjetivo + adjetivo' y 'sustantivo + sustantivo' que se encuentran en la TF y que, aunque presentes no son tan constantes en *El celoso extremeño*. La segunda es el empleo de los epítetos, mucho más rico y variado en la obra no dubitada. Por último, el uso de los adverbios terminados en -ente es de menor cantidad en la TF (8 frente a 24) incluso teniendo en cuenta la diferente extensión de los textos, aunque variado, sin repeticiones en ambos. En cuanto a las conjunciones, las obras presentan de forma semejante el uso del polisíndeton y el asíndeton sin grandes particularidades diferenciables.

11 Conclusiones

En nuestra configuración y experimentos con *stylo* procuramos que la organización de textos indubitados fuera hecha por autor y no por género, esto debería anteponer la relevancia del empleo de determinadas palabras funcionales a cualquier otro aspecto como marcador del estilo autorial. A través de los experimentos estilométricos llevados a cabo vimos que la organización de los textos por autor de las obras no dubitadas era correcta⁵¹. Al introducir la TF

⁵⁰ El título también puede verse como un guiño irónico pues no es casualidad que el protagonista sea extremeño y su defecto extremo.

⁵¹ Es necesario precisar que, aunque no se hayan incluido en el trabajo, se llevaron a cabo numerosas pruebas previas utilizando los textos sin fragmentar, con diferentes configuraciones de las MFW y utilizando diversas medidas de distancia. Es así como llegamos hasta los parámetros que se incluyen en los experimentos finales.

en el corpus después de realizar la validación y demostrar que el porcentaje de fiabilidad con las 500MFW y la distancia *Cosine Delta* era del 100% vimos que esta obra era asociada a Cervantes y más particularmente a *El celoso extremeño*. Además, llevamos a cabo una serie de experimentos estilométricos utilizando diferentes funciones para determinar la autoría. El primero de ellos fue de atribución de autoría (función *classify*) que parte de la premisa de que no sabemos quien es el autor: La TF fue asociada a Cervantes. Seguimos con el experimento de verificación de autoría (función *imposters*) dando por buenos los primeros resultados y considerando que el complutense podría ser el autor, pero aun conservando la duda y con el objetivo de utilizar esta prueba a modo de validación cruzada. Nuevamente: La TF fue asociada a Cervantes. En ninguno de los casos utilizando esta última función, vimos en otro autor resultados tan evidentes como los que asocian la obra dubitada a Cervantes. Asimismo, el empleo de diferentes distancias estadísticas como *Delta de Burrows* y *Delta de Eder* mostró que la opción de la autoría cervantina es constante. Para ver dónde se encontraban los nexos entre las obras de Cervantes y otros autores de forma más concreta se utilizaron los parámetros ya indicados, se empleó la función *stylo.network* con la que se creó un BCT (Bootstrap Consensus Tree) y se exportaron los datos a *Gephi* obteniendo así un gráfico que evidenció que la relación entre *La tía fingida* y *El celoso extremeño* era más notoria que otras.

En este punto, decidimos analizar las dos obras que presentaban la interrelación más destacada, sin embargo, vimos que existen otras *Novelas Ejemplares* que tenían relación con la TF. La única excepción fuera de las obras del alcalaíno fue una obra de Quevedo, *Sueños y discursos*, más concretamente el fragmento del texto que corresponde a la quinta parte del mismo: *El sueño de la muerte*. No obviamos este dato que creemos nos indica el camino para llevar a cabo más pruebas de determinación de autoría. A pesar de ello, consideramos que debíamos contrastar en primer lugar los textos que presentaban mayor semejanza estilométrica.

En el análisis literario nos centramos sobre todo en aspectos técnicos y expresivos que no se pueden medir cuantitativamente como son los contenidos, y las figuras retóricas. Pudimos observar que en el plano de los contenidos ambas obras pertenecen a un género, tipo y temática similar y utilizan técnicas

narrativas semejantes. En el análisis de la expresión se reconocieron también algunas semejanzas a través del uso de figuras retóricas compartidas, pero se evidenciaron divergencias notorias que están recogidas en el apartado «Síntesis del análisis literario» y que se podrían englobar en la afirmación: *El celoso extremeño* se sirve de recursos que presentan una mayor complejidad comparados con los de *La tía fingida*. Esto se observa de forma particular en la obra no dubitada por su genial uso de la ironía y la mezcla de recursos retóricos empleados metódicamente en momentos en los que la historia requiere mayor intensidad dramática. La TF hace gala de recursos irónicos, pero estos son de tipo verbal en forma de sarcasmo, carientismo y mimesis utilizados tan solo puntualmente como elementos cómicos en los que no obstante también se podría reconocer cierta crítica social⁵². No se aprecia en esta novela el empleo de la ironía dramática y la ironía del sino⁵³ que sí se encuentra en *El celoso extremeño*. También vimos que a nivel morfológico existen diferencias remarcables. En este apartado podríamos haber empleado otra herramienta digital para analizar el *Part Of Speech*⁵⁴ y profundizar así en este ámbito enfrentando los elementos morfológicos y sintácticos de ambas obras. No obstante, esto excedería el marco de esta tesis por lo que este aspecto se dejó a parte para próximos trabajos.

Si atendemos a las generalidades técnicas de las novelas cortas del complutense que exponían Díez-Echarri y Roca Franquesa (1972: 362) en base a los argumentos de Pfandl (1933) y las comparamos con la TF podemos reconocer ciertas similitudes. La finalidad ética nos parece bastante evidente a pesar de que esta está sujeta a interpretaciones. No olvidemos que Díez-Echarri y Roca Francesa (1972) ya advertían que la intención del carácter moralizador de las *Novelas Ejemplares* se ve a veces ensombrecido por «pasajes resbaladizos» (1972: 358). El final puede considerarse feliz pues hay matrimonio y Esperanza consigue huir del mal destino que sufre Claudia, personaje sobre el que recae toda la culpa moral de la historia, si bien, es una suerte poco común como advierte el narrador: «*pocas Esperanzas habrá en la vida que, de tan mala como ella la vivía, salgan al descanso y buen paradero que ella tuvo*» (Cervantes,

⁵² Nos referimos a la ignorancia de los estudiantes con respecto al arte poético.

⁵³ Recordemos que utilizamos la división tipológica de este recurso que emplea Gómez Iñiguez (1990)

⁵⁴ Por ejemplo, *Freeling*: <https://nlp.lsi.upc.edu/freeling/node/1>

La tía fingida). Por otra parte, en el uso del diálogo y el estilo directo se encuentran diferencias notables entre las obras en su modo de empleo, pero cuantitativamente la TF incluye bastantes diálogos especialmente a partir del «*Consejo de Estado y Hacienda*». En la TF, se encuentra también una buena mezcla de elementos alegres y cómicos (p.ej. el empleo de la ironía y el modo de actuar de los estudiantes) con lo serio que llega a lo escabroso. En este aspecto vemos pues también una coincidencia con el conjunto de las *Novelas Ejemplares*. Por último, aunque de un modo mucho más prosaico del que se puede observar en *El celoso extremeño* y en el resto de *Novelas Ejemplares* se puede considerar que el autor de la TF es conocedor del concepto exacto de la novela corta al actuar sobre un suceso de cierta relevancia y dirigir el relato sin accesorios ni grandes ornamentos hacia su desenlace. Vemos pues en estos aspectos generales que existen semejanzas estilísticas y temáticas entre el corpus cervantino y la TF como exponía Gallardo (1832).

Llegados a este punto solo nos queda responder a la gran pregunta ¿quién fue el autor de *La tía Fingida*? Pues bien, hay determinados aspectos a considerar antes de hacer cualquier afirmación terminante sobre la identidad del creador de esta novela. En primer lugar, la elección del corpus no incluye todos los textos que serían ideales para una investigación de este tipo ni tampoco todos los autores que se podrían considerar. Hemos seleccionado aquellos autores que fueron coetáneos a Cervantes (1547-1616) incluyendo también las obras realizadas hasta 1620. Dentro de estas obras se mostró que la fiabilidad de los textos es solo relativa pues no cumplen siempre la premisa de poseer muestras incorruptas que señalaba Juola (2008). La problemática de los textos del Siglo de Oro muestra la necesidad de seguir trabajando en la calidad, la digitalización y la puesta a disposición de los mismos para el ámbito académico.

Además de esto, la decisión de contrastar determinados textos y no otros en el experimento con *classify* no es 100% objetiva. En el conjunto de los textos a comprobar (*secondary_set*, o *test_set*) elegimos ciertas obras para ver si el algoritmo las reconocía y las asignaba a sus respectivos autores, pero ¿Qué pasa si probamos con textos que se consideren estilísticamente más lejanos a un autor? ¿Por qué elegimos una obra y no otra? Desafortunadamente la

herramienta no ofrece ninguna solución a esto más allá de realizar la prueba con diferentes conjuntos de textos en el *secondary_set*.⁵⁵

Entendidos estos escollos, podemos, ahora sí, entrar en consideraciones. En base a las pruebas estilométricas solo tenemos una respuesta posible: *La tía fingida* es con una alta probabilidad obra de Cervantes. Sin embargo, debemos tener en cuenta también que el conjunto de autores no es completo, ni lo son los textos de los representados. Por eso lo que podríamos afirmar es más bien: de entre este conjunto de autores y con el corpus seleccionado el único autor posible es Cervantes. Desde la perspectiva literaria las dos obras comparadas mostraban convergencias y divergencias. En la obra no dubitada veíamos una obra técnicamente más elaborada ¿Se puede desde el análisis literario de ciertas figuras retóricas y algunos elementos morfológicos decidir la autoría de la obra? Consideramos que no, tampoco era nuestro objetivo, pero si sirve en conjunto con los experimentos estilométricos para llegar a ciertas hipótesis.

La primera de estas hipótesis es que esta obra podría ser una novela que Cervantes compuso con anterioridad a sus otras novelas breves, un periodo en el cual el autor aún estaba puliendo su pluma en este género. No olvidemos que la fecha de escritura aproximada que manejamos es la de la compilación de los manuscritos que realizó Porras de la Cámara para el cardenal Niño de Guevara (1604-1606) no sabemos a ciencia cierta cuando fue redactada ni en qué orden con respecto a los otros textos aparecidos en el manuscrito Porras. Esto explicaría las coincidencias a la vez que visto en perspectiva podría servir como ejemplo de análisis de la evolución del alcalaíno. Así pues, aunque compartimos aquí la hipótesis de la autoría cervantina diferimos de Astrana Marín (1944: 40) en que la novela sea excelente, aunque sin desmerecerla pues como vimos emplea un vocabulario variado, numerosas técnicas expresivas y representa una buena muestra del género. La lectura evidencia además mayor complejidad de lo que pudiera aparentar a primera vista, sirviendo a numerosas interpretaciones

⁵⁵ Evidentemente, realizamos pruebas con diferentes «sets» sin embargo consideramos que la selección es siempre arbitraria. De todos modos, las diferentes configuraciones clasificaban los textos de autor conocido correctamente y atribuían la TF a Cervantes. En este contexto la utilización de textos «estilísticamente» lejanos al autor puede servir también para validar los resultados.

y aspectos dignos del análisis⁵⁶. No obstante, consideramos que no tiene el grado de complejidad y excelencia en el uso de los recursos técnicos y expresivos que hemos visto en este trabajo y de los que sí hace gala *El celoso extremeño*. Concordamos por lo tanto en parte con la opinión de Foulché-Delbosc (1899) al reconocer cierta distancia estilística con respecto a Cervantes, pero al contrario que el hispanista francés nos parece que esto no sirve como argumento para no considerarla del complutense. Dicho de otro modo, en nuestro análisis no vemos tanta distancia como para poder afirmar rotundamente que no es obra suya.

La segunda presunción es que la obra fuera de un temprano imitador como hipotetizaba Márquez Villanueva (1991). Aunque no se pueda descartar esta posibilidad de manera rotunda, pues habría que elevar el nombre de dicho imitador y contrastar sus textos, nos parece ahora poco probable en vista de los resultados estilométricos. Aunque un imitador pudiera emplear el vocabulario y las expresiones más características de Cervantes resulta poco creíble que prestara atención a todas las palabras funcionales y su frecuencia relativa. Precisamente este es el fundamento en el que se sustentan en ocasiones algunas investigaciones de la lingüística forense para resolver por ejemplo casos de plagio. Opinamos por lo tanto que la posibilidad de un imitador no se puede descartar utilizando solamente la comparación de determinadas expresiones, pasajes textuales e índices verbales. Estas, en nuestra opinión, no sirven tampoco como argumentos definitivos para desmentir u otorgar la autoría de la TF como hicieron por ejemplo Icaza (1916) y Criado de Val (1953) para rechazarla como obra de Cervantes o Madrigal (2003) para adjudicársela.

Otro aspecto relevante que nos queda por comentar es el de justificar para qué serviría considerar la TF de Cervantes. ¿Qué aporta esto a la literatura? Ya hemos tratado en parte este tema. Aunque desde nuestro punto de vista la obra nos parezca más prosaica que el resto de novelas cortas, el conocimiento de sus parecidos y diferencias nos da cierta perspectiva sobre la evolución del autor. También vemos en sus contenidos crítica social, dicotomía de espacios casa-calle, una serie de ideas sobre la mujer, imagen del mundo estudiantil, entorno

⁵⁶ Un ejemplo interesante es: del Carmen Artigas, M. (2006). Exaltada fantasía erótica en La tía fingida: el remiendo de una doncella. *Explicación de textos literarios*, 35(1-2), 6-15.

del hampa y otras múltiples posibilidades de análisis que podrían ser contrastadas con el resto de su corpus conformando en su totalidad el significado último de determinados propósitos e ideas, el empleo y desarrollo de ciertas técnicas narrativas, etc. No olvidemos que el genial Cervantes es considerado el creador de la novela breve española⁵⁷ por lo que saber más sobre el proceso y la evolución de este género es descubrir también las estrategias que se esconden tras su nacimiento.

Llegamos pues al final de esta tesis habiendo visto que nuestra hipótesis inicial no estaba desencaminada. La estilometría ha demostrado ser una herramienta útil en el proceso de atribución de autoría, que pese a sus límites nos ha ayudado a adoptar nuevas perspectivas literarias. Sin ella no habiéramos encontrado con tanta rapidez paralelismos con *El celoso extremeño*. La hipótesis de la autoría cervantina de la TF a pesar de que haya sido desdeñada por algunos críticos que no reconocían en esta obra la genialidad de Cervantes parece tener bastante sustento al menos mientras no encontremos una alternativa con textos que podamos contrastar y que presente unos resultados semejantes a los arrojados por el complutense en nuestros experimentos. Hasta entonces, si nos decantamos por la primera de las sospechas, la de la novela cervantina de composición temprana tendríamos que admitir que quizás la excelencia de los trabajos conocidos del ingenioso Miguel de Cervantes es fruto de la evolución y mejoría del mismo o dicho de otro modo que sin esfuerzo y trabajo ni el gran talento de Miguel de Cervantes Saavedra hubiera alcanzado la calidad de las *Novelas Ejemplares* ni dado vida al ingenioso caballero y a su leal escudero que viven hoy en el mundo entero.

⁵⁷ Por ejemplo, Díez-Echarry y Roca Franquesa (1972: 222) o Reguera (2006: 166)

12 Bibliografía

- Antoine, G. (1968). Stylistique des formes et stylistique des thèmes, ou le stylisticien face à l'ancienne et à la nouvelle critique. *Les chemins actuels de la critique, sous la dir. de G. Poulet*, UGE, 10(18), 240-64.
- Arellano-Ayuso, I. (1997). Las aventuras del texto: del manuscrito al libro en el Siglo de Oro. En González, A. R. F. (1997). *Unum et Diversum: Estudios en honor de Ángel-raimundo Fernández González*. Ediciones Universidad de Navarra, 41-66.
- Arellano-Torres, I. D. (2018). El entremés de Los habladores, atribuido a Cervantes. In *Anales cervantinos* (Vol. 50, pp. 299-323).
- Astrana Marín, L. (1944). «Sobre *La tía fingida*». En *Cervantinas y otros ensayos*, 37-48. Madrid: Afrodísio Aguado.
- Astrana Marín, L. (1953). *Vida ejemplar y heroica de Miguel de Cervantes Saavedra*, vol. V. Madrid: Instituto Editorial Reus.
- Avalle-Arce, J. B. (1973). Atribuciones y supercherías. In *Suma cervantina* (pp. 399-408). Tamesis Book Limited.
- Aylward, E. T. (1999). Significant Disparities in the Text of *La tía fingida* vis-à-vis Cervantes's *El casamiento engañoso*. *Cervantes: Bulletin of the Cervantes Society of America*, 19(1), 40-65.
- Bailey, R. W. (1979). Authorship Attribution in a Forensic setting. *Advances in computer-aided literary and linguistic research*, 9. Eds. Ager, D.E., Knowles, F.E. and Smith, J. Birmingham: AMLC.
- Barthes, R. (1968). La mort de l'auteur. *Manteia*, 5, 12-17.
- Barthes, R. (1987). La muerte del autor. *El susurro del lenguaje*, 65-71.
- Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open-source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and social media*. Recuperado el 10.12.21 de <https://gephi.org/users/publications/>
- Bia, A. (2016). Estilometría computacional, algunas experiencias en el marco del proyecto TRACE. In *Humanidades Digitales: Construcciones locales en contextos globales*. Asociación Argentina de Humanidades Digitales.
- Blasco Pascual, F. J. (2019). Atribuciones cervantinas desde la estilometría: el entremés de Los mirones. *Cartografía teatral*, 151-168.
- Blasco Pascual, F. J. y Ruíz Urbón, C. (2009). Evaluación y cuantificación de algunas técnicas de atribución de autoría en textos españoles. *Castilla. Estudios De Literatura*, 27-47. En línea en <https://doi.org/10.24197/cel.0.2009> (consultado el 09.10.21)
- Bonilla y San Martín, A. (1917). «Un crítico desbocado». En *De crítica cervantina. Madrid: Ruiz, 1917. 81-105*.

- Boto, B. M. A. (2017). Mapa estilométrico de la narrativa de Eduardo Mendoza: aproximación a un análisis estilístico computacional de textos literarios. *Epos: Revista de filología*, (33), 99-114.
- Bühler, K. (1934). *Teoría del lenguaje*, trad. Julián Marías. Madrid: Alianza Universidad. En Fernández Retamar, R. (1963)
- Canavaggio, J. (1992). En busca del perfil perdido. Espasa-Calpe
- Caparrós, J. D. (2002). *Teoría de la literatura*. Editorial Universitaria Ramón Areces. Págs. 303-377
- Caparrós, J. D. (2009). *Introducción a la teoría literaria*. Editorial Universitaria Ramon Areces.
- Castagnino, R. H. (1987). *El análisis literario: introducción metodológica a una estilística integral*. Editorial Nova.
- Castro, Adolfo de. Post scriptum acerca de la novela La tía fingida. *La España moderna* (April 30, 1889): 183-85.
- Cervantes, M. (2018). *Novelas ejemplares*. MVP.
- Cervantes, M. (2018). *La tía fingida*. Ediciones Cátedra.
- Chandrasekaran, S. (2013). Becoming a Data Scientist. Recuperado el 07.06.21 de <http://nirvacana.com/thoughts/becoming-a-data-scientist>.
- Corral, H. M. (1981). La mujer en las Novelas ejemplares de Miguel de Cervantes Saavedra. In *Cervantes, su obra y su mundo: actas del I Congreso Internacional sobre Cervantes* (pp. 397-408). EDI-6.
- De Val, M. C. (1953). *Análisis verbal del estilo: índices verbales de Cervantes, de Avellaneda y del autor de «La tía fingida»* (Vol. 57). Consejo superior de investigaciones científicas.
- de Cervantes Saavedra, M. (1906). *Juicio de «La tía fingida»: copia de tres ediciones raras y edición crítica de esta novela*. los sucesores de Hernando.
- de Icaza, F. A. (1916). *De cómo y por qué La tía fingida no es de Cervantes: y otros nuevos estudios cervánticos* (Vol. 4). Imp. Clásica Española.
- De la Rosa Pérez, J. (2016). Making Machines Learn. Applications of Cultural Analytics to the Humanities.
- De Leal, I. P. (1984). El análisis léxico-estilístico. *Epos: Revista de filología*, (1), 195-216.
- Desagulier, Guillaume (2017). *Corpus linguistics and statistics with R*. Springer. En Fradejas Rueda, J. M. (2019)
- Díez-Echarri, E., y Roca Franquesa, J.M. (1972). *Historia de la literatura española e hispanoamericana*. Aguilar.

Dubois, J., Giacomo, M., Guespin, L., y Domínguez, A. (1979). *Diccionario de lingüística* (No. C/410.3 D5).

Eder M, Rybicki J, Kestemont M (2016). Stylometry with R: a package for computational text analysis. *The R Journal*, 8(1), 107–121. Recuperado el 15.08.21 de <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html>.

Eder, M. (2017a). Visualization in stylometry: cluster analysis using networks. *Digital Scholarship in the Humanities*, 32(1), 50-64.

Eder, M. (2017b, August). Short Samples in Authorship Attribution: A New Approach. In *DH*. Recuperado el 06.08.21 de <https://dh2017.adho.org/abstracts/341/341.pdf>

Eder, M. (2018). Authorship verification with the package stylo. En línea, recuperado el 11.09.21 de <https://computationalstylistics.github.io/docs/imposters>

Eder, M. and Rybicki, J. (2012). Introduction to Stylometric Analysis using R. *Digital Humanities 2012*, 15. Hamburg.

Eisenberg, D. (1988). *Las Semanas del jardín de Miguel de Cervantes: estudio, edición y facsímil del manuscrito* (Vol. 3). Ediciones de la Diputación de Salamanca.

Eisenberg, D. (1990). Repaso crítico de las atribuciones cervantinas. *Nueva Revista de Filología Hispánica*, 38(2), 477-492.

Ellegard, A. (1962). A statistical method for determining authorship: the junius letters. *Gothenburg Studies in English*, 13, 1769-1772. En Holmes y Kardos (2003)

Fernández Retamar, R. (1963). *Idea de la estilística*. La Habana: Universidad de La Habana, Comisión de publicaciones.

Fernández, V. H. (2007). Diccionario práctico de figuras retóricas y términos afines: Tropos, figuras de pensamiento, de lenguaje, de construcción, de dicción, y otras curiosidades.

Ford, J. D. M., y Lansing, R. (1931). *Cervantes: A Tentative Bibliography of His Works and of the Bibliographical and Critical Material Concerning Him*. Harvard University Press.

Foulché-Delbosc, R. (1899). «Étude sur *La tía fingida*». *Revue Hispanique* 6 (19): 256-306.

Fradejas Rueda, J. M. (2016). El análisis estilométrico aplicado a la literatura española: las novelas policíacas e históricas. *Caracteres. Estudios Culturales y Críticos de La Esfera Digital*, 5(2), 196–245.

Fradejas Rueda, J. M. (2019). «Estilometría y la Edad Media castellana», en Nanette Rissler-Pipka (ed.), *Theorien von Autorschaft und Stil in Bewegung: Stilistik und Stilometrie in der Romania*, (Romanische Studien, Beiheft 6). Múnich: AVM.edition, 49-74.

Fradejas Rueda, J. M. (2020). *Cuentapalabras. Estilometría y análisis de texto con R para filólogos*. Recuperado el 08.08.21 de <http://www.aic.uva.es/cuentapalabras/>

- Fruchterman, T. M. J., y Reingold, E. M. (1991). Graph Drawing by Force-Directed Placement. *Software: Practice and experience*, 21(11), 1129-1164.
- Fumaz, R. B. (2012). Muerte del autor y literatura digital. *Eikasia: revista de filosofía*, 44, 113-128.
- Gallardo, B. J. (1835). «La tía fingida ¿es novela de Cervantes?». *El Crítico*, *papel volante de Literatura y Bellas artes* 1: 1-43.
- Genette, G. (1989). Discurso del relato. *Figuras iii*, 75-327.
- Genette, G., y Tapia, M. R. (1998). *Nuevo discurso del relato*. Madrid: Cátedra.
- Gries, Stefan (2013). *Statistics for Linguistics with R. A Practical Introduction*. De Gruyter. En Fradejas Rueda, J. M. (2019).
- Gries, Stefan (2017). *Quantitative Corpus Linguistics with R: A Practical Introduction*. Routledge. En Fradejas Rueda, J. M. (2019).
- Guiraud, P. (1969). *Essais de stylistique*.
- Guiraud, P. (1970). La Stylistique [1955]. *Paris, Puf*, 37.
- Hatzfeld, H. A. (1975). *Estudios de estilística* (Vol. 30). Editorial Planeta.
- Herrmann, J. B., van Dalen-Oskam, K., & Schöch, C. (2015). Revisiting style, a key concept in literary studies. *Journal of literary theory*, 9(1), 25-52. En Rißler-Pipka (2018)
- Holmes, D. I. (1994). Authorship attribution. *Computers and the Humanities*, 28(2), 87-106.
- Holmes, D. I., y Kardos, J. (2003). Who was the author? An introduction to stylometry. *Chance*, 16(2), 5-8.
- Icaza, F. A. D. (1917). *Supercherías y errores cervantinos*. Madrid: Imprenta Clásica Española.
- Inamoto, K. (2004). Sobre la atribución cervantina de La tía fingida o/y sobre el hiato delante del acento rítmico interno en el verso cervantino. In *Peregrinamente peregrinos: Actas del V Congreso Internacional de la Asociación de Cervantistas* (pp. 1401-1408).
- Iñiguez, L. G. (1990). Humor cervantino: «El celoso extremeño». In *Actas del II Coloquio Internacional de la Asociación de Cervantistas* (pp. 633-640). Anthropos.
- Jakobson, R. (1960). Linguistics and poetics. In *Style in language* (pp. 350-377). MA : MIT Press.
- Jakobson, R. (1963). *Essais de linguistique générale*. *Les Etudes Philosophiques*, 18(4).
- Jockers, Matthew L. (2014). *Text Analysis with R for Students of Literature*. Springer. En Soler, J. C., y Tello, J. C. (2019)
- Juola, P. (2008). *Authorship attribution* (Vol. 3). Now Publishers Inc.

- Juola, P. (2015). The Rowling case: A proposed standard analytic protocol for authorship questions. *Digital Scholarship in the Humanities*, 30(suppl_1), i100-i113.
- Kayser, W. (1968). «El estilo», Cap. IX, pp. 361-434, Investigación y análisis de la obra literaria, Editorial Gredos, Madrid, 1968. En Pérez, G. (2009)
- Kestemont, M. (2011). «Een stylometrisch onderzoek naar Jan van Boendales auteurs-chap voor de Brabantse yeesten», en *Revue belge de Philologie et d'Histoire*. 89(3-4), pp. 1019-1048. En Soler, J. C., y Tello, J. C. (2019)
- Kestemont, M., Stover, J., Koppel, M., Karsdorp, F. and Daelemans, W. (2016). Authenticating the writings of Julius Caesar. *Expert Systems with Applications*, 63: 86-96.
- Koppel, M. and Winter, Y. (2014). Determining if two documents are written by the same author. *Journal of the Association for Information Science and Technology*, 65(1): 178–87 doi:10.1002/asi.22954. <http://dx.doi.org/10.1002/asi.22954>.
- Levshina, N. (2015). How to do Linguistics with R. *Data exploration and statistical analysis*. John Benjamins. En Fradejas Rueda, J. M. (2019).
- López, F. (2011). Donde se muestran algunos resultados de atribución de autor en torno a la obra cervantina. *Revista Colombiana de Estadística*, 34(1), 15-37.
- López-Vázquez, A. R. (2013). La novela ejemplar «La tía fingida», atribuida a Cervantes. *Artifara: Revista de lenguas y literaturas ibéricas y latinoamericanas*, (13), 11.
- López-Vázquez, A. R. (2018). Cervantes y helecho de Procusto: notas críticas al uso de la Estilometría en obras de atribución dudosa y en obras apócrifas. *eHumanista: Journal of Iberian Studies*, (41), 193-201.
- Lorenzo, L. H. (2019). Poesía áurea, estilometría y fiabilidad: métodos supervisados de atribución de autoría atendiendo al tamaño de las muestras. *Caracteres: estudios culturales y críticos de la esfera digital*, 8(1), 189-228.
- Lucía Megías, J. M. (2018). El códice Porras (casi) recuperado (la copia del Cigarral del Carmen de "La tía fingida"). In *Anales cervantinos* (Vol. 50, pp. 333-351). CSIC.Reus.
- Madrigal, J. L. (2003). De cómo y por qué La tía fingida es de Cervantes. *Artifara*, 2.
- Marín, L. A. (1948). *Vida ejemplar y heroica de Miguel de Cervantes Saavedra*.
- Márquez Villanueva, F. (1995). «La tía fingida: literatura universitaria». En *Trabajos y días cervantinos*, 157-189. Alcalá: Centro de Estudios Cervantinos.
- Meregalli, F. (1992). *Introducción a Cervantes* (Vol. 111). Ariel.
- Mukherjee, J. (2005). Stylistics. *Encyclopaedia of Linguistics*, 1184-6.
- Pérez, G. (2009). Los estudios del texto literario según la Estilística: teorías y Métodos. *Contribuciones a las Ciencias Sociales*. En línea en www.eumed.net/rev/cccss/06/ypg.htm Consultado el 12/09/21

- Pfandl, L. (1933). Historia de la literatura nacional española en la Edad de Oro, trad. Jorge Rubio Balaguer, Barcelona, Sucesores de Juan Gili.
- Reguera, J. M. (2006). El nacimiento de la novela corta en España (la perspectiva de los editores). *Lectura y signo*, (1), 165-175.
- Reyes, H. E. D. C. (2018). El análisis del estilo literario: un acercamiento desde la recepción. *Lingüística y Literatura*, (74), 21-36.
- Rico, F. (2000). Imprenta y crítica textual en el Siglo de Oro, ed. al cuidado de Pablo Andrés y Sonia Garza. Valladolid, Centro para la Edición de los Clásicos Españoles/Secretariado de Publicaciones de la Universidad.
- Rissler-Pipka, N. (2016). Avellaneda y los problemas de la identificación del autor. Propuestas para una investigación con nuevas herramientas digitales. Ed. Hanno Ehrlicher. *El otro Don Quijote. La continuación de Fernández de Avellaneda y sus efectos*. Ausburgo: Institut für Spanien, Portugal-und Lateinamerikastudien (ISLA), 27-51.
- Rißler-Pipka, N. (2018). Die Digitalisierung des goldenen Zeitalters–Editionsproblematik und stilometrische Autorschaftsattribuion am Beispiel des Quijote Abstracts. *Zeitschrift Für Digitale Geisteswissenschaften*, 3, 1-46.
- Rißler-Pipka, N. (2019). Einleitung: Theorien von Autorschaft und Stil in Bewegung. Stilistik und Stilometrie in der Romania. *Romanische Studien*, 5-20.
- Romero Ferrer, A. (2000). Del rigor filológico a la falsificación cervantina: Adolfo de Castro y la literatura española de Iso Siglos de Oro.
- Rueda, J. M. F. (2019). Estilometría y la Edad Media castellana. *Romanische Studien*, 49-74.
- Rybicki, J. and Eder, M. (2012). Stylometry: Computer-Assisted Analysis of Literary Texts. "Culture & Technology" European Summer School in Digital Humanities. Leipzig University.
- Soler, J. C., y Tello, J. C. (2019). Autoría y estilo. Una atribución cervantina desde las humanidades digitales. El caso de La conquista de Jerusalén. In *Anales Cervantinos* (Vol. 51, pp. 231-250).
- Tello, J. C. (2016). Entendiendo Delta desde las humanidades. *Caracteres: estudios culturales y críticos de la esfera digital*, 5(1), 140-176.
- Todorov, T. (1970). «Los estudios del estilo. Bibliografía selectiva», *Textos y contextos. Una ojeada en la teoría literaria mundial*, 169-183.
- Winter, B. (2019). *Statistics for Linguistics. An Introdcution using R*. New York: Routledge. En Fradejas Rueda, J. M. (2019).
- Yllera, A. (1979). Estética, poética y semiótica literaria. 3ª ed., Madrid. Alianza Editorial (1ª ed. 1974)

Zahonero, M. L. (2020). Una nueva visión de la supuesta influencia de Madame Bovary en La Regenta a través de la estilometría y el análisis de sentimientos basados en lenguaje R. *Orillas: revista d'ispanística*, (9), 573-607.

Zapata, J. M. (2011). Muerte y resurrección del autor. Nuevas aproximaciones al estudio sociológico del autor. *Lingüística y literatura*, (60), 35-58.

Páginas web

Nota: los enlaces a las versiones digitales de los libros empleados para el corpus se encuentran en el anexo «Corpus».

BVMC, Guía de obras atribuidas a Miguel de Cervantes Saavedra (15.06.21). En línea http://www.cervantesvirtual.com/portales/miguel_de_cervantes/estados_atribuciones/

Los gráficos Gephi fueron creados con el software de la página: <https://gephi.org/>

Transkribus, software para la digitalización de textos: <https://readcoop.eu/transkribus/>

Los textos digitales del subcorpus «otros textos en prosa» han sido recuperados de Fradejas Rueda, 7 Partidas Digital, Github: <https://github.com/7PartidasDigital>

Freeling para POS tagging: <https://nlp.lsi.upc.edu/freeling/node/1>

13 Anexos

13.1 Anexo: Código R

Para entender mejor esta parte hemos distinguido por colores: en verde (los comentarios propios), en azul (el código), en rojo y negro (los avisos y datos que da la herramienta).

1) Código: función *classify()*

```
# Determinamos el corpus con el que queremos trabajar que hemos nombrado  
Corpus test2:
```

```
> setwd("C:/Users/ACP/Desktop/corpus cervantino - fragmentado/Corpus fragmentado  
6000W")
```

```
# Activamos la librería de stylo y las funciones para la clasificación:
```

```
> library(stylo)
```

```
> library(class, lib.loc = "C:/Program Files/R/R-4.1.0/library")
```

```
# Ejecutamos la función classify() para ello hemos creado dos carpetas un corpus  
de aprendizaje (primary_set) y uno de test (secondary set) en el que incluimos  
un texto ejemplo de cada uno de los autores (7). Estos son ejemplos de las 8  
clases que el programa debería de reconocer: sabemos que hay 8 autores y 8  
textos ejemplares por lo que lo esperado es que el algoritmo los reconozca:
```

```
> classify()
```

```
# Aparece el siguiente mensaje:
```

```
1194630 tokens will be used to create a list of features  
processing 192 text samples
```

```
# Y el resultado es el siguiente:
```

```
General attributive success: 75 of 75 (100%, sd =0%)  
MFWs from 100 to 500 @ increment 100  
Culling from 0 to 0 @ increment 20  
Pronouns deleted: FALSE
```

Para poder ver todos los resultados introducimos el siguiente código (el nombre del vector «my_cool_results» puede ser definido por el usuario:

```
> my_cool_results = classify()
```

Esto crea una tabla en la que podemos acceder a diferentes características. Nos interesa aquí ver cuales son las clases que han sido predichas y vemos que los resultados son satisfactorios con un porcentaje de acierto promedio del del 100 %. Las clases (autores) predichas son 15 (Pero hay 14 clases pues Cervantes está representado por dos partes de textos)

```
> my_cool_results[["predicted"]]
```

```
[1] "Aleman" [2] "Carvajal" [3] "Castillo" [4]"Cervantes" [5] "Cervantes" [6] "Eslava"  
[7]"Espinell" [8] "GregorioGonzalez" [9] "Lope" [10] "Montalban" [11] "Montemayor" [12]  
"Quevedo" [13] "SalasBarbadillo" [14] "SuarezFigueroa" [15] "Zayas"
```

También podemos acceder al resultado de acierto según la configuración de las MFW elegida que va desde las 100 hasta las 500 con un incremento de 100. Se observa que con 100 MFW el acierto ya es absoluto:

```
> my_cool_results[["success.rate"]]
```

```
percentage of correctly guessed samples
```

```
100 100 100 100
```

2) Código: validación cruzada

Llevamos a cabo el análisis utilizando esta vez la función stylo() con la configuración que ahora sabemos que es precisa: culling@ 0, 100-500 MFW incremento 100. Seguidamente vamos a comprobar también su efectividad a través de la validación cruzada para ello introducimos:

```
> classify(cv.folds = 10)
```

Comprobamos que los resultados son nuevamente satisfactorios pues son los siguientes:

```
General attributive success: 75 of 75 (100%, sd =0%)
```

Ahora queremos crear un vector para poder acceder a la visualización de los resultados. Hemos utilizado para ello un vector que hemos nombrado «validado»:

```
> validado = classify()
```

Para ver todas las variables que se han creado y a las que podemos tener acceso introducimos el siguiente comando:

```
> summary(validado)
```

Esto, como ya dijimos, nos da como resultado una lista de variables a las cuales podemos acceder que son las siguientes:

Available variables:

distance.table final distances between each pair of samples

expected ground truth, or a vector of expected classes

features features (e.g. words, n-grams, ...) applied to data

features.actually.used features (e.g. frequent words) actually analyzed

frequencies.both.sets frequencies of words/features accross the corpus

frequencies.test.set frequencies of words/features in the test set

frequencies.training.set frequencies of words/features in the training set

overall.success.rate average percentage of correctly guessed samples

predicted a vector of classes predicted by the classifier

success.rate percentage of correctly guessed samples

These variables can be accessed by typing e.g.:

```
validado$distance.table
```

queremos volver a comprobar nuevamente si el clasificador predijo las ocho clases que esperábamos al comienzo así que accedemos a la tabla para comprobarlo introduciendo el siguiente código que nos muestra nuevamente que se han reconocido a los 8 autores:

```
> validado[["predicted"]]
```

Si introducimos el siguiente código podemos comprobar la precisión en forma de unos y ceros. Se asigna 1 cuando el autor del texto se ha predicho y 0 cuando no se ha encontrado ninguna clase correspondiente para un determinado texto:

```
> performance.measures(predicted_classes = validado)
```

Aleman 1, Carvajal 1, Castillo 1, Cervantes 1, Eslava 1, Espinel 1, GregorioGonzalez 1, Lope 1, Montalbán 1, Montemayor 1, Quevedo 1, Salas Barbadillo 1, Suarez Figueroa 1, Zayas 1

Al repetir la prueba introduciendo el texto dubitado consultamos los resultados esperados y predichos introduciendo:

```
> validado[["expected"]]
```

```
> validado[["predicted"]]
```

3) Código: Tokenización / imposters() e imposters.optimize ()

```
# Seleccionamos la localización del directorio de trabajo con el que queremos trabajar. Este debe estar en un subdirectorío con el nombre "corpus". A continuación, activamos la librería de stylo.
```

```
> setwd("C:/Users/ACP/Desktop/corpus cervantino- fragmentado/Corpus fragmentado 6000w/corpus")
```

```
> library(stylo)
```

```
### stylo version: 0.7.4 ###
```

```
If you plan to cite this software (please do!), use the following reference:
```

```
Eder, M., Rybicki, J. and Kestemont, M. (2016). Stylometry with R: a package for computational text analysis. R Journal 8(1): 107-121
```

```
> <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html>
```

```
To get full BibTeX entry, type: citation("stylo")
```

```
Warning message: package 'stylo' was built under R version 4.1.1
```

```
# El preprocesamiento de los textos consiste en convertirlos en palabra token y crear una tabla de frecuencias de cada uno de ellos.
```

```
> tokenized.texts = load.corpus.and.parse(files = "all")
```

```
using current directory...
```

```
slicing input text into tokens...
```

```
turning words into features, e.g. char n-grams (if applicable)...
```

```
# Realizamos una tabla de frecuencias utilizando las 500MFV y creamos una tabla mediante una variable para tener acceso a los datos.
```

```
> features = make.frequency.list(tokenized.texts, head = 500)  
> cervantes = make.table.of.frequencies(tokenized.texts, features, relative = TRUE)
```

```
# hemos llamado a la tabla con los datos de las frecuencias "cervantes"
```

```
processing 207 text samples
```

```
combining frequencies into a table...
```

```
# consultamos los nombres de los textos y su indexación.
```

```
> rownames(cervantes)
```

```
[1] "Aleman_IIGuzman1"  
[2] "Aleman_IIGuzman10"  
[3] "Aleman_IIGuzman11"  
[4] "Aleman_IIGuzman12"  
[5] "Aleman_IIGuzman13"  
[6] "Aleman_IIGuzman14"  
[7] "Aleman_IIGuzman15"  
[8] "Aleman_IIGuzman16"  
[9] "Aleman_IIGuzman17"  
[10] "Aleman_IIGuzman18"  
[11] "Aleman_IIGuzman19"  
[12] "Aleman_IIGuzman2"  
[13] "Aleman_IIGuzman20"  
[14] "Aleman_IIGuzman21"  
[15] "Aleman_IIGuzman22"  
[16] "Aleman_IIGuzman23"  
[17] "Aleman_IIGuzman3"  
[18] "Aleman_IIGuzman4"  
[19] "Aleman_IIGuzman5"  
[20] "Aleman_IIGuzman6"  
[21] "Aleman_IIGuzman7"  
[22] "Aleman_IIGuzman8"  
[23] "Aleman_IIGuzman9"  
[24] "Carvajal_AmanteVenturoso"  
[25] "Carvajal_AmarSinSaberQuien1"  
[26] "Carvajal_AmarSinSaberQuien2"  
[27] "Carvajal_CelosVenganDesprecios"  
[28] "Carvajal_DichaDeDoristea"  
[29] "Carvajal_EscravoDeSuEsclavo"  
[30] "Carvajal_IndustriaVenceDesdenes1"  
[31] "Carvajal_IndustriaVenceDesdenes2"  
[32] "Carvajal_QuienBienObra"  
[33] "Carvajal_VenusDeFerrara"  
[34] "Castillo_BachillerTrapaza1"  
[35] "Castillo_BachillerTrapaza2"  
[36] "Castillo_DosDichasSinPensar1"  
[37] "Castillo_DosDichasSinPensar2"  
[38] "Castillo_FantasmaValencia1"  
[39] "Castillo_FantasmaValencia2"  
[40] "Castillo_FuerzaCastigada"  
[41] "Castillo_HarpíasMadridEstafas1"  
[42] "Castillo_HarpíasMadridEstafas2"  
[43] "Castillo_HarpíasMadridEstafas3"  
[44] "Castillo_HarpíasMadridEstafas4"  
[45] "Castillo_HarpíasMadridEstafas5"  
[46] "Castillo_HarpíasMadridEstafas6"  
[47] "Castillo_LisardoEnamorado1"  
[48] "Castillo_LisardoEnamorado2"  
[49] "Castillo_ProteoDeMadrid1"  
[50] "Castillo_ProteoDeMadrid2"  
[51] "Cervantes__FuerzaDeLaSangre"  
[52] "Cervantes__Galatea1"  
[53] "Cervantes__Galatea10"  
[54] "Cervantes__Galatea11"  
[55] "Cervantes__Galatea12"  
[56] "Cervantes__Galatea13"  
[57] "Cervantes__Galatea14"  
[58] "Cervantes__Galatea15"  
[59] "Cervantes__Galatea16"  
[60] "Cervantes__Galatea17"  
[61] "Cervantes__Galatea18"  
[62] "Cervantes__Galatea19"  
[63] "Cervantes__Galatea2"  
[64] "Cervantes__Galatea3"  
[65] "Cervantes__Galatea4"
```

[66] "Cervantes__Galatea5"
[67] "Cervantes__Galatea6"
[68] "Cervantes__Galatea7"
[69] "Cervantes__Galatea8"
[70] "Cervantes__Galatea9"
[71] "Cervantes_AmanteLiberal1"
[72] "Cervantes_AmanteLiberal2"
[73] "Cervantes_AmanteLiberal3"
[74] "Cervantes_CelosoEstremeno1"
[75] "Cervantes_CelosoEstremeno2"
[76] "Cervantes_ColoquioPerros1"
[77] "Cervantes_ColoquioPerros2"
[78] "Cervantes_ColoquioPerros3"
[79] "Cervantes_ColoquioPerros4"
[80] "Cervantes_DosDoncellas1"
[81] "Cervantes_DosDoncellas2"
[82] "Cervantes_EspanolaInglesa1"
[83] "Cervantes_EspanolaInglesa2"
[84] "Cervantes_EspanolaInglesa3"
[85] "Cervantes_Gitanilla1"
[86] "Cervantes_Gitanilla2"
[87] "Cervantes_Gitanilla3"
[88] "Cervantes_Gitanilla4"
[89] "Cervantes_IlustreFregona1"
[90] "Cervantes_IlustreFregona2"
[91] "Cervantes_IlustreFregona3"
[92] "Cervantes_LicenciadoVidriera1"
[93] "Cervantes_LicenciadoVidriera2"
[94] "Cervantes_RinconeteCortadillo1"
[95] "Cervantes_RinconeteCortadillo2"
[96] "Cervantes_SraCornelia1"
[97] "Cervantes_SraCornelia2"
[98] "Desconocido_LaTiaFingida"
[99] "Eslava_NochesInvierno1"
[100] "Eslava_NochesInvierno10"
[101] "Eslava_NochesInvierno11"
[102] "Eslava_NochesInvierno12"
[103] "Eslava_NochesInvierno13"
[104] "Eslava_NochesInvierno2"
[105] "Eslava_NochesInvierno3"
[106] "Eslava_NochesInvierno4"
[107] "Eslava_NochesInvierno5"
[108] "Eslava_NochesInvierno6"
[109] "Eslava_NochesInvierno7"
[110] "Eslava_NochesInvierno8"
[111] "Eslava_NochesInvierno9"
[112] "Espinel_EscuderoMarcosObregon2"
[113] "Espinel_EscuderoMarcosObregon3"
[114] "Espinel_EscuderoMarcosObregon4"
[115] "Espinel_EscuderoMarcosObregon5"
[116] "Espinel_EscuderoMarcosObregonI"
[117] "GregorioGonzalez_GuitonOnofre1"
[118] "GregorioGonzalez_GuitonOnofre2"
[119] "GregorioGonzalez_GuitonOnofre3"
[120] "GregorioGonzalez_GuitonOnofre4"
[121] "GregorioGonzalez_GuitonOnofre5"
[122] "GregorioGonzalez_GuitonOnofre6"
[123] "GregorioGonzalez_GuitonOnofre7"
[124] "GregorioGonzalez_GuitonOnofre8"
[125] "GregorioGonzalez_GuitonOnofre9"
[126] "Lope_DesdichaPorLaHonra1"
[127] "Lope_DesdichaPorLaHonra2"
[128] "Lope_FortunasDeDiana1"
[129] "Lope_FortunasDeDiana2"
[130] "Lope_GuzmanElBravo1"
[131] "Lope_GuzmanElBravo2"
[132] "Lope_PrudenteVenganza1"
[133] "Lope_PrudenteVenganza2"
[134] "Montalban_EnvidiosoCastigado1"
[135] "Montalban_EnvidiosoCastigado2"

[136] "Montalban_FuerzaDeDesengano1"
[137] "Montalban_FuerzaDeDesengano2"
[138] "Montalban_LaHermosaAurora1"
[139] "Montalban_LaHermosaAurora2"
[140] "Montalban_MayorConfusion1"
[141] "Montalban_MayorConfusion2"
[142] "Montemayor_LosSieteLibrosDiana1"
[143] "Montemayor_LosSieteLibrosDiana10"
[144] "Montemayor_LosSieteLibrosDiana11"
[145] "Montemayor_LosSieteLibrosDiana12"
[146] "Montemayor_LosSieteLibrosDiana2"
[147] "Montemayor_LosSieteLibrosDiana3"
[148] "Montemayor_LosSieteLibrosDiana4"
[149] "Montemayor_LosSieteLibrosDiana5"
[150] "Montemayor_LosSieteLibrosDiana6"
[151] "Montemayor_LosSieteLibrosDiana8"
[152] "Montemayor_LosSieteLibrosDiana9"
[153] "Quevedo_SuenosydiscursosI_ElJuicioFinal"
[154] "Quevedo_SuenosydiscursosII_El alguacil"
[155] "Quevedo_SuenosydiscursosIII1"
[156] "Quevedo_SuenosydiscursosIII2"
[157] "Quevedo_SuenosydiscursosIV_El mundopordentro"
[158] "Quevedo_SuenosydiscursosV1_SuenoDeLaMuerte"
[159] "Quevedo_SuenosydiscursosV2_SuenoDeLaMuerte"
[160] "SalasBarbadillo_ElSagazEstacio1"
[161] "SalasBarbadillo_ElSagazEstacio2"
[162] "SalasBarbadillo_ElSagazEstacio3"
[163] "SalasBarbadillo_ElSagazEstacio4"
[164] "SalasBarbadillo_ElSagazEstacio5"
[165] "SalasBarbadillo_ElSagazEstacio6"
[166] "SalasBarbadillo_LaIngeniosaHijac1"
[167] "SalasBarbadillo_LaIngeniosaHijac2"
[168] "SalasBarbadillo_LaIngeniosaHijac3"
[169] "SuarezFiguerola_Pasajero1"
[170] "SuarezFiguerola_Pasajero10"
[171] "SuarezFiguerola_Pasajero11"
[172] "SuarezFiguerola_Pasajero12"
[173] "SuarezFiguerola_Pasajero13"
[174] "SuarezFiguerola_Pasajero14"
[175] "SuarezFiguerola_Pasajero15"
[176] "SuarezFiguerola_Pasajero16"
[177] "SuarezFiguerola_Pasajero17"
[178] "SuarezFiguerola_Pasajero18"
[179] "SuarezFiguerola_Pasajero19"
[180] "SuarezFiguerola_Pasajero2"
[181] "SuarezFiguerola_Pasajero20"
[182] "SuarezFiguerola_Pasajero21"
[183] "SuarezFiguerola_Pasajero22"
[184] "SuarezFiguerola_Pasajero23"
[185] "SuarezFiguerola_Pasajero3"
[186] "SuarezFiguerola_Pasajero4"
[187] "SuarezFiguerola_Pasajero5"
[188] "SuarezFiguerola_Pasajero6"
[189] "SuarezFiguerola_Pasajero7"
[190] "SuarezFiguerola_Pasajero8"
[191] "SuarezFiguerola_Pasajero9"
[192] "Zayas_AventurarsePerdiendo1"
[193] "Zayas_AventurarsePerdiendo2"
[194] "Zayas_BurladaAminta1"
[195] "Zayas_BurladaAminta2"
[196] "Zayas_CastigoDeLaMiseria1"
[197] "Zayas_CastigoDeLaMiseria2"
[198] "Zayas_EscravaDeSuAmante1"
[199] "Zayas_EscravaDeSuAmante2"
[200] "Zayas_EstragosQueCausaElVicio1"
[201] "Zayas_EstragosQueCausaElVicio2"
[202] "Zayas_FuerzaDeAmor"
[203] "Zayas_InocenciaCastigada"
[204] "Zayas_PrevenidoEnganado1"
[205] "Zayas_PrevenidoEnganado2"

```
[206] "Zayas_TardeLlegaElDesengano"  
[207] "Zayas_VerdugoDeSuEsposa"
```

aplicamos el procedimiento para poder utilizar imposters(). Debemos indicar que texto queremos comparar. En este caso es la Tía Fingida que se corresponde con el número 98. Después le decimos al programa que en esta comparación quite este texto de la lista de arriba. La medida de distancia por defecto es Delta de Burrows.

```
> my_text_to_be_tested = cervantes[98, ]  
> my_frequency_table = cervantes[-c(98),]  
> imposters(reference.set = my_frequency_table, test = my_text_to_be_tested)
```

No candidate set specified; testing the following classes (one at a time):

```
Aleman Carvajal Castillo Cervantes Eslava Espinel Gregorio  
Gonzalez Lope Montalban Montemayor Quevedo SalasBarbadillo  
SuarezFigueroa  
Zayas
```

Testing a given candidate against imposters...

```
Aleman 0  
Carvajal 0  
Castillo 0.02  
Cervantes 0.98  
Eslava 0.01  
Espinel 0.02  
GregorioGonzalez 0  
Lope 0  
Montalban 0  
Montemayor 0  
Quevedo 0  
SalasBarbadillo 0.01  
SuarezFigueroa 0,02  
Zayas 0.05  
    Aleman Carvajal  
    0.00 0.00  
    Castillo Cervantes  
    0.02 0.98  
    Eslava Espinel  
    0.00 0.02  
GregorioGonzalez Lope  
    0.00 0.00  
    Montalban Montemayor  
    0.00 0.00  
    Quevedo SalasBarbadillo  
    0.00 0.01  
    SuarezFigueroa Zayas  
    0.02 0.05
```

vamos a suponer que todos los autores son candidatos y contrastar uno por uno los textos de cada autor con la TF

```
# ALEMAN  
> my_text_to_be_tested = cervantes[98,]  
> my_candidate = cervantes[1:23,]  
> my_imposters = cervantes[-c(98, 1:23),]  
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,  
candidate.set = my_candidate)  
Testing a given candidate against imposters...
```

```
Aleman 0  
Aleman 0
```

```
# CARVAJAL  
> my_candidate = cervantes[24:33,]  
> my_imposters = cervantes[-c(98, 24:33),]
```

```

> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

Carvajal          0
Carvajal
  0

# CASTILLO
> my_candidate = cervantes[34:50,]
> my_imposters = cervantes[-c(98, 34:50),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

Castillo          0.02
Castillo
  0.02

# CERVANTES
> my_candidate = cervantes[51:97,]
> my_imposters = cervantes[-c(98, 51:97),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

Cervantes          0.97
Cervantes
  0.97

# ESLAVA
> my_candidate = cervantes[99:111,]
> my_imposters = cervantes[-c(98, 99:111),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

Eslava            0.01
Eslava
  0.01

#ESPINEL
> my_candidate = cervantes[112:116,]
> my_imposters = cervantes[-c(98, 112:116),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

Espinel           0.01
Espinel
  0.01

#GREGORIO GONZÁLEZ
> my_candidate = cervantes[117:125,]
> my_imposters = cervantes[-c(98, 117:125),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

GregorioGonzalez  0.01
GregorioGonzalez
  0.01

# LOPE
> my_candidate = cervantes[126:133,]
> my_imposters = cervantes[-c(98, 126:133),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...

```

```
Lope      0
Lope
  0
```

```
# MONTALBAN
> my_candidate = cervantes[134:141,]
> my_imposters = cervantes[-c(98, 134:141),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...
```

```
Montalban      0
Montalban
  0
```

```
# MONTEMAYOR
> my_candidate = cervantes[142:152,]
> my_imposters = cervantes[-c(98, 142:152),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...
```

```
Montemayor     0
Montemayor
  0
```

```
# QUEVEDO
> my_candidate = cervantes[153:159,]
> my_imposters = cervantes[-c(98, 153:159),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...
```

```
Quevedo        0
Quevedo
  0
```

```
# SALAS BARBADILLO
> my_candidate = cervantes[160:168,]
> my_imposters = cervantes[-c(98, 160:168),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...
```

```
SalasBarbadillo      0.03
SalasBarbadillo
  0.03
```

```
# SUAREZ FIGUEROA
```

```
> my_candidate = cervantes[169:191,]
> my_imposters = cervantes[-c(98, 169:191),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...
```

```
SuarezFigueroa  0.01
SuarezFigueroa
  0.01
```

```
# ZAYAS
```

```
> my_candidate = cervantes[192:207,]
> my_imposters = cervantes[-c(98, 192:207),]
> imposters(reference.set = my_imposters, test = my_text_to_be_tested,
candidate.set = my_candidate)
Testing a given candidate against imposters...
```

```
Zayas 0.15
Zayas
0.15
```

```
# A continuación hacemos las pruebas con todo el dataset(cervantes). Comprobamos los resultados de las diferentes medidas de distancia
```

```
# WURZBURG
```

```
> my_text_to_be_tested = cervantes[98,]
> my_reference_set = cervantes[-c(98),]
> imposters(my_reference_set, my_text_to_be_tested, distance = "wurzburg")
```

```
No candidate set specified; testing the following classes (one at a time):
```

```
Aleman Carvajal Castillo Cervantes Lope Montalban SalasBarbadillo Zayas
```

```
Testing a given candidate against imposters...
```

```
Aleman 0.03
Carvajal 0
Castillo 0
Cervantes 0.96
Eslava 0
Espinel 0
GregorioGonzalez 0
Lope 0
Montalban 0
Montemayor 0
Quevedo 0.05
SalasBarbadillo 0
SuarezFigueroa 0
Zayas 0.04
    Aleman Carvajal
    0.03 0.00
    Castillo Cervantes
    0.00 0.96
    Eslava Espinel
    0.00 0.00
GregorioGonzalez Lope
    0.00 0.00
    Montalban Montemayor
    0.00 0.00
    Quevedo SalasBarbadillo
    0.05 0.00
    SuarezFigueroa Zayas
    0.00 0.04
```

```
# EDER'S DELTA
```

```
> imposters(my_reference_set, my_text_to_be_tested, distance = "eder")
No candidate set specified; testing the following classes (one at a time):
```

```
Aleman Carvajal Castillo Cervantes Lope Montalban SalasBarbadillo Zayas
```

```
Testing a given candidate against imposters...
```

```
Aleman 0.02
Carvajal 0
Castillo 0.02
Cervantes 0.95
Eslava 0.04
Espinel 0.02
GregorioGonzalez 0.01
Lope 0.01
Montalban 0
Montemayor 0
Quevedo 0
```

```

SalasBarbadillo      0
SuarezFigueroa 0.01
Zayas      0.13
  Aleman      Carvajal
    0.02      0.00
  Castillo    Cervantes
    0.02      0.95
  Eslava      Espinel
    0.04      0.02
GregorioGonzález    Lope
    0.01      0.01
  Montalban    Montemayor
    0.00      0.00
  Quevedo      SalasBarbadillo
    0.00      0.00
  SuarezFigueroa    Zayas
    0.01      0.13

```

Para entender los datos anteriores utilizamos la función `imposters.optimize()`. Comprobamos los resultados de las tres distancias utilizadas anteriormente (Wurzburg, Eder, Minmax). Esta tarea necesita bastante tiempo de procesamiento.

```

# WURZBURG
> imposters.optimize(my_reference_set, my_text_to_be_tested, distance = "wurzburg")

```

```

Aleman_IIGuzman1 ...
Aleman_IIGuzman10 ...

```

#ETC (se muestra nuevamente la lista de los 207 textos)

```
[1] 0.00 0.89
```

```

# EDER'S DELTA
> imposters.optimize(my_reference_set,my_text_to_be_tested,distance = "eder")

```

```

Aleman_GuzmanI ...
Aleman_GuzmanII ...

```

#ETC (se muestra nuevamente la lista de los 207 textos)

```
[1] 0.00 0.95
```

La distancia por defecto es Delta de Burrows. Es la que utilizamos cuando no incluimos ningún parámetro de distancia específico:

```

> imposters.optimize(my_reference_set,my_text_to_be_tested)

```

```

Aleman_GuzmanI ...
Aleman_GuzmanII ...

```

#ETC (se muestra nuevamente la lista de los 207 textos)

```
[1] 0.00 0.78
```

4) Código: función `stylo.network()`

Utilizamos la siguiente función para poder extraer

```

stylo()
stylo.network()

```

13.2 Anexo Corpus

En las siguientes páginas se encuentran las tablas con el corpus que hemos utilizado. Las tres tablas representan los tres subcorpus que configuramos en base a los datos del CORDE, a saber: prosa narrativa breve, prosa narrativa extensa y otros textos en prosa. Este último no procede de los datos del CORDE si no que ha sido añadido por la semejanza formal de los textos que incluye. Recordemos que las obras cuyo nombre y autor se encuentran en las celdas con fondo en verde son aquellas que han pasado a pasar parte de nuestro corpus. Las que tienen su título y autor con fondo rojo no han sido utilizadas debido a diversas razones, ya sea por cuestiones literarias, de edición o por no estar disponibles en formato digital. En todo caso, la tabla incluye una justificación para cada texto. Como se puede ver hay obras incluidas de las cuales no tenemos todos los metadatos. En estos casos se hará patente utilizando también el mismo sistema.

Tabla 1: Corpus 1 – Prosa narrativa breve.

Autor	Título	Fecha de escritura	Primera edición	¿Inclusión en el corpus? / Disponibilidad / Edición	Número de palabras
Capitán Salazar	<i>Respuesta de Capitán Salazar [Cartas del Bachiller de Arcadia al Capitán Salazar]</i>	1550		Excluida por ser de atribución dudosa.	3527
Cervantes Saavedra, Miguel de	<i>La gitanilla. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmc513x2	23693
"	<i>El amante liberal. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmc1c1w3	18645
"	<i>Rinconete y Cortadillo. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmcwm184	13885
"	<i>La española inglesa. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmcrv0j2	16099
"	<i>El licenciado vidriera. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmcm29s9	9572
"	<i>La fuerza de la sangre. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmch9926	7722
"	<i>El celoso extremeño. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmccj8b4	13787
"	<i>La ilustre fregona. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmc7s7m3	20176
"	<i>Las dos doncellas. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmcbk195	15034
"	<i>La señora Cornelia. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmc6w988	14483
"	<i>El casamiento engañoso. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmc348j7	5084
"	<i>El coloquio de los perros. [Novelas ejemplares]</i>		1613	BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmctm763	23244

Hurtado de Mendoza, Diego	Carta. [<i>Cartas del Bachiller de Arcadia al Capitán Salazar</i>]	c1550		Excluida por ser de atribución dudosa.	4310
Pinedo, Luis de	<i>Libro de chistes</i>	c1550		No disponible en formato digital.	11576
Quevedo y Villegas, Francisco de	<i>Sueño del Juicio</i>	1606		BVMC Edición digital a partir de Quevedo, Francisco de, Sueños y discursos de verdades descubridoras de abusos, vicios y engaños en todos los oficios y estados del mundo, Barcelona, Esteban Liberós, a costa de Juan Sapera, 1627. http://www.cervantesvirtual.com/nd/ark:/59851/bmc3t9g1	3969
"	<i>Sueño del Alguacil endemoniado</i>	1607		Ibid.	4060
"	<i>Sueño del Infierno</i>	1608		Ibid.	14101
"	<i>Sueño de la Muerte</i>	1610		Ibid.	15436
"	<i>Sueño del Mundo por de dentro</i>	1610		Ibid.	7275
Rufo, Juan	<i>Las seiscientas apotegmas</i>		1596	Versión digital en txt: https://archive.org/stream/laseiscientasapo00rufo/laseiscientasapo00rufo_djvu.txt Rufo, J. (1923). <i>Las seiscientas apotegmas</i> , ed. Agustín de Amezúa, Madrid, Sociedad de Bibliófilos Españoles.	45606
Santa Cruz de Dueñas, Melchor de	<i>Floresta española</i>		1574	Versión digital en txt: escritura no modernizada y numerosos errores. http://www.liburuklik.euskadi.eus//applet/libros/JPG/parlamento/90010597/90010597.txt	51564
Timoneda, Juan de	<i>El sobremesa y alivio de caminantes. [El sobremesa y alivio de caminantes]</i>		1562-1569	Versión digital en html: http://parnaseo.uv.es/lemir/Textos/Sobremesa/Frames.html Escritura no modernizada Timoneda, J. (1999). <i>El sobremesa y alivio de caminantes</i> . Alberto Vidal Crespo, apud. Lemir, 3. Universidad de Valencia. Valencia.	15162
Timoneda, Juan de	<i>Buen aviso y portacuentos</i>		1564	Versión digital en txt: https://archive.org/stream/elbuenavisoypor00araggoog/elbuenavisoypor00araggoog_djvu.txt Escritura no modernizada – v/b, v/u, etc. Schevill, R. (1911). <i>El Buen aviso portacuentos</i> by JuanTimoneda. Rev Hisp, 24, 171-254.	18919
Timoneda, Juan de	<i>El Patrañuelo</i>		1567	Versión digital en html: http://www.cervantesvirtual.com/nd/ark:/59851/bmc7w6b0 Texto modernizado Timoneda, J. (1979). <i>El Patrañuelo</i> , ed. Rafael Ferreres, Madrid, Castalia.	48275
Villegas, Antonio de	<i>Novela del Abencerraje y Jarifa</i>	C1550-1565		Autoría dubitada Villegas /Jiménez de Urrea	7690

Tabla 2: Corpus 2 – prosa narrativa extensa.

Autor	Título	Fecha de escritura	Primera edición	¿Inclusión en el corpus? / Disponibilidad / Edición	Número de palabras
Alemán, Mateo	<i>Primera parte de Guzmán de Alfarache</i>		1599	BVMC – versión del texto en formato txt. facilitada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung) Reproducción digital de la edición de Madrid, Imp. Renacimiento, 1912, (Biblioteca Renacimiento. Obras maestras de la literatura universal) https://sirio.ua.es/libros/BEducacion/guzman_de_alfarache/index.htm	108105
Alemán, Mateo	<i>Segunda parte de la vida de Guzmán de Alfarache.</i>		1555	Disponible en formato digital, html en: BVMC Alemán, M. (2000). <i>Guzmán de Alfarache. Segunda parte</i> . Biblioteca Virtual Miguel de Cervantes. http://www.cervantesvirtual.com/nd/ark:/59851/bmc4q7s8	143934
Cervantes Saavedra, Miguel de	<i>La Galatea</i>		1585	Disponible en formato digital en html en: BVMC Saavedra, C. (2001). Miguel de, La Galatea [1585]. Ed. Florencio Sevilla Arroyo. Alicante: Biblioteca Virtual Miguel de Cervantes. http://www.cervantesvirtual.com/nd/ark:/59851/bmcj9638	120801
Cervantes Saavedra, Miguel de	<i>El ingenioso hidalgo don Quijote de la Mancha</i>		1605	Excluiremos esta obra por su extensión y por la variedad de temas y subgéneros que aparecen incluidos en ella, es decir por su complejidad que exige un trabajo específico al respecto.	187634
Cervantes Saavedra, Miguel de	<i>Segunda parte del ingenioso caballero don Quijote de la Mancha</i>		1615	Ibid.	198691
Cervantes Saavedra, Miguel de	<i>Los trabajos de Persiles y Sigismunda</i>		1616	Disponible en html en: Los trabajos de Persiles y Sigismunda. (2021, agosto 29). Wikisource, La Biblioteca Libre. Consultado el 14:24, octubre 16, 2021 en https://es.wikisource.org/wiki/Los_trabajos_de_Persiles_y_Sigismunda Versión en formato txt. realizada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung)	145144
Duque de Estrada, Diego	<i>Comentarios del desengañado de sí mismo. Vida del mismo autor</i>	1607-1645		No disponible en formato digital libre (pago)	148733
Eslava, Antonio de	<i>Noches de invierno</i>		1609	Disponible en pdf: https://parnaseo.uv.es/Lemir/Revista/Revista24/Textos/02_Noches_Invierno.pdf Figaredo, E., S. (2020) Antonio de Eslava, «Noches de invierno». <i>Lemir</i> , (24), 133-266.	78166

Espinel, Vicente	<i>Vida del escudero Marcos de Obregón</i>		1618	Disponible en html: http://www.cervantesvirtual.com/nd/ark:/59851/bmckw5b8 Espinel, V. (1881). Vida del escudero Marcos de Obregón (Vol. 11). Biblioteca "Arte y letras".	123136
Fernández de Avellaneda, Alonso	<i>Don Quijote de la Mancha</i>		1614	Disponible en html en BVMC http://www.cervantesvirtual.com/nd/ark:/59851/bmcsx686 Versión en formato txt. realizada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung)	141150
Fernández, Jerónimo	<i>Belianís de Grecia</i>		1547	Disponible solo parcialmente en BVMC	321828
García, Carlos	<i>La desordenada codicia de los bienes ajenos</i>		1619	Versión digital (escaneo) no legible computacionalmente. En BVMC	28282
Gálvez de Montalvo, Luis	<i>El Pastor de Filida</i>		1582	Versión digital (escaneo) no legible computacionalmente. En BVMC	64863
González, Gregorio	<i>El guitón Onofre</i>	1604		Disponible en html en: http://www.vallenajerilla.com/berceo/gregoriogonzalez/guitononofre.htm González, G., y de las Rivas, M. (1995). El Guitón Onofre (No. 5). Gobierno de La Rioja Tud.	58448
Hernández de Villaubrales, Pedro	<i>Peregrinación de la vida del hombre</i>		1552	Disponible en html en: http://www.cervantesvirtual.com/nd/ark:/59851/bmcs75d1 Edición digital basada en la de Medina del Campo, Guillermo Millis, 1552 . Ejemp. de la Biblioteca Nacional (España).	129600
López de Úbeda, Francisco	<i>La pícaro Justina</i>		1605	BVMC	138977
Luján, Pedro de	<i>Coloquios matrimoniales</i>		1550	Disponible en versión digitalizada (escaneo) no legible computacionalmente en: https://trobes.uv.es/permalink/34CVA_UV/um6gse/alma991006735679706258	
Luna, Juan de	<i>Segunda parte del Lazarillo de Tormes</i>		1620	Versión del texto en formato txt. facilitada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung)	24211
Mateo Luján de Saavedra (Juan Martí)	<i>Segunda parte de la vida del pícaro Guzmán de Alfarache</i>		1602		93836
Montemayor, Jorge de	<i>Los siete libros de La Diana</i>		1559	Disponible en html. en: BVMC Edición digital basada en la de Barcelona, en casa de Jayme Cortey, 1561. http://www.cervantesvirtual.com/nd/ark:/59851/bmc7s7m7	73450
Núñez de Reinoso, Alonso	<i>Los amores de Clareo y Florisea y los trabajos de la sin ventura Isea</i>		1552	No disponible en formato digital	49813

Ortúñez de Calahorra, Diego	<i>Espejo de príncipes y caballeros. [El caballero del Febo]</i>		1555	No disponible en formato digital	463134
Romero de Cepeda, Joaquín	<i>La historia de Rosián de Castilla</i>	1586		No disponible en formato digital	52129
Salas Barbadillo, Alonso Jerónimo de	<i>La ingeniosa Elena (La hija de Celestina)</i>	1614		Disponible en pdf. En BVMC Edición digital a partir de Strasburgo, J. H. Ed. Heitz, [191-?]. http://gaudi.ua.es/uhtbin/cgisirsi/x/0/0/5/?searchdata1=(OCoLC)941955534 Versión en formato txt. realizada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung)	34868
Salas Barbadillo, A.J. de	<i>El caballero puntual, primera parte</i>		1614	No disponible en formato digital	33795
Salas Barbadillo, A.J. de	<i>El caballero puntual, segunda parte. [El caballero puntual, segunda parte]</i>	1619		No disponible en formato digital	31187
Salas Barbadillo, Alonso Jerónimo de	<i>Los prodigios del amor. [El caballero puntual, segunda parte]</i>	1619		No disponible en formato digital	13870
Salas Barbadillo, Alonso Jerónimo de	<i>El caballero perfecto</i>		1620	No disponible en formato digital	40571
Salas Barbadillo, Alonso Jerónimo de	<i>El sagaz Estacio, marido examinado</i>		1620	Versión digital en html. disponible en: BVMC Edición digital basada en la de Madrid, Espasa-Calpe, 1958, pp. 65-264. http://www.cervantesvirtual.com/nd/ark:/59851/bmc9c6v0 Versión en formato txt. realizada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung)	42431
Segura, Juan de	<i>Proceso de cartas de amores</i>		1548	Proceso de cartas de amores	15479
Sierra, Pedro de la	<i>Espejo de príncipes y caballeros, segunda parte</i>		1580	Disponible solo parcialmente en BVMC y en formato digitalizado (escaneo) en http://bdh-rd.bne.es/viewer.vm?id=0000043653	190346

Suárez de Figueroa, Cristóbal	<i>La constante Amarilis</i>		1609	Versión digitalizada (escaneo) en BVMC Edición digital a partir de Madrid, por D. Antonio de Sancha, 1781. http://biblioteca.ua.es/	52765
Suárez de Figueroa, Cristóbal	<i>El pasajero</i>		1617	Disponible en pdf. En: http://parnaseo.uv.es/lemir/Revista/Revista22/Textos/4_El_Pasajero.pdf Suárez de Figueroa, C. (2015). El pasajero. Advertencias utilísimas a la vida humana. Ed. Enrique Suárez Figaredo. Web, 25. Versión en formato txt. realizada por el departamento de Humanidades Digitales del SUB Göttingen (Forschung und Entwicklung)	145019
Urrea, Jerónimo de	<i>Primera parte del libro del invencible caballero don Clarisel de las Flores</i>	a1574		Disponible solo parcialmente en BVMC: Edición digital a partir de Antología de libros de caballerías castellanas, Alcalá de Henares (Madrid), Centro de Estudios Cervantinos, 2001, pp. 157-162 http://www.cervantesvirtual.com/nd/ark:/59851/bmchd8g0	97372
Valbuena, Bernardo de	<i>Siglo de Oro en las selvas de Erifile</i>		1608	Disponible en formato pdf en BVMC pero escritura no modernizada y procedente de un escaneo. Edición digital a partir de En Madrid, por Alonso Martin, a costa de Alonso Perez, 1608 http://www.cervantesvirtual.com/nd/ark:/59851/bmcjm2d7	49126
Valladares de Valdelomar, Juan	<i>Caballero venturoso</i>	1617		No disponible en formato digital.	136853
Vega Carpio, Lope de	<i>El peregrino en su patria</i>		1604	No disponible en formato digital	99543
Vega Carpio, Lope de	<i>Pastores de Belén, prosas y versos divinos</i>		1612	No disponible en formato digital	97410

Tabla 3: Corpus 3 - Otras novelas del Siglo de Oro.

Autor	Obra	Fecha de escritura	Primera edición	Inclusión en corpus / disponibilidad / Edición	Número de palabras
Castillo Solórzano, Alonso de	<i>La fantasma de Valencia</i>	1625		https://github.com/7PartidasDigital/NovelaBarroca/find/master contrastada con BVMC: Edición digital a partir de <i>Tardes entretenidas</i> , en Madrid, por viuda de Alonso Martín, 1625, pp. 54-85 y cotejada con la edición crítica de Evangelina Rodríguez, <i>Novelas amorosas de diversos ingenios del siglo XVII</i> , Madrid, Castalia, 1986, pp. 169-200. http://www.cervantesvirtual.com/nd/ark:/59851/bmck9356	86905 (<i>Tardes entretenidas</i>)
Castillo Solórzano, Alonso de	<i>Proteo de Madrid</i>	1625		https://github.com/7PartidasDigital/NovelaBarroca/find/master Novela incluida en <i>Tardes entretenidas</i> , Madrid: por la viuda de Alonso Martín, 1625.	13876
Castillo Solórzano, Alonso de	<i>Lisardo enamorado</i>	1628-1629		https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con BVMC: Edición digital basada en la edición de Madrid, Gráficas Ultra, 1947. http://www.cervantesvirtual.com/nd/ark:/59851/bmc2v2d4	68214
Castillo Solórzano, Alonso de	<i>Las dos dichas sin pensar</i>	1631		https://github.com/7PartidasDigital/NovelaBarroca/find/master Novela incluida en: <i>Noches de placer</i> , Barcelona: Sebastián de Cormellas, 1631	10607
Castillo Solórzano, Alonso de	<i>Las harpías en Madrid y coche de las estafas</i>		1631	https://github.com/7PartidasDigital/NovelaBarroca/find/master Dividido en 4 partes. Contrastado con BVMC: Edición digital basada en la de Barcelona, Por Sebastián Cormellas, 1631. http://www.cervantesvirtual.com/nd/ark:/59851/bmcsb408	39532
Castillo Solórzano, Alonso de	<i>Fuerza castigada</i>			https://github.com/7PartidasDigital/NovelaBarroca/find/master Noveña incluida en: <i>Noches de placer</i> , Barcelona: Sebastián de Cormellas, 1631	8655
Castillo Solórzano, Alonso de	<i>Aventuras del Bachiller Trapaza</i>		1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con BVMC: Edición digital basada en la de Zaragoza, Pedro Vergés, 1637. http://www.cervantesvirtual.com/nd/ark:/59851/bmcd21t8	71405
Carvajal y Saavedra, Mariana de	<i>Del amante venturoso</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con BVMC: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	4214
Carvajal y Saavedra, Mariana de	<i>Amar sin saber a quien</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	8855

Carvajal y Saavedra, Mariana de	<i>Celos vengan desprecios</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	3856
Carvajal y Saavedra, Mariana de	<i>La dicha de Doristeá</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	6302
Carvajal y Saavedra, Mariana de	<i>Esclavo de su esclavo</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	5019
Carvajal y Saavedra, Mariana de	<i>La industria vence desdeñas</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	11905
Carvajal y Saavedra, Mariana de	<i>Quien bien obra siempre acierta</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	3708
Carvajal y Saavedra, Mariana de	<i>La venus de Ferrara</i>		1663	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con: Navidades de Madrid y noches entretenidas / Mariana de Carvajal ; ed. anotada de Dámaso Chicharro Chamorro ; posfacio de Rosa Navarro Durán http://www.bibliotecavirtualdeandalucia.es/catalogo/es/consulta/registro.do?id=1044737	5704
Vega Carpio, Lope de	<i>Las fortunas de Diana.</i> [Novelas a Marcia Leonarda]		1621	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastada con BVMC: Edición digital a partir de Vega, Lope de La Filomena, con otras diversas rimas, prosas y versos, Madrid, en casa de la viuda de Alonso Martín, a costa de Alonso Pérez, 1621. http://www.cervantesvirtual.com/nd/ark:/59851/bmc251g5	14227
Vega Carpio, Lope de	<i>La desdicha por la honra.</i> [Novelas a Marcia Leonarda]		1623	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastada con BVMC: Edición digital a partir de Vega, Lope de La Circe, con otras rimas y prosas, Madrid, en casa de la viuda de Alonso Martín, a costa de Alonso Pérez, 1624. http://www.cervantesvirtual.com/nd/ark:/59851/bmcxd0v5	11463

Vega Carpio, Lope de	<i>Guzmán el bravo.</i> [Novelas a Marcia Leonarda]	1623	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastada con BVMC: Edición digital a partir de Lope de Vega, La Circe, con otras rimas y prosas, Madrid, en casa de la viuda de Alonso Martín, a costa de Alonso Pérez, 1624. http://www.cervantesvirtual.com/nd/ark:/59851/bmcj38n9	11847
Vega Carpio, Lope de	<i>La prudente venganza.</i> [Novelas a Marcia Leonarda]	1623	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastada con BVMC: Edición digital a partir de Lope de Vega La Circe, con otras rimas y prosas, Madrid, en casa de la viuda de Alonso Martín, a costa de Alonso Pérez, 1624. http://www.cervantesvirtual.com/nd/ark:/59851/bmcnv9d1	11670
Montalbán, Juan Pérez de	<i>El envidioso castigado</i>	1626	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con parnaseo.uv.es: Figaredo, E. S. (2014). Juan Pérez de Montalbán, Sucesos y prodigios de amor. Lemir: Revista de Literatura Española Medieval y del Renacimiento, (18), 611-800. https://dialnet.unirioja.es/servlet/articulo?codigo=4609747&orden=1&info=link	11066
Montalbán, Juan Pérez de	<i>La fuerza del desengaño</i>	1626	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con parnaseo.uv.es: Figaredo, E. S. (2014). Juan Pérez de Montalbán, Sucesos y prodigios de amor. Lemir: Revista de Literatura Española Medieval y del Renacimiento, (18), 611-800. https://dialnet.unirioja.es/servlet/articulo?codigo=4609747&orden=1&info=link	10085
Montalbán, Juan Pérez de	<i>La hermosa Aurora</i>	1626	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con parnaseo.uv.es: Figaredo, E. S. (2014). Juan Pérez de Montalbán, Sucesos y prodigios de amor. Lemir: Revista de Literatura Española Medieval y del Renacimiento, (18), 611-800. https://dialnet.unirioja.es/servlet/articulo?codigo=4609747&orden=1&info=link	11032
Montalbán, Juan Pérez de	<i>La mayor confusión</i>	1626	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con parnaseo.uv.es: Figaredo, E. S. (2014). Juan Pérez de Montalbán, Sucesos y prodigios de amor. Lemir: Revista de Literatura Española Medieval y del Renacimiento, (18), 611-800. https://dialnet.unirioja.es/servlet/articulo?codigo=4609747&orden=1&info=link	10940
Zayas y Sotomayor, María de	<i>Aventurarse perdiendo</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado en BVMC: Edición digital a partir de la edición de Agustín González de Amezúa, Novelas amorosas y ejemplares, Madrid, Real Academia Española, 1948 (Biblioteca Selecta de Clásicos Españoles. Serie II ; 8) http://www.cervantesvirtual.com/nd/ark:/59851/bmc4b317	12984
Zayas y Sotomayor, María de	<i>Burlada Aminta</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	11045
Zayas y Sotomayor, María de	<i>El castigo de la miseria</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	10757

Zayas y Sotomayor, María de	<i>La esclava de su amante</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	15190
Zayas y Sotomayor, María de	<i>Estragos que causa el vicio</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con BVMC: Edición digital a partir de la edición de Agustín González de Amezúa, Novelas amorosas y ejemplares, Madrid, Real Academia Española, 1948 (Biblioteca Selecta de Clásicos Españoles. Serie II ; 8) http://www.cervantesvirtual.com/nd/ark:/59851/bmc8p602	11466
Zayas y Sotomayor, María de	<i>Fuerza del amor</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	6995
Zayas y Sotomayor, María de	<i>Inocencia castigada</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Contrastado con BVMC: Edición digital a partir de la edición de Agustín González de Amezúa, Novelas amorosas y ejemplares, Madrid, Real Academia Española, 1948 (Biblioteca Selecta de Clásicos Españoles. Serie II ; 8) http://www.cervantesvirtual.com/nd/ark:/59851/bmcpz569	9512
Zayas y Sotomayor, María de	<i>Prevenido engañado</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	13674
Zayas y Sotomayor, María de	<i>Tarde llega el desengaño</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	9758
Zayas y Sotomayor, María de	<i>Verdugo de su esposa</i>	1637	https://github.com/7PartidasDigital/NovelaBarroca/find/master Zayas y Sotomayor, M. D. (2001). Obra narrativa completa. Ed. Estrella Ruiz Gálvez-Priego. Fundación José Antonio de Castro. No disponible en formato digital.	8029

