



Universidad de Valladolid



PROGRAMA DE DOCTORADO EN
“ESPAÑOL: LINGÜÍSTICA, LITERATURA Y COMUNICACIÓN”

FACULTAD DE FILOSOFÍA Y LETRAS

TESIS DOCTORAL

**Aplicación de los fundamentos de la Lingüística
forense a los estudios de autoría literaria: *La cárcel de
Sevilla y Los romances*, dos entremeses atribuidos a
Miguel de Cervantes Saavedra**

Presentada por
CRISTINA RUIZ URBÓN
para optar al grado de Doctor/a por la Universidad de Valladolid

Dirigida por
el DR. FRANCISCO JAVIER BLASCO PASCUAL

A mis sobrinos, Nacho y Javier

“Éste que veis aquí, de rostro aguileño, de cabello castaño, frente lisa y desembarazada, de alegres ojos y de nariz corva, aunque bien proporcionada; las barbas de plata, que no ha veinte años que fueron de oro, los bigotes grandes, la boca pequeña, los dientes ni menudos ni crecidos, porque no tiene sino seis, y éstos mal acondicionados y peor puestos, porque no tienen correspondencia los unos con los otros; el cuerpo entre dos extremos, ni grande, ni pequeño, la color viva, antes blanca que morena; algo cargado de espaldas, y no muy ligero de pies; 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”.

(Miguel de Cervantes Saavedra)

“Por mi parte, me estimula e ilumina y apasiona tanto leer a Cervantes que tener acceso a una nueva página suya es un gozo muy grande. Pero no todos piensan así. Esta línea de investigación ¿se considera todavía algo indigna, un poco como investigar la prostitución o algún otro tema supuestamente «indecente»? Espero que el mundo haya mejorado en esta hostilidad crítica. Para mí, la identificación de nuevas obras o aun nuevas supuestas o posibles obras del mayor autor de la literatura española es una tarea de cierta relevancia. Pero diga uno lo que diga, estas tesis tienen una vida llena de dificultades. Nadie se pronuncia positivamente sobre ellas. Son difíciles de publicar: una ponencia sobre la identificación de lo que considero un fragmento de la fantástica, fantasmagórica, mágica *Semanas del jardín*, la obra cuyo título hipnotizaba, fue rechazada dos veces. He encontrado la misma reticencia en cuanto a mi reconstrucción de una obra sí perdida, pero innegablemente cervantina, su *Bernardo*, «famoso» según su propia descripción”.

(Daniel Eisenberg)

“El lingüista también puede abordar el problema de la autoría cuestionada desde el principio según el cual todos los hablantes nativos tienen su propia versión individual y distinta del lenguaje que hablan y escriben, su propio *idiolecto*, y de la presunción de que dicho idiolecto se manifiesta a través de una selección textual diferenciada e idiosincrásica”.

(Malcolm Coulthard)

Índice

RESUMEN Y PALABRAS CLAVE	p. 17
ABSTRACT Y KEYWORDS	p. 19
AGRADECIMIENTOS	p. 21
INTRODUCCIÓN	p. 23

CAPÍTULO 1

EL AUTOR COMO ENIGMA

1.1. Los misterios literarios	p. 31
1.2. El enigma de la autoría	p. 33
1.2.1. Breve apunte histórico sobre el autor literario	p. 34
1.2.1.1. La conciencia literaria como cualidad inherente a la condición humana	p. 34
1.2.1.2. La Antigüedad clásica: la búsqueda de reconocimiento público a través de la escritura y el plagio como “ilícito moral” ..	p. 36
1.2.1.3. El concepto de autoría en el Medievo. La imitación, los anónimos y las falsificaciones	p. 41
1.2.1.4. La invención de la imprenta y sus consecuencias: los privilegios de impresión, el nuevo sistema editorial y la profesionalización del oficio de escritor	p. 44

1.2.1.5. El nacimiento de la propiedad intelectual como institución jurídica y el reconocimiento de las facultades patrimoniales del derecho de autor	p. 50
1.2.1.6. La gran paradoja del siglo XX	p. 54
1.2.1.7. El fin del anonimato literario: la “resurrección” del autor a través del idiolecto	p. 57
1.2.1.8. Reflexiones finales	p. 59
1.2.2. El complejo concepto de autoría en el teatro español del Siglo de Oro	p. 60

CAPÍTULO 2

LA EVIDENCIA LINGÜÍSTICA. LOS ESTUDIOS DE ATRIBUCIÓN DE AUTORÍA DE LA LINGÜÍSTICA FORENSE

2.1. La Lingüística forense	p. 69
2.1.1. Las ciencias forenses y los principios rectores de la Criminalística moderna	p. 69
2.1.2. El nacimiento de la Lingüística forense como disciplina científica	p. 72
2.1.3. Definición y campos de actuación de la Lingüística forense	p. 74
2.2. La evidencia lingüística como prueba pericial	p. 77
2.2.1. El idiolecto o estilo idiolectal	p. 77
2.2.2. La comparación forense de textos	p. 81
2.2.2.1. La identificación de locutores	p. 82
2.2.2.2. La atribución de autoría	p. 83
2.2.2.3. La detección de plagio	p. 84

2.3. Los estudios de autoría en el marco de la Lingüística forense	p. 86
2.3.1. Características de las muestras textuales	p. 86
2.3.1.1. Extensión de las muestras	p. 87
2.3.1.2. Calidad de las muestras	p. 87
2.3.1.3. Género textual y circunstancias comunicativas de las muestras	p. 88
2.3.1.4. Tema de las muestras	p. 89
2.3.1.5. Data de las muestras	p. 89
2.3.2. Metodología y tipología de los análisis	p. 90
2.3.3. Variables de análisis	p. 92
2.3.4. Herramientas de análisis	p. 95
2.3.5. Expresión de los resultados: escalas de opinión verbal vs. escalas de opinión estadística	p. 100
2.4. Aplicación de los principios teórico-metodológicos de la lingüística forense a los estudios de atribución de autoría literaria. Estado de la cuestión y perspectivas de futuro	p. 105

CAPÍTULO 3

CERVANTES, UN HOMBRE DE TEATRO

3.1. Cervantes y el teatro	p. 109
3.1.1. Un apasionado espectador teatral	p. 110
3.1.2. Un lector continuo de obras dramáticas	p. 120
3.1.3. Un teórico a caballo entre el sistema tradicional del mecenazgo y la escritura teatral como <i>modus vivendi</i>	p. 122

3.1.4	Un dramaturgo desencantado que escribe novelas	p. 130
3.2.	Tras la huella de la teatralidad cervantina	p. 136

CAPÍTULO 4.

LOS ENTREMESSES ATRIBUIDOS A MIGUEL DE CERVANTES SAAVEDRA

4.1.	Las atribuciones cervantinas	p. 137
4.2.	Historia crítica de los entremeses atribuidos a Cervantes	p. 142
4.2.1.	El “fetichismo cervantista” de la segunda mitad del XIX	p. 143
4.2.2.	Del entusiasmo al escepticismo en el asunto de los entremeses atribuidos a Cervantes	p. 148
4.2.3.	Nueva oleada de estudios sobre los entremeses atribuidos a Cervantes	p. 153
4.3.	Nómina de los diez entremeses atribuidos a Cervantes	p. 157
4.3.1.	<i>Entremés de los habladores</i>	p. 158
4.3.2.	<i>Entremés del hospital de los podridos</i>	p. 160
4.3.3.	<i>Entremés de la cárcel de Sevilla</i>	p. 161
4.3.4.	<i>Entremés de los refranes</i>	p. 162
4.3.5.	<i>Entremés de los mirones</i>	p. 164
4.3.6.	<i>Entremés de los romances</i>	p. 165
4.3.7.	<i>Entremés de Melisendra</i>	p. 166
4.3.8.	<i>Entremés de Durandarte y Belerma</i>	p. 169
4.3.9.	<i>Entremés de Doña Justina y Calahorra</i>	p. 171
4.3.10.	<i>Entremés de Ginetilla ladrón</i>	p. 172

CAPÍTULO 5.

APROXIMACIÓN AL CASO DE ESTUDIO

5.1. Objeto del estudio	p. 185
5.2. Marco metodológico	p. 186
5.3. Delimitación y descripción del corpus de análisis	p. 187
5.3.1. Delimitación del corpus de análisis y tratamiento de los textos ...	p. 187
5.3.1.1. Corpus dubitado: los diez entremeses alguna vez atribuidos a Cervantes	p. 188
5.3.1.2. Corpus indubitado del candidato a autor: las obras de Cervantes	p. 188
5.3.1.3. Corpus de respaldo: entremeses de época cervantina	p. 190
5.3.1.4. Otros textos de respaldo	p. 192
5.3.1.5. Localización y tratamiento de los textos	p. 192
5.3.2. Descripción de las muestras textuales	p. 193
5.3.2.1. La ridiculización del matrimonio y el adulterio	p. 195
5.3.2.2. El motivo del disfraz y otros engaños	p. 196
5.3.2.3. La cuestión de la limpieza de sangre	p. 197
5.3.2.4. Presencia de elementos costumbristas	p. 198
5.3.2.5. Personajes: tipos tradicionales y figuras	p. 199
5.3.2.6. Espacios de la ficción	p. 200
5.4. Operatividad del corpus de análisis	p. 201
5.4.1. Análisis de la distribución de la frecuencia de las palabras más empleadas en los textos (<i>Delta distance</i>)	p. 201
5.4.2. Análisis del porcentaje de similitud textual (<i>similarity</i>)	p. 210
5.4.3. Análisis de secuencias de palabras idénticas (<i>verbatim</i>)	p. 223

5.4.4. Análisis de la distribución de la frecuencia de n-gramas de 5 caracteres (<i>character 5-grams</i>)	p. 226
5.5. Hipótesis de trabajo	p. 230
5.6. Escala de opinión verbal para el caso de estudio	p. 253

CAPÍTULO 6.

ESTUDIO DE AUTORÍA DEL ENTREMÉS DE *LA CÁRCEL DE SEVILLA* (PROSA)

6.1. Antecedentes y estado de la cuestión	p. 255
6.2. ¿Pudo Cervantes escribir este entremés?	p. 260
6.3. Análisis de autoría basado en la comparación forense de textos	p. 261
6.3.1. Tema, argumento, personajes, segmentación y estructura	p. 264
6.3.1.1. Tema central y argumento	p. 264
6.3.1.2. Personajes	p. 269
6.3.1.3. Segmentación y estructura externa e interna	p. 270
6.3.2. Análisis cualitativo de marcas idiosincrásicas	p. 273
6.3.2.1. Presencia del superlativo en <i>-ísimo</i>	p. 274
6.3.2.2. Casos de leísmo	p. 275
6.3.2.3. Utilización de adverbios en <i>-mente</i>	p. 275
6.3.2.4. Uso del diminutivo en <i>-ico</i>	p. 276
6.3.2.5. Empleo de reduplicaciones, anáforas y repeticiones diseminadas	p. 277
6.3.2.6. Uso frecuente de figuras de dicción	p. 279
6.3.2.7. Creación de nuevas palabras	p. 281
6.3.2.8. Inclusión de expresiones latinas	p. 282

6.3.2.9. Predilección por los sinónimos coordinados	p. 283
6.3.2.10. Utilización pragmática de los diálogos de enunciados interrogativos	p. 284
6.3.2.11. Adaptación del lenguaje al nivel sociocultural de los personajes	p. 285
6.3.3. Distribución de la frecuencia de los 5-gramas de caracteres más empleados en los textos (<i>character 5-grams</i>)	p. 287
6.3.4. Distribución de la frecuencia de las palabras más empleadas en los textos (<i>Delta distance</i>)	p. 289
6.3.5. Análisis de componentes principales (PCA) aplicado a la <i>Delta</i> <i>distance</i>	p. 293
6.3.6. Clasificación supervisada por aprendizaje automático (Classify) ..	p. 294
6.3.7. Porcentaje de similitud textual (<i>similarity</i>), de vocabulario compartido una sola vez (<i>Once only in both</i>), de vocabulario compartido más de una vez (<i>Shared > 1</i>) y de vocabulario único y exclusivo (<i>Only in</i>)	p. 298
6.3.8. Clasificación secuencial de aprendizaje automático (Rolling classify)	p. 302
6.3.9. Verificación de autoría con General Imposters (GI)	p. 305
6.3.10. Longitud de secuencias de palabras idénticas (<i>verbatim</i>)	p. 306
6.3.11. Coocurrencias y redes léxicas	p. 308
6.3.12. Estudio de las palabras de función: orden de frecuencia, frecuencia relativa y PCA	p. 312
6.3.13. Ley de Zipf	p. 324
6.3.14. Función tf-idf	p. 326
6.3.15. Frecuencia de clases de palabras y bigramas y trigramas de secuencias morfosintácticas mediante etiquetadores gramaticales	

automáticos (POS tagging)	p. 328
6.4. Conclusiones sobre la autoría cervantina de este entremés	p. 335

CAPÍTULO 7.

ESTUDIO DE AUTORÍA DEL *ENTREMÉS DE LOS ROMANCES* (VERSO)

7.1. Antecedentes y estado de la cuestión	p. 339
7.2. ¿Pudo Cervantes escribir este entremés?	p. 349
7.3. Análisis de autoría basado en la comparación forense de textos	p. 351
7.3.1. Tema, argumento, personajes, segmentación y estructura	p. 554
7.3.1.1. Tema central y argumento	p. 354
7.3.1.2. Personajes	p. 359
7.3.1.3. Segmentación y estructura externa e interna	p. 359
7.3.2. Análisis cualitativo de marcas idiosincrásicas	p. 363
7.3.2.1. Presencia del superlativo en <i>-ísimo</i>	p. 363
7.3.2.2. Casos de leísmo	p. 364
7.3.2.3. Utilización de adverbios en <i>-mente</i>	p. 364
7.3.2.4. Uso del diminutivo en <i>-ico</i>	p. 365
7.3.2.5. Empleo de reduplicaciones, anáforas y repeticiones diseminadas	p. 366
7.3.2.6. Uso frecuente de figuras de dicción	p. 367
7.3.2.7. Creación de nuevas palabras	p. 367
7.3.2.8. Inclusión de expresiones latinas	p. 367
7.3.2.9. Predilección por los sinónimos coordinados	p. 367
7.3.2.10. Utilización pragmática de los diálogos de enunciados	

interrogativos	p. 368
6.4.2.11. Adaptación del lenguaje al nivel sociocultural de los personajes	p. 368
7.3.3. Análisis métrico (metros, estrofas y rimas)	p. 370
7.3.3.1. Estudio del metro	p. 371
7.3.3.1. Estudio de la estrofa y de los tipos de poema	p. 374
7.3.3.1. Estudio de la rima	p. 380
7.3.4. Distribución de la frecuencia de los 5-gramas de caracteres más empleados en los textos (<i>character 5-grams</i>)	p. 381
7.3.5. Distribución de la frecuencia de las palabras más empleadas en los textos (<i>Delta distance</i>)	p. 383
7.3.6. Análisis de componentes principales (PCA) aplicado a la <i>Delta distance</i>	p. 386
7.3.7. Clasificación supervisada por aprendizaje automático (Classify) ..	p. 388
7.3.8. Proporción palabras de contenido / palabras de función	p. 393
7.3.9. Medidas de riqueza y densidad léxica	p. 395
7.3.9.1. La ratio type/token de palabras de contenido (TTR Content)	p. 395
7.3.9.2. La riqueza léxica (<i>lexical richness</i>)	p. 398
7.3.10. Porcentaje de similitud textual (<i>similarity</i>), de vocabulario compartido una sola vez (<i>Once only in both</i>), de vocabulario compartido más de una vez (<i>Shared > 1</i>) y de vocabulario único y exclusivo (<i>Only in</i>)	p. 399
7.3.11. Verificación de autoría con General Imposters (GI)	p. 403
7.3.12. Longitud de secuencias de palabras idénticas (<i>verbatim</i>)	p. 403
7.3.13. Estudio de las palabras de función: orden de frecuencia,	

frecuencia relativa y PCA	p. 405
7.3.14. Frecuencia de clases de palabras	p. 416
7.3.15. Uso y frecuencia relativa de las conjunciones coordinadas y subordinadas adverbiales	p. 419
7.3.16. Búsqueda de palabras y expresiones del texto dubitado en corpus de referencia de la época	p. 422
7.4. Conclusiones sobre la autoría cervantina de este entremés	p. 430
CONCLUSIONES	p. 435
CONCLUSIONS	p. 444
SIGLAS	p. 453
GLOSARIO	p. 455
BIBLIOGRAFÍA	p. 461

RESUMEN Y PALABRAS CLAVE

Este proyecto de tesis doctoral plantea la aplicación de los fundamentos de la Lingüística forense a los estudios de autoría literaria, con el objetivo de dotarlos de mayor objetividad, fiabilidad y precisión. Estos fundamentos están relacionados con tres principios rectores esenciales: (i) uno teórico, fundado en el concepto del *idiolecto* como variante individual y distintiva de cada persona; (ii) otro metodológico, basado en la comparación forense de textos, que recomienda la combinación de técnicas cualitativas y cuantitativas y el estudio de variables que afecten a distintos niveles de la lengua; y (iii) otro conclusivo, que se sustenta en el principio de probabilidad, y que recomienda la expresión de nuestras conclusiones en escalas de opinión verbal. Desde este enfoque, hemos llevado a cabo el estudio de autoría de dos de los diez entremeses atribuidos a Miguel de Cervantes Saavedra (*La cárcel de Sevilla* y *Los romances*), a fin de poder determinar si son o no son suyos y en qué grado de probabilidad. La elección de estas dos obras nos lleva a su vez a reflexionar sobre el concepto de *autor*, los complejos procesos de creación, transmisión y edición textual del teatro del Siglo de Oro español y el *fetichismo cervantista* que provocó la atribución masiva de obras a Cervantes en el siglo XIX.

Palabras clave: Autoría, Entremeses, Cervantes, Lingüística forense, Idiolecto.

ABSTRACT Y KEYWORDS

This doctoral thesis proposes applying the fundamental principles of forensic linguistics to literary authorship studies in order to make them more objective, more reliable and more precise. These fundamental principles are related to three essential guiding principles: (i) a theoretical one, based on the concept of *idiolect* as an individual and distinctive variant of each person; (ii) a methodological one, based on the forensic comparison of texts, which recommends combining qualitative and quantitative techniques and the study of variables that might affect different levels of language; and (iii) a conclusive one, based on the principle of probability, which recommends the expression of our conclusions in verbal opinion scales. From this approach, we carried out a study of the authorship of two of the ten *entremeses* (interludes) attributed to Miguel de Cervantes Saavedra (*La cárcel de Sevilla* and *Los romances*) to determine whether or not they are actually his and to what level of probability. The choice of these two works leads us in turn to reflect on the concept of authorship, the complex processes of creation, transmission and textual editing of the Spanish Golden Age theatre and the *Cervantine fetishism* in the 19th century that gave rise to the massive attribution of works to Cervantes.

Keywords: Authorship, Interludes, Cervantes, Forensic Linguistics, Idiolect.

AGRADECIMIENTOS

Hoy ve la luz un proyecto de investigación al que he dedicado varios años de mi vida, muchos esfuerzos y algún que otro desvelo. Por ello, llegado este punto, quiero dar las gracias a todos aquellos que me han acompañado en el camino:

A la Consejería de Educación de la Junta de Castilla y León, que me concedió una beca de Formación de Personal Investigador —cofinanciada por el Fondo Social Europeo— con la que pude iniciar este trabajo.

A los miembros del Departamento de Literatura Española y Teoría de la Literatura y Literatura Comparada de la Universidad de Valladolid, al que he estado vinculada primero como alumna, más tarde como becaria de investigación y ahora como profesora asociada. A Pilar Celma, por su ayuda, confianza y afecto. A Teresa Gómez Trueba, Carmen Morán y Héctor Urzaiz, por contar conmigo en sus proyectos de investigación. A Patricia Marín Cepeda, por su complicidad y compañerismo. Y a todos los demás, aunque no pueda detenerme a nombrarlos uno por uno. Pero sobre todo y sobre todos, a Javier Blasco, que ha sido para mí mucho más que un tutor. Javier: gracias por contagiarme tu pasión por la Estilometría y abrirme los ojos a un nuevo modo de entender la investigación; gracias por tu generosidad, tu amistad y tus consejos a lo largo de todos estos años; gracias por enseñarme tanto y tan bien.

A mis profesores de las licenciaturas en Filología Hispánica y en Teoría de la Literatura y Literatura Comparada, a los de los programas de doctorado “El Quijote y la novela moderna” y “Español: Lingüística, Literatura y Comunicación” de la

AGRADECIMIENTOS

Universidad de Valladolid y a los del Máster en Lingüística Forense de la Universitat Pompeu Fabra, por transmitirme sus enseñanzas y conocimientos.

A los tres tutores que he tenido durante mis estancias de investigación. A María Grazia Profeti, catedrática de la Università degli Studi de Florencia, que supo guiarme por el entramado mundo del teatro español del Siglo de Oro. A María Teresa Turell —descanse en paz—, catedrática de la Universitat Pompeu Fabra y directora del Laboratorio de Lingüística forense (ForensicLab) del Instituto Universitario de Lingüística Aplicada (IULA), que me enseñó los vericuetos de la Lingüística forense en una época en la que esta disciplina aún sonaba a “cuento chino”. Y a Abraham Madroñal, catedrático de la Universidad de Ginebra y el mejor maestro que he podido tener para entender el género entremesil desde sus múltiples perspectivas.

A Agilice Digital, Empresa de Base Tecnológica de la Universidad de Valladolid, donde desarrollo mi labor como lingüista forense desde hace ya varios años; una labor que me ha dado grandes satisfacciones a nivel personal y profesional, pero que, además, me ha permitido reflexionar sobre la necesidad de establecer unas bases metodológicas fiables para el desarrollo de los estudios de atribución de autoría en el ámbito literario.

A José Manuel Fradejas Rueda, por despertar en mí el gusanillo de los análisis cuantitativos y ayudarme con la instalación de algún que otro *software*.

A todos los investigadores y especialistas que he tenido la oportunidad de conocer en los cursos, seminarios y congresos a los que he asistido durante estos años, por su ayuda y disponibilidad. Con ellos he compartido horas de trabajo, inquietudes, muchos libros y más de un café.

A mis amigos y a mi familia. A mis padres Ramón y Concha, por respetar mi vocación y creer en mí; ellos son para mí un modelo de generosidad, honradez y constancia. A mi hermana Ana, por todos sus consejos y achuchones. Y a Marco, por aguantar mis malos días y no soltarme nunca de la mano.

Y, por supuesto, a Miguel de Cervantes. A él le debo el placer de la lectura, que “aviva los ingenios de los hombres”. Sin su figura y legado, nada de esto hubiera tenido sentido.

A todos, gracias.

INTRODUCCIÓN

Uno de los grandes retos de la historia literaria sigue siendo el de delimitar y establecer qué fue lo que escribió Miguel de Cervantes Saavedra. El hecho de que no conservemos los manuscritos autógrafos originales y la afirmación del propio Cervantes de que muchas obras suyas “andan por ahí descarriadas y, quizá, sin el nombre de su dueño” ha propiciado que muchos investigadores traten de ampliar el corpus indubitado de este autor. Gracias a los trabajos de Rius [1895-1905], Ford y Lansing [1931], Astrana Marín [1948-1958 VII: 751-767], Avallé-Arce [1973], Eisenberg [1990] o Montero Reguera [1995], poseemos una serie de catálogos que nos permiten conocer todas las obras que alguna vez han sido atribuidas, con mayor o menor autoridad, al gran genio de nuestras letras: medio centenar de poemas, algunas cartas, varias narraciones en prosa y unas cuantas piezas teatrales. Pero, ¿en qué término debemos considerar tales atribuciones?

La popularidad alcanzada por el *Quijote* desde época temprana hizo que muchos se interesaran por la vida y la obra de Cervantes y que, entre otros aspectos, trataran de averiguar cuáles eran aquellas *obras descarriadas* de las que él mismo hablaba en el prólogo de sus *Novelas ejemplares*. En el siglo XVIII se imprimieron con su nombre *La Numancia* y *Los tratos de Argel* (1784) y se apuntó la posibilidad de que fueran también obra suya algunos romances de la época—como “En la Corte está Cortés” (1737)—, pero no fue hasta el siglo XIX, con el auge de los nacionalismos y la conformación del *Quijote* como héroe universal, cuando los intentos por establecer la verdadera bibliografía del autor alcalaíno cobraron una especial relevancia. Así, entre 1814 y 1884 vieron la luz una serie de estudios que defendían la paternidad cervantina de un sinfín

INTRODUCCIÓN

de piezas literarias, con argumentos tan endebles como la coincidencia temática, la similitud estilística o la genialidad de la obra; y que daban muestra del interés de ciertos investigadores por conseguir fama y prestigio, pues tenía mucha más repercusión —y la sigue teniendo— adjudicar una obra a Cervantes que a un autor de segunda o tercera fila. El afán de notoriedad hizo que incluso alguno publicara con su nombre ciertas obras apócrifas; éste fue el caso de Adolfo de Castro, responsable de la falsificación de *El Buscapié* y probablemente también de la *Carta al Cardenal Sandoval y Rojas*.

Tras el descubrimiento de la artimaña literaria de Castro y el descrédito en el que cayeron muchos de sus defensores, el interés por atribuir nuevas obras a Cervantes desapareció por completo. En los primeros años del siglo XX, varias voces críticas se alzaron contra el *fetichismo cervantista* de los estudios precedentes, proclamando la necesidad de devolver al anonimato las obras atribuidas hasta ese momento; y aunque en las décadas posteriores sí hubo interés por estudiar o editar esos textos, el asunto de la autoría continuó estando *sub iudice*. Desde las dos últimas décadas del siglo XX, y mucho más intensamente en los últimos veinte años, los estudios de autoría en general, y los de Cervantes en particular, han vuelto a ocupar un lugar destacado en la investigación literaria. Algunos especialistas han revisado las atribuciones hechas en el pasado, bien para defenderlas o para desacreditarlas; otros han apuntado la paternidad cervantina de nuevas piezas, como la *Topografía e historia general de Argel* [Eisenberg, 1994] o *La conquista de Jerusalén por Godofre de Bullón* [Arata, 1992]; y hay incluso quienes han cuestionado la originalidad de *Rinconete y Cortadillo* o de *El celoso extremeño*, según nuestro criterio con argumentos de muy poco peso [Aylward, 1982].

Sea como fuere, lo cierto es que el debate sobre qué fue lo que escribió Cervantes sigue abierto, y aunque muchas de las atribuciones —fundamentalmente las de la época decimonónica— se sustenten en argumentos débiles e intereses concretos, consideramos necesario revisarlas desde la óptica del siglo XXI. Más aún en el caso de los entremeses, a los que la crítica ha prestado poca atención; tal vez por entender, erróneamente, que teatro menor es sinónimo de teatro peor¹.

¹ Desde mediados del siglo XVI, la palabra *entremés* (junto con *paso*) hace referencia a la pieza jocosa breve que se representaba en los entreactos de la comedia para evitar que cesara el espectáculo.

INTRODUCCIÓN

Por este motivo, el presente proyecto de tesis doctoral tiene por objeto estudiar la autoría de dos de los diez entremeses que alguna vez han sido atribuidos a Cervantes (*La cárcel de Sevilla* y *Los romances*), a partir de la comparación de esas piezas dubitadas con un conjunto de obras de la época cuya autoría no se cuestiona, tanto del propio Cervantes como de otros autores; a fin de determinar si tales textos fueron escritos o no por él y con qué grado de probabilidad. La elección de estas dos piezas, lejos de ser arbitraria, responde a un doble criterio: uno estilístico (*La cárcel de Sevilla* está escrita en prosa y *Los romances* en verso) y otro temático (*La cárcel de Sevilla* está vinculada con la *Relación de la cárcel de Sevilla*, cuya tercera parte se ha atribuido también a Cervantes, y *Los romances* con la primera salida de don Quijote).

En un contexto complejo como el de la atribución de autoría literaria, la interdisciplinariedad es absolutamente necesaria; máxime cuando hablamos de textos teatrales de más de cuatrocientos años, escritos unas veces en prosa y otras en verso y sometidos a un complejo proceso de transmisión textual. Por eso, aunque nuestro estudio se enmarca dentro de la Literatura Española del Siglo de Oro, hemos querido sustentarlo en los principios teórico-metodológicos del ‘lenguaje evidencial’ de la Lingüística Forense y complementarlo, además, con otras disciplinas afines y estrechamente interrelacionadas, como son la Estilometría y las Humanidades Digitales, la Sociolingüística variacionista, la Lingüística de Corpus, la Teoría de la Literatura, la Crítica Textual, la Literatura Comparada y la Métrica. Somos conscientes de que este acercamiento interdisciplinar complica enormemente la tarea de controlar de manera exhaustiva la bibliografía pertinente, pero aún así, creemos en la conveniencia de dotar a la investigación literaria de nuevos enfoques y métodos, pues sólo de esta manera podremos seguir avanzando hacia un conocimiento integral de la materia.

Aunque inicialmente estaba ligado a la acción principal de la comedia, a fines del XVI pasa a constituirse como un género con autonomía propia. Desde el punto de vista expresivo, es importante distinguir entre una primera etapa, cuyo cauce predilecto es la prosa (siguiendo el modelo ofrecido por *La Celestina* a través de Rueda), y una segunda etapa, a partir de 1620, en la que predomina el verso, tal vez por influencia de la comedia. Los entremeses de Cervantes deben entenderse como un punto de inflexión en el género en cuanto a personajes y motivos; a partir de él, el género se bifurca en dos tendencias: la del enfoque satírico costumbrista de los entremeses de figuras y la del entremés cantado [Huerta Calvo, dir., 2008: 70-82].

INTRODUCCIÓN

El principio teórico fundamental que sustenta todos los estudios de autoría, y también los de autoría literaria, es la idea de que todo individuo posee un estilo idiolectal propio, es decir, una realización individual e idiosincrásica del lenguaje que escribe, que permite diferenciarlo y distinguirlo de los demás. Partiendo de este principio general, trabajaremos con una metodología basada en la comparación de textos escritos con la finalidad de atribuir o ‘desatribuir’ a Cervantes la autoría de esas dos piezas teatrales, en función de la cercanía o lejanía que presenten las distintas variables (léxicas, morfológicas, sintácticas, semánticas...) de análisis lingüístico: cuanto menor sea la distancia entre el estilo idiolectal de cada uno de los entremeses atribuidos (textos dubitados) y el estilo idiolectal de las obras de Miguel de Cervantes (textos indubitados), mayor probabilidad existe de que hayan sido escritas por él; y al contrario. Utilizaremos además textos de respaldo de otros autores de la época que permitan determinar si las posibles similitudes entre los textos dubitados y los indubitados cervantinos son marcas propias e idiosincrásicas de autoría o si, por el contrario, responden a usos comunes de la lengua, de la época o del género literario estudiado.

La comparación de textos escritos con fines de atribución de autoría contempla infinidad de variables, como los *n-gramas* de caracteres, la frecuencia léxica, el estudio de las palabras de función, la ratio *type/token*, la secuencia de etiquetas gramaticales (POS tagging o POST), los casos de *verbatim*, la puntuación o la estructura discursiva. Nuestro método combinará los análisis cuantitativos con los cualitativos y tendrá en consideración tanto la forma lingüística externa como el contenido y disposición de los textos. Aunque en muchas ocasiones manejaremos datos numéricos, porcentuales y estadísticos, las conclusiones finales las expresaremos en escalas de opinión verbal, tal y como postula la comunidad científica para estos casos (pero que, hasta la fecha, no se han utilizado en los estudios de autoría literaria en el ámbito hispánico); estas escalas representarán los grados de probabilidad de la conclusión alcanzada y permitirán dar una respuesta más precisa a la cuestión de la autoría.

Hemos estructurado nuestro trabajo de investigación en siete capítulos, cuatro de corte teórico y tres de aplicación práctica.

INTRODUCCIÓN

En el primer capítulo ahondaremos en el concepto de autoría y en su evolución a lo largo de la historia, centrándonos de manera específica en el teatro aurisecular español. A pesar de que la conciencia autorial es connatural al ser humano, los escritores no han tenido siempre un reconocimiento social o una compensación económica por su trabajo intelectual, y menos aún una protección jurídica. Creemos conveniente ahondar en las causas que han propiciado el nacimiento del autor como entidad, primero social y más tarde económica, jurídica, teórica e idiolectal, pues sólo así podremos afrontar la compleja tarea de identificar la función-autor en los textos teatrales de finales del XVI y principios del XVII.

Dedicaremos el segundo capítulo al nacimiento de la Lingüística forense como disciplina científica y defenderemos la necesidad de aplicar su protocolo de actuación a los estudios de atribución de autoría literaria, haciendo especial hincapié en los principios de probabilidad y certeza. A pesar de que las evidencias externas han sido consideradas por muchos la única forma fiable de atribuir un texto a una determinada persona, lo cierto es que las evidencias internas se han convertido hoy —gracias a la constatación del idiolecto como marca personal— en un factor clave para atribuir autorías, con mayor fuerza discriminante, en ocasiones, que algunas evidencias externas (imaginemos, por ejemplo, el caso de los vulgarmente llamados ‘negros literarios’ o el de las obras publicadas bajo pseudónimo). Resulta por ello sorprendente que unos métodos de identificación avalados por lingüistas de todo el mundo, empleados en el seno de las investigaciones policiales y admitidos por los tribunales de todo el mundo no hayan sido acogidos aún —o no de la manera correcta— en el ámbito de la investigación literaria.

En el tercer capítulo abordaremos la figura de Cervantes como hombre de teatro, desde sus múltiples facetas como espectador, lector, teórico y escritor. Aunque el debate entre los que defienden su éxito como dramaturgo y los que abogan por su incapacidad para adaptarse a los preceptos de la comedia nueva sigue abierto, lo que no se puede negar es la importancia que el teatro como tal tuvo en la vida y en la obra de Cervantes. Aficionado desde niño a la farándula y ávido lector del escaso teatro impreso que circulaba en esa época, Cervantes se nos presenta como un escritor con una clara conciencia dramática, incluso en su faceta como novelista. Sus obras están plagadas de

INTRODUCCIÓN

referencias a *autores* y comediantes de su tiempo, de reflexiones sobre preceptiva dramática y de recursos metateatrales; de igual modo, el empleo del estilo directo en muchas de sus novelas cortas o en ciertos pasajes de su *Quijote*, junto al dinamismo de la palabra, son fruto para muchos de la reescritura de obras dramáticas perdidas o de planes en prosa de comedias, o sino, al menos, una evidencia más de su clara conciencia teatral. En definitiva, además del idiolecto, en Cervantes el teatro es una huella identificativa de su faceta como escritor.

El cuarto capítulo estará destinado a las atribuciones cervantinas y, más en concreto, al estudio de la historia crítica de los diez entremeses que alguna vez se le han adjudicado, con mayor o menor fortuna: *Los habladores*, *El hospital de los podridos*, *La cárcel de Sevilla*, *Los refranes*, *Los mirones*, *Melisendra*, *Los romances*, *Doña Justina* y *Calahorra*, *Durandarte* y *Belerma* y *Ginetilla ladrón*. Estas atribuciones, lejos de haber sido hechas con el rigor y la cautela que precisan los estudios académicos de este tipo, se han sustentado por lo general en la valía literaria de su autor o en la coincidencia de temas o estilos. Lo que nos lleva a insistir de nuevo en la necesidad de establecer una metodología fiable para la atribución de textos literarios, que permita escapar del *fetichismo cervantista* y que se sustente en la idiosincrasia del lenguaje.

Tras estos cuatro capítulos de corte eminentemente teórico, en el quinto capítulo ofreceremos un primer acercamiento práctico a la materia. Una vez delimitado el corpus de trabajo, realizaremos una serie de análisis cuantitativos que nos permitan testar la fiabilidad del corpus indubitado y extraer unas primeras hipótesis de trabajo en relación con los textos dubitados. Asimismo, estableceremos la escala de opinión verbal en la que expresar las conclusiones, basada en grados de probabilidad, tal y como postula la Lingüística forense para estos casos, pero que, a día de hoy, no se aplica todavía en los estudios de atribución de autoría literaria.

Los capítulos seis y siete estarán dedicados al estudio de atribución de autoría del *Entremés de los romances* (en verso) y del *Entremés de la cárcel de Sevilla* (en prosa), respectivamente.

El *Entremés de los romances*, publicado por primera vez en la *Parte tercera de las comedias de Lope de Vega y otros autores* (Valencia, 1611), recoge la historia de un ladrador recién casado que, obsesionado con las historias que lee en el Romancero,

decide “dejar su esposa y tierra” para enrolarse en la guerra contra los ingleses, llevándose a su vecino como sirviente y escudero. Fue atribuido a Cervantes por Adolfo de Castro, quien consideraba que el autor de *Quijote* “hizo lo que los grandes pintores: trazó un borrón o un ligero dibujo de un gran cuadro, primitivo pensamiento que luego desarrolló en un libro admirable” [1874: 140]; una atribución tan defendida por unos [Menéndez Pelayo, 1941; Northup, 1922b o, más recientemente, Rodríguez López-Vázquez, 2019] como rechazada por otros [Márquez Villanueva, 1983; Carreira, 1998; Rey Hazas, 2006 y 2007; Pérez López, 2009; Blasco, 2013]. Esta pieza presenta la peculiaridad de poner en boca de los personajes un sinfín de composiciones del Romancero nuevo, lo que convierte el estudio de su idiolecto en todo un reto.

El *Entremés de la cárcel de Sevilla*, publicado por primera vez en la *Séptima parte de las comedias de Lope de Vega* (Madrid, 1617), aborda la esperpéntica reacción de uno de los presos de esta cárcel al conocer que ha sido condenado a muerte. Aureliano Fernández Guerra [1863], José María Asensio [1902], Adolfo de Castro [1874] y Marcelino Menéndez Pelayo [1941] defendieron la paternidad cervantina, arguyendo tanto su cercanía temática con la *Relación de la Cárcel de Sevilla*, cuya tercera parte se disputan Chaves y Cervantes, como su cercanía editorial con *El entremés de los habladores*, también atribuido a Cervantes, a quien sigue tanto en la versión impresa como en la manuscrita.

Ambos estudios de autoría contemplarán, en esencia, unas mismas variables de análisis. Ahora bien, las diferencias de extensión (3240 palabras / 1164 palabras²) y composición (prosa / verso) entre los dos textos dubitados condicionarán la elección de alguna de ellas. Así, por ejemplo, en *La cárcel de Sevilla* analizaremos los bigramas y trigramas de clases de palabras, pero no en *Los romances*, pues el orden de estos elementos podría venir condicionado por su forma versificada. En *Los romances*, sin embargo, analizaremos el metro, la estrofa y la rima y, dado que el texto “real” apenas supera las 1000 palabras, adoptaremos otras variables menos susceptibles a la extensión.

² Aunque el *Entremés de los romances* tiene una extensión total de 2243 palabras, el texto real, es decir, aquél resultante de eliminar las composiciones de otros, es de 1164 palabras.

INTRODUCCIÓN

Aunque los dos capítulos prácticos dedicados a *La cárcel de Sevilla* y *Los romances* llevan un apartado de conclusiones propias, cerramos nuestro trabajo con unas reflexiones generales sobre los estudios de autoría literaria en general y los de los entremeses del Siglo de Oro en particular. Añadimos al final un listado de siglas y un glosario, para facilitar la consulta y comprensión de ciertos conceptos técnicos.

Posiblemente Cervantes sea el autor español al que más obras se le han atribuido, las más de las veces con argumentos intuitivos de muy poco peso e, incluso, con intereses particulares. Aunque útiles en su momento y dentro de su contexto, creemos que es necesario revisar todo ese *maremágnum* de filiaciones cervantinas. Confiamos en que nuestro trabajo contribuya a esta gran empresa, pero, sobre todo, que permita empezar a sentar las bases de lo que deberían ser —y lo que no— los estudios de autoría literaria.

CAPÍTULO 1

EL AUTOR COMO ENIGMA

“Pero los discursos “literarios” ya sólo pueden recibirse dotados de la función autor: a todo texto de poesía o de ficción se le preguntará de dónde viene, quién lo escribió, en qué fecha, en qué circunstancias o a partir de qué proyecto. El sentido que se le otorga, el estatuto o el valor que se le reconoce dependen del modo como responda a estas preguntas. Y si, como consecuencia de un accidente o de una voluntad explícita del autor, nos llega el anonimato, enseguida el juego consiste en encontrar al autor. No soportamos el anonimato literario: sólo lo aceptamos en calidad de enigma” (Michel Foucault).

1.1. LOS MISTERIOS LITERARIOS

La historia del mundo de las letras, y más aún la del Siglo de Oro español, está llena de enigmas ocultos: ¿dónde estudió Miguel de Cervantes?, ¿quién se esconde en realidad tras el pseudónimo de Alonso Fernández de Avellaneda?, ¿era Góngora ludópata?, ¿y Quevedo judío?, ¿es Micaela de Luján la famosa *Celia* que nombra Lope de Vega en muchos de sus versos?, ¿la comedia *Los carboneros de Francia y reina Sevilla* es de Rojas Zorrilla o de Mira de Amescua?, ¿fue Gil Polo, autor de la *Diana enamorada*, el mismo Gil Polo que figura como catedrático de Griego de la Universidad de Valencia entre 1566 y 1573?, ¿sirvió Calderón al ejército español en Italia y Flandes entre 1626 y 1635, tal y como confirma su biógrafo y amigo Vera Tassis, o debemos hacer caso de los numerosos documentos jurídicos que lo sitúan en Madrid por esos mismos años?

Estos son sólo algunos ejemplos, pero lo cierto es que la literatura del periodo aurisecular está plagada de biografías irrastreables; documentos no transcritos; fondos mal catalogados; manuscritos perdidos, ilegibles o incompletos; obras cercenadas por la acción censora o devastadas por los procesos inquisitoriales; ediciones *contrafactas*; pies de imprenta falsificados; escritos apócrifos, anónimos o mal atribuidos; obras sin data o con una data fingida; y una larga lista de etcéteras que complica enormemente la labor de ‘reconstrucción’ de nuestra historia literaria. Una historia literaria que, dicho sea de paso, se enfrenta además al difícil reto —para algunos, imposible— de dejar de ser selectiva, parcial y subjetiva³.

En este sentido, la historia de la literatura áurea debe entenderse como un puzle imperfecto, inacabado e inacabable. La agudeza y el riguroso trabajo de los investigadores va permitiendo que se descubran y ensamblen con gran acierto muchas de sus piezas; póngase por caso el hallazgo de una copia manuscrita de la comedia *Mujeres y criados* de Lope de Vega, que hasta hace poco se daba por perdida [García-Reidy, 2013a], o la localización en el Archivo General de Simancas de la parte inicial de una de las tres cartas autógrafas de Miguel de Cervantes que se conservan en The Rosenbach Museum and Library de Filadelfia [Ruiz García, 2015]. Pero pese al esfuerzo colectivo de los investigadores, aún quedan muchos lugares por rastrear y muchos misterios por resolver. Hay piezas que han sido descartadas por los

³ Gies explica la magnitud del problema de la siguiente manera: “La historia literaria es, por supuesto, una acumulación de detalles contiguos y un acto de olvido. [...] El problema es, según mi perspectiva, ¿cuántos detalles se pueden incluir, qué detalles serán, cómo se deben estructurar y, finalmente, cuánto olvido nos parece aceptable? ¿Cómo decide uno qué se debe «olvidar»? ¿Se «olvida» por razones ideológicas? ¿Por razones estéticas? ¿Por razones de espacio o estructura o poder o mera conveniencia? Si la memoria acumulada de individuos, grupos o naciones informa el acto de la creación de la historia literaria, entonces, ¿de quién es esa memoria? ¿Es una memoria nacional y colectiva, un repertorio de las «mejores» obras canónicas? ¿Es una memoria genérica? ¿Es una memoria racial? ¿O es una red de opiniones («opinión con fechas», según Valdés [2002: 74]), una selección personalizada basada en —¿qué?— gusto, popularidad, influencia, poder estético, contenido ideológico, coherencia temática o hallazgos casuales? [...] Si la historia literaria consiste en selecciones, ¿quién selecciona? ¿Y puede la selección producir algo semejante a la objetividad? Es más: ¿es posible la objetividad? ¿Es siquiera deseable? Como escribe David Perkins, uno de los defensores más sutiles de la empresa que llamamos historia literaria, ‘La única historia literaria completa sería el mismo pasado, y esto no sería una historia, porque no sería ni interpretiva [sic] ni explicativa’ (1992: 13). Esto es lo que finalmente descubrió Funes y, como sabemos, su incapacidad de olvidar le condujo a la locura” [2004: 274].

historiadores, al considerar —no se sabe muy bien con qué criterio— que no eran merecedoras de figurar en el puzle; otras que están perdidas, en fondos no rastreados o en archivos mal catalogados; otras que están mal colocadas, en volúmenes adulterados o contrahechos; otras que han sido manipuladas y encajadas a la fuerza en lugares que no les corresponden, con rúbricas, autorías o dataciones falsas; otras que parecen iguales, pero que son en realidad piezas homónimas distintas; otras que no sabemos muy bien dónde encajar, como las sueltas sin pie de imprenta, los textos anónimos o los escritos bajo pseudónimo; y otras que no podremos recuperar, pues fueron víctimas de la biblioclastia o la censura. Un panorama ciertamente complejo, que convierte el macrorompecabezas literario del Siglo de Oro en un *collage* imperfecto plagado de claroscuros.

1.2. EL ENIGMA DE LA AUTORÍA

Una de las principales incógnitas que a día de hoy trae de cabeza a los especialistas de la literatura aurisecular es, sin duda alguna, la de las autorías. La *autoría* es la cualidad de las personas que son causa de algo, que inventan alguna cosa o que han escrito una obra científica, literaria o artística. Existe un autor siempre que existe una obra, ya que “si algo caracteriza una obra como obra, es el hecho de ser creada” [Heidegger, 1958: 91]⁴.

Parece posible afirmar que las creaciones artísticas son connaturales al ser humano y que éste ha tenido siempre, en mayor o menor medida, una cierta conciencia de autoría; de hecho, las huellas de manos que aparecen junto a las primeras

⁴ El término *autor* proviene del latín AUCTOR, que deriva a su vez verbo latino AUGERE, que significa, ‘aumentar’, ‘agrandar’ o ‘acreecer’; el sufijo *-tor* indica ‘agente’, es decir, señala al sujeto que realiza la acción (como ocurre en *inventor*, *escultor*, *lector*, etc.). De este modo, un *autor* es la persona que hace que algo aumente o crezca, entendiendo ese crecimiento como creación, pues de hecho las palabras *crecer* y *crear* tienen un origen indoeuropeo común (el término latino CREARE procede del indoeuropeo KRĒ-YĀ, que significa ‘crecimiento’). No podemos olvidar además que, en el contexto del mundo clásico, la creación *ex nihilo* sólo es propia de los dioses.

manifestaciones del arte rupestre de diferentes partes del mundo son interpretadas como un sistema arcaico de rúbrica, como un signo distintivo de identificación personal⁵. Ahora bien, esa conciencia de ser autor de ‘algo intelectual’ no ha ido acompañada siempre de la idea de ser dueño de ese algo, y menos aún de tener una serie de derechos (tanto patrimoniales como morales) por ello⁶. De igual modo, y a pesar de esa conciencia primitiva, prácticas como el plagio, la falsificación o la usurpación han existido desde tiempos remotos, y sólo hoy empiezan a poder ser probadas de manera objetiva gracias a la identificación del idiolecto, esto es, la manifestación individual y distintiva de la lengua que escribimos.

Tal y como veremos en los siguientes apartados, el reconocimiento del autor como entidad social, profesional, jurídica, teórica e idiolectal es la consecuencia gradual de una serie de circunstancias sociales, culturales, científico-tecnológicas y económicas de muy diversa índole, y que es preciso conocer para poder entender la complejidad de identificar la función-autor en los textos teatrales de finales del XVI y principios del XVII.

1.2.1. Breve apunte histórico sobre el autor literario

1.2.1.1. La conciencia autorial como cualidad inherente a la condición humana

Los arqueólogos sostienen que los primeros sistemas de escritura surgen en la Mesopotamia de finales del IV milenio a. C., como un método mnemotécnico utilizado por los mercantes sumerios en sus transacciones comerciales; así, a través de la incisión

⁵ Breuil y Berger-Kirchner afirman que "probablemente la representación de la mano es una especie de sello personal de la humanidad primitiva, algo así como una firma que sirve de unión con el grupo étnico y con las fuerzas superiores" [1962: 49].

⁶ Como es algo a lo que haremos alusión a lo largo de este capítulo, creemos importante aclarar la diferencia entre facultades patrimoniales y morales en el ámbito de los derechos de autor: las primeras son las que permiten una compensación económica, como la reproducción, distribución, comunicación pública o transformación de la obra; mientras que las segundas son las que protegen la identidad y reputación de su creador (como el derecho de paternidad o de integridad, entre otros). Las facultades patrimoniales se pueden ceder a un tercero, pero las morales son irrenunciables.

en arcilla de diferentes figuras, podían recordar datos como la cantidad de ganado que se intercambiaban o los nombres de los compradores y los vendedores. La escritura pictográfica, formada por signos de carácter iconográfico, tenía unas posibilidades comunicativas muy reducidas: servía para reproducir palabras y objetos, pero no ideas o conceptos abstractos. Poco a poco este arcaico código de escritura se fue puliendo y perfeccionando, dando paso en el III milenio a. C. a la llamada escritura cuneiforme, un nuevo sistema de expresión basado ya no sólo en iconos gráficos, sino también en símbolos fonético-silábicos con los que poder expresar al fin ideas abstractas y elementos gramaticales.

Con el nacimiento de un código que permite transcribir los más hondos pensamientos del ser humano, la tradición literaria, que hasta ese momento se había transmitido de generación en generación por vía oral, empieza a ser fijada por los escribas sumerios en tablillas de barro; los relatos épicos, los libros de sabiduría, las historias humorísticas o los poemas laudatorios se convierten en las primeras muestras escritas de la literatura universal. La aparición de la literatura escrita acrecienta la conciencia creadora de los autores mesopotámicos: si presumiblemente siempre se habían sentido creadores —del mismo modo que se sentían como tales los autores prehistóricos del arte rupestre—, el paso de la transmisión oral a la escrita constituye para ellos un primer hito en la erradicación del anonimato involuntario, la transgresión textual o la apropiación indebida de sus creaciones. De hecho, los primeros textos conservados de carácter nominal se fechan precisamente en los estertores del III milenio de la era precristiana, en torno al año 2300 a. C., y pertenecen a la princesa sumeria Enheduanna de Akkad, hija del rey Sargón I de Acadia.

La misma conciencia de autoría que presenta el pueblo sumerio del Sur de Mesopotamia aparece en la civilización egipcia del valle del Nilo, que desde el II milenio a. C. utiliza la escritura jeroglífica para fijar los textos literarios en papiros, paredes, obeliscos, tumbas y pirámides. Muchas de estas primitivas muestras son anónimas, aunque otras muchas sí recogen el nombre del creador, como ocurre con *Las enseñanzas de Ptahhotep*. Y es que son los propios escribas egipcios (¿autores o copistas?) los que difunden la importancia de preservar el nombre de los autores para

CAPÍTULO 1

“ser recordados”, tal y como queda reflejado en el siguiente texto, escrito en torno al 1300 a. C.:

¡Sé escriba! ¡Graba esto en tu corazón / para que también tu nombre sobreviva! / El papiro es mejor que la piedra tallada. / Un hombre ha muerto: su cuerpo se convierte en polvo, / y sus familiares se extinguen. / Un libro es lo que hace que sea recordado / en la boca del hablante que lo lee [Citamos por Manguel, 1998: 214-215].

En el mundo clásico no sólo sigue vigente esa conciencia de autoría, sino que, además, los autores empiezan a reflexionar sobre el origen de la creación literaria y a reclamar, aún de manera muy endeble, ciertos derechos sobre sus obras⁷.

1.2.1.2. La Antigüedad clásica: la búsqueda de reconocimiento público a través de la escritura y el plagio como “ilícito moral”

Andrew Bennett abre el segundo capítulo de su obra *The Author. The New Critical Idiom* con la siguiente pregunta: “who was the first author in the Western on European tradition?”. Tras dar una serie de posibles nombres, afirma que la respuesta depende “on which criteria one takes, whether it is sheer chronological priority or something else: individuality, print publication, transcendent and universal genius, literary self-consciousness, financial independence, or even self-obsession” [2005: 30]. Une, en definitiva, la idea de la autoría a una serie de condicionantes que poco a nada tienen que ver con lo literario.

Bennett, siguiendo a Milman Parry y Albert B. Lord, cree que la concepción moderna de autoría literaria hunde sus raíces en la Antigua Grecia, a pesar de que resulte difícil aceptar, desde la óptica actual, la originalidad de un autor como Homero, que, de existir, hubo de heredar de otros no sólo los temas y las historias, sino también sus técnicas compositivas y sus fórmulas lingüísticas. La originalidad de los *aedos*

⁷ Nuestra revisión evolutiva sobre la noción de autor fija como punto de partida las antiguas Grecia y Roma. No obstante, cabe señalar que en etapas anteriores sí existieron ciertos vestigios de propiedad artística; así, por ejemplo, tenemos constancia de que la tribu africana de los “Negritos” mantenía la idea de que sólo el que compone un poema tiene derecho a recitarlo o de que los “Boquimanos”, indígenas del África austral, prohibían tocar las pinturas rupestres de autores conocidos [Dock: 1974, 129].

(‘cantores’) sólo puede ser entendida desde las reglas de la oralidad, en donde cada representación es “original” en sí misma, en tanto que difiere del resto de representaciones; una originalidad que finaliza en el momento en el que dicha tradición se fija por escrito, pues, a partir de ese momento, sólo el texto escrito es el “original” y todos los representados pasan a ser copias de aquél. En ese sentido, y al margen del debate sobre la existencia real de Homero, la *Ilíada* y la *Odisea* deben entenderse como textos de transición entre el modo de composición oral y el escrito, con todas las implicaciones que ello conlleva:

Homeric authorship, then, is fundamentally different from that of later, literate and especially print-based cultures, cultures which rely on the possibility of a stable 'text' and, to varying degrees, on a specific relationship between text and original writer, poet or 'author' And yet the process of the retrospective 'authorization' of the oral epic poet might also be said to share something with the authorization of writers in literate and print cultures [2005: 35].

Aunque la oralidad y la escritura no deban entenderse de manera excluyente, lo cierto es que la segunda no sólo posibilita la fijación de los textos en un material duradero (mucho más que la memoria), sino que acrecienta, como indicábamos antes, la conciencia autorial y la idea de la obra como un objeto físico susceptible de comercialización. Y es que, en contra de lo que pueda parecer, y a pesar de que la inexistencia de la imprenta dificultaba enormemente la reproducción del material artístico, sí existió en el mundo clásico una activa industria editorial⁸.

Cuando los escritores de la Grecia y la Roma antiguas optaban por transmitir la cultura literaria de manera escrita, entregaban sus manuscritos a los *bibliopolae* o *librarii*, que se encargaban de reproducirlos y distribuirlos en copias manuscritas. Esta fórmula de cesión era gratuita pero no desinteresada, pues los escritores buscaban

⁸ El comercio del libro es tan viejo como el libro mismo. Los libreros comenzaron por ser manufactureros, editores y vendedores al menudeo, pero, con el desarrollo de la industria, la labor del editor se fue separando de la del vendedor. Esta división de funciones no aparece en el mundo griego, pero sí en el romano. Los titulares de los talleres de impresión tenían a su servicio a personas (en su mayoría esclavos) que copiaban los manuscritos al dictado y, aunque la reproducción era lenta y trabajosa, las tiradas podían alcanzar hasta los mil ejemplares [Reyes, 2011].

CAPÍTULO 1

obtener popularidad social y alcanzar la gloria literaria⁹. No podemos hablar de un ánimo de lucro, pero sí de cierto ánimo de notoriedad, basado en la búsqueda del honor a través del reconocimiento público¹⁰:

Horacio pinta así, en irónico contraste, el caso de una obra que promete ser todo un éxito: «He aquí un libro que hará ganar dinero a los Sosii (sus editores), y que dará fama a su autor» [Reyes, 2011: 50].

Son constantes las referencias de los escritores griegos y romanos a lo que Juvenal llama “la hueca fama”. Ovidio, desde el destierro, se consuela pensando que sus escritos recorrerán todo el mundo, de Oriente hasta Occidente; Propertio se jacta de que se le conozca hasta en las comarcas septentrionales; y Horacio, del que ya hemos hablado antes, se enorgullece de que sus poemas se lean en las riberas del Bósforo, Hispania, Las Galias y África.

El estudio de los textos conservados evidencia la inexistencia de cualquier tipo de derecho patrimonial sobre la obra intelectual, con independencia del derecho de propiedad que existía sobre el soporte de la obra en sí mismo. La mentalidad jurídica clásica presenta cierta aversión a lo abstracto, que impide la distinción entre el *corpus mysticum* (creación intelectual) y el *corpus mechanicum* (objeto que recoge dicha creación). La obra literaria es considerada un bien mueble¹¹ y, por tanto, cuando un escritor cede su manuscrito a un editor, cede también la propiedad del texto en él

⁹ Dock comenta los casos excepcionales de Cicerón y Marcial, que al parecer sí llegaron a recibir compensación económica por ceder sus escritos a los editores, a pesar de que no era lo habitual [1974: 132 y ss.]; así se desprende de la carta que Cicerón envía a Atticus, su editor, para felicitarle por las altas ventas de su discurso *Pro Ligario*, o de los epigramas de Marcial, en donde hay varias alusiones a la venta de sus obras y al comercio que de ellas hacían libreros y copistas.

¹⁰ No faltan, sin embargo, autores que lamentan que los editores se hagan ricos a su costa. Marcial, que siempre andaba escaso de dinero, habla de este asunto en varios de sus epigramas: “El discreto lector podrá comprar un hermoso volumen con todas mis *Xeniae* por cuatro sestercios. La verdad es que cuatro es mucho. Ya estaría bien pagar dos. Y aun entonces mi editor habría hecho un buen negocio” [Citamos por Reyes, 2011: 50]

¹¹ La clasificación de *bienes muebles* e *inmuebles* viene de antiguo. Los *bienes muebles* son aquellos que tienen portabilidad, es decir, aquellos en los que su naturaleza no dificulta su desplazamiento (como un coche, un ordenador portátil o los muebles de una casa). Los *bienes inmuebles*, también denominados *bienes raíces*, en cambio, son aquellos anclados al suelo (por ejemplo, nuestra casa). En ambos casos, hablamos de propiedades físicas (no intelectuales) que se aplican a las cosas materiales.

incorporado. La *scriptura* sigue la misma suerte jurídica que la materia y, en consecuencia, el dueño del objeto físico es dueño también de su contenido intelectual, teniendo potestad para modificarlo a su antojo. Lucio Anneo Séneca escribe a este respecto lo siguiente:

Decimos que los libros pertenecen a Cicerón; el librero Dorus llama suyos a los mismos libros y la verdad es de doble faz. Uno los reivindica como autor y el otro como comprador; y es con razón que se dice que los libros pertenecen al uno y al otro. En efecto, aquellos pertenecen a ambos, pero no de la misma manera [Séneca, *Sobre los beneficios*: VII, 6. Citamos por Espín Cánovas, 1995: 43].

Tampoco existen en el mundo clásico lo que hoy denominamos derechos morales de propiedad intelectual, a pesar de que nos hayan llegado varios testimonios de autores que alzan su voz contra la publicación no consentida de sus obras o la copia ilícita de las mismas. Sirva a modo de ejemplo la anécdota que recoge Aulio Gelio en las *Noches Áticas*, en donde acusa a Platón de plagiar a Pitágoras: “Tú, Platón, puesto que ansiabas el saber, / por una suma desorbitada compraste un librito / que te enseñó a escribir el *Timeo*” [Citamos por Perromat Augustín, 2010: 29]. O la del poeta Teognis, que decide poner un sello a todas sus obras “para que nunca [más] pase inadvertido si alguien las roba [...] y así todo el mundo diga «Estos son versos de Teognis de Mégara»” [Citamos por Banks, 1856: 218].

Pero sin duda, las quejas más sonoras en este sentido son la de Marcial, que en varios de sus epigramas acusa a Fidentino de recitar versos suyos, atribuyéndose la autoría; textos que reproducimos a continuación y que han sentado las bases del concepto moderno de plagio literario [Marcial, 1997: 1, epigramas 29, 38, 53 y 72]:

AL PLAGIARIO FIDENTINO: La fama dice que tú, Fidentino, recitas mis escritos / a la gente como si fueran tuyos. Si quieres / que se digan míos, te enviaré gratis los versos: si quieres / que se digan tuyos, cómpralos para que no sean míos (Epigrama 29).

A FIDENTINO, POETA PLAGIARIO: Una sola página tuya, Fidentino, está en mis escritos, / pero sellada con la firma segura de su dueño, / la que lleva a tus versos a un robo manifiesto. / Así una capa lingónica, colocada en medio, contamina con / su lana grasienta a los vestidos color violeta de la ciudad; / así las vasijas de Arretio desvirtúan las copas de vidrio; / así, cuando acaso el negro cuervo vaga por las orillas / del Caístro, hace el ridículo entre los cisnes de Leda; / así, cuando el bosque sagrado se anima con el variado canto / del ruiseñor, la insolente urraca desafina con las quejas / cecropias. Mis libros no necesitan de un delator o de un juez, pero tu página se levanta contra ti y te dice: “eres un ladrón” (Epigrama 53).

CAPÍTULO 1

AL PLAGIARIO FIDENTINO: Lo que recitas, Fidentino, es mi libro, / pero cuando recitas mal, empieza a ser tuyo (Epigrama 38).

A FIDENTINO, UN POETA PLAGIARIO: ¿Crees, Fidentino, que tú eres poeta / por mis versos y deseas que se te tome como tal? / También Egle se cree con dentadura / con dientes postizos y marfil de la India; / también la que es más negra que la mora madura, / Licoris, se agrada maquillada de blanco. / Y tú, por la misma razón que eres poeta, / tendrás melena, cuando eres calvo (Epigrama 72).

La ley Fabia romana establecía la pena “ad plagas” —es decir, de latigazos— para los ladrones de niños, esclavos u hombres libres, pero en ningún caso para el robo de palabras o ideas. Marcial es el primero en usar metafóricamente este término (*plagiarius*) para señalar a aquél que usurpa un bien intelectual a otro, denotando así la importancia de la identidad en la relación del creador con su obra (relación paterno-filial): si el ladrón de niños es un plagiario, de igual modo lo es que el roba una obra (hijo) a su autor (padre)¹².

Aunque cierto sector crítico se ha basado en las palabras de Marcial para defender la existencia en Roma de un cierto derecho moral de los autores, nosotros coincidimos con Miró Llinares en negar la existencia de cualquier tipo de protección jurídica, tanto en la esfera de lo patrimonial como de lo moral [2007: 113]. Desde nuestro punto de vista, no se puede defender la existencia de un derecho de paternidad cuando su principal negación, es decir, el robo literario, no era castigado. Tampoco se puede defender la existencia de un derecho de integridad de la obra cuando, como hemos visto antes, la *actio furti* protegía el manuscrito como soporte, pero no como idea. Ni la existencia de unos derechos de divulgación basándose en la *actio injuriarum*, cuya finalidad era la protección del honor, pero no la publicación de una obra sin el permiso de su autor. El rechazo social al plagio y a las publicaciones no autorizadas puede entenderse como un “ilícito moral”, pero lo cierto es que el sistema jurídico romano no protege, en ninguno de sus aspectos, el derecho de propiedad intelectual:

¹² El término *plagio*, acuñado por Marcial, cayó en desuso durante la Antigüedad tardía y toda la Edad Media, y no volverá a registrarse hasta finales del siglo XV, de la mano del humanista Lorenzo Valla [Perromat Agustín, 2010: 668].

Este derecho existía *in abstracto*, se manifestaba en las relaciones de los autores con los bibliópolas y los organizadores de los juegos; pero las necesidades sociales de la época no habían impuesto que éste entrara a formar parte de la esfera del derecho [Dock, 1974: 154].

En definitiva, en las antiguas Grecia y Roma hay conciencia del valor de la obra en sí misma, pero no existen los derechos patrimoniales (más allá del soporte físico), ni se reconocen los derechos morales de su autor. Los escritores buscan un reconocimiento público y social y, a falta de un ordenamiento jurídico protector, se ven obligados a defenderse de los *ladrones de ideas* por sí mismos, a través de la denuncia explícita en sus propias creaciones intelectuales. Los plagios no eran perseguidos por la ley, aunque, eso sí, los escritos inculpatorios de los autores originarios los convertían en objeto de crítica y burla.

1.2.1.3. El concepto de autoría en el Medievo. La imitación, los anónimos y las falsificaciones

La ausencia de reconocimiento de los derechos de propiedad intelectual se mantiene durante toda la Edad Media, en donde se produce además un importante retroceso en la conciencia de autoría. A partir del siglo II d. C., los escritores desarrollan una especial reverencia por determinados autores del mundo helénico y del siglo de Augusto, a los que consideran infinitamente mejores que ellos. La originalidad está proscrita y los escritores entienden que para alcanzar las más altas cotas deben imitar —e incluso parafrasear— las obras de los grandes clásicos de la Antigüedad (*auctoritas*)¹³; autores a los que muchas veces no se refieren expresamente, pero cuyos textos son tan conocidos por sus contemporáneos que en caso alguno podemos entender esta práctica como un intento de copia subrepticia. El famoso aforismo de Bernardo de Chartres (“Somos como enanos a hombros de gigantes”) sirve para ilustrar esta idea y

¹³ El término *auctoritas* en la Edad Media hace referencia a la noción de un patrimonio colectivo, una fuente de la que estaba permitido sacar ideas, temas, conceptos, palabras o textos sin necesidad de una justificación explícita. La *auctoritas* aporta valor a las obras, ya que para el hombre medieval la veracidad está por encima de la originalidad.

CAPÍTULO 1

poner de manifiesto la importancia de la *imitatio* y el papel secundario que asumen los autores del medievo.

En los primeros siglos de la Edad Media, la cultura se circunscribe a los monasterios, en donde la creación literaria, por influencia de la doctrina cristiana, ya no es entendida como una actividad de carácter individual, sino como un proceso del que participa toda la comunidad. Los monjes, condicionados por sus votos monásticos de humildad y pobreza, no sólo renuncian a la obtención de un beneficio económico por la producción de sus obras literarias, sino también a la obtención de fama o prestigio, lo que hace que apenas encontremos en ellas marcas individuales de autoría. Muchas de las obras que conservamos de este periodo son anónimas, e incluso en las que figura un patronímico resulta difícil saber si éste hace referencia al copista, al recitador, al compilador o al autor propiamente dicho, como ocurre con el famoso “Per Abbat” del *Cantar de Mio Cid*. La autoría entendida en el sentido moderno se desdibuja y da paso a una idea de autoría compartida entre los diferentes agentes responsables de la transmisión textual.

Los medievalistas a menudo llaman la atención sobre un comentario del monje franciscano San Buenaventura (siglo XIII), quien establece cuatro grados de intervención textual, cada uno de ellos relacionado con una forma diferente de concebir a su creador: el primero es el copista o amanuense (*scriptor*), que no añade ni cambia nada; el segundo es el compilador (*compiler*), que junta varios pasajes de textos que no son suyos; el tercero es el comentarista (*commentator*), que agrega palabras propias a los textos de otros autores; y el último es el autor (*auctor*), que escribe con sus propias palabras, pero añade también las de otros para apoyar sus ideas. Lo interesante es que San Buenaventura no concede ningún privilegio particular a la última categoría, la que puede parecer más cercana a nuestro propio sentido de autoría, pues, dentro del contexto medieval, el autor está configurado como parte de un continuo que se extiende desde el proceso “simple” de copiar, en un extremo, hasta el acto de composición “original”, en el otro. Si bien las dos funciones pueden parecernos fundamentalmente diferentes, para este teólogo franciscano no hay una separación clara entre una y la otra [Bennett, 2005: 38].

La anonimidad voluntaria es una característica esencial de las obras del alto medievo y la búsqueda de la gloria literaria deja de ser una motivación legítima para sus autores. En este contexto, la práctica más denostada ya no es la del plagiario, que hace pasar por suyo un texto ajeno, sino justamente la contraria, la del falsificador, que adscribe un texto suyo a un autor canónico para dotarlo de autoridad y verdad. Son de sobra conocidos los textos apócrifos atribuidos a Aristóteles durante la Edad Media; y contamos incluso con el testimonio del traductor Adelardo de Bath, quien llega a reconocer que ha atribuido ideas propias a autores árabes para aprovecharse de su prestigio.

Con el paso de la Alta a la Baja Edad Media se va produciendo una paulatina restauración del concepto de autoría, primero con la irrupción escrita de las lenguas romances y la creación de la Escuela de Traductores de Toledo y, más tarde, con la aparición de las primeras universidades y la instauración de un nuevo sistema de citación por parte de la Escolástica. Las universidades posibilitan una nueva difusión secular de la cultura escrita —ajena a los monasterios—, así como una importante revalorización de los autores como entes individuales. Escritores como Gonzalo de Berceo; Juan Ruiz, el Arcipreste de Hita; don Juan Manuel; Íñigo López de Mendoza, el Marqués de Santillana; Pero López de Ayala; Juan Fernández de Heredia; Ausias March o Ramón Llull empiezan a proclamar con orgullo la nobleza de su trabajo y a exigir el reconocimiento de sus obras, aunque éstas sean herencia de las lecturas de otros:

Bienaventurado príncipe, podría ser que algunos, los quales por aventura se fallan más prestos a las reprehensiones e a redargüir e emendar que a fazer nin ordenar, dixessen yo haver tomado todo o la mayor parte d'estos Proverbios de las dotrinas e amonestamientos de otros, assí como de Platón, de Aristóteles, de Sócrates, de Virgilio, de Ovidio, de Terencio e de otros philótophos e poetas, lo qual yo no contradiría, antes me plaze que assí se crea e sea entendido. Pero estos que dicho he de otros lo tomaron, e los otros de otros, e los otros de aquellos que por luenga vida e sutil inquisición alcançaron las experiencias y las causas de las cosas [Íñigo López de Mendoza, el Marqués de Santillana, Proemio a sus *Proverbios*. Citamos por Perromat Agustín, 2010: 87-88]

Aunque durante la Baja Edad Media la conciencia de autoría parece restituirse, las relaciones entre autores y editores son mucho más endeble que las que existían en la antigua Roma, algo que no favorece en absoluto el reconocimiento de unos derechos de autor. No conservamos ningún testimonio que acredite el desarrollo de una doctrina

CAPÍTULO 1

jurídica relacionada con lo que hoy conocemos como propiedad intelectual, a excepción del derecho de *Tabula Picta*, destinado a dirimir los conflictos entre el autor y el propietario de una obra pictórica. Lo que vuelve a poner de manifiesto un tipo de mentalidad jurídica incapacidad de discernir entre la creación artística y el soporte físico en el que se presenta.

1.2.1.4. La invención de la imprenta y sus consecuencias: los privilegios de impresión, el nuevo sistema editorial y la profesionalización del oficio de escritor

En algún momento de la década de 1440, el joven herrero alemán Johannes Gutenberg revoluciona el mundo editorial con la presentación de una nueva técnica tipográfica, que permite realizar un alto número de copias en un corto periodo de tiempo y con un coste bastante reducido. La imprenta de tipos móviles se expande con gran celeridad por toda Europa y antes del 1500 ya existen en España más de una treintena de talleres de impresión. Si el paso de la literatura oral a la literatura manuscrita permitió fijar los textos, el paso del manuscrito al impreso favorece su permanencia e inalterabilidad:

In the system of manuscript transmission, it was normal for lyrics to elicit revisions, corrections, supplements, and answers, for they were part of an ongoing social discourse. In this environment texts were inherently malleable, escaping authorial control to enter a social world in which recipients both consciously and unconsciously altered what they received [Marotti, 1995: 35].

La reproducción masiva del material escrito tuvo consecuencias inmediatas. La multiplicación del número de ejemplares de una misma obra posibilitó que se empezase a distinguir entre el continente y el contenido, entendiéndose al fin que el soporte material en que se recoge un texto no es más que un instrumento para exteriorizar una creación intelectual. Pero, además, la aparición de un alto raudal bibliográfico de fácil acceso y mucho más asequible desde el punto de vista económico propició la democratización de la cultura y, con ella, la difusión del libre pensamiento; basta recordar que es por estos años cuando Lutero traduce e imprime la *Biblia* a lengua vernácula y propugna la libre interpretación de las Sagradas Escrituras.

La imprenta no es bien acogida por todos. Los representantes del poder público advierten que este nuevo sistema de copia facilita la rápida propagación de ideas contrarias a sus propios intereses; los editores, por su parte, se ven incapaces de controlar la competencia desleal que ejercen todos aquellos que reeditan sus mismos textos, a pesar de haber sido ellos los que han negociado la compra o cesión del manuscrito con los autores. Circunstancias todas ellas que hacen necesaria la regulación y el control de los libros impresos, y que son el origen de los llamados *privilegios*¹⁴ (autorizaciones otorgadas por el poder real mediante las cuales se concede a alguien la posibilidad de imprimir, publicar y vender una determinada obra de forma exclusiva y temporal).

El régimen de privilegios se da en la mayoría de países europeos, entre ellos, España. Los Reyes Católicos, que habían defendido inicialmente la libre circulación de ideas y la importación de libros de otros Reinos (Pragmática de Toledo de 1480), pronto empiezan a advertir la necesidad de controlar y regular la impresión documental, y lo hacen primero con la imposición de una alcabala para los libros extranjeros (Pragmática del 8 de julio de 1482) y poco después con la redacción de una disposición en la que se prohíbe imprimir, introducir o vender libros sin especial licencia del Rey o de persona autorizada por él (Pragmática de 8 de julio de 1502)¹⁵. Un control que se va a ir

¹⁴ El instituto jurídico del privilegio viene del mundo romano y es una norma aplicable a un caso específico (*lex privata*) y no a una situación general (*lex generalis*). A partir de la Edad Moderna, y por los motivos ya expuestos, el privilegio se convierte en una “gracia real” mediante la que el monarca otorga un poder específico para dar solución a una determinada situación, pero no puede considerarse en sentido estricto el reconocimiento de ningún tipo de derecho subjetivo [Miró Llinares: 2007, 117].

¹⁵ “Otro sí mandamos, y defendemos, que ningún librero, ni impressor de molde, ni mercaderes, ni factor de los susodichos, no sea osado de hazer imprimir de molde de aquí adelante por vía directa ni indirecta ningún libro de ninguna facultad, o lectura, o obra que sea pequeña, o grande, en Latín, ni en Romance, sin que primeramente tenga para ello nuestra licencia y especial mandado, o de las personas siguientes: en Valladolid, y Granada los Presidentes que residen, o residieren en cada una de las nuestras Audiencias que allí residen: y en la Ciudad de Toledo, el Arzobispo de Toledo: y en la Ciudad de Seuilla, el Arçobispo de Seuilla: y en la ciudad de Granada, el Arçobispo de Granada: y en Burgos, el Obispo de Burgos: y en Salamanca, y Zamora, el Obispo de Salamanca; ni sean asimismo osados de vender en los dichos nuestros Reynos ningunos libros de molde, que truxeren fuera dellos de ninguna facultad, ni materia que sea, ni otra obra pequeña, ni grande, en Latín, ni en Romance, sin que primeramente sean vistos, y examinados por las dichas personas, o por aquellas a quienes ellos lo cometieren, y ayan licencia dellos para ello, so pena, que por el mismo hecho ayan los que las imprimieron sin licencia, o los que

CAPÍTULO 1

recrudeciendo en los años sucesivos, bajo los reinados de Felipe II, Felipe III, Felipe IV, Carlos II, Felipe V y Fernando VI, y que sólo empezará a cambiar con la llegada al poder de Carlos III, ya en el siglo XVIII.

Además, con el edicto de los Reyes Católicos de expulsión de los judíos (1492), la labor del Tribunal del Santo Oficio de la Inquisición se recrudece y, a partir de 1551, se empiezan a publicar índices de libros prohibidos por razones de ortodoxia religiosa. Ya sea por vía directa (censura) o indirecta (autocensura), la Inquisición condiciona la labor de los escritores, independientemente de su condición civil o estado. De hecho, hasta Santa Teresa de Jesús llega a confesar en su *Libro de la vida* que “iban a mí con mucho miedo a decirme que andaban los tiempos recios, ya que podía ser que me levantasen algo y fuesen a los Inquisidores” [Citamos por Pascual, 1998: 405].

El régimen de privilegios promueve el crecimiento de la imprenta y hace rentable el oficio de impresor. Tal y como explica Miró Llinares, se limita la competencia mediante “la concesión de un monopolio, durante un tiempo limitado, y sobre un producto relacionado con la actividad que se quiere fomentar, de forma que se hace rentable esa actividad, se incita a la competencia y, pasado un tiempo, se liberaliza el comercio sobre ese bien concreto” [2007: 118]. Un buen ejemplo de este nuevo sistema editorial lo encontramos en un valiosísimo documento notarial de 1605, por el que el mercader de libros Diego Pérez, que asegura haber comprado los derechos de impresión y venta de *La pícaro Justina* a fray Baltasar Navarrete¹⁶, otorga “dejación, cesión, renunciación y traspaso” de tales derechos al también mercader de libros Jerónimo Obregón:

vendieron los que truxeren de fuera del Reyno sin licencia, perdido, y pierdan todos los dichos libros, y sean quemados todos públicamente [...], y más pierda el precio que huieren recibido, y se les diere, y paguen en pena otras tantos maravedís como valieren los dichos libros, que assí fueren quemados” [Pragmática de 8 de julio de 1502. Citamos por Simón Díaz, 1983: 6].

¹⁶ Rojo Vega presupone que fray Baltasar de Navarrete, que aparece en el documento como propietario de los derechos de edición, es el “autor probable” de *La pícaro Justina*, y no el dominico fray Andrés Pérez o el médico toledano Francisco López de Úbeda, como se había defendido hasta la fecha [2004].

EL AUTOR COMO ENIGMA

Sean quantos esta carta de poder en causa propia y zesión y lo que de yuso será conthenido bieren, como yo Diego Pérez, mercader de libros vezino de la villa de Medina del Campo, residente en esta ciudad de Valladolid, Corte de su majestad, digo que por quanto yo tengo derecho y action para imprimir y bender un libro intitulado la pícara, que le compré del padre presentado fray Baltasar Navarrete de la orden de señor santo domingo, según consta de la escritura de compra, asiento y concierto que con él hize, que está otorgada ante Cristóbal de Santiago, escribano real vecino desta dicha ciudad, ques notoria y a que me refiero, por tanto agora yo el dicho Diego Pérez de mi boluntad y por causas y justos respectos que a ello me mueben, otorgo y conozco por esta carta que hago y otorgo dejazón, çesión, renunziacion y traspaso en bos Gerónimo Obregón, mercader de libros vezino desta dicha ciudad de todo el derecho que tengo adquirido en virtud de la sobredicha escritura y conforme a ella para poder inprimir y bender el dicho libro intitulado la pícara y de oy en adelante podais bos el dicho Gerónimo Obregón o quien vuestro poder tubiere inprimir el dicho libro y benderle a la persona o personas que quisiéredes y vuestro gusto y boluntad fuere, en esta Corte o fuera de ella y en las partes que buestra voluntad fuere y cobreis las quantias de maravedís que prozederen de la tal benta o bentas o hiziéredes de qualquier cantidad o cantidades de cuerpos del dicho libro, husando para todo ello de la misma çesión benta y derecho que en mi favor tiene fecha y otorgada el dicho padre maestro fray Baltasar Navarrete... y yo el dicho Gerónimo Obregón, que estoy presente, aceto esta dicha escritura y conforme a ella me obligo que durante el tiempo de la dicha escritura de benta de la inpresión del dicho libro cunpliré y pagaré al dicho fray Baltasar Navarrete todo aquello a que vos el dicho Diego Pérez estais en ella obligado y de todo os sacaré a paz y a salvo e indemne [Citamos por Rojo Vega, 2004: 214-215].

Aunque algunos ven en el sistema de privilegios el origen del derecho de autor moderno, lo cierto es que las autorizaciones se otorgan al editor o al impresor, rara vez a los autores, que siguen estando tan desamparados como antes; tanto es así que, una vez vendida la obra, éstos pierden todo derecho patrimonial, moral y estético sobre ella. A pesar de esto, el nuevo sistema sí beneficia de algún modo a los autores, que empiezan a tener la posibilidad de ganarse la vida “escribiendo”, al margen de las reglas del clientelismo y el mecenazgo. Las relaciones contractuales entre escritores y editores se fortalecen¹⁷, ya que éstos, al tener garantizada la fácil recuperación de los gastos de impresión, empiezan a estar más predispuestos a pagar cierta cantidad de dinero por la

¹⁷ La relación entre un autor y su editor se mantiene normalmente a lo largo del tiempo, aunque también puede romperse y pasar a establecerse con otro editor: “Ejemplos de esta continuidad cambiante nos la dan Cervantes y Lope de Vega. Miguel de Cervantes está unido editorialmente con los Robles, familia de librerros. Blas de Robles le edita en 1585 *La Galatea*; su hijo Francisco las dos partes del *Quijote*, 1605 y 1615, y las *Novelas ejemplares*, 1613. En 1615 cambia Cervantes de editor, eligiendo al joven Juan de Villarroel, que publica en este año las *Ocho comedias y ocho entremeses nuevos*, y en 1617 su obra póstuma, *Los trabajos de Persiles*. Hasta 1605, las obras no dramáticas de Lope son generalmente editadas por Juan de Montoya, que es substituido por el gran librero-editor Alonso Pérez, el padre de Juan Pérez de Montalbán” [Moll, 2003].

CAPÍTULO 1

adquisición de obras literarias. Surge así el derecho pecuniario de los autores, que “es tutelado mediante la protección arbitrada a los intereses de sus editores” [Bondía Román, 1988: 74].

Como ha estudiado muy acertadamente Carlos M. Gutiérrez, siguiendo la teoría del *campo literario* del sociólogo francés Pierre Bourdieu¹⁸, la profesionalización de la escritura es uno de los elementos que permiten la configuración en el Madrid del 1600 del primer campo literario de la historia de España; esto es, un espacio social definido, con sus jerarquías y sus estructuras internas, en el que sus agentes (los autores) luchan por distinguirse —literaria o socialmente— en el ámbito de la producción restringida (literatura de élite) o en el de la producción de masas (literatura popular):

Esto ocurrió en un momento histórico en el que la literatura se había convertido en una actividad específica y reconocida por amplias capas sociales; en una actividad, en suma, que devengaba beneficios directos (venta de libros, encargos, comisiones de obras, premios materiales) o indirectos (reconocimiento simbólico por parte del campo de poder). Uno de los argumentos principales para llegar a esta conclusión fue constatar que Lope y Góngora comenzaron a ser percibidos por sus contemporáneos como estandartes de los dos espacios simbólicos en los que, según Bourdieu, se divide el campo literario: el de la producción masiva (Lope) y el de la producción restringida (Góngora) [Gutiérrez, 2005a: 265]

Con el auge del mercado editorial y el teatro comercial, el oficio de escritor empieza a sufrir una progresiva profesionalización. Si Cervantes es un “hombre que escribe” pero que también “trata negocios”, Lope de Vega se nos presenta ya como un escritor que vive exclusivamente de su pluma, sin necesidad de tener una profesión secundaria como fuente constante de ingresos¹⁹. En su epistolario privado, Lope

¹⁸ El *campo literario* para Bourdieu es un “campo de fuerzas que actúan sobre todos los que entran en ese espacio y de maneras diferentes según la posición que ellos ocupan en él (sea, para tomar puntos muy distantes entre sí, la del autor de piezas de éxito o la del poeta de vanguardia), a la vez que de un campo de luchas que procuran transformar ese campo de fuerzas” [1989-1990: 2]; en relación con su estructura, considera que “la oposición principal se establece entre los productores más autónomos (campo de producción restringida), poseedores de un gran capital específico, pero de escaso éxito en el público (al menos en la fase inicial de su empresa), y los productores más heterónomos y de gran éxito económico (campo de gran producción)” [1989-1990: 21]. No deja de resultar paradójico que el valor de una obra literaria sea mayor conforme aumenta la distancia que dicha obra mantiene con el campo de producción masiva; algo que Bourdieu calificó como “mundo económico invertido”.

¹⁹ La escritura teatral era la base fundamental de los ingresos económicos de Lope de Vega. García-Reidy calcula que obtuvo unos 500 escudos anuales por la venta de comedias que “supusieron al

reconoce que depende económicamente de la venta de comedias a las compañías del momento (poeta popular), aunque no por ello renuncia a escribir para la élite del poder (poeta cortesano), lo que permite advertir en él, en palabras de García-Reidy, una “conciencia autorial en tensión, entre el orgullo de la obra creada para el mercado literario y el ansia por pertenecer a una idealizada concepción aristotélica de la literatura, plasmada a través de la búsqueda del mecenazgo nobiliario y regio” [2013b: 15 y 16].

El sistema de privilegios, por tanto, favorece la industria del libro y beneficia a los agentes implicados en ella (autores, editores, impresores, traductores...), pero no logra poner freno a la existencia de circuitos comerciales paralelos, en donde se imprimen ediciones sin las licencias legalmente exigidas o con licencias falsificadas, ediciones contrahechas o ediciones subrepticias de difusión clandestina. Así lo denuncian muchos autores del momento, como fray Antonio de Guevara, que se queja de la existencia de copias impresas no autorizadas de su *Relox de Príncipes de Marco Aurelio*, que ni siquiera llevan su nombre, o San Juan de Ávila, que asegura que su *Audia filia* se vende sin sus correcciones, ni su permiso, ni el de su editor. Y es de sobra conocido el caso de la falsificación de la primera edición de la *Tercera parte de las comedias de Lope de Vega y otros autores*, que no fue impresa en Barcelona en 1609 sino en Sevilla por Gabriel Ramos Bejarano [Moll, 1979]²⁰.

menos dos terceras partes de sus ingresos, si no más” [2013b: 181]; y concluye: “En ese sentido, Lope fue más allá que muchos escritores del campo literario moderno, quienes necesitaron tener una profesión secundaria a la literaria con la que obtener los ingresos necesarios para su mantenimiento. El Fénix, por el contrario, fue un escritor que, si bien participó en una relación clientelar con la nobleza, tuvo siempre en su pluma literaria su principal fuente de ingresos” [2013b: 199].

²⁰ Los escritores de esta época siguen aceptando la imitación como un acto legítimo, siempre que la copia sea un verso o sentencia breve. Así lo propugna Luis Alfonso de Carvallo en el *Cisne de Apolo* (1600), y también Cervantes en la *Adjunta* de su *Viaje al Parnaso* (1614), donde advierte “que no ha de ser tenido por ladrón el poeta que hurtare algún verso ajeno y lo encajare entre los suyos, como no sea todo el concepto y toda la copla entera, que en tal caso tan ladrón es como Caco”. En definitiva, la diferencia entre la imitación como práctica legítima (intertextualidad) o denostada (plagio) era simplemente cuestión de tamaño.

CAPÍTULO 1

1.2.1.5. *El nacimiento de la propiedad intelectual como institución jurídica y el reconocimiento de las facultades patrimoniales del derecho de autor*

Con el advenimiento del pensamiento ilustrado, son muchas las voces que se alzan en defensa de la libertad de imprenta, como una manifestación más del libre ejercicio de la actividad económica. Las ideas de Milton, Locke y Blackstone en Inglaterra, Pütter en Alemania o D’Hericourt y Diderot en Francia son el caldo de cultivo para el desarrollo de la idea de propiedad intelectual, que, como veremos, no se materializa de igual modo ni al mismo tiempo en todos los países.

Los primeros vestigios de cambio los encontramos en Inglaterra. Ya en el año 1644, el escritor y político John Milton, por encargo de la Sociedad de Libreros de Londres, había pronunciado un discurso ante el Parlamento inglés en el que se oponía a toda forma de censura o licencia previa. Pero no será hasta medio siglo más tarde, en 1710, cuando se apruebe el *Estatuto de la Reina Ana*, por el se reconoce a los autores —o a aquellos que hubieran tenido la cesión por parte de los mismos— el derecho exclusivo de reproducción de sus obras, por un tiempo de catorce años para los libros inéditos (prorrogable por otros catorce si el autor seguía con vida) y de veintiuno para los ya publicados. Inglaterra se convierte así en el primer país en reconocer unos derechos de autor, aunque estos sólo afecten a la esfera de lo patrimonial.

En la Europa continental, Francia es el primer país en regular este tipo de derechos. Allí, al igual que ocurrió en Inglaterra, hubo una serie de autores que fueron encendiendo el debate en torno a este asunto; especialmente significativa nos parece la *Carta sobre el comercio de la librería* de Diderot (1763), una memoria encomendada por los libreros parisinos al enciclopedista francés para legitimar la permanencia del sistema de privilegios y con la que éste logra defender no sólo el oficio de los editores, sino también el de los autores, sin comprometer la idea de libertad que propugna el liberalismo²¹. Pero el verdadero reconocimiento de unos derechos de autor lo

²¹ Diderot defiende a los libreros frente a la censura, la piratería, las restricciones comerciales y la intromisión económica del Estado; pero, al tiempo, defiende también a los autores como creadores intelectuales: “Lo digo y lo repito: el autor es dueño de su obra o nadie en la sociedad es dueño de sus

encontramos en la promulgación de los decretos franceses de 1777, por los que se establecen dos tipos distintos de privilegio: uno concedido al editor por un periodo determinado de tiempo y otro concedido al autor de manera exclusiva y perpetua²². De este modo, mientras en Inglaterra la propiedad intelectual es entendida como un instrumento de interés colectivo, en Francia lo es como un derecho inalienable de la persona, es decir, como un interés individual; y en esta discrepancia está el origen de la diferenciación entre el modelo anglosajón del *copyright* y el continental de los derechos de autor, que aún hoy persiste [Miró Llinares, 2007: 123-135].

En España, el cambio empieza a gestarse durante el reinado de Carlos III, quien establece que “de aquí en adelante no se conceda a nadie privilegio exclusivo para imprimir ningún libro, sino al mismo autor que lo haya compuesto” (Real Orden de 22 de marzo de 1763), y que dichos privilegios puedan transmitirse *mortis causa* a los herederos (Real Orden de 20 de octubre de 1764). Pero no será hasta el siglo XIX cuando, por influencia de la normativa francesa, se reconozca la propiedad de los autores sobre sus obras, primero con el Decreto CCLXV, de 10 de junio de 1813²³, y

bienes. El librero entra en posesión de aquella, como antes fue posesión del autor: tiene el derecho incontestable de sacar el partido que mejor le convenga de sus reiteradas ediciones, y tan insensato sería impedirselo como condenar a un agricultor a dejar sus tierras yermas, o al propietario de una casa a conservar las estancias vacías” [1997: 59]. Con esta argumentación, Diderot diferencia entre el derecho natural de los autores sobre sus obras y los derechos adquiridos por los libreros tras su compra, respaldando el privilegio no como gracia real sino como “salvaguardia” de un contrato privado.

²² Estos decretos recogen al fin la idea de que el autor es el propietario de su obra, pero falta el reconocimiento total del derecho. Éste se produce con la abolición del sistema de privilegios en 1789 y la promulgación de los Decretos de 1791 y 1793, por los que se concede al autor derechos exclusivos de representación y de reproducción durante toda su vida, así como derechos de representación por un periodo de cinco años y de reproducción por uno de diez a sus herederos.

²³ Este decreto, no obstante, tampoco acaba de convencer a todos. En un artículo publicado el 11 de noviembre de 1820 en el periódico *El censor*, titulado la “Propiedad literaria”, leemos lo siguiente: “No deja de ser admirable que en un tiempo en que todo el mundo clama, porque se establezcan las bases necesarias para asegurar á cada uno el goce libre y tranquilo de su propiedad, haya una ley en España, por la cual no solo no se protege la propiedad de los individuos, sino que se les priva legalmente de ella al cabo de un cierto y determinado tiempo [...]. El decreto de las Cortes de 10 de junio de 1813, relativo á la propiedad de las obras literarias, asegura á sus autores el dominio exclusivo de ellas durante toda su vida, y diez años después de su muerte en beneficio de sus herederos, de modo que el hijo de un autor célebre pierde la herencia de su padre á los diez años y un dia despues de la muerte de aquel. Es de advertir que esta disposición de las Cortes, como todas las que contiene el preciado decreto de 10 de junio, fueron dadas con el único obgeto de proteger la propiedad de los autores, abandonada hasta entonces al pillage,

CAPÍTULO 1

después con la llamada Ley Calatrava de 1823, en el marco de la Constitución de 1812 y las Cortes de Cádiz.

Es importante advertir que la vigencia de estas primeras regulaciones fue puramente testimonial: el decreto de 1813 fue derogado tras la vuelta de Fernando VII a España en 1814 y la Ley Calatrava, promovida por el Gobierno constitucional de la Revolución de 1820, fue abolida por este mismo monarca tres años después, al recobrar definitivamente el poder y reinstaurar el absolutismo. El propio Larra —en un artículo que publica en *El pobrecito hablador*, bajo el pseudónimo del Bachiller don Juan Pérez de Munguía— lamenta que en 1832 España siga a la zaga respecto de otros ordenamientos europeos en cuanto a derechos de autor se refiere: “¿Qué cosa es el derecho de propiedad? Si nosotros no lo decimos, ¿quién lo dirá? Y si ninguno lo dice, ¿quién lo sabrá? Y si ninguno lo sabe, ¿quién lo remediará?” [1832: 15].

Las palabras de Larra parecen tener rápida respuesta y, tras restablecerse el régimen constitucional con la muerte de Fernando VII, se suceden una serie de órdenes, decretos y proyectos que confluyen primero en la aprobación de la Ley de Propiedad Literaria de 1847²⁴ y, poco después, de la Ley de Propiedad Intelectual de 1879, vigente hasta 1987²⁵. Estas nuevas leyes ponen fin al antiguo sistema de privilegios, y

digámoslo así, del primer librero ó impresor que se aventuraba á imprimirlas. Sin embargo no entendemos qué especie de proteccion es aquella que lleva envuelto en si misma un despojo legal al cabo de diez años de uso” [Anónimo, 1820: 223-225].

²⁴ La Ley de Propiedad Literaria de 1847 no reconoce la obra como creación inmaterial sino como objeto físico, algo que cambia con la Ley de Propiedad Intelectual de 1879, donde se extiende la propiedad a todas las formas posibles, y no sólo a su materialidad. Como advierte Marco Molina, la Ley de Propiedad Literaria de 1847 “no consigue todavía abstraer el derecho sobre la obra como creación inmaterial del derecho de propiedad ordinaria que recae sobre el objeto físico. Es por ello que, al igual que los antiguos privilegios, en realidad sólo protege plenamente a la obra escrita (artículo 1º de la ley) en forma de libro” [1995: 22].

²⁵ Es importante señalar que, en 1876, el jurista Manuel Danvila Collado había presentado en el Congreso de los Diputados una proposición de Ley en la que se recogía que la propiedad intelectual, como el resto de propiedades, es perpetua y que “no admite más limitaciones que las impuestas por la ley o por a voluntad de los que la tienen”. Esta tendencia a favor de la perpetuidad no se materializa en la Ley de 1879, posiblemente por influencia de Víctor Hugo, impulsor de la Asociación Literaria y Artística Internacional (ALAI), quien un año antes se había mostrado partidario de la temporalidad: “El derecho de propiedad intelectual pertenece a dos unidades: el autor y la sociedad. Antes de la publicación, el autor tiene un derecho incontestable, ilimitado; pero, desde que la obra se da a la luz, ya el autor no es dueño de

reconocen por primera vez la existencia de un derecho de autor como facultad patrimonial (pero no como facultad moral o personal). Si la ley de 1847 otorgaba a los autores el derecho exclusivo para reproducir o autorizar la reproducción de sus obras durante toda su vida y a sus herederos durante cincuenta años, la ley de 1879 amplía el derecho de los herederos hasta los ochenta años (que se rebajará a los setenta con la Ley de Propiedad Intelectual de 1987).

El reconocimiento jurídico de la propiedad intelectual supone un gran avance para los escritores, pero no llega a tiempo en el caso de José Zorrilla, que en el mismo año de su aprobación lamenta que ésta no tenga carácter retroactivo:

No puedo hacer al tiempo volver atrás: no puedo quitarme de encima ni uno solo de mis sesenta y cuatro años: no puedo hacer volver á mis manos el capital pagado por las deudas de mi herencia paterna, ni lo por mí gastado en vivir bien o mal: no puedo rescindir los contratos de venta de mi *Don Juan* ni de mi *Zapatero y el Rey*, escritos cuando la ley de propiedad no existía: esta ley no tiene efecto retroactivo ni protege mi propiedad por lesión enorme: y no puedo pedir limosna en España, sinó poniéndome al pecho un cartel que diga: “este es el autor de *Don Juan Tenorio*, que mantiene en la primera quincena de Noviembre todos los teatros de verso de España y América”; pero para esto sería preciso que yo explicase cómo el autor de tal obra podía pedir limosna; cosa muy fácil de explicar, pero muy difícil de comprender [1880: 3].

Pero ni incluso la llegada de la ley de 1879 parece contentar a los escritores, y muchos de ellos siguen criticando las condiciones abusivas de los contratos de edición: Valle-Inclán se queja de que sólo le ofrezcan el 5% por algunas de sus obras; Pío Baroja critica que la editorial de V. Prieto y Compañía —más conocida como Biblioteca Renacimiento— le pague derechos muy bajos (500 pesetas) por tiradas muy superiores a las manifestadas (10.000 ejemplares); y Blasco Ibáñez lamenta que *La Barraca* le dé más dinero en Francia “a los dos meses de traducirse, que en los ocho años de vida que lleva en España” [Sánchez García, 2002: 1018].

ella... El heredero no tiene derecho a hacer ni una raspadura ni a suprimir un solo renglón; sólo tiene un derecho: el de vivir de la parte de herencia que le ha legado su ascendiente. El escritor trabaja en primer lugar para los hombres, después para sus hijos. Hay que conciliar los derechos de los tres personajes, el autor, el dominio público y el heredero... Durante su vida nade tiene derecho de publicar sus obras; después de su muerte no puede impedir a la posteridad que las recobre” [Citamos por Rogel Vide, 2006: 79].

CAPÍTULO 1

A pesar de estas discrepancias de índole económica, tanto el auge de los nacionalismos como la idea de la existencia de una propiedad sobre las creaciones intelectuales propicia que muchos filólogos y críticos pongan el acento tanto en editar las obras de los grandes clásicos (como hace, por ejemplo, Rivadeneyra en la famosa *Biblioteca de autores españoles*) como en escribir la historia de la literatura y de la crítica literaria, siguiendo, por lo general, el método biográfico de Charles de Sainte-Beuve, que parte de la premisa de que en toda obra hay una estrecha relación autor-texto. También surgen por estos años muchos estudios de atribución de autoría, primero como un modo de devolver a los autores del pasado esos derechos usurpados, pero, más tarde, como una manera de obtener fama y notoriedad por parte de ciertos investigadores; recordemos, por ejemplo, el caso de Adolfo de Castro, que en 1948 hizo pasar por cervantina una obra —*El Buscapié*— que había escrito él mismo.

En el último tercio del siglo XIX, tras el reconocimiento masivo de los derechos de los autores dentro de cada territorio, se advierte la necesidad de que éstos tengan dimensión internacional, pues los múltiples avances tecnológicos estaban facilitando la realización de actos lesivos. Es en este contexto en el que surgen una serie de acuerdos bilaterales, que darán lugar, poco después, al primer gran movimiento de cooperación internacional: el Convenio de Berna para la Protección de Obras Literarias (1886), gracias al cual los titulares de los derechos de propiedad intelectual se ven respaldados, al fin, tanto dentro como fuera de su país²⁶.

1.2.1.6. La gran paradoja del siglo XX

A lo largo del siglo XX, el texto del tratado de Berna es revisado en varias ocasiones para adaptarlo a las nuevas formas de creación, utilización y difusión de las

²⁶ El texto de este primer tratado es suscrito por diez países; entre ellos, España (completan la nómina Alemania, Bélgica, Gran Bretaña, Francia, Haití, Liberia, la Confederación Helvética, Italia y Túnez). A día de hoy están adheridos un total de 180 países [Véase listado completo en https://wipo.int/es/treaties/ShowResults?start_year=ANY&end_year=ANY&search_what=C&code=ALL&treaty_id=15. Fecha de consulta: 10 de diciembre de 2021]

obras literarias y artísticas. De todas las revisiones, la más interesante para el objeto que nos ocupa es la que se produce en 1928, pues en ella se reconoce, por primera vez, la existencia de unos “derechos morales” de autor:

Independientemente de los derechos patrimoniales del autor, y lo mismo después de la cesión de dichos derechos, el autor conserva el derecho de reivindicar la paternidad de la obra, así como el derecho de oponerse a toda deformación, mutilación u otra modificación de dicha obra, que fuere perjudicial a su honor o su reputación [Citamos por Marco Molina, 1995: 135].

Los derechos de autor también aparecen recogidos en el artículo 27.2 de la primera Declaración Universal de los Derechos Humanos, adoptada por la Asamblea General de la Naciones Unidas en 1948, en donde se especifica que “toda persona tiene derecho a la protección de los intereses morales y materiales que le correspondan por razón de las producciones científicas, literarias o artísticas de que sea autora” [Citamos por Serrano Gómez, ed., 2014: 156].

Pero, sin duda, el hito más importante en la conformación del derecho de autor, tal y como lo conocemos actualmente, lo constituye la primera declaración de la Convención Universal de los Derechos de Autor (1952), encaminada a armonizar el modelo anglosajón con el continental, que responden —como hemos explicado antes— a dos formas diferentes de entender la protección de los autores (interés colectivo *vs.* derecho individual)²⁷. Como advierte Miró Llinares, esta declaración fue “el primer gran paso dado en el ámbito internacional en el camino hacia una legislación unificada en materia de protección de derechos de propiedad intelectual” [2007: 31], que se completará poco después con la creación, en 1967, de la Organización Mundial de la Propiedad Intelectual (OMPI).

²⁷ España, a pesar de ser uno de los primeros países en comprender la urgencia de ampliar la protección de los autores al ámbito internacional, no es tan diligente en el reconocimiento de los derechos morales de autor, que no aparecen legislados en el derecho nacional hasta casi entrado el siglo XX. Será la Ley de Propiedad Intelectual de 1987, caparazón y corazón de la ley actual (1996), la que reconozca por primera vez de manera expresa que el derecho de autor comprende tanto facultades patrimoniales como morales, y que éstas últimas son irrenunciables e inalienables.

CAPÍTULO 1

Resulta paradójico advertir que es precisamente cuando la institución jurídica se fortalece e internacionaliza, cuando la figura del autor como entidad real empieza a sufrir una profunda crisis de identidad, que tendrá su punto álgido con el Postestructuralismo. A finales de los años sesenta, Jacques Derrida, Roland Barthes y Michel Foucault proclaman la destrucción del autor en aras de una postura menos individualista de la creación, dando así por finalizado el matrimonio autor-texto que había caracterizado la tradición crítica desde Sainte-Beuve. Derrida critica el “falocentrismo” como una obsesión no justificada del ser humano por tratar de fijar un origen para todo, un patriarca que ejerza de creador universal²⁸. Barthes, mucho más tajante, defiende que la escritura es la destrucción de toda voz y que el nacimiento del lector se paga con la muerte del autor [1968]. Y Foucault, por su parte, sostiene que el autor debe ser despojado de su rol de artífice para pasar a ser analizado como una entidad discursiva —presente sólo en algunos textos y en algunos momentos de la historia— que se mueve entre los límites de lo real y lo ficticio (función-autor); se distancia así de la idea del autor como sujeto de carne y hueso para pasar a entenderlo como una entidad teórica y conceptual, gracias a la cual se ordenan y agrupan cierto número de textos bajo unos criterios de unidad²⁹:

El autor no funciona como un nombre propio: la relación entre el autor y lo que nombra no es isomorfa con la relación entre el nombre propio y el individuo que designa. Foucault nos ofrece varios ejemplos. Descubrir que Pierre Dupont no es médico o no vive en París no modifica el nexo de designación. Del mismo modo, si se descubre que Shakespeare no

²⁸ En una entrevista a Lucette Finas, traducida al español por Cristina de Peretti della Rocca, Jacques Derrida explica el *falocentrismo* del siguiente modo: “Con este término —*falocentrismo*— trato de absorber, de hacer desaparecer el guion mismo que une y vuelve pertinentes el uno para con la otra aquello que he denominado, por una parte, logocentrismo y, por otra, allí donde opera, la estratagema *falocéntrica*. Se trata de un único y mismo sistema: erección del logos paterno (el discurso, el nombre propio dinástico, rey, ley, voz, yo, velo del yo-la-verdad-hablo, etc.) y del falo como «significante privilegiado» (Lacan)” [de Peretti della Rocca, 1989: 35].

²⁹ Foucault distingue cuatro características fundamentales de la función-autor: (i) está ligada a un sistema jurídico-institucional que rodea, determina y articula el universo de los discursos que son objeto de apropiación; (ii) no actúa en todos los discursos —ni lo hace con las mismas características en todos ellos—, ni en todas las épocas o civilizaciones; (iii) está definida por una serie de operaciones específicas y complejas que confieren unidad a los discursos; (iv) no remite pura y simplemente a un individuo real, pues, por un lado, puede ser ocupada por varios individuos, y, por otro, un mismo individuo puede ocupar diferentes posiciones-sujeto [1969].

nació en la casa que se visita como su lugar natal, ello no altera el funcionamiento del nombre del autor. Por el contrario, si se revela que no es el autor de los *Sonetos* o que es el autor del *Novum organum* de Bacon, que Shakespeare y Bacon son la misma persona, entonces se modifica por completo el funcionamiento del nombre del autor [Castro, 2018: 135]

El nacimiento de la función-autor está vinculado a la idea de “apropiación penal”, pues es en el momento en el que los textos empiezan a ser transgresivos cuando las autoridades advierten la necesidad de identificar y castigar a sus autores:

Les textes, les livres, les discours ont commencé à avoir réellement des auteurs (autres que des personnages mythiques, autres que de grandes figures sacralisées et sacralisantes) dans la mesure où l'auteur pouvait être puni, c'est-à-dire dans la mesure où les discours pouvaient être transgressifs. Le discours, dans notre culture (et dans bien d'autres sans doute), n'était pas, à l'origine, un produit, une chose, un bien ; c'était essentiellement un acte —un acte qui était placé dans le champ bipolaire du sacré et du profane, du licite et de l'illicite, du religieux et du blasphématoire. Il a été historiquement un geste chargé de risques avant d'être un bien pris dans un circuit de propriétés [1969: 84].

Desde este punto de vista, Foucault concede a la Literatura un *status* especial. Para el filósofo francés, la función-autor actúa de pleno en las obras literarias de la Edad Moderna, de las que, entre otras cosas, necesitamos saber quién las ha escrito. El anonimato literario (ya sea accidental o voluntario) no es socialmente soportable, y por eso la búsqueda de autores se convierte en un *jeu*, en un *énigme* por resolver.

Mais les discours «littéraires» ne peuvent plus être reçus que dotés de la fonction-auteur: à tout texte de poésie ou de fiction on demandera d'où il vient, qui l'a écrit, à quelle date, en quelle circonstance ou à partir de quel projet. Le sens qu'on lui accorde, le statut ou la valeur qu'on lui reconnaît dépendent de la manière dont on répond à ces questions. Et si, par suite d'un accident ou d'une volonté explicite de l'auteur, il nous parvient dans l'anonymat, le jeu est aussitôt de retrouver l'auteur. L'anonymat littéraire ne nous est pas supportable ; nous ne l'acceptons qu'à titre d'énigme. La fonction auteur joue à plein de nos jours pour les oeuvres littéraires [1969: 85].

1.2.1.7. *El fin del anonimato literario: la “resurrección” del autor a través del idiolecto*

A pesar de que el Postestructuralismo, con Barthes a la cabeza, anuncie a bombo y platillo la “muerte del autor”, lo cierto es que dicha muerte no acaba con el autor en sí mismo, sino con la concepción romántica que de él se tenía. Coincidimos con Zapata en

CAPÍTULO 1

la idea de que la declaración de Barthes no es más que una manera de culminar una larga historia de negaciones sobre la idea del autor como autoridad cuasi divina³⁰:

La idea de un sujeto propietario y el responsable de la obra parecía incomodar a toda una generación para la cual la autoridad no era más que un simple tropiezo en la historia. “La escritura es la destrucción de toda voz, de todo origen”, protestaba Barthes, y al hacerlo, proclamaba la liberación de la palabra, la emancipación de un lenguaje anónimo, democrático, redimido de la autoridad del “Autor-Dios”. Dadas así las cosas, reducir la interpretación de una obra a la vida del autor resultaba no sólo disparatado e irrealizable, sino que era políticamente incorrecto, pues era una manera más de plegarse a una autoridad a la que toda una generación de estudiantes, escritores y profesores parecían oponerse tajantemente [2011: 37].

La obra de Foucault, además, ha servido para fortalecer la idea del autor como entidad teórica. No son pocos los que desde los años 80 hasta nuestros días han profundizado en el análisis de la figura autorial desde múltiples perspectivas; no podemos detenernos ahora en todos ellos, pero basta recordar los trabajos de Pierre Bourdieu, Roger Chartier, Jacques Dubois, Dominique Maingueneau o Alain Vaillant, entre otros, para darse cuenta de su alcance.

Al margen de estos estudios teóricos, el autor ha resurgido, con más fuerza que nunca, bajo la idea del idiolecto. El idiolecto (del latín *idiolectus*, ‘habla individual’) hace referencia —como veremos con detalle en el segundo capítulo— a la variedad lingüística de un determinado individuo, que lo distingue y diferencia de los demás; y es precisamente este carácter diferenciador lo que imposibilita desligar a una persona de los textos que produce [Baldwin, 1979; Coulthard, 1998 y 2004; McMenamín, 2002; Coulthard y Johnson, 2010]³¹. Desde los años 80 del siglo XX, pero mucho más

³⁰ Nietzsche había declarado la muerte de Dios, Hegel y Marx la muerte del arte, Mallarmé la necesidad de suprimir al poeta en beneficio de la escritura misma o Valéry la impertinencia de la intención del autor para describir el valor y significación de sus textos.

³¹ Recordemos, por ejemplo, el caso de Stephen King, que siendo ya un autor conocido decidió publicar varios libros con el pseudónimo de Richard Bachman. El engaño fue descubierto por el librero norteamericano Steve Brown, que advirtió similitudes lingüísticas entre los escritos de Bachman y los de King. En 1966, la editorial Dutton, en una acción de marketing sin precedentes, publicó *Poseción*, anunciando que su autor, Richard Bachman, había fallecido de “cáncer de pseudónimo, una extraña variedad de esquizofrenia”. El propio King, años más tarde, reconoció que, aunque es un alivio poder escribir desde el anonimato, cada vez es más difícil mantener el secreto a lo largo del tiempo, pues él fue

intensamente en el siglo XXI, los estudios de autoría de la Lingüística forense han empezado a dar respuesta a esa idea foucaultiana de la identificación como apropiación penal; el lenguaje, por lo que tiene de idiosincrásico y personal, se constituye como una evidencia que permite perseguir no sólo delitos relacionados con la propiedad intelectual —como el plagio³²— sino todas aquellas infracciones que se produzcan bajo la falsa idea de la anonimidad. Un aspecto al que, por la importancia que adquiere en el contexto de esta investigación, dedicaremos un capítulo aparte.

1.2.1.8. Reflexiones finales

Este breve repaso por la historia nos permite advertir que la institución de la autoría literaria, tal y como la conocemos hoy día, ha estado determinada por una serie de factores sociales, culturales, económicos e ideológicos de muy diversa índole.

El interés individual de los escritores por ser reconocidos (origen de la facultad moral del derecho de autor) y por obtener un beneficio económico de sus obras (origen de la facultad patrimonial) es muy anterior en el tiempo al reconocimiento jurídico del hecho de la creación y de su reproducción. Es más, a pesar de que los escritores tienen conciencia de la existencia de cierto derecho de índole personal sobre sus obras incluso

desenmascarado “by someone who recognized my style, and in these days of the Internet, that becomes more and more likely”.

³² A pesar de ser el plagio una violación del derecho moral de autor (en concreto, de la paternidad de la obra), el término como tal no aparece definido en ninguna de las leyes de propiedad intelectual ni en nuestro Código Penal, si bien este último tipifica como delito la reproducción, plagio, distribución, comunicación pública u por otro medio de obras protegidas por la propiedad intelectual, con ánimo de lucro y en perjuicio de terceros, sin autorización de sus legítimos titulares. La falta de definición del concepto jurídico de plagio, tanto en el ámbito civil como en el penal, ha convertido la jurisprudencia en la principal fuente legal; en este sentido, la definición más citada hasta la fecha es la contenida en la sentencia dictada el 28 de enero de 1995 por la Sala Primera del Tribunal Supremo, que en sus Fundamentos de Derecho indica que “por plagio hay que entender, en su acepción más simplista, todo lo que supone copiar obras ajenas en lo sustancial” y que “ha de referirse a las coincidencias estructurales básicas y fundamentales y no a las accesorias, añadidas, superpuestas o modificaciones no trascendentales”. Pero a pesar de que a día de hoy tenemos una legislación que protege los derechos morales y patrimoniales de los autores y que tipifica como delito el plagio y la piratería, lo cierto es que estas prácticas ilícitas siguen *in crescendo*, alentadas en parte por la facilidad de acceso a la información que proporcionan las nuevas tecnologías.

antes de comprender el interés económico individual derivado de su explotación, el reconocimiento de las facultades patrimoniales es anterior en el tiempo al de las facultades morales. La razón estriba en el hecho de que la concesión de autorizaciones exclusivas de explotación —bien en forma de privilegios regios, bien ya como derechos de propiedad— responde no sólo a necesidades de tipo individual relacionadas con la retribución al impresor y al autor por su trabajo, sino también (y sobre todo) a necesidades supraindividuales relacionadas con el control de las ideas y el desarrollo monopolizado del mercado editorial.

Los abusos contra los derechos de autor han existido siempre, pero sólo empiezan a ser perseguidos y castigados cuando estas malas praxis atentan contra la moral o el sistema económico. Sólo la existencia de una necesidad social conlleva la aparición de una solución jurídica, y ésta evoluciona conforme lo hace dicha necesidad; en palabras de Macrobio, “leges bonae ex malis procreantur”.

De un tiempo a esta parte, y gracias los estudios de atribución de autoría de la Lingüística forense, el lenguaje se ha convertido en garante de la autoría, en un medio de prueba válido para identificar a los autores de los textos (incluso cuando ellos no desean ser identificados). Gracias al idiolecto —o por su culpa—, la idea de la anonimia empieza a ser una entelequia. Los escritores, al fin y al cabo, siempre dejan su huella.

1.2.2. El complejo concepto de autoría en el teatro español del Siglo de Oro

Si la desprotección a la que han estado sometidos los autores a lo largo de la historia ha propiciado que se cometan numerosas infracciones en contra de la paternidad e integridad textual de sus obras, esta desprotección ha sido aún más notoria en el caso del teatro aurisecular. Aunque es de sobra conocido, recordaremos, a través de las palabras de Profeti, el complejo proceso de transmisión textual de las obras dramáticas del Siglo de Oro:

El comediógrafo escribe la comedia para una compañía, que la compra e inmediatamente después la representa; el autor de comedias (el director-empresario de la compañía) adquiere todos los derechos sobre el texto: lo puede reducir, arreglar, le puede añadir

fragmentos, etc. [...] Sólo después de que la comedia se haya utilizado en el tablado se puede vender de nuevo a un impresor para su impresión [1996: 131-132].

El escritor vende la pieza a un *autor de comedias*³³ que, automáticamente, se convierte en el propietario exclusivo de la misma, teniendo pleno derecho para manipular, seccionar, alterar o aumentar ciertos pasajes del texto; sólo cuando considera que ha explotado suficientemente la obra en los teatros, o bien se la vende a otra compañía —que adquiere su propiedad y puede introducir a su vez nuevos cambios sobre el texto—, o bien se la vende a un librero para que la edite e imprima³⁴. El autor de esta época “never quite *owned* a literary work, or a least not a literary work as we now somewhat abstractly conceive it”, ya que, en sentido estricto, el texto teatral sólo adquiere valores abstractos de propiedad en el momento en el que alguien decide comprarlo [Loewenstein, 1985: 102].

Conservamos algunos testimonios, no muchos, de los contratos de compraventa de comedias, que a veces se producen incluso antes de la composición de las obras. En estos contratos se estipula la cesión de la exclusividad por parte del dramaturgo al empresario, unas veces de manera indefinida y otras con carácter limitado en el tiempo. Así, por ejemplo, tanto en el contrato de la comedia *El relincho*, suscrito entre Damián Salucio del Poyo y Nicolás de los Ríos (1593), como en el de *La descendencia de los condes de Orgaz*, entre Agustín Castellanos y Alonso de Heredia (1602), los dramaturgos renuncian para siempre a los derechos sobre sus obras; sin embargo, en la

³³ En el mundo teatral barroco, la palabra “autor” puede referirse al “que escribe libros, y compone y saca a luz obras literarias”, pero también al “que es cabeza y principal de la farsa, que representa las Comedias en los corrales o teatros públicos, en cuyo poder entra el caudal que adquieren para su mantenimiento, y para repartirlo entre los cómicos” (*Diccionario de autoridades*). Para evitar la confusión, utilizaremos la redonda para referirnos al autor como dramaturgo, y la cursiva para referirnos al *autor* como empresario.

³⁴ El principal fin de una obra dramática es la representación, y la edición textual es tan solo una posibilidad que apenas interesa a los comediógrafos, principalmente porque la compensación económica es mucho menor (Se calcula que el valor de una comedia vendida a una compañía es seis veces superior que el de su venta a un librero para su impresión). Calderón se hace eco de este desequilibrio económico entre la ‘comedia representada’ y la ‘comedia impresa’ en el prólogo a su *Cuarta parte*, al afirmar —hiperbólicamente— que “dada a la estampa, la que ayer valía cien ducados en casa del Autor, vale hoy un real en casa del Librero”.

CAPÍTULO 1

escritura de compraventa de comedias entre Cervantes y Gaspar de Porres (1588), el autor del *Quijote* se compromete a no darlas ni entregarlas “a ningún *autor de comedias* destes reinos ni fuera dellos dentro de dos años cumplidos primeros siguientes”, y en el firmado entre Agustín Moreto y Gaspar Fernández Valdés en torno al año 1650, el plazo se amplía hasta los diez años. Sea como fuere, en el contexto del teatro barroco, el objeto protegido no es la obra creada por el ingenio del autor (*corpus mysticum*) sino el producto comprado por el empresario (*corpus mechanicum*), y esa protección va destinada a garantizar la rentabilidad de su inversión.

Conservamos también escritos de compraventa o alquiler de obras dramáticas entre *autores de comedias*, sobre los que el dramaturgo ya no tiene ninguna voz. Tal y como ha estudiado García-Reidy, el *autor* Andrés de la Vega compra al dramaturgo sevillano Andrés de Claramonte una comedia titulada *San Carlos* o *Las dos columnas de Carlos*, y tras representarla con su compañía durante cerca de dos décadas, en 1642 decide copiarla y vendérsela al *autor* Bartolomé Romero. Así figura en la copia manuscrita, donde puede leerse:

Yo, Andrés de la Vega, autor por el rey nuestro señor, digo que esta comedia es de Bartolomé Romero por cuanto me la ha pagado, quedándome yo con el original, en que queda la aprobación del Sr. Gregorio López Madera, y porque, como dicho he, me la ha pagado y ser verdad, lo firmé en Madrid a 21 de marzo de 1642. Andrés de la Vega [Citamos por García-Reidy, 2008: 179-180].

En este escrito se deja claro que la propiedad del manuscrito original no es de su legítimo creador, Andrés de Claramonte, a quien ni se nombra en el acuerdo, sino del empresario Andrés de la Vega, y que éste, como propietario exclusivo del texto, tiene absoluta legitimidad para vendérselo a Bartolomé Romero sin tener que dar cuentas a nadie. Estos acuerdos entre compañías a veces se circunscriben a un lugar concreto, como ocurre con la comedia *La ventura por el pie*, que al *autor* Baltasar de Pinedo entrega en 1614 a los mayordomos de la cofradía del Santísimo Sacramento de Hervás, con la condición de que “la representen en dicho pueblo y nada más que en él” [Citamos por Pérez Pastor, 1901: 142].

Los manuscritos dramáticos funcionan como instrumentos de trabajo, y es habitual que los *autores* modifiquen los textos para actualizarlos o para adaptarlos a las

circunstancias de sus propias compañías, a los espacios de representación o a los gustos del público. Es entonces, y sólo entonces, cuando el texto escrito se convierte en texto espectacular. Aunque por lo general estas modificaciones se limitan a pequeñas supresiones, y rara vez a añadidos³⁵, lo cierto es que son una prueba más del derecho de modificación que poseían los *autores* y de la absoluta desprotección de los dramaturgos, sin derecho alguno sobre sus textos (ni siquiera de índole personal o moral).

Las compañías muestran especial cuidado en proteger los manuscritos teatrales, guardándolos incluso bajo llave, ya que mantener el monopolio de representación de los textos que han comprado es su única salvaguardia para recuperar la inversión y ganar dinero. A pesar de sus precauciones, la piratería de textos teatrales está presente de muy diversas formas. Es habitual que los actores de las compañías vendan subrepticamente textos a la competencia, o que se lleven copias en secreto al cambiar de formación. De igual modo, ciertas compañías logran hacerse con los textos de otras gracias a los llamados poetas duende o memorillas, personas con facultades nemotécnicas excepcionales que acuden a las representaciones y se aprenden de memoria los textos, para luego transcribirlos —eso sí, a su modo— y vendérselos a las compañías rivales, que cambian su título y empiezan a representarlos como si fueran textos nuevos³⁶. Lope de Vega es especialmente sensible a este problema, tal vez por ser uno de los mayores

³⁵ Ruano de la Haza explica cómo el autor de comedias Antonio de Escamilla, famoso por hacer papeles de gracioso, añade veinticuatro versos a las intervenciones de este personaje en la comedia *Cada uno para sí*, de Calderón de la Barca, con el objeto de potenciar sus habilidades cómicas [1982: 60]. Pero ésta no es la práctica habitual, pues como advierte Presotto tras estudiar minuciosamente los manuscritos autógrafos de Lope, “los *autores* parecen dedicarse a lo largo de los años a «eliminar» y casi nunca «añadir»” [1997: 162].

³⁶ Los dos memoriones más famosos de la época son los hermanos Ramírez de Arellano: Luis, apodado *La Feliz Memoria* o *Gran memoria*, y Juan, el *Memorilla*, por estar algo peor dotado para esta práctica que su hermano. En la *Plaza universal de todas ciencias y artes*, Cristóbal Suárez de Figueroa explica el método nemotécnico desarrollado por Luis Ramírez de Arellano: “Él toma de memoria una comedia entera de tres veces que la oye, sin discrepar un punto en traga y veros. Aplica el primer día a la disposición; el segundo a la variedad de la composición; y el tercero a la puntualidad de las coplas. De este modo encomienda a la memoria las comedias que quiere. En particular tomó así la *Dama Boba*, el *Príncipe Perfecto*, y la *Arcadia*, sin otras” [Citamos por Cáseda Teresa, 2019: 302].

CAPÍTULO 1

damnificados³⁷. Así desacredita la labor de estos sujetos en el prólogo a su parte decimotercera³⁸:

A esto se añade el hurtar las comedias estos que llama el vulgo, al uno *Memorilla*, y al otro *Gran Memoria*, los cuales, con algunos versos que aprenden, mezclan infinitos suyos bárbaros, con que ganan la vida vendiéndolas a los pueblos y autores extramuros: gente vil, sin oficio, y que muchas veces han estado presos. Yo quisiera librarme de este cuidado de darlas a luz, pero no puedo, porque las imprimen con mi nombre y son de los poetas duendes que arriba digo. Reciba, pues, el lector esta Parte, lo mejor que ha sido posible corregirla, y con ella mi voluntad, pues sólo tiene por interés que lea estas comedias menos erradas y que no crea que hay en el mundo quien pueda tomar de memoria una comedia viéndola representar, y que si le hubiera, yo le alabara y estimara por único en esta potencia, aunque le faltara el entendimiento, porque raras veces se hallan juntas por opinión del Filósofo, confirmada de la experiencia [Citamos por Cáseda Teresa, 2019: 303].

Y lo mismo encontramos en la dedicatoria de *La Arcadia*, dirigida al doctor Gregorio López Madera, miembro del Consejo Supremo de Su Majestad:

Espero entre, otras cosas, que quien ha escrito e impreso (si bien en tan distintas y altas materias), se dolerá, de los que escriben, y que ahora tendrá remedio lo que tantas veces se ha intentado, desterrando de los teatros; unos, hombres, que viven, se sustentan y visten de hurtar a los *autores* las comedias, diciendo que las toman de memoria de sólo oírlas y que este no es hurto, respecto de que el representante las vende al pueblo, y que se puede valer de su memoria que es lo mismo que decir que un ladrón no lo es, porque se vale de su entendimiento dando trazas, haciendo llaves rompiendo rejas, fingiendo personas cartas, firmas y diferentes, hábitos. Esto no sólo es en daño de los autores por quien andan perdidos y empellados; pero lo que es más: de sentir de los ingenios que las escriben porque yo he hecho diligencia para saber de uno de estos, llamado el de la gran memoria, si era verdad que la tenía, y he hallado, leyendo sus traslados que para un verso mío hay infinitos suyos, llenos de locuras, disparates e ignorancia,, bastantes a quitar la honra y opinión al mayor ingenio en nuestra nación y las, extranjeras, donde ya se leen con tanto gusto; pues si aquel antiguo poeta quebró al ollero los vaso, con el báculo, porque cantaba mal sus versos, ¿qué harán los que ven contrahacer los suyos de oro en barro? [Citamos por Cáseda Teresa, 2019: 305].

³⁷ Especialmente notorio es el caso del testimonio manuscrito de *Peribáñez y el Comendador de Ocaña* conservado en la Biblioteca de la Melbury House (en Dorset, Inglaterra), poco fiel al texto original y que parece corresponderse con una versión recompuesta por algún memorilla [Ruano de la Haza, 1983].

³⁸ Los dos poetas duende más conocidos de la época eran Luis Ramírez de Arellano, apodado el *Memorión* o *Gran Memoria*, y su hermano Juan, el *Memorilla*. En la *Plaza Universal de todas ciencias y artes*, Cristóbal Suárez de Figueroa explica así el *modus operandi* del *Gran Memoria*: “Él toma de memoria una comedia entera de tres veces que la oye, sin discrepar un punto en traga y veros. Aplica el primer día a la disposición; el segundo a la variedad de la composición; y el tercero a la puntualidad de las coplas. De este modo encomienda a la memoria las comedias que quiere. En particular tomó así la *Dama Boba*, el *Príncipe Perfecto*, y la *Arcadia*, sin otras” [Citamos por Cáseda Teresa, 2019: 302].

Las copias ilícitas también afectan a los entremeses. Así, en 1618, el representante Sancho de Paz publica *El ingenioso entremés del examinador Miser Palomo*, de Antonio Hurtado de Mendoza, precedido de un prólogo en el que explica lo siguiente: “Lo que me ha obligado a imprimir este ingenioso entremés ha sido el saber que en diez o nueve veces que lo he representado en esta ciudad, muchos a quien no he querido dar traslado dél, lo han ido sacando, ya de memoria o ya escribiéndolo, mientras yo le representaba” [Citamos por Asensio, 1971: 68].

La proliferación de copias secretas, hurtadas o elaboradas por memoriones llega a ser tan habitual que Gregorio López de Madera, protector de Comedias en Madrid, dicta en 1636 en auto en el que establece que “ahora y de aquí alante [*sic*] ninguno de los dichos *autores* ni los representantes de sus compañías puedan hacer ni representar comedia alguna, baile o entremés que no sea propio del *autor* que lo representare y los haya comprado al poeta” [Citamos por García-Reidy, 2013b: 306].

La integridad de los textos teatrales también puede verse afectada por la acción censora, pues cualquier obra, tanto para ser representada como para ser impresa, debe disponer de las oportunas aprobaciones y licencias. Los documentos conservados permiten establecer que los criterios que rigen las aprobaciones de los tomos colectivos o sueltas de comedias no son los mismos que los que promueven las censuras de representación: si por un lado la censura previa de impresión es más flexible y permisiva en cuanto que parece menos grave aquello que se lee que aquello que se representa ante los ojos de los espectadores como recreación verdadera³⁹, por otro, la censura previa de representación es menos estricta y rigurosa por el hecho de que su cumplimiento tiene un seguimiento mucho menos minucioso por parte de las autoridades pertinentes. Sea como fuere, lo cierto es que —al igual que ocurría en el caso de los *autores*— los censores suelen suprimir aquellos pasajes que atentan contra la Santa fe y las buenas costumbres, pero pocas veces optan por reescribir lo atajado, lo

³⁹ Como dice Wilson, “the word spoken in public was always thought to be more scandalous than might quite legitimately be read in private” [1961: 170]. Este crítico recuerda el caso de *El purgatorio de San Patricio*, cuya representación fue prohibida por un censor en 1652, pero que luego circula impresa en pliegos sueltos durante los siglos XVIII y XIX.

CAPÍTULO 1

que atenta gravemente contra la calidad de los textos. Un buen ejemplo para entender el funcionamiento de la censura lo encontramos en uno de los manuscritos conservados de la comedia *Sólo en Dios la confianza*, de Pedro Rosete Niño, en el que figura el encargo que, en octubre de 1689, el protector de comedias Juan Yseca Alvarado hace a tres personas para que “informen en orden a su contenido”: al censor y fiscal oficiales, Francisco Lanini Sagredo y Fermín de Sarasa, tal y como estipulan las Ordenanzas de los teatros, y también a Juan de la Vera Tassis, que actúa como su hombre de confianza, pues por aquel entonces aún no es fiscal de comedias. El análisis de las censuras permite advertir no sólo discrepancias de criterio entre los examinadores, sino también diferentes modos de ejecutar las prohibiciones: mientras Lanini Sagredo opta por ‘cortar por lo sano’, sin importarle romper el sentido del pasaje o la métrica de los versos, Vera Tassis pone un especial empeño en enmendar lo atajado, para suavizar la expresión sin desvirtuar el sentido ni dañar su belleza estética [Ruiz Urbón, 2014].

Pero si en el recorrido que va de la creación a las distintas puestas en escena, el texto dramático va sufriendo paulatinas transformaciones (de mano de los *autores*, los copistas, los apuntadores que elaboran los ‘papeles de actor’, los censores, los memorillas....), no más cercano al original resulta el texto que aparece impreso; ni más fiables son sus autorías. Cuando el teatro comercial alcanza su madurez, son muchos los *autores* que disponen de obras que ya han agotado su ciclo de representación en las tablas y, ante la imposibilidad de poder vendérselas a otras compañías, deciden entregárselas a algún editor para que las dé a la stampa y obtener, al menos, una mínima rentabilidad por ellas. La práctica de imprimir volúmenes coleccionados de comedias arranca en 1603 con la publicación de las *Seis comedias de Lope de Vega y Carpio y de otros autores* (en Lisboa, por Pedro Crasbeeck), en donde —a pesar del título— sólo se recoge una comedia del fénix de los ingenios (*El perseguido*). A partir de este momento, y tal y como estudia Cruickshank [2000], empiezan a proliferar las ediciones piratas y los tomos con autorías inventadas y con textos llenos de erratas o

cortes, que ni sus propios creadores son capaces de identificar⁴⁰. Conocida es la falsificación de la *Tercera parte de comedias de Lope de Vega y otros autores*, hecha en Sevilla por Gabriel Ramos Verejano, o todas las ediciones piratas de las partes de Calderón de la Barca; y muy recientemente se ha localizado en la Biblioteca Nacional de España una suelta fraudulenta de *El castigo sin venganza* [García-Reidy, Valdés Gázquez y Vega García-Luengos, 2021].

Los dramaturgos alzan la voz contra todas estas prácticas, y empiezan a reclamar un mayor control de los textos publicados. Es lo que García-Reidy ha definido como “reapropiación autorial”, dejando claro que la idea de reapropiación evidencia “una pérdida previa” [2013b: 302]. Lope de Vega advierte a los que leen sus escritos con afición que “no crean que aquellas son mis comedias, aunque tengan mi nombre” [Citamos por Couderc, 2009: 120], y llega a presentar una reclamación ante el licenciado Baltasar Gilimón de la Mota, miembro del Consejo de su Majestad, para exigir que nadie imprima sus obras “sin que él por lo menos las vea y corrija” [Citamos por García-Reidy, 2013b: 333]; Juan Pérez de Montalbán asegura que le atribuyen “muchas [comedias] que no son mías”, lo cual es muy enojoso, “porque si son buenas, les usurpo la gloria a sus dueños, y si malas me desacredito con quien las compra” [Citamos por Couderc, 2009: 125]; Antonio Enrique Gómez opta por publicar una lista con sus comedias “para [que] se conozcan por mías, pues todas ellas ó las mas que se imprimen en Seuilla les dan los impressores el titulo que quieren y el dueño que se les antoja” [Citamos por Cruickshank, 2000: 130]; el hermano de Calderón de la Barca decide publicar las obras de éste, por “el pesar de aver visto impressas algunas dellas antes de ahora por hallarlas todas erradas, mal corregidas, y muchas que no son suyas en su nombre, y otras que lo son en el ageno” [Citamos por Cruickshank, 2000: 130]; y lo mismo denuncia el propio Calderón en el prólogo de su *Cuarta parte*:

⁴⁰ A pesar de que los entremeses gozan de notable éxito en los corrales de comedias, en un primer momento sólo se imprimen dentro de los volúmenes de partes de comedias. Las primeras colecciones impresas de teatro breve son de 1640 (*Flor de sainetes*, de Francisco Navarrete y Ribera, y *los Entremeses nuevos, de diversos autores*), aunque tenemos constancia de que Luis Quiñones de Benavente ya tenía en mente publicar sus entremeses desde al menos 1638 [Madroñal Durán, 1994].

CAPÍTULO 1

No solo hallé en sus impresiones, que ya no eran mías las que lo fueron; pero muchas que no lo fueron, impresas como mías, no contentándose los hurtos de la prensa con añadir sus yerros a los míos, sino con achacarme los ajenos, pues, sobre estar, como dixé (las ya no mías) llenas de erratas, y por el ahorro de papel, aun no cabales (pues donde acaba el pliego, acaba la jornada, y donde acaba el cuaderno, acaba la Comedia), hallé, ya adocenadas, y ya sueltas, todas éstas que no son mías, impresas en mi nombre [...] Pues, bien mirada al primer viso esta materia, ¿qué le importa a la República que la comedia de Juan ande en nombre de Pedro, ni la de Pedro esté cabal o adulterada? Y, aunque mirada a segunda luz, tiene considerables inconvenientes en daño tercero, ¿quién quiere Vuestra Merced que se meta en advertirlos el día que no los advierte la conciencia de quien, no pudiendo ignorar que una comedia en su primera estimación cuesta al autor cien ducados (y si le sale mala, no vale el papel en que está escrita, y si buena, no hay precio con que pagarla, porque es un crédito abierto en todos los lugares donde llega nueva), y no pudiendo (digo otra vez) ignorar tampoco el ser hurtada, pues no es su dueño el que la vende, sino el apuntador que la traslada o el compañero que la estudia o el ingenio que la contrahace, con todo esto se la compra? Con que, dada a la estampa, la que ayer valía cien ducados en casa del Autor, vale hoy un real en casa del Librero, cuyo menoscabo lleva tras si el no averiguable precio de mañana [Citamos por Couderc, 2009: 127].

El texto del teatro aurisecular es un producto abierto, en cuya elaboración interviene en primer término el dramaturgo, pero poco después los *autores de comedias*, los actores, los memorillas, los censores, los copistas, los editores, los impresores..., según van teniendo acceso a él⁴¹. El texto sólo adquiere la idea de propiedad en el momento en que se convierte en “mercadería vendible”, y su creador, aunque lo pretenda, no dispone de ningún derecho como autor, ni siquiera en la esfera de lo moral. Todas estas circunstancias deben hacernos desconfiar tanto de los textos heredados como de los nombres que figuran en ellos, elevando la incógnita de la autoría a un enigma de grandes dimensiones.

⁴¹ A todos estos problemas relacionados con la integridad y la paternidad de las obras, hay que sumar el de las autorías colectivas, un fenómeno muy habitual en el teatro del Siglo de Oro español, y que surge de la urgencia de dar respuesta a la demanda creciente de piezas teatrales.

CAPÍTULO 2

LA EVIDENCIA LINGÜÍSTICA: LOS ESTUDIOS DE AUTORÍA DE LA LINGÜÍSTICA FORENSE

“Forensic linguistics responds to legal questions that involve language. Who wrote a ransom note or a bomb threat letter? What is the meaning of a word in a contract? Just what constitutive «consenting» to a search? If an interrogator indirectly makes a promise, does that invalidate a confession? Is «Mc» part of the McDonalds trade name or just part of English (and can I name my restaurant McHamburgers)? Could a fourth-grade dropout actually have written a confession with the phrase, «he approached the vehicle and I raised my weapon?»” (Robert Andrew Leonard).

2.1. LA LINGÜÍSTICA FORENSE

2.1.1. Las ciencias forenses y los principios rectores de la Criminalística moderna

A día de hoy siguen siendo muchos los que asocian la palabra “forense” con un cadáver o con la investigación de un asesinato, influidos tal vez por series como *CSI: Las Vegas*, *Mentes criminales* o *Dexter*. En realidad, el término *forense* viene del latín FORENSIS, que significa “perteneciente o relativo al foro”; a su vez, FORENSIS deriva de FORUM, que hace referencia al “foro” o “plaza pública”. En la Antigua Roma, los juicios se celebraban en esas plazas públicas y de ahí que el adjetivo *forense* haga referencia, en

CAPÍTULO 2

la actualidad, a las disciplinas científicas que sirven de soporte a la Policía y a la Justicia para que puedan determinar la existencia cierta de un hecho delictivo, así como las circunstancias en que se ha producido y la identificación de sus autores.

Las ciencias forenses, por tanto, son materiales auxiliares del Derecho, y también imprescindibles, que aplican métodos especializados a las investigaciones policiales o a los procedimientos legales con el fin de descubrir a la verdad, entendiendo por “verdad” lo más próximo o lo más probable. Existe una gran variedad de campos de interés en relación con las ciencias forenses, desde la identificación de una persona por su ADN, su voz o los rasgos caligráficos de su firma hasta la datación de un cadáver a través del estudio de la fauna sarcosaprófaga. Aunque la más conocida es la Medicina forense, cualquier ciencia puede ser forense en el momento en que sirve al procedimiento judicial: la Genética, la Psicología, la Toxicología, la Economía, la Arquitectura, la Biología, la Entomología, la Informática... y, por supuesto, también la Lingüística.

Siete son los principios fundamentales que conforman el paradigma de la Criminalística moderna y que también adopta la Lingüística en el ámbito forense⁴²:

1. El *principio de uso*, que propugna que toda persona que comente un hecho delictivo utiliza algún agente para llevarlo a cabo (físico, químico, biológico, mecánico, informático, tecnológico o lingüístico). En un caso de asesinato, por ejemplo, el agente puede ser un arma de fuego, un martillo o una sustancia venenosa; en un caso de plagio, el agente lo constituirá el propio acto de escritura.

2. El *principio de producción*, basado en la idea de que en la comisión de un hecho delictivo se producen una serie de elementos materiales de gran variedad morfológica y estructural, y que dichos materiales son susceptibles de ser analizados como elementos reestructores del hecho e identificadores de los agentes causantes del mismo (indicios o evidencias). Los indicios más habituales son el ADN o las huellas dactilares, pero, como veremos, el lenguaje también puede constituirse en evidencia

⁴² Cuatro de ellos fueron enunciados en 1910 por el médico francés Edmond Locard, uno de los principales precursores de la Criminalística moderna: el principio de intercambio o transferencia, el principio de correspondencia, el principio de reconstrucción de los hechos y el principio de probabilidad.

científica para identificar a una persona, detectar un plagio o descubrir la mentira en una declaración policial.

3. El *principio de intercambio o transferencia*, que parte de la idea de que siempre que dos objetos entran en contacto transfieren parte del material que incorporan al otro; aunque teóricamente el intercambio es bidireccional, basta con encontrar la impronta de uno de ellos en el otro para demostrar el contacto. De este modo, siempre va a existir un intercambio de elementos materiales (indicios) entre el autor, la víctima y el lugar de los hechos, o, en su caso, entre el autor y el lugar de los hechos. En un caso de agresión sexual, por ejemplo, podremos encontrar indicios del autor en la víctima (semen, piel bajo las uñas), indicios del lugar en el autor (restos de tierra en el zapato), indicios del autor en el lugar (huellas en el suelo), indicios de la víctima en el autor (arañazos), etc. También podemos ejemplificarlo a través de un caso literario: en *El nombre de la rosa*, de Umberto Eco, la evidencia que permite vincular la muerte de varios frailes con la lectura prohibida de la *Retórica* de Aristóteles es el resto visible que el veneno impregnado en las páginas del manuscrito dejaba en los dedos y la boca de las víctimas.

4. El *principio de correspondencia de características*, según el cual la acción de determinados agentes puede dejar una señal en ciertos lugares o elementos. De este modo, la marca de un estrangulamiento en el cuerpo de la víctima permitirá identificar el grosor de la soga; la huella en un terreno, el tipo de zapato y pisada de su autor; o la composición de la tinta en un documento manuscrito, el tipo de esferográfica con el que ha sido redactado.

5. El *principio de reconstrucción de los hechos*, que parte de la idea de que gracias al estudio de las evidencias materiales asociadas al hecho analizado es posible reconstruir sus mecanismos y las formas de acción del presunto autor. En un caso de asesinato, las huellas dactilares, los rastros de sangre, los casquillos de las balas o el estudio del cadáver permitirán establecer la fecha probable de la muerte, el lugar del crimen o el mecanismo de comisión del delito.

6. El *principio de probabilidad*, que establece que la acumulación de datos o evidencias permite inferir probabilidades en relación con los hechos delictivos, que van desde la probabilidad nula al grado máximo de probabilidad.

7. Y, por último, el *principio de certeza*, que está relacionado con la valoración cualitativa y cuantitativa de las evidencias, a través de los métodos, técnicas y procedimientos adecuados para ello.

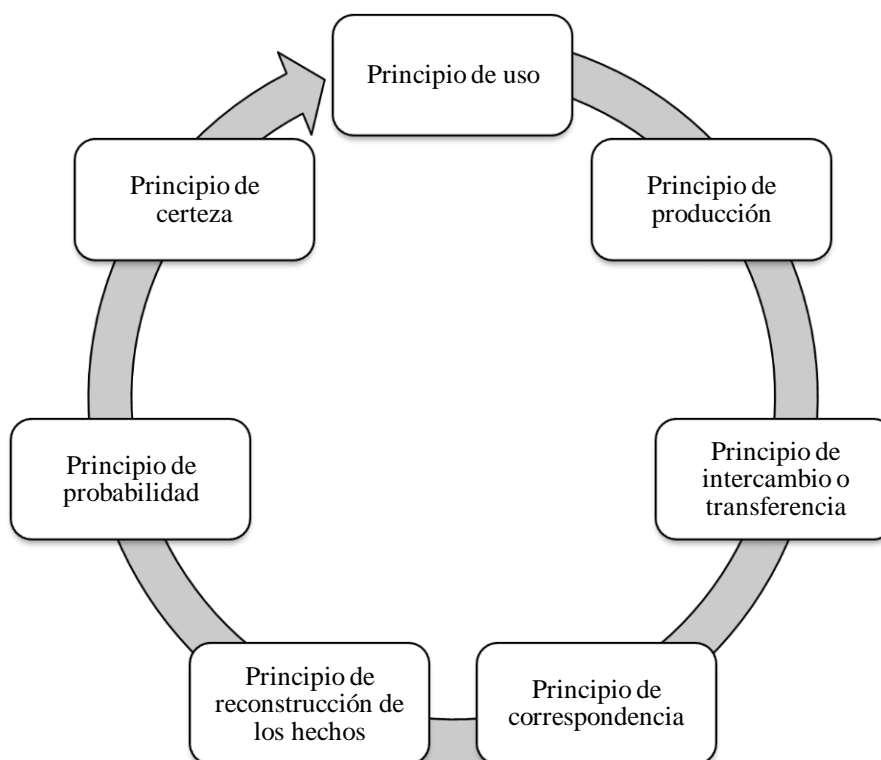


Figura 1: Principios rectores de la Criminalística moderna. Fuente: Elaboración propia.

2.1.2. El nacimiento de la Lingüística forense como disciplina científica

La Lingüística forense, como rama de la Lingüística aplicada, puede definirse de manera global como una ciencia multifacética e interdisciplinar que pone en conexión los campos de la Lingüística y el Derecho con el fin de dar respuesta a las cuestiones legales que, de un modo u otro, implican la lengua. Aunque aún gran desconocida en

nuestro país y de muy reciente aplicación, lo cierto es que la Lingüística forense posee un amplio bagaje en el mundo anglosajón, en donde su practicidad y fiabilidad están fuera de toda duda. Los orígenes de esta disciplina se sitúan a mediados del siglo pasado, primero en Estados Unidos con la aparición de la voz *forensic english* [Philbrick, 1949] y poco después en Europa con *forensic linguistic* [Svartvik, 1968], aunque uno y otro término no son empleados en el mismo sentido.

En el año 1949, el filólogo americano Frederick A. Philbrick escribe una obra sobre inglés legal titulada *Language and the Law: the Semantics of Forensic English*, en la que analiza el uso del lenguaje jurídico y judicial en el habla inglesa. Philbrick emplea el término *forensic english* para referirse al lenguaje especializado que utilizan los actores del proceso penal: “The book is accordingly a study of *forensic English*, that is, the English used by advocates and judges in courts of the law” [1949: vi]. Casi dos décadas después, en 1968, el lingüista sueco Jan Svartvik publica *The Evans statements. A case for forensic linguistics*, en el que utiliza el lenguaje como evidencia para probar la inocencia del joven galés Timothy John Evans, que había sido injustamente ejecutado por los asesinatos de su mujer y su hija en 1950 (pero que cometió en verdad su vecino John Christie). Según esta investigación, las partes de las declaraciones policiales en las que Evans se autoinculpaba de los hechos (dubitadas) presentaban un estilo gramatical muy diferente del resto (indubitadas)⁴³, lo que permitía inferir que esas partes habían sido manipuladas por la persona encargada de transcribir el interrogatorio. Aunque para Evans este estudio llegó tarde,

this sally into the relatively uncultivated field of “forensic linguistics” has been interesting for a number of reasons, but two in particular: Firstly, it has provided the linguist with one of those rare opportunities of making a contribution that might be directly useful to society. Such a statement does not imply the utility should be a goal of linguistic research, only that,

⁴³ Llegado este punto, consideramos necesario aclarar la diferencia entre *dubitado* e *indubitado*, términos propios de las ciencias forenses y del ámbito pericial. Según el artículo 608 de la Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil, un documento indubitado es aquel “sobre cuya veracidad no cabe duda alguna y, en consecuencia, puede servir para comparar con otro documento y demostrar la autenticidad de éste”. Ciñéndonos al contexto de la atribución de autoría, hablaremos de material dubitado para referirnos a las muestras lingüísticas cuya autoría se cuestiona y de material indubitado para referirnos, por el contrario, a las muestras lingüísticas cuya autoría no se cuestiona.

CAPÍTULO 2

once in a while, it is good for the linguist to know that he can be useful, and that applied linguistics need not be identical with language teaching or machine translation. Secondly, it has highlighted our present inadequate knowledge of how language is used in various situations [Svartvik, 1968: 6].

Tras la publicación de la obra de Svartvik, que hoy es considerada el punto de partida de la Lingüística forense como tal [Coulthard y Johnson, 2007: 5], empiezan a sucederse una serie de trabajos, generalmente escritos por lingüistas, que analizan confesiones dubitadas, investigan incoherencias lingüísticas en declaraciones policiales o estudian la autenticidad de ciertos registros conversacionales [Shuy, 1993; Levi, 1994; Eades, 1994]. Todos estos trabajos constituyen los primeros pasos de una disciplina en ciernes, que no empezará a desarrollar unos principios metodológicos fiables hasta la llegada del nuevo siglo.

2.1.3. Definición y campos de actuación de la Lingüística forense

La International Association of Forensic Linguists (AIFL)⁴⁴ define la Lingüística forense como la “interfaz entre lengua y derecho”. Aunque es una definición comúnmente aceptada, lo cierto es que a día de hoy todavía existe una falta de consenso entre los lingüistas y los juristas para delimitar lo que esta disciplina realmente engloba⁴⁵:

There is no consensus among linguists and legal experts as to a definition of the term forensic linguistics. Some adhere to a narrow definition such as “the use of linguistic techniques to investigate crimes in which language data constitute part of the evidence.” Others subscribe to a broad interpretation of forensic linguistics as the study of the intersection between language and the law [Gibbons, 1999].

⁴⁴ A finales de septiembre de 2021, tras una votación en la que hemos podido participar todos los socios, la International Association of Forensic Linguists (AIFL) ha pasado a llamarse International Association for Forensic and legal Linguistics (AIFLL).

⁴⁵ Para una revisión de las distintas definiciones y su evolución en el tiempo, véase Garayzábal Heinze, Jiménez Bernal y Reigosa Riveiros [2014: 29-34].

Los problemas para establecer una definición exacta residen en el hecho de que el maridaje entre Lengua y Derecho puede dar respuesta a muchas y muy variadas incógnitas, tal y como expresa Leonard en la cita con la que abrimos este capítulo [2006: 3].

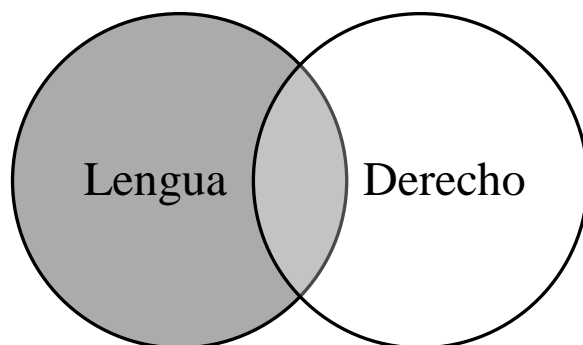


Figura 2: Definición de Lingüística forense. Fuente: Elaboración propia.

Tratando de sistematizar todas estas posibilidades, Gibbons y Turell establecieron los tres campos de actuación que engloba la Lingüística forense en su más amplio sentido del término [2008]:

1. *The language of the law / El lenguaje jurídico-administrativo*⁴⁶. Este ámbito de actuación examina el lenguaje que se emplea en la legislación, en el ordenamiento jurídico y en los documentos administrativos desde muy diversas ópticas. Así, por ejemplo, analiza la legibilidad y comprensibilidad de documentos relacionados con el Derecho, tanto públicos como privados (leyes, decretos, normas, reglamentos, subvenciones, contratos, testamentos, etc.); trata de solventar los problemas de interpretación derivados de la complejidad gramatical o de la ambigüedad terminológica de este tipo de textos; o intenta sentar las bases de una buena traducción, que tenga en consideración la asimetría de los diferentes sistemas jurídicos.

⁴⁶ En los manuales suele traducirse como “lenguaje de la ley”, “lenguaje legal” o “lenguaje jurídico”, pero nosotros optamos por “lenguaje jurídico-administrativo” porque creemos que es la etiqueta que mejor se adapta al tipo de textos que realmente se engloban en esta categoría (desde leyes, decretos y normas hasta subvenciones públicas o contratos privados).

CAPÍTULO 2

2. *The language of the court / El lenguaje del procedimiento judicial.* En esta segunda dimensión, la Lingüística forense se centra en el análisis del lenguaje utilizado por los diferentes agentes que intervienen en el procedimiento judicial, lo que engloba tanto el lenguaje utilizado en la investigación (interrogatorios, entrevistas policiales, etc.) como en los juzgados (intervenciones de jueces, abogados, acusados, peritos, testigos, etc.). Un ejemplo de este tipo de estudios sería el análisis de los mecanismos lingüísticos utilizados por los abogados para tratar de convencer al juez o al jurado popular.

3. *The language as evidence / El lenguaje evidencial o probatorio.* En este ámbito de actuación, el lenguaje se usa como evidencia para elaborar perfiles lingüísticos, identificar a una persona por su voz o sus producciones escritas, detectar casos de plagio, establecer la fiabilidad de la transcripción de una declaración policial, determinar la presencia de mecanismos lingüísticos propios de la mentira o establecer las consecuencias derivadas de la ambigüedad textual. Es el campo de actuación que más interés despierta y, también, el más mediático.

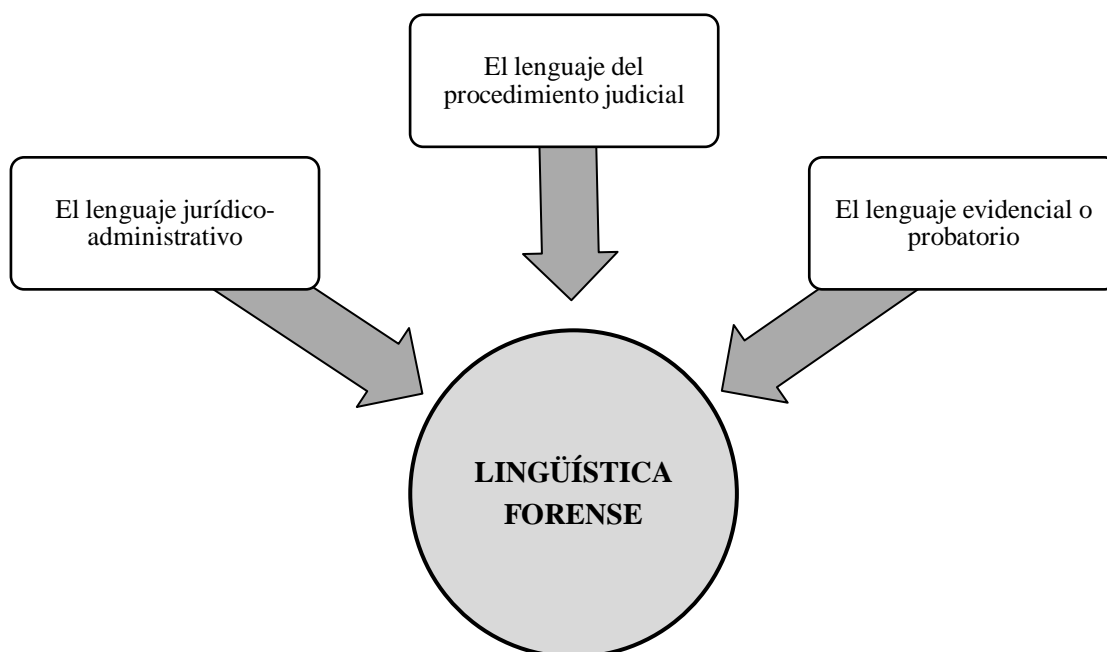


Figura 3: Campos de actuación de la Lingüística forense. Fuente: Elaboración propia.

Aunque esta clasificación se ha ido revisando y matizando con el paso de los años [Fernández Trillo, 2015: 19-34; McMenamín, 2012: 69-80 y 2017: 34-38; Coulthard, Johnson y Wright, 2017: 4], lo cierto es que, en líneas generales, podemos hablar de dos grandes posiciones: la de aquellos que por Lingüística forense entienden cualquier interacción entre Lengua y Derecho, lo que englobaría las tres áreas de estudio, o aquellos que, adoptando una postura mucho más estricta, entienden que la Lingüística forense sólo debe hacer referencia al lenguaje evidencial o probatorio.

Dado que el objeto de análisis de la presente tesis es la aplicación de los fundamentos teórico-metodológicos de los estudios forenses de atribución de autoría al ámbito literario, nos centraremos únicamente en este campo de actuación (*The Language as evidence*) y, en concreto, en la comparación forense de textos escritos para discriminar autorías.

2.2. LA EVIDENCIA LINGÜÍSTICA COMO PRUEBA PERICIAL

2.2.1. El idiolecto o estilo idiolectal

El concepto *idiolecto* está estrechamente relacionado con el de *variación lingüística*⁴⁷, introducido en el ámbito de la Sociolingüística por William Labov a mediados del siglo pasado. Según la perspectiva laboviana, toda producción lingüística es variable —es decir, puede manifestarse de diferentes modos—, lo que implica que incluso los individuos que comparten un mismo código lingüístico produzcan en la práctica estructuras superficiales distintas para una misma estructura profunda [Labov, 1966, 1972, 1994, 2001].

⁴⁷ Una *variable lingüística* es un conjunto de manifestaciones de un mismo elemento, y cada una de estas manifestaciones recibe el nombre de *variante* [Moreno Fernández, 1998: 18]. Ejemplos de variantes lingüísticas en español pueden ser las distintas realizaciones del fonema /s/ implosivo ([s], [h] o [Ø]); la elección del morfema *-ra* o *-se* en el imperfecto de subjuntivo (*cantara* o *cantase*); la presencia o ausencia del sujeto pronominal; o la manera de escribir los años (2020 o 2.020).

CAPÍTULO 2

La variación lingüística es inherente a todas las lenguas y se puede producir en todos sus niveles, desde el fonético-fonológico al pragmático, pasando por el morfológico, el sintáctico y el léxico-semántico. Los factores que determinan la aparición de unas variantes lingüísticas en ciertas circunstancias y de otras variantes en circunstancias diferentes pueden ser externos e internos; los primeros hacen referencia a variaciones sociales, tales como el sexo, la edad, la clase social, el nivel educativo, la etnia o la profesión, mientras que los segundos hacen referencia a las variaciones propiamente lingüísticas [Moreno Fernández, 1998; Turell, 2003].

A pesar de que la Sociolingüística de la variación contempla ambos factores, la mayor parte de los estudios de esta disciplina se han centrado en la variación en comunidades lingüísticas o comunidades de habla (sociales) y muy pocos en la variación que atañe a las estructuras internas de las propias lenguas (lingüísticas). Estos últimos demuestran que, incluso dentro de comunidades que comparten el mismo contexto social y la misma variedad lingüística, emergen diferencias significativas entre los distintos sujetos que sólo pueden explicarse como elecciones individuales; una perspectiva clave para la Lingüística forense, que parte del fundamento de que cada persona tiene un uso particular de la lengua que habla y escribe, que lo separa y distingue del resto de usuarios de esa misma lengua. De este modo, el *idiolecto* debe entenderse como la variante lingüística característica de un determinado individuo, al igual que el *dialecto* lo es de una determinada zona geográfica, el *sociolecto* de un determinado estatus social, el *cronolecto* de un determinado grupo de edad⁴⁸, etc. Nuestro idiolecto, por tanto, está conformado por una serie de rasgos sociolectales, geolectales y dialectales, cronolectales, sexolectales..., pero también por una serie de rasgos propios e individuales, que nada tienen que ver con todas estas variantes.

⁴⁸ Aunque a día de hoy nadie parece cuestionar la existencia de un habla individual, es importante advertir que no siempre fue así. Jakobson creía que “la propiedad privada, en el campo del lenguaje, no existe” y que “todo está socializado”, incluso en el nivel individual, pues todo emisor adapta, en mayor o menor medida, su mensaje al receptor, especialmente en lo que concierne a su vocabulario [1975: 21] y Barthes llegó a afirmar que el idiolecto era en gran medida una ilusión [1977: 21]

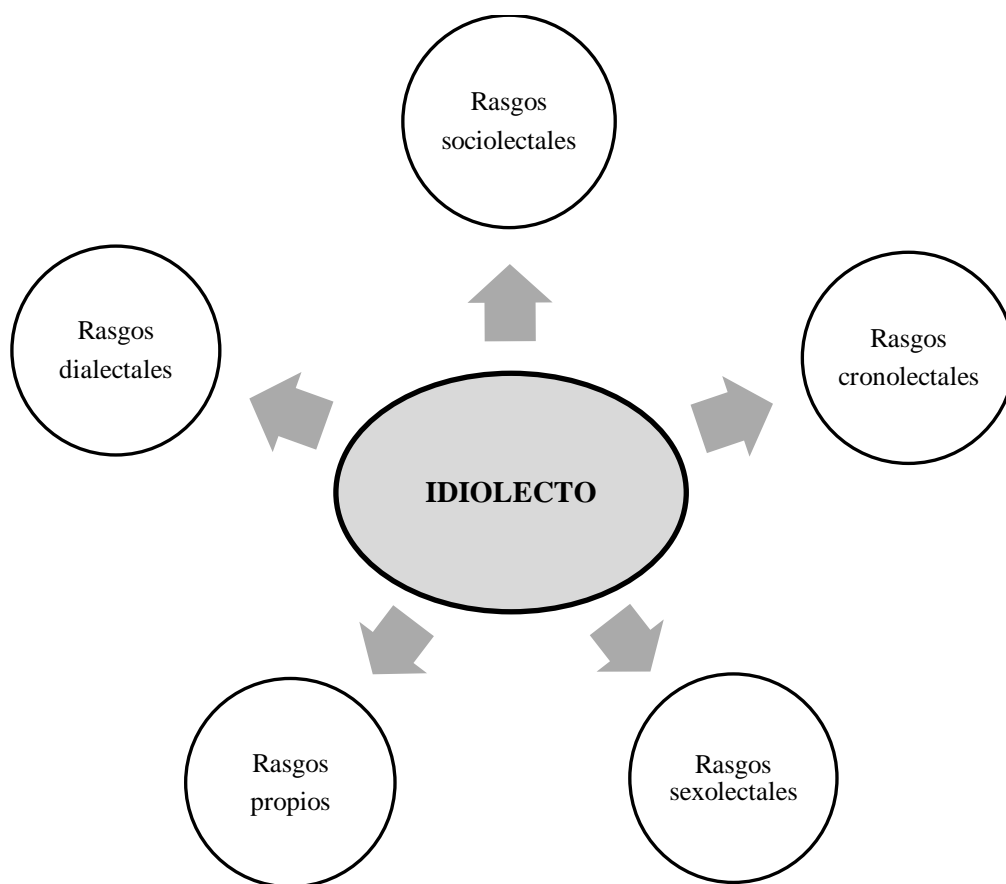


Figura 4: La conformación del idiolecto. Fuente: Elaboración propia.

Uno de los primeros autores que habló del *idiolecto* desde esta perspectiva fue Baldwin, al definirlo como una forma única que tipifica a un individuo particular y que está constituido por sus preferencias lingüísticas, como, por ejemplo, la predilección por ciertas palabras o patrones de entonación [1979: 231]. Brown, poco después, precisó que el idiolecto está conformado por *rasgos idiosincrásicos*, es decir, marcas lingüísticas individuales, y únicas en su combinación, que no guardan correlación con factores de grupo (sexo, edad, origen, posición social, salud, etc.) [1982: 16]. Johnstone, recordando a Saussure, advirtió que hay dos modos de entender la variación individual: tratar la idiosincrasia como desviación o hacerlo como “the set of strategic adaptations he or she makes from a closed set of conventional possibilities, in the inter-actions in which he or she takes part” [1996: 12]. Burridge y Mulder lo explicaron como la

CAPÍTULO 2

“variation within a language that is associated with individual speakers” [1998: 302]. Y ya en el siglo XXI, Coulthard advierte las enormes posibilidades de identificación que facilita la idea de la existencia de una versión individual y distintiva de la lengua:

El lingüista también puede abordar el problema de autoría cuestionada desde el principio según el cual todos los hablantes nativos tienen su propia versión individual y distinta del lenguaje que hablan y escriben, su propio idiolecto, y de la presunción de que dicho idiolecto se manifiesta a través de una selección textual diferenciada e idiosincrásica [2005a: 253].

Es importante advertir que la variación lingüística afecta a todos los niveles y que, por tanto, los rasgos lingüísticos caracterizadores del idiolecto de una persona también mostrarán cierta variación, dependiendo del contexto social en que se produzca el mensaje, el registro que se emplee, el género textual que se adopte o el momento vital y personal en que se encuentre su emisor. Y por eso algunos han visto en esta variabilidad la principal diferencia entre la evidencia lingüística y otros tipos de evidencia no cambiantes, como la huella dactilar o el ADN⁴⁹.

El idiolecto es una consecuencia de la variación lingüística, de eso no cabe duda⁵⁰. Pero lo cierto es que los estudios de los últimos quince años han demostrado que la

⁴⁹ Foster equipara el idiolecto a la huella dactilar o al ADN, al afirmar que “after the crime, the words remain; like fingerprints and DNA” [2000: 4]. Coulthard, por el contrario, desaconseja usar el término “linguistic fingerprint” (“huella lingüística”) al entender que es una metáfora poco fiable, pues “the value of the physical fingerprint is that every sample is both identical and exhaustive, that is, it contains all the necessary information for identification of an individual, whereas, by contrast, any linguistic sample, even a very large one, provides only very partial information about its creator’s idiolect” [2005c: 14]; y en la misma línea se han manifestado muchos otros. Tal vez sea pretencioso comparar el poder discriminatorio de la evidencia lingüística con el de la evidencia dactiloscópica, pero lo cierto es que cada vez son más los estudios orientados a determinar la fuerza idiosincrásica de ciertas marcas lingüísticas. No creemos que el problema resida tanto en la variabilidad de la muestra lingüística (pues a día de hoy parece estar demostrado que las marcas que conforman la huella dactilar de un individuo también presentan cierta variación con el paso del tiempo, aunque ésta sea mínima) como en el peso identificativo de la misma: tanto una huella dactilar como un ADN contienen toda la información de un determinado individuo, mientras que una evidencia lingüística sólo contiene una ínfima parte de los rasgos idiosincrásicos que componen el idiolecto.

⁵⁰ Como bien explica Gerald R. McMenamin en *Forensic Linguistics. Advances in Forensic Stylistics*, la idea de variación no se opone ni al orden ni a la regularidad; el lenguaje no tiene por qué carecer de polimorfismo para ser estructurado, ni necesita ser homogéneo para ser regular; y en este sentido, “a natural language cannot be successfully observed outside a theoretical paradigm, but the paradigm cannot be constructed without the observation of language as it is used” [2007: 65].

variación interpersonal (entre diferentes individuos de una misma lengua) es muy superior a la variación intrapersonal (entre muestras lingüísticas de un mismo individuo), también en español [Turell y Gavaldà 2013; Cicres y Gavaldà, 2014], por lo que sí es posible identificar a una persona por los rasgos idiosincrásicos que caracterizan su voz o sus escritos.

The analysis of variation is important in the forensic sciences because the variation left in trace evidence can be associated with the individual or class characteristics of the instruments or persons involved in a crime. The language left by a speaker in a recorded message or by a writer in an anonymous letter can likewise be associated with the language known to be used by groups and individuals [McMenamin, 2002: 45].

2.2.2. La comparación forense de textos

Como explicábamos anteriormente, “the language as evidence” tiene su fundamento en el principio teórico fundamental de que cada hablante/escritor tiene un *idiolecto* propio —es decir, una versión marcada, distintiva, individual e inconsciente de su lengua materna— que le lleva a producir una serie de muestras orales o escritas substancialmente diferentes de las producidas por otros hablantes/escritores de esa misma lengua. Por este motivo, la comparación forense de textos (entendiendo como textos tanto las muestras orales como las escritas) es el método en el que se sustentan tanto la identificación de hablantes como los estudios de autoría y plagio.

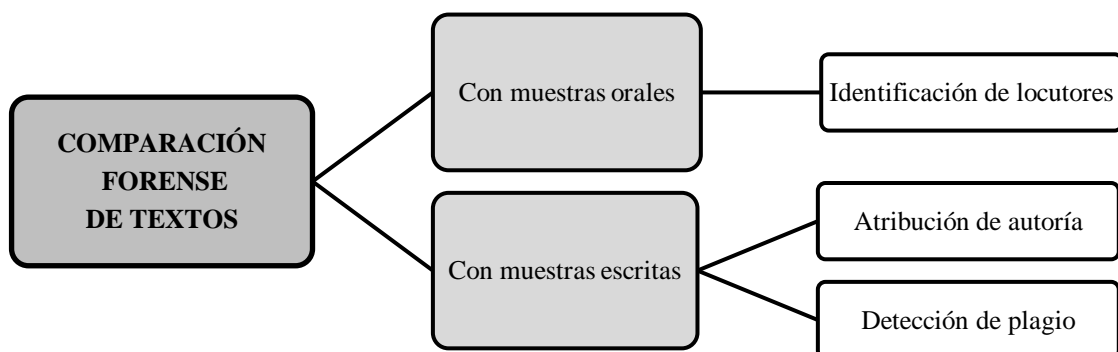


Figura 4: Aplicaciones de la comparación forense de textos. Fuente: Elaboración propia.

2.2.2.1. La identificación de locutores

Una de las aplicaciones más fructíferas de la Acústica forense es el estudio contrastivo de muestras de habla con el fin de determinar si han sido producidas o no por la misma persona. Los hablantes de una lengua muestran diferencias no sólo en las propiedades acústicas de sus producciones (cualidad de voz, frecuencia fundamental, estructura espectral, patrones de duración e intensidad de vocales y consonantes, etc.), sino también en las decisiones —conscientes o inconscientes— del contenido lingüístico de las mismas y en otra serie de factores no estrictamente lingüísticos, como las pautas de respiración, el ritmo, la velocidad de articulación o el uso de pausas y de sonidos no léxicos. Por este motivo, el protocolo de análisis de la comparación forense de voces combina el estudio de las características fisiológicas del hablante con el de toda una serie de fenómenos lingüísticos y paralingüísticos [Cicres, 2014: 367].

La identificación de locutores mediante muestras de voz es una de las áreas de actuación con más larga trayectoria dentro de la Lingüística forense, tanto a nivel internacional como nacional. De sobra conocido es el caso del español Óscar Sánchez, víctima de una suplantación de identidad y condenado en Italia a 14 años de prisión por tráfico de drogas, sobre la base de un dictamen pericial erróneo elaborado por el perito italiano Roberto Porto, que aseguraba que la voz que aparecía en las grabaciones telefónicas interceptadas por la policía era la de Sánchez; finalmente, y gracias a varios contraperitajes posteriores, se logró demostrar que las muestras de voz dubitadas eran sustancialmente diferentes de las muestras indubitadas tomadas al sospechoso, que acabó siendo absuelto y puesto en libertad [Cicres y Gavaldà, 2014: 64].

La comparación de muestras de habla también ha sido determinante en la investigación de la trama Gürtel o en la operación Pokemon [Garayzábal Heinze, Queralt Estévez y Reigosa Riveiros, 2019: 182-184]. Como bien apunta Fernández Trillo, no es aventurado afirmar que la acústica forense seguirá incrementando su relevancia dentro de las indagaciones policiales y judiciales, pues la “creciente proliferación de elementos de comunicación a través de telefonía móvil e Internet convierten el análisis de las conversaciones orales en una herramienta imprescindible para la resolución de muchos casos” [2015: 31-32].

2.2.2.2. *La atribución de autoría*

Cuando un individuo escribe un mensaje, produce un texto único e idiosincrásico que contiene una serie de marcas o recursos lingüísticos que lo hacen irrepetible. Este principio nos permite presuponer que cuando dos textos tienen un alto grado de similitud, o bien tienen un autor común, o bien uno ha copiado al otro; y, en este sentido, puede decirse que los estudios de autoría y plagio constituyen dos caras de una misma realidad.

Junto al caso Evans, del que hemos hablado antes, uno de los primeros casos —y también de los más mediáticos— es el del terrorista estadounidense Unabomber, que durante casi dos décadas envió paquetes bomba a diferentes universidades, aerolíneas y empresas norteamericanas, causando la muerte de hasta tres personas. Después de muchos años de búsqueda infructuosa, el agente James R. Fitzgerald, experto en análisis de textos, pudo probar que Unabomber era Ted Kaczynski, tras comparar un manifiesto anónimo que éste había enviado a los periódicos (texto dubitado) con varias cartas que había escrito a sus familiares (textos indubitados) [Fitzgerald, 2004: 193-222].

Otro caso en el que la prueba lingüística sirvió para resolver un asesinato fue el de Jenny Nicholl, una joven inglesa que en junio de 2005 desapareció de su casa sin dejar rastro; días más tarde, sus familiares y amigos recibieron varios mensajes (SMS) desde el número de teléfono de la joven en los que ésta asegura encontrarse bien, lo que dirigió las primeras hipótesis de la investigación hacia la huida voluntaria. Tres años más tarde, gracias al lingüista Malcolm Coulthard, la familia pudo demostrar que la desaparición de Jenny no había sido voluntaria, ya que los mensajes que supuestamente había enviado tras su desaparición presentaban diferencias lingüísticas notables con los que había enviado antes de la misma (por ejemplo, “i am” vs. “im”, “me” vs. “my”, “meself” vs. “myself”, “phone” vs. “fone”, “have2 go” vs. “have2get”, “of” vs. “off” o “cya” vs. “cu”); al tiempo, esos mensajes dubitados guardaban gran similitud con los mensajes indubitados de David Hodgson, un hombre casado con el que la joven había mantenido una relación sentimental. La prueba lingüística fue suficiente para condenar a Hodgson por asesinato, a pesar de que el cuerpo de la joven nunca fue localizado [Coulthard y Johnson, 2010: 508-522].

2.2.2.3. *La detección de plagio*

La comparación de muestras escritas también puede emplearse para determinar si un texto es original o si, por el contrario, se ha apropiado intencionadamente de las palabras o el contenido de otro. La naturaleza del plagio lingüístico es pluridimensional y muy compleja; puede afectar sólo al lenguaje (forma), sólo al contenido (fondo) o tanto al lenguaje como al contenido (forma + fondo) y puede darse entre textos de una misma lengua o entre traducciones. Una de las definiciones más completas que se han dado hasta la fecha es la de Turell, quien advierte que el plagio “implica la usurpación de una idea y/o la copia intencionada del texto (lingüístico, musical, etc.) usado para expresar esa idea, con el objetivo de ocultar la falta de originalidad” (Turell, 2005b: 278). Desafortunadamente, el plagio es una práctica muy común en la sociedad actual, que se produce en contextos muy diversos, desde el ámbito educativo y académico hasta el periodístico o literario.

Hay casos de plagio textual muy sonados, como el del escritor peruano Alfredo Bryce Echenique, acusado en 2008 de haber copiado varios artículos periodísticos de *La Vanguardia*, el *Periódico de Extremadura* y la revista literaria *Jano* (Tabla 1). El escritor se defendió primero esgrimiendo que el plagio es una forma de halago, para confesar más tarde su culpa; eso sí, respaldándose en una compleja trama de errores informáticos cometidos por su secretaria. El Instituto Nacional de Defensa de la Competencia y de la Protección Intelectual de su país le impuso una multa de 50 Unidades Impositivas Tributarias (equivalentes a unos 42.000 euros).

Otros escritores sobre los que ha planeado la sombra de la sospecha han sido Manuel Vázquez Montalbán, José Saramago, Lucía Etxebarria, Arturo Pérez Reverte, Fernando Aramburu, Luis Alberto de Cuenca o Camilo José Cela, contra el que se querelló Carmen Formoso al entender que *La Cruz de San Andrés*, ganadora del Premio Planeta en 1994, guardaba grandes parecidos con *Carmen*, *Carmela*, *Carmiña*, la novela que ella había presentada a ese mismo certamen (aportando, como prueba, un estudio comparativo de ambas novelas) [Reina, 2012]. Y tampoco se han librado de esta sospecha un buen número de políticos, acusados de plagiar sus tesis doctorales.

FRAGMENTO DE “SÉGOLÈNE, DE CORAZÓN” (FRANCES –MARC ÁLVARO, <i>LA VANGUARDIA</i> , 2006)	FRAGMENTO DE “UN LATIDO LLAMADO SÉGOLÈNE” (BRYCE ECHENIQUE, <i>NEXOS</i> , 2007)
<p>[P2 S1] Escribe Lluís Uria desde París que Ségolène ha sabido llegar al corazón de la gente. [P2 S2] A veces, pendientes de la mercadotecnia, olvidamos que la gran política requiere tanto de corazón como de estrategia. {P3 S2} [P2 S3] El mensaje ilusionado de esta señora me recuerda, salvando todas las distancias, el aire fresco que representó en 1974 el discurso del centrista Valéry Giscard d'Estaing, que tenía como uno de sus lemas clave una frase dirigida a la izquierda: Vosotros no tenéis el monopolio del corazón. {P3 S3} [P2 S4] Gran frase. {P3 S4} [P2 S5] Válida todavía en muchos lugares, empezando por la Catalunya tramposa donde lo que se lleva es hablar de políticas sociales mientras se llevan los hijos a un colegio de élite. [P2 S6] Lo de Giscard d'Estaing era una forma valiente y bella de plantar cara al típico complejo de superioridad moral gauchista, que en la Francia de aquel momento bebía de la resaca de Mayo del 68. {P3 S7} [P2 S7] Giscard d'Estaing venció por la mínima al socialista Mitterrand (que había pactado un programa conjunto con el poderoso Partido Comunista de Marchais) y trató de desplegar un programa reformista avanzado para superar la presidencia del fallecido Pompidou. {P3 S8} [P2 S8] Con los años, su empuje se marchitó (también a causa de una oposición nacida a su derecha y liderada por Chirac) y en 1981, finalmente, Mitterrand alcanzó el Eliseo. {P3 S9}</p>	<p>{P3 S1} Me escribe Jean-Marie Saint-Lu, mi traductor y viejo colega parisino, desde su retiro en el país vasco francés, contándome que Ségolène realmente ha sabido llegar al corazón de la gente. {P3 S2} A veces, pendientes de la mercadotecnia, olvidamos que la gran política requiere tanto de corazón como de estrategia. [P2 S2] {P3 S3} El mensaje ilusionado de esta señora me recuerda, salvando, claro esta, todas las distancias, el aire fresco que representó en 1974 el discurso del centrista Valéry Giscard d'Estaing, que tenía como uno de sus lemas clave una frase dirigida a la izquierda: Ustedes no tienen el monopolio del corazón. [P2 S3] {P3 S4} Gran frase. [P2 S4] {P3 S5} Válida todavía en muchos países, empezando por el Perú, tan complejo que ya le resulta a muchos hasta ininteligible y, se diría, incluso políticamente engañoso. {P3 S6} Ese Perú donde lo que se pregona es una tras otra política social, mientras se lleva a los hijos a un colegio de élite, por ejemplo. {P3 S7} Lo de Giscard d'Estaing fue una forma entonces valiente y bella de plantar cara al típico complejo de superioridad gauchista, que en la Francia de aquel momento bebía de la resaca de mayo del 68. [P2 S6] {P3 S8} Giscard d'Estaing venció por la mínima al socialista Mitterrand (que había pactado un programa común con el entonces poderoso Partido Comunista al mando del estalinista Georges Marchais) y trató de desplegar un programa reformista avanzado para superar la presidencia del fallecido Georges Pompidou. [P2 S7] {P3 S9} Con los años, el empuje del vanidoso Giscard d'Estaing (solía sentarse en una silla mas alta que sus interlocutores) se marchitó, y, en 1981, Mitterrand finalmente alcanzó el Eliseo, aunque es cierto también que en esta victoria de Mitterrand mucho tuvo que ver el actual presidente Chirac, con una oposición surgida a la derecha de Giscard d'Estaing. [P2 S8]</p>

Tabla 1. Ejemplo de análisis contrastivo para la detección de plagio lingüístico realizado con CopyCatch (“MarkUp by Sentence. File with File”). Fuente: Elaboración propia.

2.3. LOS ESTUDIOS DE AUTORÍA EN EL MARCO DE LA LINGÜÍSTICA FORENSE

De todas las aplicaciones de la Lingüística forense, la que más nos interesa para el objeto de esta investigación es la comparación de textos con fines de atribución de autoría. Aunque algunos autores han propuesto diferentes etiquetas en función de la tipología de los textos (sólo dubitados o dubitados e indubitados) y del número de candidatos a autor (ninguno, uno o varios), a día de hoy la etiqueta “atribución de autoría” se utiliza de manera genérica para referirnos a cualquier proceso mediante el cual se analizan las características lingüísticas de unos determinados textos para extraer conclusiones sobre su autoría [Zheng, Qin, Huang y Chen, 2003]⁵¹.

2.3.1. Características de las muestras textuales

El primer paso para llevar a cabo una comparación forense de textos es valorar el material objeto de estudio, principalmente en relación a su extensión y tipología, pues, aunque “no one has ever speculated about, let alone tested, how much text and of what kind would be required to allow one to successfully and exhaustively specify an individual's idiolect” [Coulthard, 1998: 141], parece lógico pensar que cuanto mayor longitud tengan las muestras y cuanto más rasgos genéricos compartan entre sí, más variables podremos analizar y más fiables serán los resultados alcanzados. De igual modo, debemos tener en cuenta su temática y datación.

⁵¹ Abbasi y Chen hablan de “identification task” cuando conocemos de antemano a los candidatos a autor y de “similarity detection” cuando no es así [2008]. Juola distingue entre “open-class problem” y “closed-class problem”, en función de si en el conjunto de sospechosos se encuentra con seguridad o no el autor real [2008]. Koppel, Schler y Argamon plantean tres posibles escenarios: “profiling” cuando disponemos de un texto dubitado pero no de posibles candidatos a autor, “authorship verification” cuando disponemos de un texto dubitado y de textos indubitados de un solo candidato y “needle-in-a-haystack” (es decir, “aguja en un pajar”) cuando disponemos de un texto dubitado y de textos indubitados de infinidad de posibles candidatos [2009]. Turell diferencia entre “determinación de autoría” (varios candidatos) y “atribución de autoría” (un único candidato) [2011]. Y Fernández Trillo aboga por añadir la “estilometría”, esto es, el análisis estadístico de los textos, como herramienta de gran valor en la tarea de atribución de autoría [2015: 8 y 21-22].

2.3.1.1. Extensión de las muestras

No hay una longitud mínima establecida y es bastante habitual que los lingüistas forenses deban enfrentarse a textos extremadamente breves, como *whatsapps*, *tuits*, *posts* o correos electrónicos que rara vez superan las 200 palabras [Coulthard, 2005b]. Ahora bien, es evidente que cuanto mayor sea la extensión del texto, mayor número de variables podremos analizar. Como hemos visto al explicar el caso de Jenny Nicholl, un análisis comparado de textos breves puede arrojar resultados concluyentes desde el punto de vista cualitativo, pero habrá ciertas variables lingüísticas, especialmente de tipo cuantitativo, que no podremos contemplar, ya que requieren de una porción mínima de texto. Aunque hay estudios que han alcanzado resultados prometedores con textos muy breves [Sanderson y Guenter, 2006; Luyckx y Daelemans, 2010], lo cierto es que la precisión de los análisis cuantitativos decrece significativamente cuando la longitud del texto es inferior a las 1000 palabras [Hirst y Feiguina, 2007]. También es importante tener en cuenta, como advierte Stamatatos, que la longitud, aunque es importante, no es el único requisito esencial:

It is not yet clear whether other factors (beyond text length) also affect this process. For example, let a and b be two texts of 100 words and 1,000 words, respectively. A given authorship attribution tool can easily identify the author of a, but not the author of b. What are the properties of a that make it an easy case, and what makes b so difficult, albeit much longer, than a? On the other hand, what are the minimum requirements in training text we need to be able to identify the author of a given text? [Stamatatos, 2009: 553].

2.3.1.2. Calidad de las muestras

La calidad de las muestras está relacionada con la *contaminación* textual o autorial que éstas hayan podido sufrir⁵².

En relación con la ‘contaminación textual’, es importante considerar si los textos nos han llegado completos o sesgados y si, además, la versión de que disponemos es

⁵² Tomamos este término de la ecdótica, en donde la *contaminatio* hace referencia a las posibles alteraciones que presenta un texto cuando su transmisor se ha basado en varias copias [Blecula, 1983; Pérez Priego, 1997].

CAPÍTULO 2

absolutamente fiable (como sería un texto autógrafo) o si, por el contrario, es una copia del original (como sería, por ejemplo, la transcripción de una grabación de voz o de las declaraciones policiales de un acusado) y susceptible, por tanto, de contener errores de transmisión textual.

En relación con la ‘contaminación autorial’, como advierte Love, es necesario distinguir lo que podríamos llamar la autoría puramente *ejecutiva* (en la que el texto nos llega en su versión más primigenia) de otro tipo de autorías, como la autoría *de revisión*, que se produce cuando el autor repasa el texto; la autoría *colaborativa*, cuando son terceros los que revisan y corrigen; la autoría *declarativa*, que se da cuando un tercero asume la autoría de un texto escrito por otro (es el caso de los famosos “negros literarios”); o la autoría *precursora*, en la que el autor reproduce de manera más o menos sistemática los textos de otros, ya sea en forma de cita textual o de parafraseo [2002: 32-50]. Para Love, “the attributionist who sees furthest is the one with the richest sense of the possibilities in the given case” [2002: 50].

2.3.1.3. Género textual y circunstancias comunicativas de las muestras

Otro requisito importante a la hora de enfrentarnos a una comparación de textos con fines de atribución de autoría —aunque no imprescindible— es el de disponer de muestras indubitadas del mismo género textual que las dubitadas y que, a ser posible, todas ellas compartan unas mismas circunstancias comunicativas. Es evidente que un mismo sujeto (emisor) no se expresará de igual modo si escribe un informe sobre el impacto económico de la COVID-19 en la hostelería o una poesía amorosa (mensaje), ni si el mensaje se lo dirige a su jefe o a su novia (receptor), ni si lo hace en un entorno laboral o familiar (contexto); tampoco se expresará de la misma manera en su lengua materna que en una adquirida (código), ni si lo hace a través de un correo electrónico o un tuit, en el que hay un número limitado de caracteres que fuerzan la escritura abreviada (canal).

En un caso idílico de comparación forense con fines de atribución de autoría, las muestras dubitadas y las indubitadas compartirán un mismo género textual, irán

dirigidas a una misma persona o tipo de persona, estarán escritas en la misma lengua, tendrán un mismo canal y habrán sido producidas en un contexto similar; pero en la práctica, pocas veces ocurre así y, a menudo, los peritos lingüistas debemos enfrentarnos a textos de diversa naturaleza. En estos casos, en base a nuestra experiencia y nuestro leal saber y entender, deberemos decidir la viabilidad de dicha comparación, teniendo en cuenta que “genre effects generally will supersede authorial features in the discrimination process” [Holmes, 1998: 111].

2.3.1.4. Tema de las muestras

Junto al género textual, la temática es otra de las variables a tener en cuenta a la hora de realizar un estudio de atribución de autoría. Los textos que comparten unos mismos temas van a compartir también ciertas analogías léxicas, que poco o nada tienen que ver con el idiolecto de sus autores, y que debemos tener en cuenta a la hora de discriminar autorías. Si analizamos dos noticias periodísticas sobre un mismo hecho (por ejemplo, la contaminación acústica en las ciudades) van a darse, inevitablemente, una serie de palabras coincidentes relacionadas con la propia materia de estudio (*ruido, decibelios, tráfico, volumen, etc.*), pero sólo si han sido escritas por la misma persona encontraremos, además, otras variables coincidentes (como la presencia compartida de una secuencia de 5 o más palabras iguales o de una falta de ortografía determinada).

2.3.1.5. Data de las muestras

Por último, es importante considerar la fecha de las muestras textuales, pues está demostrado que el idiolecto de una persona varía con el paso del tiempo, aunque no siempre al mismo ritmo, ni con la misma intensidad [Koppel, Schler y Argamon, 2009: 26]. En las obras de ficción narrativa de Miguel Delibes, por ejemplo, es fácil advertir diferencias estilísticas en función de la fecha de composición de sus obras, aunque también es cierto que, al menos en este autor, las diferencias genéricas se imponen sobre las temporales [Celma Valero y Ruiz Urbón, 2021].

Por todos estos motivos, como advierte Holmes, “with attributional problems, it is certainly wise to work within the same genre when faced with a list of candidate authors for a disputed work and to work within as close a time period as possible when selecting appropriate 'control' authors” [1998: 111]. No obstante, aunque la alternancia genérica o la evolución estilística de un autor pueda condicionar los resultados, es importante destacar que los estudios parecen demostrar que la variación intrapersonal siempre será menor que la variación interpersonal:

A pesar de que nuestros patrones lingüísticos cambian con el tiempo, debido a muchas circunstancias —el nivel cultural, el medio social en el que nos desenvolvemos, el tema sobre el que hablamos— los estudios forenses del habla y la escritura precisan que algo permanezca constante en nuestra producción verbal y escrita [Fernández Trillo, 2015: 49].

2.3.2. Metodología y tipología de los análisis

Como vimos en el apartado 2.1.1., uno de los principios esenciales de la Criminalística moderna es el de la certeza, que propugna que para llegar a conclusiones robustas debemos realizar análisis cualitativos y cuantitativos de las evidencias. Los análisis cualitativos, basados en la lectura minuciosa y la observación, ofrecen mucha información y muy valiosa, pero los datos que manejan son subjetivos y menos controlables que los cuantitativos. Por su parte, los análisis cuantitativos, basados en la medición, proporcionan datos objetivos y replicables⁵³, pero a veces insuficientes para dilucidar el complejo proceso de la autoría. Por eso, un buen estudio de atribución de autoría debe contemplar —siempre que sea posible— una comparación de textos que combine ambos tipos de análisis [Jeffries y McIntyre, 2010: 12].

Hasta hace relativamente poco tiempo, los estudios lingüísticos (y también los literarios) eran eminentemente cualitativos; de un tiempo a esta parte, sin embargo, se

⁵³ En relación con la objetividad y precisión de los análisis cuantitativos, Fernández Trillo advierte que, en la práctica, “no es exactamente así, pues la decisión crucial de qué rasgos lingüísticos van a ser empleados como elemento discriminador sigue siendo decisión del investigador, por lo que el elemento humano, con su carga inherente de subjetividad, continúa presente en proceso, aunque sea en menor medida” [2015: 51].

han ido adoptando técnicas cuantitativas para analizar el lenguaje y entender su variación en relación con factores como el periodo histórico, el contexto, el género o el autor [Savoy, 2020]. Así, gracias a la Estilometría, podemos realizar análisis estadísticos del estilo de un determinado autor, entendiendo por *estilo* el idiolecto escrito que caracteriza su forma de escribir en una situación concreta:

By style I mean the set of linguistic features that, taken together, uniquely identify a language user. The notion presupposes that there has been a choice—that someone has opted for Feature P rather than Feature Q (or R, S, ...). In this view, if we have no choice in the use of a feature, that feature cannot be a part of our style [Crystal, 2008: 16].

El uso de métodos estilométricos en los estudios de atribución de autoría es relativamente reciente. Como recuerdan Holmes [1998], Juola [2006] y Stamatatos [2009], los primeros intentos por cuantificar el estilo de escritura se remontan al siglo XIX, cuando el físico estadounidense Thomas Mendenhall midió la extensión de varios cientos de miles de palabras de las obras de Bacon, Marlowe y Shakespeare, tratando de demostrar que la longitud de palabra no es una variable discriminatoria entre autores. Ya en el siglo XX, George Under Yule en Inglaterra y George Kingsley Zipf en Estados Unidos intentaron establecer fórmulas matemáticas para discriminar autorías, basadas en la frecuencia léxica, como la “Característica K de Yule”, ineficaz según los estudios posteriores, y la “Ley Zipf”, que a día de hoy se sigue utilizando. Pero no fue hasta el trabajo de los estadísticos estadounidenses Mosteller y Wallace sobre la autoría de los *Papeles Federalistas*, del año 1964, cuando los análisis estadísticos o computacionales empezaron a ganar adeptos entre los lingüistas: su método, basado en el análisis estadístico bayesiano de la frecuencia de palabras funcionales, como “y” o “a”, ofreció resultados de discriminación significativos entre los distintos candidatos a autor. Otras investigaciones estilométricas posteriores, sin embargo, dieron lugar a múltiples controversias, como el estudio sobre la autoría de *El Libro del Mormón*, que generó un intenso debate entre los defensores de la autoría de Joseph Smith, su firmante, y los defensores de una autoría compartida por al menos cuatro personas; un debate que puso sobre la mesa la necesidad de testar algunos de los métodos propuestos. Desde finales de los años 80 del siglo XX, y más intensamente en el siglo XXI, los esfuerzos de la Estilometría han ido encaminados en esta dirección; así, se han ido descartando algunos

parámetros propuestos para discriminar autorías, como la “Característica K de Yule” (antes comentada), y validando otros nuevos, como la distribución de la frecuencia de las palabras más utilizadas (*Delta distance*), el análisis de componentes principales (PCA) de las palabras de función o el estudio de las clases de palabras mediante etiquetadores gramaticales automáticos (Pos tagging), entre otros [Savoy, 2020]⁵⁴. No obstante, y retomando lo que decíamos al principio de este epígrafe, “a traditional authorship study looking at all of the external and traditional internal evidence must be completed before a non-traditional study is undertaken”, pues “non-traditional attribution studies provide only a few of the tools for the attribution scholar. And these tools are by no means the most important” [Rudman, 1998: 359].

2.3.3. Variables de análisis

Las variables lingüísticas son las elecciones habituales e inconscientes que una persona realiza al escribir y que pueden, por tanto, constituirse como prueba o indicio para discriminar autorías. Como advierte McMenamin, hay dos tipos generales de variables: la elección de formas opcionales o la desviación de una norma [2017: 134]⁵⁵.

Hay infinidad de variables lingüísticas susceptibles de poder ser analizadas cualitativa y cuantitativamente. Rudman ya indicó hace unos años que se habían propuesto mas de mil variables diferentes para los estudios de atribución de autoría y que, desgraciadamente, no había consenso en el establecimiento de las más mismas, siendo los especialistas los encargados de seleccionar las más útiles para cada caso [1998: 360]⁵⁶; sin duda las preferidas, tal y como fueron caracterizadas por Labov, son

⁵⁴ La sigla MFW deriva del inglés “most frequent words”; PCA de “principal components analysis” y POS tagging de “Part-Of-Speech Tagger”.

⁵⁵ McMenamin ofrece los siguientes ejemplos: para la elección de formas opcionales, “te doy mi corazón / a ti te doy mi corazón / te doy a ti mi corazón / te doy mi corazón a ti”; y para la desviación de una norma, “ella quería que la acompañaras / ella quería que la acompañes” [2017: 134].

⁵⁶ Afortunadamente, la investigación en atribución de autoría ha avanzado mucho desde el estudio de Rudman y ya disponemos de varios estudios que revisan los distintos parámetros utilizados y dan

aquellas de “alta frecuencia, inmunes a la supresión consciente, codificables (y contables), y ampliamente distribuidas a lo largo de una cierta población (o individuo)” [McMenamin, 2017: 134]. Incidiendo en este aspecto, Garayzábal Heinze, Queralt Estévez y Reigosa Riveiros establecen las distintas características que debemos contemplar a la hora de seleccionar una variante en contextos forenses de atribución de autoría [2019: 62 y 63]:

- (i) que sea muy frecuente y esté estratificada;
- (ii) que muestre alta variabilidad interescritor y baja variabilidad intraescritor;
- (iii) que sea relativamente fácil de extraer y calcular;
- (iv) que pueda ser intercambiable en algunos contextos;
- (v) que sea lo más independiente posible de otras variables;
- (vi) que su producción sea lo más inconsciente posible;
- (vii) que esté definida lingüísticamente de manera clara y precisa;
- (viii) y que tengan en cuenta la inestabilidad del rasgo.

Las variables lingüísticas se dan en todos los niveles de la lengua, desde el fonético-fonológico hasta el pragmático-discursivo. Aunque en un texto escrito el nivel fonético-fonológico, a priori, no es relevante, deberemos tenerlo en cuenta a la hora de estudiar aspectos relacionados con el acento o la rima (en el caso de textos en verso) o con ciertas figuras retóricas vinculadas a los sonidos y la musicalidad, como la aliteración, la onomatopeya, la similitudencia o la paranomasia. Pero, sin duda, las variables que debemos contemplar en el nivel más bajo del texto escrito serán todas

cuenta de la efectividad de cada uno de ellos [Juola, 2006; Grieve, 2007; Koppel, Schler y Argamon, 2009; Stamatatos, 2009; Luyckx, 2010; Jockers y Witten, 2010; Savoy, 2020]. Algo similar hicimos hace unos años para la lengua española [Blasco y Ruiz Urbón, 2009].

CAPÍTULO 2

aquellas relacionadas con el empleo de grafemas y caracteres⁵⁷ y, en especial, con el uso de n-gramas de caracteres. Grieve defiende la efectividad de los n-gramas de dos caracteres, con un 65% de precisión cuando deben discriminar entre 40 autores y de un 94% cuando deben hacerlo entre dos [2007]; no obstante, para el español, y tal y como estudiamos hace unos años, la precisión aumenta progresivamente a medida que lo hace también el número *n* del n-grama [Blasco y Ruiz Urbón, 2009].

Muy interesantes resultan también las variables morfológicas, como, por ejemplo, la alternancia del morfema de tiempo-modo-aspecto en el pretérito imperfecto de subjuntivo (*-ra / -se*), el estudio de los mecanismos de formación de palabras o la frecuencia relativa de las distintas categorías gramaticales (una variable muy efectiva y que ahora, gracias a los programas de etiquetado automático, ha cobrado una especial dimensión).

En los estudios de comparación de textos con fines de atribución de autoría también deberemos tener en cuenta las variables relacionadas con el nivel sintáctico, como el estudio del tipo de oración, la utilización de bigramas y trigramas de clases de palabras, el uso de ciertas construcciones sintácticas, la posición relativa de una palabras con respecto a otras o los pronombres utilizados para introducir las oraciones subordinadas [Svartvik, 1968; Baayen, Van Halteren y Tweedie, 1996; Hirst y Feiguina, 2007; Spassova, 2009; Turell, 2011; Stamatatos, Fakotakis y Kokkinakis, 2001].

Pero sin duda, uno de los niveles que más páginas de investigación ha ocupado hasta ahora es el del léxico⁵⁸. Dentro de las variables léxicas, se analizan las palabras más frecuentes, la riqueza de vocabulario, las erratas ortográficas, el porcentaje de ciertos radicales o la frecuencia relativa de las palabras funcionales, entre otras; en este

⁵⁷ Tomando como referencia a Grieve [2007], usamos *grafema* para referirnos a una letra del alfabeto y *carácter* para cualquier tipo de unidad gráfica, ya sea una letra, un dígito, un signo de puntuación o un espacio en blanco.

⁵⁸ Durante mucho tiempo, las marcas léxicas han sido las más utilizadas en los análisis cuantitativos. De hecho, Tallentire llegó a afirmar en 1973 que “no potential parameter of style below or above that of the word is equally effective in establishing objective comparison between authors and their common linguistic heritage” [Citamos por Holmes, 1998: 112].

nivel también han resultado ciertamente interesantes los nuevos métodos estadísticos, como el análisis clúster, el análisis de componentes principales (PCA) o el análisis multivariante [Mosteller y Wallace, 1964; Wools y Coulthard, 1998; Holmes, 1985; Holmes y Forsyth, 1995; Burrows, 1989, 2002 y 2003; Hoover, 2004a y 2004b; Argamon y Levitan, 2005; Statamatos, 2009; Savoy, 2020]. Aunque todos estos análisis poseen gran fuerza discriminante, las palabras de función se han constituido en la marca preferida por muchos investigadores:

Este hecho se debe a tres principales características propias de este tipo de marcas: en primer lugar, a su alta frecuencia de uso, que facilita su cuantificación y posterior análisis estadístico; en segundo lugar, a su papel gramatical, que restringe el control consciente en su producción por parte del sujeto-autor, lo cual presupone un indicio claro de su aptitud como marcas identificativas. Además, el contenido semántico de las palabras funcionales está limitado a la relación gramatical o a la propiedad genérica que designan y no suelen estar ligadas al contexto concreto, lo que es indicio claro de su pertenencia al grupo de rasgos de estilo autorial [Spasova, 2009: 45].

De un tiempo a esta parte están cobrando también cierto protagonismo las variables semánticas relacionadas con la expresión de ciertos sentimientos, como la obligación, la condición, la alegría, la tristeza... Y también las variables pragmático-discursivas, como las fórmulas de salutación y despedida, los marcadores discursivos, la puntuación o el número de palabras por frase o párrafo [Chaski, 1996; Mannion y Dixon, 1997; Gains, 1999; Wright, 2013].

En todo caso, es importante advertir, que no hay una variable única y definitiva, sino que lo importante es la acumulación de rasgos:

We must strive to identify all of the markers that make up “style” - to map style the way biologists are mapping the gene. Function words, type/tokens, word lengths, hapax legomena, and other specific style markers may not in themselves be an indicator of a unique style, but when used in conjunction with all of the other quantifiable indicators that make up style, they become important [Rudman, 1998: 360].

2.3.4. Herramientas de análisis

Para analizar las distintas variables lingüísticas objeto de estudio, podemos ayudarnos de herramientas informáticas que permitan realizar análisis automáticos o semi-automáticos de los textos.

A día de hoy disponemos de muchas y muy variadas herramientas de análisis textual [Woolls y Coulthard, 1998; Woolls, 2003; McEnery y Hardie, 2012; Anthony, 2013; Garayzabal Heinze, Queralt Estévez y Reigosa Riveiros, 2019: 65-66]. En nuestra opinión, las más útiles e interesantes para la atribución de autoría son las siguientes⁵⁹:

- *WordSmith Tools*.- Esta herramienta de análisis textual, diseñada por Mike Scott (Universidad de Liverpool), permite explorar grandes conjuntos de textos mediante búsquedas de carácter contextual o estadístico. Ofrece, por ejemplo, análisis de concordancias (*Concord*), de frecuencia léxica (*WordList*) y de colocaciones (*KeyWords*). Disponible en <https://www.lexically.net/wordsmith/>. Es de pago.
- *AntLab*.- Está compuesto por varios software de análisis textual desarrollados por Laurence Anthony (Universidad de Waseda). Tiene funcionalidades similares a las de WordSmith Tools, con la ventaja de que son gratuitas. Permite detectar concordancias (*AntCont*), analizar secuencias de palabras (*AntGram*), medir el nivel de complejidad léxica (*AntWordProfiles*), identificar clases de palabras (*TagAnt*), analizar la estructura de los textos (*AntMover*), localizar variantes de una palabra (*VariAnt*), etc. Disponible en <http://www.laurenceanthony.net/software.html>.
- *Contawords*.- Es una herramienta en línea y de acceso gratuito creada por el Institut Universitari de Lingüística Aplicada (IULA) la Universitat Pompeu Fabra, en el marco de la iniciativa europea CLARIN, destinada a la creación de Infraestructuras para la investigación en Humanidades y Ciencias Sociales. Este software permite conocer la frecuencia de las palabras, los lemas o las etiquetas gramaticales de un texto (*types*, *tokens*, POS...), distinguir clases de palabras, reconocer “entidades nombradas” y combinaciones de palabras, etc. Disponible en <http://contawords.iula.upf.edu/>.

⁵⁹ Para acceder a un repertorio más amplio, véase la página web “Estilometría”, dirigida por el profesor Javier Blasco (Disponible en <https://www.estilometria.com/herramienta/>).

- *Voyant Tools*.- Es una herramienta en línea, de código abierto y gratuita, para analizar y visualizar textos, desarrollada por Stéfan Sinclair (Universidad McGill) y Geoffrey Rockwell (Universidad de Alberta). Ofrece diferentes posibilidades de análisis, como la creación de nubes de palabras o la frecuencia de un término a lo largo de la diacronía de un texto. Disponible en <https://voyant-tools.org/>.
- *Linguistic Inquiry and Word Count (LIWC)*.- Es un programa, de pago, desarrollado por James W. Pennebaker, Roger J. Booth, y Martha E. Francis (Universidad de Texas en Austin y Universidad de Auckland), que permite analizar varios de los componentes emocionales, cognitivos, estructurales y de procesamiento presentes en las muestras textuales. La versión LIWC2001 incluye diccionario en español. Disponible en <http://www.liwc.net/liwcspanol/>.
- *CopyCatch Gold*.- Esta herramienta de análisis textual, diseñada por David Woolls (CFL Software Development / Universidad de Birmingham), permite realizar diferentes análisis contrastivos entre dos textos dados (umbral de similitud, vocabulario compartido, vocabulario compartido una sola vez, vocabulario exclusivo, frases coincidentes, etc.) [Woolls y Coulthard, 1998; Woolls, 2003 y 2005]. Es un software de pago. Aunque suele aplicarse a casos de detección de plagio, también ofrece posibilidades muy interesantes para la atribución de autoría. Disponible en <http://glotta.ntua.gr/IS-Social/CopyRight/CopyCatch%20Versions%20Review.htm>.
- *Vocalyse Toolkit*.- Este software, también diseñado por David Woolls (CFL Software Development / Universidad de Birmingham), permite calcular la proporción de palabras de contenido y de función, los *hapax legomena* y *dislegomena*, la densidad léxica (*type/token ratio*), la característica K de Yule y otras medidas relacionadas con la riqueza léxica [Woolls y Coulthard, 1998; Woolls, 2003 y 2005]. Es de pago. Disponible en <http://glotta.ntua.gr/IS-Social/CopyRight/Welcome%20to%20the%20home%20of%20powerful%20text%20analysis%20tools.htm>.

- *Java Graphical Authorship Attribution Program (JGAAP)*.- Es una herramienta de código abierto, desarrollada por el equipo de Patrick Juola (Universidad de Duquesne), que trabaja con técnicas de aprendizaje automático. Se hizo bastante popular hace unos años, ya que gracias a él se descubrió que la autora de la novela *El canto del cuco*, publicada bajo el pseudónimo de Robert Galbraith, era en realidad J. K. Rowling [Juola, 2013 y 2015]. Funciona fácilmente con Java y no requiere conocimientos de codificación. Disponible en <http://evllabs.github.io/JGAAP/>.
- *Python*.- Python, que debe su nombre al grupo de humoristas *Monty Python*, es un lenguaje de programación de código abierto, creado por el holandés Guido van Rossum en 1989 [Karsdorp, Kestemont y Riddell, 2021]. Ha servido de base a la creación de programas mundialmente conocidos, como Instagram, Dropbox, Youtube o Pinterest. Python permite trabajar Delta a través de PyDelta [Jannidis, Pielström, Schöch y Vitt, 2015]. Disponible en <https://github.com/fotisj/pydelta>⁶⁰.
- *R y RStudio*.- R es un lenguaje de programación para el análisis estadístico y gráfico de datos, creado por Ross Ihaka y Robert Gentleman (Universidad de Auckland), que gracias a él obtuvieron la Medalla de Pickering, otorgada por la Sociedad Real de Nueva Zelanda (disponible en <https://cran.r-project.org/>). Para poder trabajar con R de manera mucho más gratificante, recomendamos la instalación de RStudio, un entorno de desarrollo integrado para este lenguaje de programación, creado por Joseph J. Allaire, que incluye una consola, un editor de sintaxis que apoya la ejecución de código, así como herramientas para el trazado, la depuración y la gestión del espacio de trabajo (disponible en <https://rstudio.com/products/rstudio/>). RStudio permite trabajar con el paquete Stylo, desarrollado en el año 2013 por Eder, Rybicki y Kestemont [2016].

⁶⁰ Nosotros, en lugar de utilizar PyDelta, trabajaremos el algoritmo Delta con el paquete Stylo de RStudio.

- Etiquetadores gramaticales automáticos (*Treetagger*, *CoreNLP*, *Udpipe* y *FreeLing*).- Hay varios paquetes que permiten etiquetar morfológicamente un texto dado, como *koRpus*, que depende de *Treetagger* (disponible en <https://cran.r-project.org/web/packages/koRpus/index.html>); *coreNLP*, que depende de la librería Java “Stanford CoreNLP” (disponible en <https://cran.r-project.org/web/packages/coreNLP/index.html>); *udpide* (disponible en <https://cran.r-project.org/web/packages/udpipe/index.html>) o *FreeLing* (disponible en <http://nlp.lsi.upc.edu/freeling/index.php/node/1>). Todos ellos pueden desarrollarse dentro del entorno de R (también *FreeLing*, aunque el proceso de instalación sea algo más complejo).
- *Statistical Package for the Social Sciences (SPSS)*.- Es una de las herramientas de análisis estadístico de datos más utilizada en las Ciencias Sociales y las Humanidades. Fue creada en 1968 por Norman H. Nie, C. Hadlai Hull y Dale H. Bent, tres estudiantes de posgrado de la Universidad de Stanford, y desde el año 2009 está en manos de IBM (*IBM SPSS Statistics*). Es de pago. Disponible en <https://www.ibm.com/analytics/spss-statistics-software>.
- Corpus de referencia o respaldo.- La Real Academia Española ofrece acceso a un Corpus Diacrónico del Español (CORDE, disponible en <http://corpus.rae.es/cordenet.html>) y a un Corpus de Referencia del Español Actual (CREA, disponible en <http://corpus.rae.es/creanet.html>), ya cerrados, y también a un Corpus del Español del Siglo XXI (CORPES XXI, disponible en <https://webf1.rae.es/CORPES>), aún en fase de desarrollo⁶¹. También resultan de interés el Corpus del Español (<https://www.corpusdelespanol.org/>) y Sketch Engine (<https://www.sketchengine.eu/>).

⁶¹ El CORDE permite el acceso a cerca de 300 millones de formas, desde los orígenes del idioma hasta 1974; el CREA, a algo más de 160 millones de formas, desde 1975 hasta 2004; y la última versión del CORPES, de julio de 2021, a 350 millones de formas, desde 2001 a 2021.

2.3.5. Expresión de los resultados: escalas de opinión verbal vs. escalas de opinión estadística

Las conclusiones de los informes periciales en Lingüística forense suelen expresarse en escalas de opinión, también llamadas escalas de probabilidad, para cumplir así con el sexto principio de la Criminalística moderna, según el cual —recordemos— los investigadores pueden inferir probabilidades en relación al objeto de su análisis, que van desde la probabilidad nula al grado máximo de probabilidad.

Como en atribución de autoría se trabaja con metodologías combinadas, puede resultar ciertamente complejo traducir todos esos resultados a escalas de opinión. Las primeras escalas adoptadas por los lingüistas forenses europeos —basándose en escalas similares de la American Society for Testing and Materials (ASTM) o del Scientific Working Group for Forensic Document Examination (SWGDOC)— fueron estrictamente verbales, tal y como puede verse en los ejemplos que proponen Coulthard (Tabla 2) y McMenamín (Tablas 3 y 4):

Más positivo	
5	“Estoy personalmente <i>convencido</i> de que X es el autor”
4	“En mi opinión, es <i>muy probable</i> que X sea el autor”
3	“En mi opinión, es <i>probable</i> que X sea el autor”
2	“En mi opinión, es <i>bastante probable</i> que X sea el autor”
1	“En mi opinión, es <i>algo probable</i> que X sea el autor”
0	“En mi opinión, es <i>posible</i> que X sea el autor”
-1	“En mi opinión, es <i>algo probable</i> que X no sea el autor”
-2	“En mi opinión, es <i>bastante probable</i> que X no sea el autor”
-3	“En mi opinión, es <i>probable</i> que X no sea el autor”
-4	“En mi opinión, es <i>muy probable</i> que X no sea el autor”
-5	“Estoy personalmente <i>convencido</i> de que X no es el autor”
Más negativo	

Tabla 2. Traducción de la escala propuesta por Coulthard para el caso de Danielle Jones.
Fuente: Coulthard, Johnson y Wright, 2017: 197 y 198.

¿Este texto lo escribió el sospechoso?	Escala	Conclusión
SÍ	9	Identificación
	8	Altamente probable: lo escribió
	7	Probable: lo escribió
	6	Indicios: lo escribió
INCONCLUSO	5	No concluyente
NO	4	Indicios: no lo escribió
	3	Probable: no lo escribió
	2	Altamente probable: no lo escribió
	1	Eliminación

Tabla 3. Traducción de la escala propuesta por McMenamin para expresar las conclusiones relativas a la similitud entre el texto dubitado y los textos indubitados. Fuente: McMenamin, 2002: 125

¿Estos textos los escribió un único autor?	Escala	Conclusión
SÍ	9	Definitivo: un escritor
	8	Altamente probable: un escritor
	7	Probable: un escritor
	6	Indicios: un escritor
INCONCLUSO	5	No concluyente
NO	4	Indicios: más de un escritor
	3	Probable: más de un escritor
	2	Altamente más de un escritor
	1	Definitivo: más de un escritor

Tabla 4. Traducción de la escala propuesta por McMenamin para expresar las conclusiones relativas a la similitud a entre varios textos dubitados. Fuente: McMenamin, 2002: 125

Las escalas de opinión verbal, como puede verse en los ejemplos propuestos, recogen los distintos niveles de probabilidad dentro de los cuales debe enmarcarse el resultado global de nuestro análisis comparado de textos. Las escalas se han ido perfeccionando con el tiempo, pues si la de Coulthard no contemplaba la posibilidad de que el análisis condujera a conclusiones no concluyentes, las de McMenamin, sí. Mcmenamin, además, ofrece un listado de los criterios que permiten determinar los distintos niveles de conclusión, tratando así de objetivar al máximo las escalas propuestas (Tabla 5):

CAPÍTULO 2

DUBITADO Vs. INDUBITADO	CRITERIOS	DUBITADO Vs. DUBITADO
9 IDENTIFICACIÓN (lo escribió)	<ol style="list-style-type: none"> 1. Hay similitudes significativas sustanciales. 2. No hay disimilitudes significativas. 3. No hay limitaciones asociadas a la no ocurrencia de variables, diferencias, cantidad de escritura. 	9 DEFINITIVO (un escritor)
8 ALTAMENTE PROBABLE (lo escribió)	<ol style="list-style-type: none"> 1. Hay similitudes significativas sustanciales. 2. No hay disimilitudes significativas. 3. Hay limitaciones asociadas a la no ocurrencia de variables, diferencias, cantidad de escritura. 	8 ALTAMENTE PROBABLE (un escritor)
7 PROBABLE (lo escribió)	<ol style="list-style-type: none"> 1. Hay algunas similitudes significativas. 2. No hay disimilitudes significativas. 3. Hay limitaciones asociadas a la no ocurrencia de variables, diferencias, características individualizadoras, cantidad de escritura. 	7 PROBABLE (un escritor)
6 INDICIOS (lo escribió)	<ol style="list-style-type: none"> 1. Hay pocas similitudes significativas. 2. No hay disimilitudes significativas. 3. Puede haber limitaciones asociadas a la no ocurrencia de variables, diferencias, características individualizadoras, cantidad de escritura. 	6 INDICIOS (un escritor)
5 INCONCLUSO (No concluyente)	<ol style="list-style-type: none"> 1. Las similitudes significativas son insuficientes 2. Las disimilitudes significativas son insuficientes 3. Puede haber limitaciones asociadas a la no ocurrencia de variables, características individualizadoras, cantidad de escritura. 4. Puede haber similitudes y disimilitudes. 	5 INCONCLUSO (No concluyente)
4 INDICIOS (no lo escribió)	<ol style="list-style-type: none"> 1. Hay pocas diferencias significativas. 2. Puede haber limitaciones asociadas a la no ocurrencia de variables, características individualizadoras, cantidad de escritura. 3. Puede haber similitudes. 	4 INDICIOS (más de un escritor)
3 PROBABLE (no lo escribió)	<ol style="list-style-type: none"> 1. Hay algunas disimilitudes significativas. 2. Puede haber limitaciones asociadas a la no ocurrencia de variables, características individualizadoras, cantidad de escritura. 3. Puede haber similitudes. 	3 PROBABLE (más de un escritor)
2 ALTAMENTE PROBABLE (no lo escribió)	<ol style="list-style-type: none"> 1. Hay disimilitudes significativas sustanciales 2. Hay limitaciones asociadas a la no ocurrencia de variables, características individualizadoras, cantidad de escritura. 3. Puede haber similitudes. 	2 ALTAMENTE PROBABLE (más de un escritor)
1 ELIMINACIÓN (no lo escribió)	<ol style="list-style-type: none"> 1. Hay disimilitudes significativas sustanciales 2. No hay limitaciones asociadas a la no ocurrencia de variables, características individualizadoras, cantidad de escritura. 3. Puede haber variables no recurrentes. 4. Puede haber similitudes. 	1 DEFINITIVO (más de un escritor)

Tabla 5. Traducción de los criterios propuestos por McMenamín. Fuente: McMenamín, 2002: 126-127.

Aunque las escalas de probabilidad verbal son la opción más utilizada por los especialistas en Lingüística forense para expresar sus conclusiones, no están exentas de crítica; sobre todo desde que el Tribunal Supremo de Estados Unidos estableciera en 1993 unos estándares para determinar si los métodos científicos utilizados por los peritos eran fiables (entre los que incluía la necesidad de contemplar una tasa de error real o probable sobre la técnica aplicada). En 1999, Broeders denunció el carácter subjetivo de las propuestas verbales al considerar que dos expertos podrían establecer diferentes niveles de la escala basándose en unos mismos datos [1999: 237 y ss.]; algo que también fue comentado por Champod y Evett [2000], que defendían un enfoque bayesiano. Pero precisamente Evett, junto a Weir, había propuesto unos años antes [1998] una correspondencia verbal para la escala de razón de verosimilitud utilizada en la identificación de personas mediante el ADN (Tabla 6), poniendo así de manifiesto la importancia de emplear escalas verbales que ayuden a comprender los valores de LR (*Likelihood Ratio*):

Razón de verosimilitud (LR)	Correspondencia verbal
De 1 a 100	Respaldo limitado
De 10 a 100	Respaldo moderado
De 100 a 1000	Respaldo fuerte
Más de 1000	Respaldo muy fuerte

Tabla 6. Traducción de la escala propuesta por Evett y Weir. Fuente: Evett y Weir, 1998: 244.

En relación con este asunto, Delgado Romero, Jefe del Laboratorio de Acústica Forense de la Comisaría General de la Policía Científica, comentó las ventajas de utilizar el teorema de Bayes para establecer la relación de probabilidades sobre dos hipótesis competitivas y que, a su vez, son excluyentes entre sí (“el sospechoso ha realizado la voz dubitada vs. la voz dubitada no ha sido realizada por el sospechoso”), pero también las dificultades de hacerlo en el entorno lingüístico [2005: 122-127]⁶². En

⁶² Aunque Delgado Romero se centró en la comparación forense de voces, sus opiniones son perfectamente extrapolables a la atribución de autoría.

CAPÍTULO 2

el caso del ADN, cada genotipo suele comprender 15 marcadores genéticos y a cada uno de ellos se le otorga un valor de verosimilitud, y, puesto que cada marcador es independiente del resto, el valor final del LR para cada individuo (producto de las probabilidades que definen el LR de cada uno de los 15 marcadores) puede llegar a expresarse en términos de billones o trillones; pero en el caso de la voz —y añadimos nosotros, más aún en el del texto—, el cálculo del LR está claramente influenciado por las características de las muestras que, a diferencia de las del ADN, son fractales y variables, lo que dificulta, además, la determinación de los ejes que dimensionan la población de referencia. Por este motivo, Delgado Romero sigue apoyando el “uso de las denominadas escalas de probabilidad verbal”, pues la propuesta de interpretación del teorema de Bayes, aunque interesante, “no acaba de ofrecer el grado de satisfacción que sería deseable”; y concluye:

En el caso de la Identificación Forense de Locutores (IFL), la interpretación de conclusiones de estudio a través de un entorno bayesiano resulta difícil de concretar. Existen aspectos, no sólo referidos a la población de referencia utilizada, sino también a la naturaleza de la variable del habla de un mismo individuo, que aparecen como serios obstáculos al carácter objetivo que dicho ámbito pretende conferir a las tareas asignadas a los científicos. Por otra parte, y haciéndolo ya extensivo a otras disciplinas forenses, es complicado evitar el uso de equivalencias verbales para expresar de una forma más comprensible los niveles de certeza alcanzados por los expertos. La intervención humana, tanto en los procesos de selección y análisis de muestras, como en el de traducción de resultados, aunque subjetiva, se antoja ineludible en la actualidad [2005: 128].

En la misma línea se manifestaron un grupo de fonetistas forenses del Reino Unido [French y Harrison, 2007], que si bien aceptaban la conveniencia de aplicar proporciones de verosimilitud —incluidas las bayesianas— para la expresión de las conclusiones en grados de convicción, advertían que la falta de datos poblacionales de referencia lo imposibilitaba. Una declaración que dio lugar a otras manifestaciones a favor, pero también en contra, como la de Rose y Morrison [2009], que consideraban que el marco conceptual apropiado es el de la razón de verosimilitud bayesiana, que es en el que se están desarrollando otras ciencias forenses. Como explican estos dos autores, un proceso judicial consiste en tomar decisiones sobre incertidumbres, y la mejor manera de cuantificar dichas incertidumbres es utilizar el teorema de Bayes, que permite calcular la probabilidad de una hipótesis dada la evidencia:

Bayes' Theorem is of paramount importance when one wants to know the probability of a hypothesis given the evidence, and this is what gives it its special status in forensic identification. Bayes' Theorem states informally that the probability of the hypothesis, given the evidence, can be estimated from two things: firstly, how probable the hypothesis is, before the evidence is adduced; and secondly, the strength of the evidence [2009: 144].

Aunque el debate teórico sigue abierto, lo cierto es que, en la práctica, la opción más utilizada por los lingüistas forenses de todo mundo sigue siendo la de la escala de probabilidad verbal, ya que la falta de estudios poblacionales, indispensables para el cálculo de la razón de verosimilitud, hace aún muy difícil la aplicación de escalas bayesianas [Gold y French, 2011: 6; Coulthard, Johnson y Wright, 2017: 213 y ss.].

2.4. APLICACIÓN DE LOS PRINCIPIOS TEÓRICO-METODOLÓGICOS DE LA LINGÜÍSTICA FORENSE A LOS ESTUDIOS DE ATRIBUCIÓN DE AUTORÍA LITERARIA. ESTADO DE LA CUESTIÓN Y PERSPECTIVAS DE FUTURO

El futuro de la Lingüística forense parece realmente prometedor [Shuy, 2000]. Cada vez son más las universidades de todo el mundo que imparten formación especializada en esta materia⁶³ y los trabajos de investigación relacionados con el lenguaje jurídico-administrativo, el lenguaje del procedimiento judicial o el lenguaje evidencial se han incrementado de manera exponencial en los últimos años, gracias principalmente a la plataforma que ofrece *The International Journal of Speech, Language and the Law*, pero también a los eventos de difusión científica organizados por la *International Association for Forensic and Legal Linguistics* (IAFLL), *The International Association for Forensic Phonetics and Acoustics* (IAFPA) y otras asociaciones y grupos de investigación nacionales e internacionales. Más allá del ámbito

⁶³ A día de hoy podemos encontrar másteres especializados en Lingüística forense en las universidades de Aston y Cardiff en Reino Unido, Hofstra y California of Pennsylvania en Estados Unidos, Umea en Suecia o VU Amsterdam en Países Bajos, si bien ya no hay ninguno en España (ya que el que ofrecía la Universitat Pompeu Fabra, y que tuvimos la oportunidad de cursar, desapareció hace unos años), ni en ningún país de habla hispana.

CAPÍTULO 2

académico o científico, la presencia de la Lingüística forense se ha ido viendo reforzada en el entorno de los Cuerpos y Fuerzas de Seguridad del Estado y en el de la Justicia, y cada vez son más los policías, jueces, abogados o fiscales que precisan de la colaboración de expertos lingüistas para resolver, defender o juzgar casos en los que, de un modo u otro, interviene el lenguaje. Además, de un tiempo a esta parte, los medios de comunicación se han ido interesado por esta disciplina, lo que ha contribuido a su difusión y equiparación con otras especialidades forenses más arraigadas en la sociedad, como la Medicina, la Grafología o la Psicología⁶⁴.

Pero si la Lingüística forense ha dado pasos de gigante en el establecimiento de unos métodos atributivos eficaces y fiables que permitan identificar al autor de un determinado texto, los estudios de autoría literaria no parecen haber avanzado al mismo ritmo, a pesar de que su preocupación por el anonimato sea la misma. Aunque los estudios de atribución son tan antiguos como la Literatura misma, siempre han sido vistos con cierto recelo en el ámbito académico. Es cierto que aún son muchos los que consideran que las evidencias externas son la única forma fiable de atribuir un texto a un determinado autor, pero también que cada vez son más los que se convencen de la fuerza discriminante de la evidencia lingüística, gracias, principalmente, a la solvencia que están demostrando los análisis estilométricos. Sin embargo, y a pesar de estos importantes avances, creemos necesario dotar a los estudios de atribución de autoría literaria de todos los principios que rigen las ciencias forenses, pues sólo así podrán alcanzar un mayor reconocimiento y prestigio.

El protocolo de actuación de un estudio de atribución de autoría debe ser siempre el mismo, con independencia de si el contexto en el que se circunscribe es forense o

⁶⁴ A pesar de la amplia expansión que ha vivido la Lingüística forense en los últimos veinte o treinta años, no podemos obviar que aún quedan muchos problemas por resolver. Algunos, como el del establecimiento de un método único o el de la conveniencia de avanzar en la expresión de las conclusiones en escalas bayesianas ya los hemos comentado antes; pero hay más. Tal vez el más preocupante, en nuestra opinión, sea el del intrusismo profesional, derivado de la falta de reconocimiento y regulación de la figura del perito lingüista forense en nuestro país (a día de hoy no existe ningún Colegio Nacional de Lingüistas y la figura del perito judicial experto en Lingüística forense sólo está regulada en Cataluña y Valencia).

literario. Por este motivo, creemos conveniente adoptar los principios de la comparación forense de textos al ámbito de la investigación literaria, haciendo especial hincapié en los principios de certeza y probabilidad (véase el apartado 2.1.1):

- El principio de certeza respalda el análisis cualitativo, cuantitativo y comparado de las evidencias. Tradicionalmente, los estudios de atribución de autoría literaria han prestado atención a los análisis cualitativos, pero, de un tiempo a esta parte, y gracias al avance de las nuevas tecnologías, han pasado a adoptar un enfoque eminentemente cuantitativo. Lo que proponemos, siguiendo este principio de la Criminalística moderna, es que estos análisis no sean excluyentes sino complementarios, siempre y cuando el objeto de estudio lo posibilite.
- De igual modo, consideramos necesario incorporar el principio de probabilidad. Los estudios de atribución de autoría literaria en el ámbito hispánico ponen mucho interés en la metodología de estudio, pero poco o muy poco en la expresión de las conclusiones. Estamos acostumbrados a leer que Fulanito “puede ser el autor” o que Menganito “muy probablemente” no lo sea, pero esa expresión semántica no aparece respaldada por una escala de opinión que permita cuantificar el nivel real de la misma. Por este motivo, proponemos que los estudios de atribución de autoría literaria incluyan escalas de opinión que permitan encuadrar las conclusiones en grados de certeza, desde la probabilidad nula al grado máximo de probabilidad de cada una de las hipótesis posibles (SÍ / NO).

CAPÍTULO 3

CERVANTES, UN HOMBRE DE TEATRO

“Sólo un hombre de teatro pudo haber escrito el *Quijote*” (José Martínez Ruiz, Azorín).

3.1. CERVANTES Y EL TEATRO

El teatro de Cervantes ha sido mirado siempre con cierto recelo. Ya en época del autor, como él mismo confiesa en el prólogo de sus *Ocho comedias y ocho entremeses nuevos*, había quienes afirmaban que de su prosa se podía esperar mucho, pero que del verso, nada. Esa misma idea de negatividad ha perdurado a lo largo de los siglos y los juicios de Bernardo de Iriarte, Pedro Estala, Leandro Fernández de Moratín o Alberto Lista son una buena muestra de ello. No ha sido hasta el siglo XX, gracias a los trabajos de Cotarelo Valledor, Casaldueiro, Marrast, Canavaggio o Zimic, cuando el teatro de Cervantes ha empezado a recibir mayor atención (aunque no siempre mejor opinión) por parte de la crítica; atención que se ha ido intensificando desde entonces hasta ahora. Así, mientras unos defienden “que la garantía mayor de su éxito [como dramaturgo] es la actual penuria de comedias conservadas” [Granja, 1995: 232], otros no dudan en afirmar que Cervantes fue “un fracasado en las tablas, que no quiso, no supo o no pudo subirse al carro de la comedia nueva, que puso en circulación Lope de Vega” [Díez Borque, 2016: 11]. Sea como fuere, lo que no se puede negar es la importancia que el teatro como tal tuvo en la vida y la obra de Cervantes. Aficionado desde niño a las tablas, fue un ávido lector de comedias y un hombre que supo moverse y relacionarse en el

ambiente farandulesco de la época. Así lo demuestran sus obras, toda una lección del teatro del momento, con referencias constantes a compañías teatrales y a *autores* y comediantes de su tiempo [Arco y Garay, 1951]; sin olvidar, por supuesto, el recurso del “metateatro”, al que Cervantes recurre en infinidad de ocasiones para hablar del teatro desde dentro del teatro. El teatro, en el sentido más amplio del término, es una constante en la obra cervantina y, sin duda, uno de sus grandes sellos de identidad.

3.1.1. Un apasionado espectador teatral

El interés de Cervantes por el mundillo teatral surgió a edad muy temprana. En el undécimo capítulo de la segunda parte del *Quijote*, recuerda, por boca de su personaje, que desde muchacho fue “aficionado a la carátula” y que en su mocedad se le “iban los ojos tras la farándula”; y en el prólogo de sus *Ocho comedias y ocho entremeses*, se enorgullece de haber visto representar a Lope de Rueda, cuando las comedias “eran unos coloquios, como églogas, entre dos o tres pastores y alguna pastora”, aderezadas con “con dos o tres entremeses, ya de negra, ya de rufián, ya de bobo y ya de vizcaíno”⁶⁵:

Los días pasados me hallé en una conversación de amigos, donde se trató de comedias y de las cosas a ellas concernientes [...] Tratóse también de quién fue el primero que en España las sacó de mantillas, y las puso en toldo y vistió de gala y apariencia; yo, como el más viejo que allí estaba, dije que me acordaba de haber visto representar al gran Lope de Rueda [...] ni entonces ni después acá ninguno le ha llevado ventaja; y, aunque por ser muchacho yo entonces, no podía hacer juicio firme de la bondad de sus versos, por algunos que me quedaron en la memoria, vistos agora en la edad madura que tengo, hallo ser verdad lo que he dicho [...] En el tiempo deste célebre español, todos los aparatos de un autor de comedias se encerraban en un costal, y se cifraban en cuatro pellicos blancos guarnecidos de guadamecí dorado, y en cuatro barbas y cabelleras y cuatro cayados, poco más o menos. Las comedias eran unos coloquios, como églogas, entre dos o tres pastores y alguna pastora; aderezábanlas y dilatábanlas con dos o tres entremeses, ya de negra, ya de rufián, ya de bobo y ya de vizcaíno: que todas estas cuatro figuras y otras muchas hacía el tal Lope con la mayor excelencia y propiedad que pudiera imaginarse.

⁶⁵ Todas las citas de Cervantes que aparecen a lo largo de estas páginas están tomadas de las ediciones de Florencio Sevilla Arroyo, disponibles en la Biblioteca Virtual Miguel de Cervantes.

CAPÍTULO 3

No tenemos ningún documento que atestigüe cuándo y dónde pudo Cervantes haber visto representar al batihoja sevillano, aunque sí datos más o menos fiables que parecen confirmar que ambos coincidieron en un mismo momento y en una misma ciudad en al menos tres ocasiones⁶⁶. Los biógrafos de Cervantes no han logrado ponerse de acuerdo: unos sitúan el encuentro en Valladolid entre 1551 y 1553, en donde Lope de Rueda actuó primero en la ceremonia organizada con motivo del regreso del príncipe Felipe II de su viaje a Flandes y después en varias fiestas de celebración del *Corpus Christi*; otros en Madrid en 1563; y otros un año más tarde en Sevilla [Palacín, 1952]⁶⁷. Pero sea como fuere, lo cierto es que aquellos cuatro pellicos blancos, aquellas cuatro barbas y cabelleras y aquellos cuatro cayados quedaron grabados fuertemente en su memoria⁶⁸. De hecho, la “máscara” del bobo de los *Pasos* de Rueda se advierte en el personaje de Cristina de *La guarda cuidadosa*, en Pancracio de *La cueva de Salamanca*, en el soldado de *El juez de los divorcios*, en los rústicos de *El retablo de las maravillas* o en los candidatos de *La elección de los alcaldes de Daganzo* [Pérez de León, 2005: 71]; de igual modo, el recurso estructural final de *Las aceitunas* y *El convidado*, en donde un personaje irrumpe en escena para denunciar lo absurdo de la situación —y que Johnson ha definido como “intimidación sorprendida” [1981]—, tiene su correlato en *El retablo de las maravillas*.

⁶⁶ Abordamos esta cuestión en profundidad en un trabajo publicado hace varios años [Ruiz Urbón, 2009].

⁶⁷ Alonso Cortés se aventura a situar el encuentro entre Cervantes y Lope de Rueda en Valladolid, pero a finales de la década de los 50, a pesar de no tener constancia documental de que los Cervantes siguieran viviendo en esta ciudad por aquel entonces. Su teoría, bastante endeble, se basa en la posibilidad de que el personaje de “don Lope Méndez de Almendárez” que aparece en *El casamiento engañoso* esté inspirado en un tal Lope de Almendárez, corregidor del Ayuntamiento de Valladolid desde 1558; y escribe: “¿Daría nombre Cervantes al personaje de su novela por reminiscencia a este otro? ¿Estaría aún con su familia en Valladolid, desde su traslado en 1551, el cirujano Rodrigo de Cervantes? Muy fácil lo creo” [1917b: 610]. Esta posibilidad, que más que “fácil” resulta arriesgada, pierde entidad si tenemos en cuenta que *Lope* era un nombre común en la época y *Almendárez* un apellido de procedencia arábiga bastante usual.

⁶⁸ Cotarelo Valledor conjetura la participación de Cervantes como actor en alguna “gangarilla” o “garnacha”; de ser así, podría haber observado “el bullicioso vivir del histrionismo, tantas veces aludido, y tomar de memoria pales de comedias, como aquel de Lope de Rueda, mucho después recordado, y que no se corresponde, por cierto, a ninguna de las farsas conocidas del célebre batihoja” [1943: 74].

También recuerda Cervantes en ese mismo prólogo el avance que supuso la llegada de Navarro, quien modificó el vestuario y el *atrezzo*, sacó la música a la escena e incorporó las primeras tramoyas⁶⁹:

Sucedió a Lope de Rueda, Navarro, natural de Toledo, el cual fue famoso en hacer la figura de un rufián cobarde; éste levantó algún tanto más el adorno de las comedias y mudó el costal de vestidos en cofres y en baúles; sacó la música, que antes cantaba detrás de la manta, al teatro público; quitó las barbas de los farsantes, que hasta entonces ninguno representaba sin barba postiza, y hizo que todos representasen a cureña rasa, si no era los que habían de representar los viejos o otras figuras que pidiesen mudanza de rostro; inventó tramoyas, nubes, truenos y relámpagos, desafíos y batallas, pero esto no llegó al sublime punto en que está agora.

Pero además de todas estas representaciones de carácter profano, es muy probable que Cervantes conociera también en sus años jóvenes las representaciones de carácter religioso del teatro escolar⁷⁰. Si aceptamos la hipótesis de que en 1564 se encontraba estudiando en el colegio de los jesuitas de Sevilla —donde conocería a Mateo Vázquez⁷¹—, debemos considerar también su participación (ya fuera como espectador, ya como actor) en las representaciones en lengua latina que, como ejercicio escolástico, los jesuitas empleaban en sus clases para instruir intelectual y moralmente a los alumnos. En el colegio de Sevilla se hallaba por esos años el padre Acevedo, máximo exponente

⁶⁹ Este “Navarro, natural de Toledo” es sin duda el mismo “Navarro, natural de Toledo” al que se refiere Agustín de Rojas Villandrando en su *Viaje entretenido*, quien también ensalza el importante papel que desempeñó en la evolución de la técnica escénica. Aunque pueda hacer referencia al *autor* Pedro Navarro —citado por Lope de Vega en su *Arte nuevo* o por Juan de la Cueva en el *Exemplar poético*—, consideramos más acertada la propuesta de Schevill y Bonilla, quienes lo identifican con Cristóbal Navarro, un representante posterior a Rueda que llevó autos al Corpus de Toledo.

⁷⁰ Para ahondar en esta cuestión, véanse los trabajos de García Soriano [1945], González Gutiérrez [1997], Sierra de Cózar [1998] y Ortés [2002].

⁷¹ El diálogo de Cipión y Berganza en *El coloquio de los perros* sobre la labor docente de los jesuitas y el funcionamiento de sus colegios nos permite suponer que Cervantes se formó bajo la doctrina pedagógica de la Compañía de Jesús y que, por tanto, conoció el método de aprendizaje de la lengua latina a través de representaciones teatrales. Aunque Cervantes pudo asistir a los colegios de Alcalá de Henares, Valladolid, Córdoba o Sevilla (ya que la inclusión de alumnos seculares en estos centros se da en 1547, 1548, 1553 y 1554, respectivamente), el hecho de que la acción de *El coloquio de los perros* se sitúe en Sevilla ha llevado a varios investigadores a decantarse por esta ciudad [Navarro y Ledesma, 1944; García Soriano, 1945; Cascón, 1947]. Por otro lado, la aparición en 1863 de una carta en verso dirigida a Mateo Vázquez, que no pocos atribuyen a Cervantes, podría probar que ambos se conocían ya desde sus años de juventud, tras haber estudiado juntos en el colegio de los jesuitas de Sevilla.

CAPÍTULO 3

de esta práctica teatral; y algunos han querido ver al autor del *Quijote* detrás del “Miguel” que aparece en uno de los manuscritos que se conservan de estas representaciones, como actor de la tragedia *Lucifer furens*.

Pero más allá de estos primeros contactos, tenemos datos que acreditan que durante los años que Cervantes pasó en Italia antes de alistarse en la milicia como soldado, conoció —directa o indirectamente— las tragedias de corte senequista de Cinthio o de Dolce, así como las farsas populares de la *commedia all'improvviso* o del teatro de marionetas de los *puppi*. Aunque todas las experiencias vividas a través del teatro italiano dejaron su huella en la obra cervantina, sin duda las que más atrajeron su interés fueron las representaciones de los cómicos del arte, que por aquellos años triunfaban en Italia gracias a un modo nuevo de entender el concepto teatral. La crítica ha visto influencias de este teatro popular italiano en los aspectos escénicos de *La cueva de Salamanca*; en los personajes de *El Viejo celoso*, con un Cañizares viejo, avaro e impotente (como *Pantalone*) y una Lorenza joven, inteligente y fogosa (como *Columbina*); también en el paralelismo entre los discursos de don Quijote y Arlequín o en las semejanzas entre la caracterización de don Quijote y los *scenari* de la compañía *I Gelosi* (Imagen 1)⁷². De hecho, como bien advierte Vélez-Sainz, el siguiente pasaje del *Quijote* (II, 12) sólo puede entenderse desde la óptica de los personajes de la *commedia dell'arte*, en la que *Arlecchino* es un rufián; *Brighella*, un embustero;

⁷² Parece probable que Cervantes conociese de cerca las actuaciones de *I Gelosi*. Entre los documentos iconográficos de esta compañía se conserva el *Recueil Fossard*, “una colección de grabados de madera donde se recogen partes de tramas teatrales, *lazzi* [...] y representaciones de personajes con las ropas, posturas, formas y los instrumentos que les caracterizaban”, fechados antes de la publicación del *Quijote*, en donde ya aparece el personaje de *Arlecchino*, montado sobre un asno, que porta una vieja armadura, una espada rota, una olla de hierro por yelmo y una lanza en alto en la mano izquierda [Sito Alba, 1983: 23]. Al observar el grabado parecemos estar viendo al mismo personaje descrito por Cervantes en las primeras páginas del *Quijote*: “un hidalgo de los de lanza en astillero, adarga antigua, rocín flaco y galgo corredor. Una olla de más de vaca que carnero” (Imagen 1). El grabado viene acompañado del siguiente texto: “Pues tengo por casco esta buena olla de hierro / y este arnés a la espalda, montado en mi mulo / quiero, como Hércules, asaltar todo el infierno/ por mantener la belleza de mi querida amante: / la bella Francisquina es mi corazón y mi bien/ es mi felicidad, es mi sola esperanza. / Para conseguirla emplearé a mis amigos, mis medios / y mi fuerza en culmen para dar un golpe de lanza; / no faltará después el filo de mi espada / como caballero andante y de la tabla redonda, / enamorado / hasta el fondo, porque no se diga / que haya nadie como Arlequín para triunfar en el mundo” [Vélez Sanz, 2000: 36].



Imagen 1. Grabado XV de *Recueil Fossard* (Arlequin chevalier errant).

Fuente: Vélez Sanz, 2000: 45.

CAPÍTULO 3

Pantalone, un mercader; *Capitano*, un soldado; *Pullicinella*, un simple discreto; y el *Innamorato*, un enamorado simple:

¿no has visto tú representar alguna comedia adonde se introducen reyes, emperadores y pontífices, caballeros, damas y otros diversos personajes? Uno hace el rufián, otro el embustero, éste el mercader, aquél el soldado, otro el simple discreto, otro el enamorado simple; y acabada la comedia y desnudándose de los vestidos della, quedan todos los recitantes iguales [2000: 31].

Muy probablemente Cervantes siguió en contacto con el teatro durante sus años de cautiverio en Argel, pues “era lógico que en los ‘baños’, en las prisiones argelinas, los esclavos italianos también hicieran representaciones”, sirviéndose de *soggetti* del tipo de la *commedia dell’arte* [Sito Alba, 1983]. De hecho, el propio Cervantes nos presenta al comienzo de la tercera jornada de *Los baños de Argel* a un grupo de cristianos cautivos escenificando un coloquio pastoril de Lope de Rueda:

OSORIO: Antes que más gente acuda,
 el coloquio se comience,
 que es del gran Lope de Rueda,
 impreso por Timoneda,
 que en vejez al tiempo vence.

Aunque Cervantes, tras ser liberado por los frailes trinitarios, volvió a pisar suelo español en octubre 1580, no podrá retomar su contacto con el teatro representado al menos hasta noviembre de 1581, momento en el que se levanta la prohibición de representar comedias, decretada con motivo de la muerte de la reina Ana de Austria trece meses antes. Sin duda hubo de sorprenderle la profunda transformación que había vivido la realidad escénica española durante sus años de ausencia: si de joven había asistido a las arcaicas representaciones que se celebraban en iglesias, calles, plazas, patios o casas particulares durante las fiestas y los espectáculos públicos, ahora podrá presenciar unas funciones mucho más evolucionadas, que, además de en estos escenarios improvisados, se representan sobre teatros fijos —el corral de la Cruz ya se había inaugurado (1574) y el del Príncipe está a punto de hacerlo (1582)— administrados por cofradías.

En la década de los 80, Cervantes, dedicado en exclusiva al cultivo de las letras, se integró en los círculos cortesanos y frecuentó los teatros de la capital, apadrinado por

dos personas muy vinculadas a su familia: Alonso Getino de Guzmán, antiguo bailarín y músico de la compañía de Lope de Rueda, que ocupa en ese momento el cargo de alguacil de la Villa y Corte⁷³, y el comediante Tomás Gutiérrez de Castro⁷⁴. Cervantes asistió muy probablemente a las representaciones de Juan de la Cueva, Cristóbal de Virués, los hermanos Argensola, Agustín de Tárrega o Andrés Rey de Artieda, a los que nombra con halago en sus obras, e incluso a las de su propia pluma, que, como él mismo afirma en el prólogo de sus *Ocho comedias y ocho entremeses*, se vieron representar “con general y gustoso aplauso de los oyentes [...] sin que se les ofreciese ofrenda de pepinos ni de otra cosa arrojadiza”.

Aunque sólo conservamos dos contratos que constaten la actividad teatral de Cervantes en esta época (uno con Gaspar de Porres en 1585⁷⁵ y otro con Rodrigo Osorio siete años después⁷⁶), existen otras referencias, tanto documentales como literarias, que

⁷³ “Alonso Getino de Guzmán había sido, en el pasado, miembro ocasional, como danzante, de la compañía de Lope de Rueda, y en la época a la que nos referimos, ocupaba el puesto de alguacil. A este empleo había llegado después de actuar en diversas ocasiones como encargado de la organización de los festejos de la ciudad de Madrid. Desde su cargo, Getino de Guzmán parece que tuvo bastante que ver con la apertura del Corral de la Cruz” [Blasco, 2007: 213].

⁷⁴ Tomás Gutiérrez de Castro fue íntimo amigo de Cervantes. Sabemos que en 1589 éste le perdonó algunas deudas al autor del *Quijote* y que en 1593, cuando ya había abandonado el oficio de cómico y se encontraba regentando una posada en Sevilla, Cervantes declaró a su favor durante el proceso que había iniciado contra la cofradía del Santísimo Sacramento del Sagrario de la Iglesia Mayor de Sevilla, que se negaba a aceptarlo en su seno por no tener “las calidades que se requiere”. Para profundizar en este asunto, véase Cruz Casado, 2005.

⁷⁵ “Sepan cuantos la presente escritura de obligación y concierto vieren como yo, Miguel de Cervantes, residente en esta corte, de la una parte, y de la Otra Gaspar de Porres, autor de comedias, estante al presente en esta corte decimos que es así que yo el dicho Miguel de Cervantes estoy convenido y concertado con Gaspar de Porres en que le tengo de dar dos comedias, la una llamada la *Confusa* y la otra *El Trato de Constantinopla y muerte de Celín*, y la comedia la *Confusa* le he de dar dentro de quince días de la fecha de esta carta, y la otra del dicho *Trato de Constantinopla y muerte de Celín* para ocho días antes de pascua de flores, primera que vendrá de la fecha de ésta; y por ellas el dicho Gaspar de Porres me ha de dar cuarenta ducados en reales” [Citamos por Blasco, 2007: 451].

⁷⁶ “Sepan cuantos esta carta vieren como yo, Miguel de Cervantes Saavedra, vecino de la villa de Madrid, residente en esta ciudad de Sevilla, otorgo y conozco que soy convenido y concertado con vos, Rodrigo Osorio, autor de comedias, vecino de la ciudad de Toledo, estante al presente en esta ciudad de Sevilla, que estáis presente, en tal manera: que yo tengo de ser obligado y me obligo de componer, de hoy en adelante, y entregaros, en los tiempos que pudiere, seis comedias, de los casos y nombres que a mí me pareciere, para que las podáis representar, y os las daré escritas con la claridad que convenga, una a una, como las fuere componiendo, con declaración que dentro de veinte días primeros siguientes, que se

CAPÍTULO 3

permiten probar las relaciones más o menos estrechas que mantuvo con diferentes empresarios teatrales, como Jerónimo Velázquez, que aparece como testigo en una carta de censo impuesta por su mujer a favor de Gaspar Maldonado⁷⁷, o Angulo el Malo, al que nombra en *El coloquio de los perros* y en la segunda parte del *Quijote*; y también con actores de reconocido prestigio, a los que cita en sus obras y a los que muy probablemente vio representar sobre las tablas, como Pedro de Morales, Pedro Montiel, Nicolás de los Ríos o Catalina Hernández, esposa de Gaspar de Porres. Y tampoco podemos olvidar que fue en una taberna de la calle Tudescos, a la que asistía con frecuencia la gente del teatro, donde conoció a Ana Franca (o Villafranca), que se convertiría poco después en la madre de su hija Isabel.

Pero si Cervantes ya hubo de conocer los mecanismos de operación de la *commedia all'improvviso* durante sus años en Italia, parece probable que retomara su contacto con este tipo de representaciones a través de las compañías italianas que llegaron a España por esos años, tal y como se desprende de la documentación que sitúa en Madrid, por ejemplo, a la compañía de Alberto Naselli, más conocido como Ganassa, entre 1580 y 1584, y cuya influencia queda patente en la onomástica de Alonso Quijada (Ganassa es “il tipo dalla mascella lunga”, es decir, “el de la quijada larga”).

Y de igual manera, parece probable que conociera el teatro de títeres, volatines y marionetas, al que ataca con fiereza por boca de Tomás Rodaja en *El licenciado vidriera* y de Berganza en *El coloquio de los perros*⁷⁸. De hecho, una de las escenas más

cuenten del día que se os entregare cada comedia, habéis de ser obligado de la representar en público; y pareciendo que es una de las mejores comedias que se han representado en España, seáis obligado de me dar e pagar por cada una de las dichas comedias cincuenta ducados, los cuales me habéis de dar e pagar el día que la representarais o dentro de ocho días de cómo las hubiereis representado” [Citamos por Blasco, 2006: 467].

⁷⁷ Cristóbal Pérez Pastor comenta que “dicha firma puede ser indicio de relaciones artísticas, tan naturales y propias como las que ha habido siempre entre los autores cómicos y los representantes” y que puede incluso “significar que Jerónimo Velázquez hubiera puesto en escena alguna o algunas de las comedias de Cervantes” [1897: vol I, 257].

⁷⁸ “De los titereros decía mil males: decía que era gente vagamunda y que trataba con indecencia de las cosas divinas, porque con las figuras que mostraban en sus retratos volvían la devoción en risa, y que les acontecía envasar en un costal todas o las más figuras del Testamento Viejo y Nuevo y sentarse sobre él a comer y beber en los bodegones y tabernas. En resolución, decía que se maravillaba de cómo

divertidas de la segunda parte del *Quijote* tiene lugar cuando el falso titerero Maese Pedro, con su mono al hombro, representa la historia de Gaiferos y Melisendra a través de un retablo de figuras. Los italianismos presentes en estos capítulos⁷⁹, que aluden sin duda a estos volatineros itinerantes, llevan a Díaz Plaja a conjeturar que Cervantes se pudo inspirar también en los recuerdos de su viaje a Italia, donde, como hemos señalado antes, muy probablemente vio representar a los llamados *puppi* [1963a y 1963b].

La asistencia de Cervantes a los corrales de comedias hubo de mantenerse también durante los años que ejerció como comisario en Andalucía, pues por aquel tiempo Sevilla era una ciudad con grandísima actividad teatral. Y también tras la llegada del nuevo siglo y la irrupción en escena de una nueva oleada de dramaturgos *paniaguados*, con Lope de Vega a la cabeza. Cervantes, fuera ya del circuito comercial como dramaturgo, presenció durante los últimos años de vida ese “arte nuevo” de hacer comedias con “cien mil impropiedades y disparates” (*Quijote*, II, 26); y así lo deja patente también en estos versos de su *Viaje del Parnaso*:

Adiós teatros públicos, honrados
por la ignorancia, que ensalzada veo
en cien mil disparates recitados.

No cabe duda de que el teatro fue para Cervantes su gran pasión. En varias de sus obras hace alusión a cuestiones referentes al espectáculo dramático, que sólo había de conocer aquel que hubiese vivido de cerca la realidad de los tablados; habla, por ejemplo, del trabajo de los empresarios teatrales⁸⁰, del papel de los corregidores como

quien podía no les ponía perpetuo silencio en sus retablos, o los desterraba del reino” (*El licenciado vidriera*); “que esto de ganar del comer holgando tiene muchos aficionados y golosos; por esto hay tanto titereros en España, tantos que muestran retablos [...] Toda esta gente es vagamunda, inútil y sin provecho” (*El coloquio de los perros*).

⁷⁹ El ventero se refiere a maese Pedro en términos italianos al decir que es “hombre galante” y “bon compañero”, y don Quijote le pregunta en un defectuoso italiano: “¿Qué peje pillamo?” (*Quijote*, II, 25).

⁸⁰ “El trabajo de los autores es increíble, y su cuidado, extraordinario, y han de ganar mucho para que al cabo del año no salgan tan empeñados, que les sea forzoso hacer pleito de acreedores. Y, con todo esto, son necesarios en la república, como lo son las florestas, las alamedas y las vistas de recreación, y como lo son las cosas que honestamente recrean” (*El licenciado vidriera*).

CAPÍTULO 3

censores teatrales⁸¹, del privilegio que tienen los dramaturgos al poder acudir gratuitamente a los teatros⁸², de los requisitos que han de tener los comediantes⁸³ y de la pluralidad de personajes a que dan vida⁸⁴, de los bailes que se representan durante el espectáculo dramático⁸⁵ o de la distinta suerte que sufre una misma comedia en dos ciudades distintas⁸⁶.

⁸¹ “Desta manera [...] llegaron a Badajoz, donde ya tenía el Corregidor castellano nuevas de Lisboa, cómo por allí habían de pasar los nuevos peregrinos, los cuales, entrando en la ciudad, acertaron a alojarse en un mesón do se alojaba una compañía de famosos recitantes, los cuales aquella misma noche habían de dar la muestra para alcanzar la licencia de representar en público, en casa del Corregidor” (*Los trabajos de Persiles y Segismunda*).

⁸² “Que todo poeta cómico, que felizmente hubiere sacado á luz tres comedias, pueda entrar sin pagar en los teatros, si ya no fuere la limosna de la segunda puerta, y aun esta, si pudiese ser, la escuse” (*Adjunta al Parnaso*).

⁸³ “De gran memoria, primero; / segundo, de suelta lengua; / y que no padezca mengua / de galas es lo tercero. Buen talle no le perdono, / si es que ha de hacer los galanes; / no afectado en ademanes, / ni ha de recitar con tono. / Con descuido cuidadoso, / grave anciano, joven presto, / enamorado compuesto, / con rabia si está celoso. / Ha de recitar de modo, / con tanta industria y cordura, / que se vuelva en la figura / que hace de todo en todo. / A los versos ha de dar / valor con su lengua experta, / y a la fábula que es muerta, / ha de hacer resucitar. / Ha de sacar con espanto / las lágrimas de la risa, / y hacer que vuelvan con [p]risa / otra vez al triste llanto. / Ha de hacer que aquel semblante / que él mostrare, todo oyente / le muestre, y será excelente / si hace aquesto el recitante” (*Pedro de Urdemalas*); “lo que menos ha menester la farsa es personas bien nascidas; galanes sí, gentiles hombres y de expeditas lenguas” (*El licenciado vidriera*); “digo, en fin, que este poeta, a quien la necesidad había hecho trocar los Parnasos con los mesones y las Castalias y las Aganipes con los charcos y los arroyos de los caminos y ventas, fue el que más se admiró de la belleza de Auristela, y al momento la marcó en su imaginación y la tuvo por más que buena para ser comedianta, sin reparar si sabía o no la lengua castellana” (*Los trabajos de Persiles y Segismunda*).

⁸⁴ “Ya podré ser patriarca, / pontífice y estudiante, / emperador y monarca: / que el oficio de farsante / todos estos abarca” (*Pedro de Urdemalas*); “el que servía a una comedianta, en sola una servía a muchas damas juntas, como era a una reina, a una ninfa, a una diosa, a una fregona, a una pastora, y muchas veces caía la suerte en que sirviese en ella a un paje y a un lacayo: que todas estas y más figuras suele hacer una farsanta” (*El licenciado vidriera*).

⁸⁵ “Le bailasen al modo que se canta y baila en las comedias” (*La ilustre fregona*); “comenzaron a entrar por diversas partes de la enramada muchas y diferentes danzas, entre las cuales venía una de espadas, de hasta veinte y cuatro zagales de gallardo parecer y brío, todos vestidos de delgado y blanquísimo lienzo, con sus paños de tocar, labrados de varias colores de fina seda; y al que los guiaba (*Quijote*, II, 20).

⁸⁶ “—Créame vuesa merced —dije yo— que las comedias tienen días, como algunas mujeres hermosas; y que esto de acertarlas bien va tanto en la ventura como en el ingenio: comedia he visto yo apedreada en Madrid que la han laureado en Toledo, y no por esta primer desgracia deje vuesa merced de proseguir en componerlas, que podrá ser que, cuando menos lo piense, acierte con alguna que le dé crédito y dineros” (*Adjunta al Parnaso*).

Pero más allá de todas estas referencias directas, su papel de espectador en la vida real se refleja también —y de manera muy significativa— en el recurso del teatro dentro del teatro, que aparece de manera recurrente tanto en su producción dramática como en su novelística [Maestro, 2004]. Ya hemos hablado antes de la representación de los cristianos cautivos en *Los baños de Argel* o del espectáculo de títeres de Maese Pedro en el *Quijote*, aunque podemos citar muchos más ejemplos de esta metateatralidad: Dorotea disfrazada de princesa Micomicona en la segunda Parte del *Quijote*, la función fingida de Chanfalla en *El retablo de las maravillas*, el entremés escenificado por los criados Muñoz, Torrente y Ocaña en *La entretenida* o la autorepresentación de Pedro de Urdemalas, entre otros muchos.

3.1.2. Un lector continuo de obras dramáticas

Cervantes da muestras de su afición por los libros y la lectura a través de muchos de sus personajes. El narrador del *Quijote* dice ser “aficionado a leer, aunque sean los papeles rotos de las calles” (*Quijote*, I, 9), y su famoso personaje reconoce que los más de trescientos libros que posee son el *regalo* de su alma y el *entretimiento* de su vida (*Quijote*, I, 24), al tiempo que asegura que “el que lee mucho y anda mucho, ve mucho y sabe mucho” (*Quijote*, II, 25)⁸⁷. Vivaldo hace una defensa apasionada de “algunos libros y muchos papeles abiertos y cerrados” que rodean el cuerpo sin vida de Crisóstomo, por ser un ejemplo para los vivientes “en los tiempos que están por venir” (*Quijote*, I, 13). Tomás Rodaja posee “muchos libros” a los que atiende más “que a otros pasatiempos” (*El licenciado vidriera*) y Diego Miranda “hasta seis docenas [...], cuáles de romance y cuáles de latín, de historia algunos y de devoción otros” (*Quijote*, II, 16). Luscinda es muy aficionada a las novelas de caballerías y solicita a Cardenio que le preste el *Amadís de Gaula* (*Quijote*, I, 24). Y Juan Palomeque nos cuenta que un

⁸⁷ Esta reflexión aparece de modo muy similar en el sexto capítulo del segundo libro del *Persiles*, donde Auristela postula: “El ver mucho y el leer mucho aviva los ingenios de los hombres” (*Los trabajos de Persiles y Segismunda*).

CAPÍTULO 3

huésped desmemoriado se ha dejado en la venta una maleta con papeles de “buena letra”, entre los que se encuentran la *Novela del curioso impertinente* (*Quijote*, I, 32) y la *Novela de Rinconte y Cortadillo* (*Quijote*, I, 47)⁸⁸.

Cervantes era un bibliófilo, de eso no cabe duda. Cotarelo Valledor, en el discurso que pronuncia en 1940 en el Instituto de España para conmemorar la Fiesta Nacional del libro, asegura que, aunque fue un “lector continuo”, Cervantes carecía de libros o tenía pocos, porque éstos eran “caros” y él “pobre” [1943: 5-9]. Eisenberg, oponiéndose a esta creencia tan largamente traída, advierte que Cervantes “no era rico pero tampoco pobre”, pues “disponía de los recursos para comprar libros”, y establece los 210 títulos a los que hubo de tener acceso [1987]⁸⁹, bien porque los compró en librerías o subastas⁹⁰, bien porque los heredó, se los regalaron o se los dejaron prestados⁹¹.

Entre los libros que señalan uno y otro encontramos escritos religiosos, tratados de filosofía, geografía, medicina o ciencias naturales, poéticas, libros de refranes, manuales de gramática, cancioneros y romanceros, novelas picarescas, novelas de caballerías..., pero pocos o muy pocos libros de teatro (que, dicho sea, no eran muy habituales en la época)⁹². Basándonos en las referencias directas e indirectas que

⁸⁸ Para profundizar en el papel de los libros como objeto material en Cervantes —que guarda relación con esa distinción entre *corpus mysticum* y *corpus mechanicum* de la que hablábamos en el capítulo 1— véase el trabajo de Adrián J. Sáez titulado “El bueno, el feo y el malo: los libros en Cervantes” [2019].

⁸⁹ En realidad, en la edición de “La biblioteca de Cervantes” de 1987, Eisenberg da cuenta de 202 entradas; nos referimos aquí a 210 porque son las que figuran en su versión de 2002, disponible en <<http://bigfoot.com/~daniel.eisenberg>>.

⁹⁰ Astrana Marín refiere un documento que acredita que en 1590 Cervantes compró en una subasta la *Historia de Santo Domingo* —que él identifica como la de Hernando del Castillo—, por 30 reales, y “cuatro libritos dorados, de letra francesa”, por 18 reales [1889-1959: tomo 4, 463-464].

⁹¹ Es probable que Cervantes heredara la biblioteca de su padre y que durante su viaje a Italia satisficiera sus ansias de lectura de la mano del cardenal Acquaviva, en Roma, y de Pedro Laínez, en Nápoles, que parece haberlo presentado en varios cenáculos [Canavaggio, 2003: 81]. También hubo de tener acceso a los volúmenes que publicaron sus amigos (Cristóbal de la Mesa, Salas Barbadillo, Pedro de Padilla, Vicente Espinel, Juan de Jáuregui...) o, al menos, a los que llevaban versos proemiales suyos.

⁹² Es importante significar que el mercado editorial caminaba por derroteros muy distintos. Las ediciones de obras dramáticas eran escasas, pues los lectores preferían los libros de caballería o de épica culta [Chevalier, 1976].

aparecen en sus obras, podemos presuponer que Cervantes leyó alguna de las múltiples ediciones que se imprimieron de la *Celestina*, a la que alude en los preliminares de la primera parte del *Quijote*, así como la *Clariana*, la *Tragicomedia de Lisandro y Roselía*, la *Tragedia Policiana*, la *Comedia Selvagia* y la *Comedia Vidriana*, herederas de ésta. Igualmente hubo de leer las obras de Lope de Rueda impresas por Timoneda⁹³; la *Primera parte de las Comedias y Tragedias* de Juan de la Cueva, publicada en Sevilla por Andrea Pescioni, a costa de Jácome López y Antonio de Acosta, en 1583; varias de las comedias impresas en la segunda mitad del XVI, como la *Florinea* de Rodríguez Florián (Medina, 1554) o la *Tidea* de Francisco de Natas (1550); la tragedia *Los amantes* (Valencia, 1581); la traducción que hizo Cristóbal Suárez de Figueroa de la tragedia *El pastor Fido* de Guarini (Nápoles, 1602, o Valencia, 1609); las *Obras trágicas y líricas* de Cristóbal de Virués, publicadas en Madrid por Alonso Martín, a costa de Esteban Bogía, en 1609; la *Propalladia* de Torres Naharro, bien en la edición príncipe de Nápoles (1517) o en cualquiera de las reimpressiones de Sevilla, Burgos, Amberes o Madrid (1534-1590); y *El viaje entretenido* de Agustín de Rojas Villandrando, que aunque no es propiamente una obra dramática, tiene mucho de *teatro*.

3.1.3. Un teórico a caballo entre el sistema tradicional del mecenazgo y la escritura teatral como *modus vivendi*

Mucho se ha dicho y escrito sobre la preocupación de Cervantes por la teoría literaria. A través de las palabras de sus personajes o de las voces de sus narradores, va esgrimiendo diversas opiniones sobre la novela, el romance en prosa, la poesía o el género dramático; y es precisamente “la paradoja dada por la multiplicidad de voces que los enuncian y el dialogismo implícito que esta forma de concebir la crítica literaria supone” lo que dificulta su interpretación [Alcalá Galán, 2002: 255].

⁹³ Las quintillas de un coloquio pastoril “del gran Lope de Rueda, / impreso por Timoneda” que Cervantes reproduce en *Los baños de Argel* no se corresponden con ninguna de las obras del autor impresas por Timoneda. Barrera y Leirado cree que esas quintillas pertenecen al *Coloquio pastoril* publicado por Pedro Mey en Valencia en 1567, y que Pastor Fuster atribuye a Rueda [1860: 392].

CAPÍTULO 3

Hay varios momentos en los que Cervantes reflexiona sobre la teoría de la comedia⁹⁴, aunque son esencialmente tres los que han centrado una mayor atención de la crítica: la conversación que el canónigo toledano y el cura Pedro Pérez mantienen en el capítulo 48 de la primera parte del *Quijote*, el diálogo alegórico que se produce entre la Curiosidad y la Comedia en la segunda jornada de *El rufián dichoso* y el “Prólogo al lector” que introduce las *Ocho comedias y ocho entremeses nuevos, nunca representados*.

En el capítulo 48 de la primera parte del *Quijote* (1605), Cervantes —a través de las palabras del canónigo y el cura— parece identificarse con una preceptiva teatral clásica, marcada fuertemente por la influencia de la *Poética* de Aristóteles, el *Ars poetica* de Horacio y la *Philosophía antigua poética* de López Pinciano. El canónigo ensalza a los “entendidos poetas” que compusieron comedias y tragedias que “guardaban bien los preceptos del arte”, como *La Isabela*, *La Filis* y *La Alejandra*, de Lupercio Leonardo de Argensola, *La ingratitud vengada*, de Lope de Vega, o la *Numancia*, del mismo Cervantes. El cura, siguiendo la misma postura, rechaza cualquier alteración del principio lógico de verosimilitud, pues “habiendo de ser la comedia [...] espejo de la vida humana, ejemplo de costumbres e imagen de la verdad”, las que en esos últimos años se representan son “espejo de disparates, ejemplo de necedades e imágenes de lascivia”; y se lamenta de la falta de decoro, de la ruptura de las unidades de tiempo y lugar y de la mezcla de historia y ficción⁹⁵:

⁹⁴ El entremés, entendido como forma dramática inserta en la tradición de los géneros cómicos, no llega a constituirse nunca como doctrina teórica independiente; de hecho, la consideración que tiene en el *Arte nuevo de hacer comedias en este tiempo* es meramente subsidiaria y funcional: rellenar el tiempo que transcurre entre las jornadas de las comedias [Huerta Calvo, 1995 y 2008]. A pesar de ello, Cervantes nos ofrece en *La entretenida* un interesantísimo diálogo entre Marcela, que opina “que sea / el baile y el entremés / discreto, alegre y cortés / sin que haya en él cosa fea”, y Cristina, su criada, que defiende el “baile cantado, / al modo y uso moderno”. Según Maestro, este texto pone de manifiesto “la ridiculización del gusto del público de la época, así como también el de los autores comerciales que, sólo por fama y dinero, estarían dispuestos a escribir los mismos tópicos, y a repetir las mismas mixtificaciones, sólo por estar de acuerdo con el gusto del vulgo” [2013: 100].

⁹⁵ Encontramos opiniones muy similares en los versos finales de *Pedro de Urdemalas*, cuya fecha de composición se ha establecido entre 1608 y 1611 [González, 1997: 81]: “Mañana, en el teatro, se hará una, / donde por poco precio verán todos / desde principio al fin toda la traza, / y verán que no acaba en casamiento, / cosa común y vista cien mil veces, / ni que parió la dama esta jornada, / y en otra tiene el

Porque ¿qué mayor disparate puede ser en el sujeto que tratamos que salir un niño en mantillas en la primera escena del primer acto, y en la segunda salir ya hecho hombre barbado? ¿Y qué mayor que pintarnos un viejo valiente y un mozo cobarde, un lacayo retórico, un paje consejero, un rey ganapán y una princesa fregona? ¿Qué diré, pues, de la observancia que guardan en los tiempos en que pueden o podían suceder las acciones que representan, sino que he visto comedia que la primera jornada comenzó en Europa, la segunda en Asia, la tercera se acabó en África, y aun, si fuera de cuatro jornadas, la cuarta acababa en América, y, así, se hubiera hecho en todas las cuatro partes del mundo? Y si es que la imitación es lo principal que ha de tener la comedia, ¿cómo es posible que satisfaga a ningún mediano entendimiento que, fingiendo una acción que pasa en tiempo del rey Pepino y Carlomagno, el mismo que en ella hace la persona principal le atribuyan que fue el emperador Heraclio, que entró con la Cruz en Jerusalén, y el que ganó la Casa Santa, como Godofre de Bullón, habiendo infinitos años de lo uno a lo otro; y fundándose la comedia sobre cosa fingida, atribuirle verdades de historia y mezclarle pedazos de otras sucedidas a diferentes personas y tiempos, y esto no con trazas verisímiles, sino con patentes errores, de todo punto inexcusables? Y es lo malo que hay ignorantes que digan que esto es lo perfecto y que lo demás es buscar gullurías. Pues ¿qué, si venimos a las comedias divinas? ¡Qué de milagros falsos fingen en ellas, qué de cosas apócrifas y mal entendidas, atribuyendo a un santo los milagros de otro! Y aun en las humanas se atreven a hacer milagros, sin más respeto ni consideración que parecerles que allí estará bien el tal milagro y apariencia, como ellos llaman, para que gente ignorante se admire y venga a la comedia (*Quijote*, I, 48).

Pero tal vez lo más interesante de este pasaje sea la reflexión que hacen ambos sobre las causas que han motivado estos “disparates”, que no están en la ignorancia de los poetas, “porque algunos hay dellos que conocen muy bien en lo que yerran, y saben estremadamente lo que deben hacer”, sino en las exigencias de los representantes, “que no se las comprarían si no fuesen de aquel jaez”, porque “así las quiere el vulgo”.

La relación entre este capítulo y el *Arte nuevo de hacer comedias en este tiempo* —publicado en Madrid en 1609 por Alonso Martín, a costa de Alonso Pérez— ha sido advertida desde antiguo por los investigadores, que han supuesto que Cervantes hubo de conocer la obra de Lope de Vega cuando estaba escribiendo el final del *Quijote*, bien porque existió una edición anterior a la de conocida de 1609, o bien, y más probable, porque circuló manuscrita antes de ser impresa. Sea como fuere, lo cierto es que

niño ya sus barbas, / y es valiente y feroz, y mata y hiende, / y venga de sus padres cierta injuria, / y al fin viene a ser rey de un cierto reino / que no hay cosmografía que le muestre. / Destas impertinencias y otras tales / ofreció la comedia libre y suelta / pues llena de artificio, industria y galas, / se cela del gran Pedro de Urdemalas” (*Pedro de Urdemalas*). Y también en el capítulo 26 de la segunda parte del *Quijote*, por boca de Maese Pedro (“mil comedias llenas de mil impropiedades y disparates”), y en el *Viaje del Parnaso* (“cien mil disparates recitados”).

CAPÍTULO 3

Cervantes, a través del canónigo y el cura, parece dar respuesta punto por punto a las ideas lopianas, utilizando incluso sus mismos términos (*vulgo, pagar, gusto, necio, confuso, arte, pocos, fama, ignorante, ingenios, bárbaros*, etc.) [Martín Jiménez, 2006: 259-277]. La alusión a Lope de Vega se hace patente en la parte final del capítulo, bajo la figura de un “felicísimo ingenio” acomodado al gusto de los representantes:

Y que esto sea verdad véase por muchas e infinitas comedias que ha compuesto un felicísimo ingenio destos reinos, con tanta gala, con tanto donaire, con tan elegante verso, con tan buenas razones, con tan graves sentencias, y, finalmente, tan llenas de elocución y alteza de estilo, que tiene lleno el mundo de su fama; y por querer acomodarse al gusto de los representantes, no han llegado todas, como han llegado algunas, al punto de la perfección que requieren.

En la jornada segunda de *El rufián dichoso*, cuya fecha de composición se ha establecido entre 1597 y 1610 [González, 1997: 81] —si bien nosotros, al igual que Riley⁹⁶, nos inclinamos a pensar que es posterior a la redacción de la primera parte del *Quijote*—, Cervantes nos ofrece un diálogo alegórico entre la Curiosidad y la Comedia, en el que esta última admite la necesidad de que la realidad teatral evolucione con el paso del tiempo:

CURIOSIDAD: Comedia.
COMEDIA: Curiosidad
¿qué me quieres?
CURIOSIDAD. Informarme
qué es la causa por que dejas
de usar tan antiguos trajes,
del coturno en las tragedias,
del zueco en las manuales
comedias, y de la toga
en las que son principales;
cómo has reducido a tres
los cinco actos que sabes
que a un tiempo te componían
ilustre, risueña y grave;
ahora aquí representas,

⁹⁶ “No se sabe si el diálogo entre Curiosidad y Comedia en dicha obra se escribió antes o después de la primera parte del *Quijote*. Conjeturo que después, a causa de la extensión y el tono defensivo de la apología de la Comedia. Es de suponer que Cervantes no se hubiera extremado tanto en disculparse si antes ni había propagado ideas contrarias” [Riley, 1973: 303].

CERVANTES, UN HOMBRE DE TEATRO

y al mismo tiempo en Flandes;
truecas sin discurso alguno tiempos, teatros, lugares.

COMEDIA. Los tiempos mudan las cosas
y perfeccionan las artes,
y añadir a lo inventado
no es dificultad notable.
Buena fui pasados tiempos,
y en éstos, si los mirares,
no soy mala, aunque desdigo
de aquellos preceptos graves
que me dieron y dejaron
en sus obras admirables
Séneca, Terencio y Plauto,
y otros griegos que tú sabes.
He dejado parte de ellos,
y también guardado parte,
porque lo quiere así el uso,
que no se sujeta al arte.
Ya represento mil cosas,
no en relación, como de antes,
sino en hecho; y así, es fuerza
que haya de mudar lugares;
que, como acontecen ellas
en muy diferentes partes,
voime allí donde acontecen,
disculpa del disparate.
Ya la comedia es un mapa,
donde no un dedo distante
verás a Londres y a Roma,
a Valladolid y a Gante.
Muy poco importa al oyente
que yo en un punto me pase
desde Alemania a Guinea
sin del teatro mudarme;
el pensamiento es ligero:
bien pueden acompañarme
con él doquiera que fuere,
sin perderme ni cansarse [...].

CURIOSIDAD. Aunque no lo quedo en todo,
quedo satisfecha en parte.

Un tercer pasaje que conviene recordar a la hora de repasar el pensamiento teórico de Cervantes en relación con el teatro es el del prólogo a las *Ocho comedias y ocho entremeses*, en el que alude —esta vez de manera directa— al “monstruo de la

CAPÍTULO 3

naturaleza, al gran Lope de Vega”, cuyas comedias, no sin cierta ironía, califica ahora de “felices y bien razonadas”⁹⁷:

alzóse con la monarquía cómica; avasalló y puso debajo de su jurisdicción a todos los farsantes; llenó el mundo de comedias propias, felices y bien razonadas, y tantas, que pasan de diez mil pliegos los que tiene escritos, y todas (que es una de las mayores cosas que puede decirse) las ha visto representar, o oído decir, por lo menos, que se han representado; y si algunos, que hay muchos, han querido entrar a la parte y gloria de sus trabajos, todos juntos no llegan en lo que han escrito a la mitad de lo que él sólo.

Cervantes ya no sólo no cuestiona la configuración de una nueva preceptiva dramática, alejada de los imperativos clásicos, sino que además se incluye a sí mismo en la lista de poetas que han contribuido al desarrollo evolutivo de la misma, al haber sido el primero en reducir las comedias a tres jornadas y llevar a la escena las figuras morales⁹⁸:

Y esto es verdad que no se me puede contradecir, y aquí entra el salir yo de los límites de mi llaneza: que se vieron en los teatros de Madrid representar *Los tratos de Argel* que yo compuse; la destrucción de la *Numancia* y *La batalla naval*; donde me atreví a ir las comedias a tres jornadas, de cinco que tenían; mostré, o por mejor decir, fui el primero que representase las imaginaciones y los pensamientos del alma, sacando figuras morales al teatro, con general y gustoso aplauso de los oyentes.

⁹⁷ Agustín de la Granja interpreta este prólogo como una “respuesta irónica” de Cervantes al prólogo de *El peregrino*, en el que Lope se atribuía hasta doscientas treinta comedias: “¿es posible que sólo en cinco o seis años hubiese doblado su producción? La cifra redonda de diez mil pliegos no es más que una cómica cantidad hiperbólica, que hacía sonreír a los entendidos” [1995: 233].

⁹⁸ Estas palabras de Cervantes también han generado controversia, pues hubo otros autores antes que él que ya habían reducido las comedias a tres jornadas (de cuatro que tenían, y no cinco) e introducido personajes alegóricos en sus textos. Alonso Cortés, en el capítulo sobre Cervantes que escribe para la *Historia general de las literaturas hispánicas* de Díaz-Plaja, recuerda que algunas obras anteriores a las de Cervantes, como el *Auto de Clarindo* de Antonio Díez (¿1535?), la *Comedia Florisea* de Francisco de Avendaño (1553) o algunas comedias de Cristóbal de Virués (1550-1614) tenían una estructuración tripartita de las jornadas, pero que fue con *La batalla naval* cuando este uso dramático aislado “vino a prevalecer”. Además, Alonso de la Vega, Juan de la Cueva o Lupercio Leonardo de Argensola, entre otros, habían creado personajes alegóricos como la Verdad, el Consuelo, el Remedio o la Fama, si bien es cierto que éstos carecían aún de la complejidad interna y la profundidad psicológica de los cervantinos [Díaz-Plaja, 1968: 832-833]; en palabras de Canavaggio, “las «figuras morales» de *El trato de Argel* ilustran por entero las palabras de nuestro poeta. No son, como las alegorías tradicionales, obstáculos personalizados que materialicen las fuerzas antagónicas con que se supone que el individuo se enfrenta; intervienen en la acción para exteriorizar el debate interior de un personaje presente en escena, pero sin dialogar nunca con él ni intentar sustituirlo: técnica original de la que comprendemos que Miguel se haya preocupado por reivindicar la paternidad” [2003: 178-179].

¿Cómo debemos leer y entender estos tres pasajes? Desde que Cotarelo Valledor asegurara en su monografía sobre el teatro cervantino que hay en él un “cambio de estética”, un “desacuerdo entre la teoría y la práctica, entre sus ideas de 1604 y las de 1615” [1915: 45], han sido muchos los que han retomado esta idea, no sin ciertas matizaciones⁹⁹. Juliá Martínez habla de la “falta de fijeza de ideas” de Cervantes y de su incapacidad para “ordenar sus pensamientos sobre el teatro” [1948: 357]; Marrast cree que su teoría de la comedia es “muy variable” [1961: 15]; Riley considera que “modificó sus ideas, pero que las discrepancias entre las opiniones publicadas en 1605 y las publicadas en 1615 no es tan grande como generalmente se ha creído” [1973: 310], pues las críticas del canónigo y el cura van dirigidas únicamente a los abusos extremos y, además, se pueden encontrar críticas igual de severas en alguna de las obras posteriores al *Quijote* de 1605; Wardropper niega que su teoría sea inestable y defiende que la clave del cambio debe buscarse en su “experimentación continua en el plano práctico” [1973: 156]; Sevilla Arroyo y Rey Hazas abogan por una “teoría dramática vacilante y confusa” y una “experiencia literaria diversa y múltiple, que se modifica con cada nueva obra que sale del taller teatral” [1987: 24 y 285]; Alcalá Galán explica que “nos deja una atomizada y espontánea ‘poética de la comedia’ en permanente conflicto con su tiempo y de la que disenterá ligeramente la propia práctica teatral llevada a cabo en sus *Ocho comedias y entremeses*” [2002: 256]; y Maestro defiende que sus “vacilaciones se deben a la doble limitación que experimenta la creación teatral cervantina, entre la ‘lógica’ de la preceptiva clásica y los ‘códigos’ de la comedia lopista” [2000: 276].

En nuestra opinión, la raíz del problema —si es que puede considerarse como tal— está en la consolidación del hecho teatral como *mercadería vendible*, esto es, un

⁹⁹ Ya antes de Cotarelo Valledor, Menéndez Pelayo explicaba la antonimia entre las teorías estéticas del canónigo y la palinodia de la Comedia del siguiente modo: “primero, el respeto á una tradición literaria tenida por infalible, respeto más bien habitual y mecánico que nacido del propio convencimiento; segundo, el mal humor contra los poetas noveles, que habían arrojado del teatro á sus predecesores naturales, la escuela de Juan de la Cueva y de Virués, á la cual pertenecía cervantes; tercero, el buen gusto ofendido por dislates evidentes, no tanto por la inobservancia de lugar y de tiempo como por la monstruosa confusión de tiempos y lugares que en el breve espacio de tres jornadas abarcaba una crónica entera; cuarto, la preocupación del fin moral del teatro” [1884: 420-421].

CAPÍTULO 3

negocio institucionalizado del que dependen económicamente todos los agentes implicados (poetas, *autores*, actores, arrendatarios, Hospitales...), y que tiene su razón de ser en la existencia de un público masivo y urbano que lo paga y consume [Maravall, 1975: 174-265]. Tal y como explicamos detalladamente en el primer capítulo, en torno al año 1600 se conforma en nuestro país el primer *campo literario*, que es la consecuencia final de una serie de reacciones en cadena: la expansión de la imprenta posibilita la democratización de la cultura, y esta democratización (o vulgarización, según se mire) obliga a las autoridades a establecer un sistema de privilegios que controle tanto la impresión documental como la representación escénica, lo que da lugar, a su vez, al nacimiento de la industria editorial, del teatro profesional y del autor como entidad económica [Bourdieu, 1989-1990 y 1995; Gutiérrez, 2005a]. Con la aparición de un nuevo mercado consumidor de literatura, los espacios de producción se desdobl原因an y los autores pueden orientar su escritura hacia la obtención de capital económico o simbólico, que es el que realmente determina su posición dentro del campo literario y sus interacciones con el campo de poder.

En este sentido, la postura de Cervantes —y esto es algo en lo que conviene insistir— no es muy distinta a la de otros escritores del momento, entre ellos el propio Lope de Vega, que como bien ha estudiado García-Reidy mostró siempre una conciencia autorial en tensión entre los códigos del mecenazgo y los del teatro comercial [2013b: 201-299]. En varios de los documentos conservados, Lope manifiesta su dependencia económica respecto de la escena; así, en una carta de 1604 dirigida a un amigo suyo asegura que “escribe por dinero”; en otra que envía al duque de Sessa en torno a 1612 lamenta los problemas que tiene en un ojo por “escribir disparates para vivir” (y es importante advertir que, al igual que Cervantes, califica a sus propias comedias de “disparates”); y en otra que escribe al duque de Lemos casi una década después, asegura que ha podido estar “un año sin ser poeta de *pane lucrando*” gracias a los quinientos ducados que le envió el duque de Osuna. Pero sin duda la carta en la que Lope mejor ilustra esta diatriba entre la realidad emergente del mercado teatral y el sistema tradicional del mecenazgo nobiliario y regio es la que dirige al duque de Sessa en 1616, y que comienza así:

Señor: para trasladar los romances no he tenido lugar; mañana me dejarán las musas, a que me obliga la pura necesidad, porque en mí no son damas, sino rameras [Citamos por García-Reidy: 2013: 11].

Con esta poderosa imagen de las musas convertidas en rameras describe Lope el ejercicio de la escritura teatral, una profesión que le aporta beneficios económicos rápidos y directos, pero que no le permite incrementar su prestigio dentro de la jerarquía interna del campo literario, ni tampoco trasladar dicho prestigio al campo de poder; y que por eso debe complementar con una escritura dirigida a la élite, que le aporte valor simbólico.

Entendidas dentro de este contexto, las supuestas vacilaciones de Cervantes en torno a la preceptiva teatral no son ajenas a las de otros escritores; y más en concreto, a las de Lope. Sus opiniones sobre la comedia deben englobarse dentro de un debate superior y mucho más profundo: el de las musas como damas o como rameras, o, dicho de otro modo, el de la escritura como prestigio o como *modus vivendi*. El prólogo a sus *Ocho comedias*, dedicado al Conde de Lemos, debe entenderse como un deseo de Cervantes de acercarse a la literatura de élite a través del patronazgo, tal vez porque el mercado teatral le había dado ya definitivamente la espalda.

3.1.4. Un dramaturgo desencantado que escribe novelas

Toda la producción dramática de Cervantes se ha explicado siempre desde dos códigos contrapuestos: de un lado, el de un teatro *viejo y representado*, destinado a los espectadores, con el que triunfa en la década de 1580; y, del otro, el de un teatro *nuevo y nunca representado*, destinado a los lectores, que decide dar a la estampa en el verano de 1615, cuando roza ya los setenta años de edad, porque ningún autor de comedias quiere llevarlo a las tablas. Y entre ambos códigos —el de la representación y el de la lectura— planea siempre la sombra alargada de Lope de Vega [García Aguilar, Gómez Canseco y Sáez, 2016].

Nuestro objetivo en este epígrafe no es ahondar en todas estas cuestiones, ya tantas veces estudiadas, ni valorar la calidad de su dramaturgia, tan defendida por unos

CAPÍTULO 3

como denostada por otros. Nuestro objetivo —mucho más modesto— es el de incidir en la importancia que el Cervantes-dramaturgo tuvo en la conformación de toda su producción literaria, y no sólo en la de sus comedias y entremeses, pues no puede entenderse el éxito de su novela sin advertir que parte de éste reside precisamente en el fracaso de su teatro.

Cervantes supo “transformar la materia dramática en materia narrativa con gran soltura” [Madroñal, 2008: 274]. Más allá de las numerosas referencias a *autores*, comediantes o dramaturgos, de las que ya hemos hablado antes, su conciencia de autor teatral se plasma en la estructuración de algunos pasajes del *Quijote* y de *Los trabajos de Persiles y Segismunda*, en donde aparecen breves historias intercaladas, apartadas de la acción principal, que parecen responder a la organización del espectáculo teatral en el Siglo de Oro, en el que, entre jornada y jornada, se insertaban entremeses, jácaras, mojigangas o bailes. En el *Quijote*, la aventura de los cueros de vino (I, 35), la disputa sobre el baciuelmo (I, 44-45) o los pasajes de Sancho en la ínsula de Barataria (II, 45, 49, 51 y 53) se configuran como tres “entremeses” intercalados entre las “comedias” formadas por la novela de *El curioso impertinente*, el episodio de los amores de doña Clara y don Luis y la historia de don Quijote con los duques, respectivamente; y lo mismo sucede en el *Persiles*, en donde la acción principal se ve interrumpida con frecuencia por historias intercaladas, algunas de las cuales siguen directrices muy próximas a las del teatro breve, como la del matrimonio de los hijos de los alcaldes Tozuelo y Cobeño, la de los falsos cautivos o la de Isabela Castrucha.

Además de estas comedias y entremeses intercalados, son muchos los que han advertido la teatralidad intrínseca de la obra magna cervantina. Hace ya varias décadas, Sito Alba señalaba que “toda investigación seria sobre la gestación del *Quijote* debe tener muy en cuenta no sólo la literatura escrita, sino también la oral, y concretamente las manifestaciones escénicas” [1983: 23]. Y más recientemente, Madroñal ha llamado la atención sobre la presencia de acciones cómicas integradas y asimiladas en el propio conjunto narrativo, como, por ejemplo, la conversación sobre la cena que mantienen Sancho y un ventero en el capítulo 59 de la segunda parte del *Quijote*, que recuerda ineludiblemente a *La tierra jauja* de Lope de Rueda:

Cervantes está convirtiendo en materia narrativa un episodio que perfectamente podría haberse dramatizado en un entremés, tanto por su extensión, como por sus personajes o por el diálogo que mantienen, plagado de exclamaciones, giros populares y frases hechas, que funcionarían perfectamente en escena [2008: 269].

Pero además de estas piezas intercaladas o integradas en el *Quijote* y el *Persiles*, y de aparente origen dramático, la crítica también ha llamado la atención sobre la teatralidad de las *Novelas ejemplares*. Hace ya casi ocho décadas, Pfandl, al escribir su *Historia de la literatura nacional española en la Edad de Oro*, advertía la cercanía que existía entre la novela corta de Cervantes y su teatro, hermanados por la predilección por el diálogo, el estilo directo y la agilidad de palabra:

Cervantes me parece haber sido el primero que se percató del próximo parentesco de la novela con el drama, y creo que percibió también claramente que una diferencia esencial entre la novela y la novela corta se da en el dinamismo de la palabra y de la acción, en la viveza preponderante de esta última, en su adecuación a ser narrada más bien que leída. La abundancia del estilo directo no es ningún mero recurso para animar externamente la obra, sino que brota verdadera e íntimamente de aquella fuerza plástica que convierte la narración de las cosas en descripción de ellas, y que transforma la simple relación en el arte de encarnar, *representar* y hacer revivir lo que se evoca, que en lugar del anuncio introduce la visión directa de las cosas [1933: 351].

Partiendo de esa misma idea, Herrero García apuntó años más tarde la posibilidad de que las *Novelas ejemplares* de Cervantes fueran fruto de la reescritura de alguna de sus comedias perdidas, identificando *El amante liberal* con *La gran turquesa*, *La gitanilla* con *La única*, *La señora Cornelia* con *La bizarra Arsinda* y *La ilustre fregona* con *La Amaranta o la del Mayo* [1950]. Díaz Plaja también advirtió la teatralidad presente en las novelas cortas cervantinas y se preguntaba: “¿Pensaba, pues, Cervantes en *comediógrafo* al escribir sus novelas?” [1963a: 48]. E Yndurain llegó a desarrollar la hipótesis de que *Rinconete* y *Cortadillo* fuera en su origen un entremés, convertido en novela “por causa de la no petición de obras teatrales cervantinas por parte de los «autores»” [1966: 333].

Cierto es que en el siglo de Oro era bastante habitual que los dramaturgos redactaran un bosquejo o traza en prosa antes de componer sus comedias en verso, algo que ya apuntaba Lope de Vega en su *Arte nuevo* (“el sujeto elegido, escriba en prosa”) y de lo que se han conservado algunos ejemplos [Machado, 1925; Ferrer Vals, 1991; Vaccari, 2003]. Por eso, es posible pensar que Cervantes preparara algunos de estos

CAPÍTULO 3

bosquejos en prosa con la intención de venderlos a los autores de comedias, y que, al no poder hacerlo, los reconvirtiera en novelas cortas; bien para intercalarlas en otras de mayor envergadura (el *Quijote* o el *Persiles*), bien para darlas a la estampa en un tomo colecticio (*Novelas ejemplares*)¹⁰⁰.

No hay pruebas que permitan probar esta hipótesis, aunque la cronología de los hechos parece respaldarla. La última vinculación de Cervantes con el mercado teatral la encontramos en el documento contractual que firma el 5 de septiembre de 1592 con Rodrigo Osorio, por el que se compromete a componer y entregarle seis comedias; no sabemos si llegó a hacerlo, aunque parece poco probable, pues tan sólo quince días después es arrestado y encarcelado en Castro del Río por la supuesta venta ilegal de trigo. Cervantes va perdiendo poco a poco el favor de los autores de comedias, y, siendo así, parece lógico pensar que decidiese aprovechar todo el material dramático de que disponía (entremeses en prosa, que ya empezaban a estar en desuso, y algunas comedias, probablemente en forma de bosquejos en prosa) para aderezar la primera parte de las aventuras del hidalgo manchego (1605), ya fuera como historias intercaladas ajenas a la acción principal (como *El curioso impertinente* o la historia del capitán cautivo) o como parte de la misma (la disputa del baciuelmo, la aventura de los cueros de vino, etc.). Y también parece lógico pensar que al advertir el éxito que había cosechado *El curioso impertinente*, que ya en 1608 circulaba traducido al francés al margen del *Quijote*, decidiese llevar a la estampa doce novelas cortas más (1613)¹⁰¹, cuya teatralidad —plasmada esencialmente en el dinamismo y el (ab)uso del diálogo— también nos lleva a presuponer su género primigenio¹⁰²:

¹⁰⁰ Son numerosos los estudios que profundizan en las similitudes entre la comedia y la novela, al entender que van dirigidas a un mismo público y que reflejan la “curiosidad inextinguible del pueblo por la ficción [Morñigo: 1957: 48].

¹⁰¹ *Rinconete y Cortadillo* ya había sido nombrada en el *Quijote* de 1605 y apareció, junto a *El celoso extremeño*, en el conocido como manuscrito Porras, una miscelánea de textos copiados a mano por Francisco Porras de la Cámara, racionero de la Catedral de Sevilla, para entretener a su amo, el cardenal Niño de Guevara.

¹⁰² Cabe hacer notar, además, que el 3 de octubre de 1611 se había decretado el cierre de los corrales de comedias por la muerte de Margarita de Austria, esposa de Felipe II (que no volverán a abrir

En el caso de las *Novelas ejemplares* —pero en realidad de toda la producción cervantina— la relación entre el teatro y la novela no se limita a un nivel más o menos abstracto de modelos, sino que se concreta en elementos tan específicos como los recursos compositivos o estructurantes, o incluso en composiciones o construcciones textuales complejas [González, 2015, 34].

En Cervantes, los límites entre novela y teatro empiezan a desdibujarse, cobrando fuerza la hipótesis de que llegó a la primera a través del segundo. Doce son sus *Novelas ejemplares*, como doce eran también las comedias que contenían las famosas *Partes*, cuyas primeras ediciones datan de principios del XVII. Esta relación entre la comedia y la nueva narrativa cervantina fue rápidamente advertida por sus contemporáneos, que aseguraron que sus *Novelas ejemplares* eran en realidad “comedias en prosa” (Avellaneda, Suárez de Figueroa), pues seguían “los mismos preceptos que las comedias” (Lope de Vega).

Cervantes entiende entonces que ha perdido para siempre el tren de la comedia, tal y como confiesa su “alter-ego” a Pancracio de Roncesvalles en la *Adjunta al Parnaso* (1614), al reconocer que tiene seis comedias y seis entremeses que ningún autor quiere comprar¹⁰³ y que piensa dar a la estampa “para que se vea de espacio lo que pasa apriesa y se disimula, o no se entiende cuando las representan”; y así lo hace en 1617, ampliando el número de doce a dieciséis¹⁰⁴. Entre estas piezas encontramos algunas que

hasta el verano de 1613), por lo que cobra aún más sentido la idea de que Cervantes quisiera dar salida a su excedente dramático bajo el molde de la novela corta.

¹⁰³ En esa fecha, Nicolás de los Ríos había muerto y Gaspar de Porres, que en los últimos años había estado más próximo a Lope de Vega, ya se había retirado. En palabras de Florencio Sevilla y Antonio Rey Hazas, “nadie le aceptó para las tablas sus nuevas comedias, sin duda porque no se adaptaban a las exigencias del público, hecho ya al gusto del teatro lopeveguesco” [1998: X].

¹⁰⁴ Los motivos que da Cervantes para publicar sus textos dramáticos recuerdan a la que, por boca de Cide Hamete Berengeli, plantea en la segunda parte del *Quijote* para sacar las novelas interpoladas de su narración: “También pensó, como él dice, que muchos, llevados de la atención que piden las hazañas de don Quijote, no la darían a las novelas, y pasarían por ellas, o con priesa o con enfado, sin advertir la gala y artificio que en sí contienen, el cual se mostrara bien al descubierto cuando, por sí solas, sin arrimarse a las locuras de don Quijote ni a las sandeces de Sancho, salieran a luz. Y así, en esta segunda parte no quiso ingerir novelas sueltas ni pegadizas, sino algunos episodios que lo pareciesen, nacidos de los mismos sucesos que la verdad ofrece; y aún estos, limitadamente y con solas las palabras que bastan a declararlos; y, pues se contiene y cierra en los estrechos límites de la narración, teniendo habilidad, suficiencia y entendimiento para tratar del universo todo, pide no se desprecie su trabajo, y se le den alabanzas, no por lo que escribe, sino por lo que ha dejado de escribir” (II, 44). Esta misma idea aparece

CAPÍTULO 3

guardan ciertas similitudes temático-argumentales con otras obras suyas, lo que nos lleva a pensar, una vez más, que Cervantes trabajaba con borradores en prosa, que lo mismo transformaba en una novela (*El celoso extremeño*) que en un entremés (*El viejo celoso*) [Moríñigo, 1957: 41-61]. Al no poder novelizar el teatro, su solución fue hacer teatral la novela [Bergamín, 1959: 83].

Cervantes alcanza su plenitud como novelista gracias a su fracaso como dramaturgo. Un fracaso que le permitió, además, advertir las posibilidades de desarrollar un nuevo tipo de teatro, ya no pensado para las tablas sino para la lectura individual y silenciosa. Como advierte Blasco, el autor del *Quijote* “era ya consciente de la diferencia que empezaba a existir entre lector y espectador, de la misma manera que era consciente de que, en su tiempo, estaba naciendo un nuevo tipo de lector” [2006: 360].

Los elementos de teatralidad de la narrativa de Cervantes quedan patentes cuando advertimos que varias obras dramáticas de principios del XVII están inspiradas en sus novelas, como algunas comedias de Guillén de Castro (*Don Quijote de la Mancha*, *La fuerza de la sangre* y *El curioso impertinente*), el *Entremés de los invencibles hechos de don Quijote* de Francisco de Ávila o el anónimo *Entremés de los romances* (que algunos —en nuestra opinión, equivocadamente— han fechado con anterioridad a la obra cervantina). Igualmente, son muchas las adaptaciones teatrales que se han hecho del *Quijote* desde poco después de su publicación hasta ahora¹⁰⁵, así como de algunas de sus *Novelas ejemplares*, como *Rinconete y Cortadillo* o *El celoso extremeño*.

también en el *Persiles*, en donde explica: “porque las lecciones de los libros muchas veces hacen más cierta experiencia de las cosas, que no la tienen los mismos que las han visto, a causa que el que lee con atención, repara una y muchas veces en lo que va leyendo, y el que mira sin ella no repara en nada, y con esto excede la lección a la vista” (III, 8). Cabe advertir que Lope de Vega llega a la misma reflexión años más tarde, pues en el prólogo a su *Parte XII* leemos: “Quedo consolado, que no me pudrirá el vulgo como suele; pues en tu aposento, donde la has de leer, nadie consentirás que te haga ruido, ni que te diga mal de lo que tú sabrás conocer, libre de los accidentes del señor que viene tarde, del representante que se yerra, y de la mujer desagradable por fea y mal vestida”.

¹⁰⁵ “Felipe Pérez Capo cita *The Knight of the burning pestle*, estrenada en 1611 en Londres, como la primera obra teatral basada en el *Quijote*, y se tiene constancia de unos festejos en 1607 en honor del virrey en la localidad de Pausa (Perú) en donde aparecieron el hidalgo y su escudero caricaturizados; dos

3.2. TRAS LA HUELLA DE LA TEATRALIDAD CERVANTINA

El teatro fue para Cervantes su gran obsesión [Florit Durán, 2009]. Por eso, aunque el idiolecto debe ser la evidencia clave para determinar si una obra dubitada pudo haber sido escrita o no por Cervantes (y con qué grado de probabilidad), la *teatralidad* de toda su producción se erige como otra de las huellas identificativas de la autoría cervantina.

Cuando hablamos de teatralidad no nos referimos únicamente a obras que aborden el tema del teatro desde dentro del teatro (metateatralidad), que debatan cuestiones de preceptiva dramática o viertan opiniones sobre empresarios teatrales o actores y actrices del momento; cuando hablamos de *teatralidad* nos referimos también —y sobre todo— al concepto del ‘teatro menos el texto’ que defendió Barthes y que Villegas Morales define como un discurso en el que se privilegia “la construcción y percepción visual del mundo” [2000: 50]. Más allá de la posibilidad de que el origen de su narrativa pueda derivar de piezas teatrales perdidas o borradores en prosa de comedias, lo cierto es que la condición teatral de Cervantes invade toda su producción literaria: desde los recursos cómicos presentes en sus entremeses, heredados de los espectáculos de Lope de Rueda y de los cómicos del arte, hasta la predilección por el estilo directo de sus *Novelas ejemplares*, pasando por la visualidad de varias escenas de su obra magna o por las historias intercaladas que aderezan las aventuras de Quijote y Sancho y de Persiles y Segismunda, que recuerdan a la organización del espectáculo teatral del Siglo de Oro.

La función-autor en Cervantes está íntimamente ligada a su naturaleza dramática y, por ello, ha de ser una seña de identidad complementaria en la labor de identificación de la posible autoría cervantina de los textos dubitados.

años antes, los personajes cervantinos ya habían intervenido en unas fiestas en Valladolid con motivo del nacimiento del príncipe Felipe” [Fernández Ferreiro, 2012, 187].

CAPÍTULO 4

LOS ENTREMESSES ATRIBUIDOS A MIGUEL DE CERVANTES SAAVEDRA

“y otras obras que andan por ahí descarriadas y, quizá, sin el nombre de su dueño” (Cervantes, Prólogo de las *Novelas Ejemplares*).

4.1. LAS ATRIBUCIONES CERVANTINAS

Uno de los grandes retos de la historia literaria sigue siendo el de delimitar y establecer qué fue lo que escribió Miguel de Cervantes Saavedra. La crítica acepta como obras indubitadas la *Galatea*, las dos partes del *Quijote*, el *Viaje del Parnaso*, las *Novelas ejemplares*, las *Ocho comedias y ocho entremeses nuevos, nunca representados*, la novela póstuma *Los trabajos de Persiles y Segismunda* y los poemas encomiásticos que figuran al frente de las obras de otros escritores (y amigos), como Pedro de Padilla, Juan Rufo Gutiérrez, Alonso de Barros o Gabriel López Maldonado¹⁰⁶. También admite como suyas las comedias perdidas que figuran en el contrato que firma con el *autor* Gaspar de Porres en 1585 (*La Confusa* y el *Trato de Constantinopla y muerte de Selim*)

¹⁰⁶ La aceptación de la paternidad cervantina de todas estas obras es general, pero no unánime. Aylward defendió —a nuestro parecer con argumentos de muy poco peso— la idea de que Cervantes había plagiado *Rinconete y Cortadillo* y *El celoso extremeño* de un manuscrito de la Biblioteca Colombina, hoy perdido, compilado por Francisco Porras de la Cámara, y que se conoce comúnmente como el manuscrito Porras [1982].

o las que nombra en la Adjunta de su *Viaje del Parnaso* y en el prólogo a sus *Ocho comedias y ocho entremeses* (*Los tratos de Argel, La destrucción de Numancia* o *La Numancia, La gran Turquesca, La batalla naval, La Jerusalén, La Amaranta o la del Mayo, El bosque amoroso, La única y La bizarra Arsinda*)¹⁰⁷, así como los proyectos de futuro a los que hace referencia en la dedicatoria del *Persiles*, y que probablemente dejó inconclusos (*Las semanas del jardín, El famoso Bernardo* y la segunda parte de *La Galatea*)¹⁰⁸.

Pero al margen de esta nómina cierta, y basándose en la afirmación del propio Cervantes de que hay obras suyas que “andan por ahí descarriadas y, quizá, sin el nombre de su dueño”, son muchos los que le han atribuido otras tantas obras, engrosando el listado con más de medio centenar de poemas, varias narraciones en prosa, algunas cartas y unas cuantas piezas teatrales [Rius, 1895-1905, Ford y Lansing, 1931, Astrana Marín, 1948-1958, VII: 751-767, Avalor-Arce, 1973, Eisenberg, 1990; Montero Reguera, 1995] Ahora bien, ¿qué fiabilidad podemos otorgar a tales atribuciones?

Empecemos por la poesía. Aunque la gran mayoría de críticos dan por segura la autoría cervantina de varios poemas sueltos —como el soneto de juventud “Serenísima reina, en quien se halla”, que debió componer en 1567 con motivo del nacimiento de la Infanta Catalina Micaela, segunda hija de Felipe II e Isabel de Valois—, lo cierto es que hay otros tantos de dudosa atribución, entre los que se encuentran varias canciones,

¹⁰⁷ Fitzmaurice-Kelly sospechó que *La batalla Naval* se conservaba manuscrita en la librería del conde-duque Gaspar de Guzmán bajo el rótulo *La batalla naval y el suceso de la guerra del sultán Amurates*, manuscrito que desgraciadamente se ha perdido [Cotarelo Valledor, 1915: 680]. Por su parte, Cotarelo Valledor planteó la posibilidad de que Cervantes recopilara en su volumen de 1615 algunas de las comedias que había compuesto en la década de los 80, basándose en la ‘cercanía’ entre los títulos de tres de las publicadas y tres de las que figuran en la nómina de la *Adjunta del Parnaso*; según esta hipótesis (ciertamente arriesgada) *La casa de los celos, La gran sultana* y *El laberinto de amor* serían la refundición de *El bosque amoroso, La gran turquesa* y *La confusa* [1947-1948].

¹⁰⁸ Era una práctica habitual que los dramaturgos del Siglo de Oro utilizaran los prólogos, las dedicatorias y otros paratextos tanto para dar la nómina de sus composiciones, y reivindicar así la autoría de los textos perdidos o usurpados, como para desvincularse de las falsas atribuciones. Lo vemos, además de en Cervantes, en Lope de Vega, Rojas Zorrilla, Ruiz de Alarcón, Calderón de la Barca, Tirso de Molina o Antonio Enrique Gómez, entre otros.

algunos sonetos y unos pocos romances [Montero Reguera, 1995: 67-71]. La gran mayoría de estas adjudicaciones datan de la segunda mitad del siglo XIX y pocas han sido revisadas en los últimos años, a excepción de la *Epístola a Mateo Vázquez*, que, a la luz de los últimos estudios, ya no debe ser considerada una falsificación [Stagg, 2003; Sánchez-Molero, 2007]. Recientemente, Alvar Ezquerro ha apuntado la posible paternidad cervantina del soneto “Debajo de esta dura piedra helada” y de la copla castellana “Aquí yace sepultada” [2014: 315], y Gracia la del soneto mutilado “Ocupa breve término de tierra” [2016: 200-202], aunque ambos con argumentos lingüísticamente endebles.

Mayor atención han recibido las atribuciones en prosa. El primer texto en ser vinculado a Cervantes fue *La tía fingida*, y aunque esa filiación ha sido defendida por muchos, también ha sido rechazada por otros [Sáez, 2018]¹⁰⁹. Y lo mismo ocurre con la *Tercera parte de las cosas de la Cárcel de Sevilla*, como continuación a las dos de Cristóbal de Chaves [Eisenberg, 1990: 489]; la *Relación de lo sucedido en la ciudad de Valladolid desde el punto del felicísimo nacimiento del Príncipe Don Felipe Dominico Víctor*, con más defensores que detractores [Eisenberg, 1990: 490-492]¹¹⁰; el *Diálogo de Cilenia y Selanio sobre la vida en el campo*, que Adolfo de Castro identificó con la segunda parte de la *Galatea* y Schevill y Bonilla con las *Semanas del jardín* [Blasco, 2010]; o la *Topografía e historia general de Argel*, con cuatro posibles candidatos a autor [Marín Cepeda, 2010]. Aunque todas estas atribuciones han despertado el interés reciente de muchos cervantistas, lo cierto es que a día de hoy no hay una opinión unánime al respecto.

Lo que ya nadie hoy parece cuestionar es que Cervantes no es el autor del *Quijote* apócrifo, a pesar de no tener claro quién se esconde en realidad tras el pseudónimo de

¹⁰⁹ Sáez ha preparado la edición de este texto para la colección ‘Letras Hispánicas’ de Cátedra. En el estudio introductorio da cuenta de los argumentos esgrimidos por la crítica a favor y en contra de la paternidad de Cervantes, así como de las alternativas propuestas (desde Francisco de Porras hasta Alonso de Salas Barbadillo o Luis Vélez de Guevara) [2018: 13-23].

¹¹⁰ Véase la edición de Marín Cepeda [2006].

Alonso Fernández de Avellaneda¹¹¹, ni el de los *Exemplos de Catón, sacados a la luz agora de nuevo por el licenciado Miguel de Cervantes*, publicados en 1609 por Sebastián de Cormellas, y que son en realidad obra de un autor homónimo [Glaser, 1954]. Y que tampoco le pertenecen una serie de textos en prosa de muy dudosa procedencia, que resultaron ser burdas falsificaciones; nos referimos a *Unos borradores cervantescos, transcritos y comentados por Manuel Gómez-Moreno* [Astrana Marín, 1948-1958: VII, 763], a los *Capítulos de mi don Quijote de la Mancha, no publicados en España* [Eisenberg, 1990: 483] y al *Buscapié*, escrito en realidad por Adolfo de Castro¹¹².

También se han atribuido a Cervantes algunas cartas, como las dirigidas a Antonio Veneziani, al cardenal don Bernardo de Sandoval y Rojas, a don Diego de Astudillo Carrillo o don Juan de Acevedo y Fuentes, así como una misiva, sin fecha ni destinatario, editada por primera vez en Nantes en 1901 [Avalle-Arce, 1973: 405]; de todas ellas, la que más atención ha generado es la supuesta carta autógrafa dirigida al Cardenal Sandoval y Rojas, publicada por primera vez en el periódico *La ibérica* en 1961, y cuya originalidad fue refutada con argumentos muy sólidos por Rodríguez-Moñino [1962: 89]. E incluimos también en este apartado —aunque no se trate de una

¹¹¹ Desde el siglo XVIII se han barajado infinidad de propuestas: los dominicos fray Luis de Aliaga y Juan Blanco de Paz, Quevedo, Lope de Vega, Tirso de Molina, Agustín de Rojas, Pedro Liñán de Rianza Francisco López de Úbeda, Alfonso Lamberto, Juan Ruiz de Alarcón, Alonso Castillo de Solórzano, fray Luis de Granada, Alonso Fernández Zapata, fray Cristóbal de Fonseca, Alonso Jerónimo de Salas Barbadillo, fray Alonso Fernández, Juan Martí, Alonso Pérez de Montalbán, el clérigo cordobés Juan Valladares, Alonso de Ledesma, el conde de Lemos en colaboración con Mira de Amescua y Gabriel Leonardo de Albión y Argensola, Ginés Pérez de Hita, Baltasar Eliseo de Medinilla, fray Hortensio Félix Paravicino, Cristóbal Suárez de Figueroa, alguno de los hermanos Leonardo de Argensola e incluso el propio Miguel de Cervantes. De un tiempo a esta parte, el debate se centra en el soldado aragonés Jerónimo de Pasamonte [Riquer, 1988 o Martín Jiménez, 2005], el dominico vallisoletano Baltasar de Navarrete [Blasco, 2005] y el poeta arriáicense José de Villaviciosa [Rodríguez López-Vázquez, 2011].

¹¹² De todas las falsificaciones cervantinas, la del *Buscapié* es la que más comentarios ha generado. En 1847, Adolfo de Castro dice haber encontrado este ensayo sobre la primera parte de el *Quijote* en un baratillo de Cádiz, procedente de la biblioteca del duque de Lafoes. A pesar de que George Ticknor, Cayetano Alberto de la Barrera o Bartolomé José Gallardo, entre otros, apuntaran tempranamente que se trataba de una falsificación —reconocida más tarde por el propio Castro—, lo cierto es que el *Buscapié* se incluye como apéndice en las ediciones del *Quijote* hasta 1976, alcanzando enorme popularidad y trascendencia internacional. Para profundizar en este asunto, léase Romero Ferrer y Vallejo Márquez [2003-2004].

carta— la aprobación de la segunda parte del *Quijote*, firmada por Márquez Torres, tras la que algunos han creído ver la pluma y el ingenio de Cervantes [Eisenberg, 1990: 479].

En relación con el teatro, el caso de las atribuciones cervantinas también ha generado múltiples estudios y controversias. A día de hoy nadie parece cuestionar que sean de Cervantes la tragedia *El cerco de Numancia* y la comedia *El trato de Argel*, descubiertas y publicadas por Antonio de Saca en el siglo XVIII, a pesar de que el argumento —la semejanza de título con dos de sus obras perdidas— no soportaría las exigencias de los estudios actuales¹¹³. Y también es bastante aceptada la tesis de Stefano Arata de que la comedia *La conquista de Jerusalén por Godofre de Bullón* sea en realidad la *Jerusalén* perdida que nombra Cervantes en la Adjunta de su *Viaje del Parnaso*, pues, además del título, existen similitudes en la métrica, los personajes, las acotaciones y el uso de las figuras morales y alegóricas [1992 y 1997]¹¹⁴. Menor unanimidad encontramos en el resto de atribuciones teatrales: un auto sacramental titulado *Las cortes de la muerte*; las comedias *María la de Esquivias* y *La toledana*, atribuidas a Cervantes a finales del siglo XIX y a las que la crítica posterior apenas ha prestado atención; la comedia *La soberana Virgen de Guadalupe, y sus milagros y grandezas de España*, cuya paternidad fue rebatida hace unos años por Montero Reguera [1998]; y una decena de entremeses, tanto en prosa como en verso, de los que daremos detalle en los siguientes epígrafes.

En definitiva: la nómina de atribuciones cervantinas se conforma como un *totum revolutum* que integra varias falsificaciones (*El buscapié* o la “Carta al cardenal D. Bernardo de Sandoval y Rojas”), ciertas asignaciones ridículas (el *Quijote* de Avellaneda o los *Exemplos de Catón*), algunas atribuciones dudosas (la *Epístola a Mateo Vázquez* o el *Entremés de Los habladores*), otras más probables (como *La tía*

¹¹³ Estas atribuciones quedan en entredicho si pensamos en obras del Siglo de Oro con títulos homónimos, como las comedias *Allá se verá* (de Rosete y Matos Frago) y *La rosa de Alejandría* (de Rosete y Vélez de Guevara) o el entremés *El retablo de las maravillas* (de Cervantes y Quiñones de Benavente).

¹¹⁴ La autoría cervantina de esta comedia ha sido refrendada recientemente por Cerezo Soler y Calvo Tello, tras compararla con otras diecisiete piezas teatrales en verso de siete autores distintos (entre ellos, Cervantes) desde las posibilidades metodológicas que ofrecen los análisis estilométricos [2020].

fingida, La Conquista de Jerusalén o el Diálogo entre Cilenia y Selenio) y dos unánimemente aceptadas (El cerco de Numancia y El trato de Argel).

4.2. HISTORIA CRÍTICA DE LOS ENTREMESSES ATRIBUIDOS A CERVANTES

En la *Adjunta al Parnaso*, publicada en 1614, Cervantes confiesa tener en sus cajones seis comedias “con otros seis entremeses” que pensaba dar a estampa “para que se vea de espacio lo que pasa apriesa”. Tan sólo un año después saca a la luz sus *Ocho comedias y ocho entremeses nuevos, nunca representados*, quejándose en el prólogo de que se ha visto obligado a vendérselos a un librero porque ningún autor de comedias ha querido representarlos. Esa edición, de la que se conservan varios ejemplares, contiene los siguientes entremeses: *El juez de los divorcios, El rufián viudo, La elección de los alcaldes de Daganzo, La guarda cuidadosa, El vizcaíno fingido, El retablo de las maravillas, La cueva de Salamanca y El viejo celoso*¹¹⁵.

En 1617, sólo un año después de la muerte de Cervantes, se publica en Madrid la *Séptima parte de las comedias de Lope de Vega* (Imagen 2); de entre los entremeses anónimos que contiene, llama notablemente la atención el de *Los habladores* por su extraordinaria fuerza cómica y su brillantez de estilo. Fueron varios los que en el siglo XIX aseguraron tener o haber visto ediciones de este entremés en las que figuraba Cervantes como autor indiscutible¹¹⁶, lo que desató las primeras polémicas en torno a su autenticidad y los primeros deseos por delimitar el verdadero corpus entremesil cervantino.

¹¹⁵ Sobre el número definitivo de entremeses, léase el apartado “Ocho eran ocho” de la introducción que hace Sáez a su edición de los *Entremeses* de Cervantes [2020].

¹¹⁶ Fernández de Navarrete cita una impresión suelta hecha en Sevilla en 1624, que nadie ha logrado ver, y Fernández Guerra dice poseer una edición de Cádiz de 1646, aunque todo parece indicar que se trata de un apócrifo.

En las siguientes páginas daremos noticia de los principales estudios que han abordado este asunto, desde los que en época decimonónica atribuyeron a Cervantes una serie de entremeses anónimos, basándose fundamentalmente en la coincidencia temática o de estilo, hasta los que tiempo después revisaron estos escritos y apuntaron la falta de validez de los razonamientos empleados.

4.2.1. El “fetichismo cervantista” de la segunda mitad del XIX

En el verano de 1845, Aureliano Fernández Guerra —siguiendo los consejos de José María de Álava y José Fernández y Velasco— decide estudiar un códice de la Biblioteca Colombina de Sevilla con el objetivo de encontrar “algo nuevo relativo a Quevedo y Cervantes” [1863: 1245]. El manuscrito, escrito en la primera mitad del siglo XVII por una única mano (de la que desconocemos su identidad), lleva por título *Poesías y relaciones varias*, y está formado por siete volúmenes en 4º español (signatura 82-3-38, antiguamente AA-141-4)¹¹⁷. Tras analizar las trece obras de que se compone el cuarto tomo de este códice, Fernández Guerra no duda en atribuir a Cervantes *La tía fingida*, la “Carta a D. Diego de Astudillo Carrillo” y la *Tercera parte de las cosas de la Cárcel de Sevilla* [1863: 1257]. Además, al considerar que la tercera parte de la relación la compuso Cervantes durante el tiempo que estuvo en prisión en la capital andaluza, le parece lógico pensar que también escribiera en ese tiempo el entremés titulado *La cárcel de Sevilla*, inserto en la *Séptima parte de las comedias de Lope de Vega*, y que, en consecuencia, también le pertenezcan el resto de entremeses que se incluyen en este volumen (*Los habladores* y *El hospital de los podridos*):

A toda luz le pertenecen los tres sainetes anónimos que aderezan esta *Séptima parte*, a saber: *Los habladores*, *La cárcel de Sevilla* y *El Hospital de los Podridos*. Espíritu, genio, estilo, son unos mismos en todos: por algo se pusieron juntos. Impreso en Cádiz, año de 1646, con el ilustre nombre de su autor, poseo yo *Los Habladores*; ¿quién sabe si allí entonces,

¹¹⁷ En su día Bartolomé José Gallardo creyó leer en el lomo de este cuarto tomo "*Poesías Palacio varias*" en vez de "*Poesías y Relaciones varias*", y esta lectura errónea lleva a Fernández Guerra a creer equivocadamente que el colector de estos tomos fue el racionero Francisco Porras de la Cámara.

sueltos también, y también con la marca de su verdadero dueño (reproduciendo ediciones más ó menos antiguas y ya agotadas), volvieron a darse a la estampa los otros dos sainetes? Que el primero es hijo legítimo del númen de Cervantes, hoy no ofrece ocasión de litigio: siendo esto así, y estando el fallo ejecutoriado, forzosamente hay que hacerlo extensivo a las otras dos piecitas [1863: 1371].

Aunque los argumentos que da Fernández Guerra para atribuir estos tres entremeses al autor del *Quijote* son muy endebles, sus hipótesis llaman pronto la atención de otros cervantistas del momento. Así, el 6 de agosto de 1867, José María Asensio le escribe una carta para informarle que en el sexto tomo de ese mismo códice (signatura 82-3-40, antiguamente AA-141-6) se recogen catorce piezas entremesiles de gran calidad literaria¹¹⁸ y, recordando el orden en que fueron editados los entremeses de la *Séptima parte de las comedias de Lope de Vega*, arguye que “es raro, en verdad, que a continuación [de *Los habladores*] venga también en la copia el entremés de *La cárcel de Sevilla*” [1902: 40]¹¹⁹, lo que parece confirmar que ambos sean de Cervantes. Tras analizar el resto de entremeses del códice, Asensio también aprecia rasgos estilísticos propios de Cervantes en *Melisendra*, *Durandarte y Belerma* y *Los refranes*.

No quiero yo incurrir en exageración al atribuir á Cervantes obras que no llevan su nombre. Sin embargo, si se reconocen, como creo, en los entremeses de *Melisendra*, y de *Durandarte y Belerma*, aunque burlescos ambos, algunos rasgos de su pluma, y se suponen escritos durante la permanencia de autor en Sevilla (á cuya suposición daría margen y probabilidad de acierto el encontrarlos incluidos en este códice sevillano), importantísimos serían para indicar el aprecio que en el ánimo de Cervantes iban teniendo las invenciones caballerescas, y el punto de vista del ridículo bajo el cual empezaba á considerarlas. No quiero hacer interminable esta carta, y por eso no me decido á entrar en otras apreciaciones,

¹¹⁸ Este códice sevillano en cuarto consta de 112 folios, de 201 x 145 mm. La letra es del siglo XVII y se caracteriza por la abundancia de acotaciones y el andalucismo lingüístico del copista. Contiene catorce entremeses sin nombre de autor, siete de los cuales han sido atribuidos alguna vez a Cervantes: (1) *El ingeniosos y ridículo entremés del examinador Miser Palomo*, fols. 1r-7v; (2) *Entremés de los Habladores*, fols. 8r-13r; (3) *Entremés famoso de La cárcel de Sevilla*, fols. 13r-19v; (4) *Entremés de los mirones*, fols. 20r-33v; (5) *Entremés del sacristán Soguijo*, fols. 34r-40v; (6) *Entremés de La villana de Getafe y carreteros de Madrid*, fols. 41r- 44v; (7) *Entremés de la endemoniada fingida y chistes de Bacallao*, fols. 45r-52v; (8) *Famoso entremés de Melisendra*, fols. 53r-62r; (9) *Entremés del rey Cachumba de Motril y la Infanta Palancona*, fols. 62v-73v; (10) *Entremés de Durandarte y Belerma*, fols. 74r-91v; (11) *Entremés de Doña Justina y Calahorra* fols. 92r-97r; (12) *Entremés del doctor Zarabulleque*, fols. 98r-101v; (13) *Entremés del zurdo*, fols. 102r-107r; y (14) *Entremés de refranes*, fols. 107v-112r (Imágenes 5-11).

¹¹⁹ Citamos por la reimpresión que se recoge de esta carta en *Cervantes y sus obras* [1902].

limitándome a incluir á V. copia exactísima del *Entremés de refranes*, que no solamente en mi concepto, sino también en el de nuestros buenos amigos D. José María de Alava y D. Juan E. Hartzenbusch, en nada desdice del giro y distribución de los de Cervantes [1902: 41].

Unos años más tarde, concretamente en 1874, el gaditano Adolfo de Castro publica un trabajo titulado *Varias obras inéditas de Cervantes*, en donde vuelve a reflexionar sobre la autoría de los entremeses que figuran en el sexto tomo del código colombino; según su criterio, *Los habladores*, *La cárcel de Sevilla*, *Los mirones* y *Los refranes* son de Cervantes; *Doña Justina* y *Calahorra* “parece también obra suya, pero escrita en los últimos años” [1874: 20]. Dice tener constancia de que José María Asensio ha estudiado ese mismo tema, pero que expuso sus conclusiones en una publicación tan rara que no ha podido hacerse con ella. Ahora bien, si comparamos las conclusiones a las que llegan uno y otro, advertimos que:

- los dos consideran que *Los habladores*, *La cárcel de Sevilla* y *Los refranes* han salido de la pluma de Cervantes (coincidencia muy significativa si tenemos en cuenta que ambos estudios fueron hechos de manera independiente);
- Asensio apunta la posibilidad de que *Melisendra* y *Durandarte* y *Belerma* sean también cervantinos; Castro, por su parte, recuerda que *Melisendra* no tiene por qué ser de Lope de Vega, aunque aparezca impreso junto a sus comedias¹²⁰, y que *Durandarte* y *Belerma* puede ser de Mira de Amescua;
- Castro considera cervantinos *Los mirones* y *Doña Justina* y *Calahorra*; Asensio no dice nada relevante del primero, pero sí del segundo, que considera obra de Quevedo.

Castro también reflexiona sobre otro entremés anónimo que, aunque no viene recogido en el código de la Colombina, no duda en atribuir a Cervantes. Se trata del

¹²⁰ “Se halla en las comedias famosas del poeta Lope de Vega Carpio, recopiladas por Bernardo Grassa, dirigidas á D. Valeriano Boyl. Van añadidos en esta impresión muchos entremeses [...] No es decir que eran de Lope los entremeses, sino de los que solian representarse con ellas” [1874: 10].

Entremés de los romances, publicado en la *Parte tercera de las comedias de Lope de Vega y otros autores*, y que para él es el bosquejo de la primera salida del ingenioso hidalgo: “Cervantes hizo lo que los grandes pintores: trazó un borrón o un ligero dibujo de un gran cuadro, primitivo pensamiento que luego desarrolló en un libro admirable” [1874: 140].

Tan solo un mes después de que viese la luz el libro de Adolfo de Castro, un jovencísimo Marcelino Menéndez Pelayo escribe cinco artículos para la revista estudiantil *Miscelánea científica y literaria*, en los que también reflexiona sobre las obras de ese códice de la Biblioteca Colombina¹²¹. En materia de entremeses sigue fielmente las conclusiones de Castro, adjudicando a Cervantes *Los habladores* y *El hospital de los podridos* (primer artículo), *La cárcel de Sevilla*, *Los mirones*, *Doña Justina* y *Calahorra* y *Los refranes* (segundo artículo) y *El entremés de los romances* (tercer artículo); además, al igual que éste, dice que *Durandarte* y *Belerma* parece ser obra de Mira de Amescua y que *Melisendra* no tiene autor conocido.

Para terminar con la nómina de entremeses que fueron atribuidos al autor del *Quijote* en esta época de ciega adoración a todo lo cervantino, tenemos que hablar de *Ginetilla ladrón*, del que da noticia José María Asensio en unas cartas dirigidas a los cervantistas José de Palacio Vitery y Mariano Pardo de Figueroa, que se publicaron el 25 de octubre de 1874 en la *Revista Europea* [1874]. Según cuenta Asensio a sus dos destinatarios, tiene constancia de que José María de Álava poseía un cuaderno manuscrito en el que figuraba este entremés con letra del siglo XVII, y que éste identificaba con el bosquejo de Sancho como gobernador de la ínsula de Barataria:

El argumento capital de Pepe Álava, para sospechar que se debiese á la pluma de Cervantes, estaba (á más del nombre del protagonista, que le recordaba a Pasamonte) en un cuento

¹²¹ Los cinco artículos, titulados “Obras inéditas de Cervantes”, aparecen publicados por primera vez en la revista *Miscelánea Científica y Literaria* entre junio y septiembre de 1874, y más tarde se reimprimen en el primer tomo de sus *Estudios y discursos de crítica histórica y literaria* (Madrid, CSIC, 1941), por el que citaremos. Menéndez Pelayo es consciente de que son muchos los que miran con recelo las afirmaciones de Castro, pues su falsificación del *Buscapié* le ha hecho perder autoría y credibilidad; sin embargo, en su opinión, el códice de la Biblioteca Colombina “no se trata de un supuesto manuscrito, como el del *Buscapié*, sino de obras auténticas conservadas en nuestras bibliotecas públicas” [1941: 300].

cuya estructura, lenguaje y versos encontraba iguales en todo á otro de *La elección de los alcaldes de Daganzo* [1874: 557-558].

Lamentablemente esta pieza no ha llegado hasta nosotros (probablemente esté perdida entre los papeles del Seminario de Vitoria), y sólo tenemos el fragmento que transcribe Asensio para probar su cercanía con otro de *La elección de los alcaldes de Daganzo*. Asensio no apoya la teoría de Álava, pues “así como podrían ésta y otras obrillas ser primera idea ó comienzo de otras mayores, también podrían ser imitaciones, traslados, recuerdos más o menos fieles, copias mejor ó peor disfrazadas de Cervantes” [1902: 76], y aprovecha este mismo argumento para desmentir la paternidad cervantina del *Entremés de los romances* —defendida por Castro— y para replantearse la del *Entremés de refranes*, que él mismo había apuntado años atrás:

Esta duda me asaltaba cuando en 1867 hice imprimir el de *Refranes*, y por eso no me atreví á atribuirle decididamente á Cervantes. ¿Quién será capaz de asegurar que esas obrillas son bocetos, y no copias de cuadros anteriores? [...] Prudente es, y aún necesario, andarse con piés de plomo en estas adopciones y porfijamientos [1874: 558].

Y en esta misma línea se manifiesta Ramón León Maínez, al asegurar que en la atribución de ciertos entremeses a Cervantes “influiría más la opinión de algún recitante, el criterio equivocado de algún editor, o el capricho de algún apasionado de nuestro escritor preclaro, que la razón, la lógica o el pleno convencimiento” [1876: 330].

A partir de este momento, el tema de los entremeses atribuidos a Cervantes cae en descrédito, y los intelectuales del siglo XX ya no parecen comulgar con ruedas de molino. Emilio Cotarelo y Mori, Armando Cotarelo Valledor, Dámaso Alonso, Agustín del Campo o Eugenio Asensio, entre otros, consideran que es inaceptable seguir atribuyendo un texto a Cervantes por las semejanzas de tema y estilo con alguna de sus obras reconocidas, pues bien es sabido que los autores de nuestra literatura barroca bebieron de fuentes comunes y emplearon la refundición temático-argumental en gran parte de sus creaciones; y más aún hacerlo por su calidad literaria, pues no todos los buenos textos son obra del autor del *Quijote*. El fervor por la figura de Cervantes y por los descubrimientos literarios, que había imperado durante todo el siglo XIX, y más intensamente en la segunda mitad de la centuria, dan paso a un escepticismo propio de

la crisis de fin de siglo, alentado por el descrédito de las falsificaciones y el interés de muchos investigadores por conseguir fama y prestigio.

4.2.2. Del entusiasmo al escepticismo en el asunto de los entremeses atribuidos a Cervantes

En 1911, Emilio Cotarelo y Mori publica su famosa *Colección de entremeses, loas, bailes, jácaras y mojigangas*, una antología de consulta obligada para todos aquellos que deseen adentrarse en el estudio del teatro menor español de los siglos XVI, XVII y XVIII. Al referirse a *Los romances*, *Los mirones*, *Los refranes* y *Doña Justina y Calahorra*, argumenta que “han sido adjudicados [...] a Cervantes sin más razón que las semejanzas de lenguaje y estilo que los que lo intentaban descubrían en ellos” y que “mezclados ahora con otros anteriores y posteriores, se ve que no se diferencian gran cosa de éstos, como tampoco unos y otros se diferencian mucho de los auténticos de Cervantes en las cualidades señaladas, porque en aquel tiempo era comunísimo [sic] escribir bien” [1911: vol. I, p. LXIX].

Tan solo cuatro años después, Armando Cotarelo Valledor publica su famoso estudio sobre el teatro de Cervantes. Al hablar de los entremeses, asegura que “es muy probable que algunos de los anónimos de fines del siglo XVI, como *el famoso de Mazalquivi*¹²², sean suyos, y verosímil que igualmente le pertenezcan otros más modernos, como el de *Los habladores*” [1915: 70], pero no cree hayan salido de su pluma casi ninguno de los que se le han atribuido hasta la fecha:

Salvo dos de estas piecillas, *Los dos habladores* y *Los mirones*, las demás no vacilamos en calificarlas de indignas del Príncipe de los Ingenios [1915: 73]

Cotarelo Valledor considera que ha sido “la ciega adoración que el *fetichismo cervantista* produjo en sus adeptos” la que ha hecho que muchos estudiosos adjudiquen

¹²² En otro momento señala que el entremés de *Mazalquiví* es “semejante por su corte al *Rufián viudo*” [1915: 597], pero no lo hace con el afán de atribuírselo a Cervantes, sino para demostrar que son muchos los entremeses que comparten similitudes entre sí.

a Cervantes cualquier obra anónima de mérito que se halle, como si éste fuese el único autor notorio de nuestras letras [1915: 685]. Por esto, en varios pasajes de su obra reclama un mayor rigor académico a la hora de hacer este tipo de elucubraciones:

No negaré yo que en una obra de cierta extensión, y tratándose de determinados autores, el atento estudio del estilo pueda servir de mucho a semejantes efectos; tampoco desconozco que las similitudes de la forma contribuyen muy poderosamente a reforzar otros argumentos; pero leer el nombre de un autor por fortuitas semejanzas accidentales y por el uso de frases y palabras que eran patrimonio de todos, no me parece cosa seria ni aceptable [1915: 687].

Un año más tarde, cuando Adolfo Bonilla y San Martín edita los entremeses de Cervantes para la Asociación de la Librería de España, decide incluir el *Entremés de los habladores* a modo de complemento de los auténticos, a pesar de que reconoce que es de “dudosa autenticidad” [1916: XXXIX]. Menor credibilidad le merecen las atribuciones de *El hospital de los podridos* y *La cárcel de Sevilla*, pues “las razones esgrimidas por Fernández Guerra, suponiendo que mereciesen este nombre, distan mucho de ser convincentes” [1916: XXXVII], así como las de Asensio y Castro. Y finaliza su introducción con una reflexión interesantísima (además de premonitoria) sobre el camino que deben seguir este tipo de estudios en el futuro:

Afirmar que los datos estilísticos, científicamente recogidos, bastan para resolver el problema de si una determinada producción debe atribuirse a tal o cual autor, es también aventurado, porque en la obra literaria, además de estilo, hay lenguaje, ritmo, pensamiento y estructura, amén de otros elementos, todos los cuales deben entrar en cuenta. Para que filológicamente pudiera llegarse a algún resultado en este sentido, respecto de las obras cervantinas, sería preciso empezar por un escrupuloso trabajo de estadística lingüística, análogo al que han realizado, por ejemplo, respecto de Platón, Lewis, Campbell, Dittenberg y Lutoslawski. Y este trabajo apenas ha comenzado a hacerse en nuestros días [1916: XXXIX]

Tras unos años de silencio en torno a las atribuciones entremesiles cervantinas, Dámaso Alonso decide reunir en una nueva edición cinco de los entremeses que alguna vez habían sido prohijados a nuestro escritor más universal: *El hospital de los podridos*, *Los habladores*, *La cárcel de Sevilla*, *Los mirones* y *Los romances*. En el prólogo de su edición, que ve la luz en 1936, arguye que tal vez el de *Los habladores* sea de Cervantes, pero que los otros cuatro deben considerarse anónimos, pues no hay pruebas definitivas que indiquen lo contrario:

CAPÍTULO 4

No hay, repito, ninguna prueba de que estos entremeses sean de Cervantes. Uno se puede atribuir a Cervantes; los otros cuatro ha que considerarlos, hoy por hoy, como anónimos. ¿Será posible sustraernos, sin embargo, a la creencia casi universal, de que el mejor de ellos, el de *Los habladores*, es obra de Cervantes? Aun para esta atribución habría que esperar una prueba definitiva; creo que hasta hoy no se ha dado, y es lo más probable que no se dé nunca [1936: 17].

Dámaso Alonso no estudia a fondo la posible adjudicación de esas piezas a Cervantes, pero marca al fin la pauta que debe regir todo estudio atributivo serio: el rigor académico. Las atribuciones no pueden basarse en intuiciones o semejanzas de estilo, sino en pruebas sólidas y definitivas. Y esta es la misma idea que defiende Agustín del Campo en su edición de los entremeses de Cervantes, publicada en 1948, y en la que incluye, además de los ocho publicados en vida del autor, los cinco dubitados de la edición de Alonso. En el prólogo, haciendo referencia a la posible autoría cervantina de estas piezas, insiste en que “son puras sugerencias, sin ningún valor demostrativo”, y que no identifica en ellas “la huella o la inspiración del glorioso escritor”:

No pretendemos, en modo alguno, resucitar el problema de la paternidad de estas piezas; para ello habría que renovar los términos de la controversia y entregarse a rebuscas minuciosas sobre lengua y estilo. Pero ni siquiera entonces se hubiera conseguido mucho, dado lo engañoso de todo cotejo puramente formal. Basta decir que en ninguna de las obritas atribuidas reconocemos la huella o la inspiración del glorioso escritor. Es una convicción que mana espontáneamente de una comparación ideal y que sólo puede justificarse desde un punto de vista subjetivo [1948: 38-39].

Aunque reconoce que el *Entremés de los habladores* le genera dudas, pues la alegría por la creación literaria y por los personajes nacidos de la fantasía son características muy representativas de la obra cervantina, echa en falta en él una mayor maestría en el manejo de la lengua hablada. *La cárcel de Sevilla*, al igual que *El rufián viudo*, retrata los bajos fondos de la sociedad; sin embargo, “de una a otra pintura no hay más parecido que el adjetivo”, pues mientras “Cervantes depura, acendra, hace materia estética la realidad”, el autor anónimo “pone en primer plano la mueca descarnada de las almas abyectas” [1948: 45]. No cree que *Los romances*, el único de los cinco escrito en verso, sea de Cervantes, pero sí que éste tuvo en cuenta la pieza para elaborar los primeros capítulos del *Quijote*. Y tampoco cree que lo sea *El hospital de los podridos*, aunque le recuerde a *El juez de los divorcios* y a *Los alcaldes de Daganzo*:

En las de Cervantes, la dramatización del problema particular —conflictos conyugales, administración de la justicia— era lo importante, y la resolución final que se diera al mismo quedaba a gusto del espectador; en el Hospital, al contrario, la conclusión de la trama garantiza el máximo acierto. La abierta indecisión de aquellas piezas se cambia ahora en línea perfectamente cerrada, merced a la sorprendente infición de los regentes del hospital [1948: 39-40].

De *Los Mirones* dice que “no puede llamarse entremés, ni aun pieza dramática”, porque se trata más bien “de una serie de anécdotas, cuentos y observaciones que se entretejen en la conversación de los distintos personajes” [1948: 48], y que quien lo escribió debió ser persona de agudo ingenio y de gran sutileza espiritual. Y concluye con la siguiente reflexión:

Esperemos que se haya hecho visible la radical diferencia que separa entre sí a los entremeses no auténticos: ¿no parece extraño que puedan ser considerados de un mismo autor? [1948: 50].

Sorprende enormemente que Agustín del Campo, que critica que estos entremeses hayan sido adjudicados a Cervantes por razones de estilo o puramente formales, utilice esos mismos razonamientos para rechazar su autoría, y también que se base en las diferencias argumentales entre ellos para negar que puedan ser de un mismo autor (como si *El juez de los divorcios*, *La cueva de Salamanca* o *El retablo de las maravillas*, por poner un ejemplo, tuvieran todos ellos una misma trama)¹²³.

En la segunda mitad del siglo XX, los estudios sobre los entremeses atribuidos a Cervantes caen en el olvido, volviendo a cobrar fuerza la idea de la anonimia. Tanto es así que ni Joaquín Casalduero [1951], ni Jean Canavaggio [1977], ni Zimic [1992] abordan este asunto en las monografías que dedican al teatro cervantino¹²⁴.

¹²³ Desde las ediciones parciales de Dámaso Alonso [1936] y Agustín del Campo [1948], no ha habido ningún intento por editar de manera conjunta la nómina de entremeses atribuidos a Cervantes.

¹²⁴ Canavaggio, en *Cervantès dramaturge. Un théâtre à naître un théâtre*, sólo hace una breve referencia en nota al pie [1977: 29]. En 2003, al participar en una de las sesiones plenarias del Congreso Internacional dedicado al teatro de Cervantes, que coincidía además con el vigesimoquinto aniversario de la publicación de su obra, explica los motivos: “A partir de tales premisas, consideré que convenía considerar el teatro de Cervantes en tanto que conjunto, y no examinar una tras otra cada obra que lo integra, al estilo de lo que se había hecho hasta la fecha. Desde esta perspectiva, mantenida a lo largo de los trece años que dediqué a mis investigaciones, quise incluir los entremeses en esta aproximación, en

Eugenio Asensio, en la introducción que figura el frente de su edición de los entremeses de Cervantes, asegura que no hay ningún argumento sólido para apoyar la paternidad cervantina de los *partos anónimos* que “algunos cervantistas andaluces [...] pusieron a la puerta de Cervantes, cual si fuesen hijos suyos legítimos” [1970: 14]; y en la misma línea se manifiesta primero en su famoso *Itinerario del entremés*, en donde asegura que la atribución de *La cárcel de Sevilla* y *El hospital de los podridos* “ha de tomarse como juicio de valor” [1971: 85], y más tarde en el capítulo que escribe para *Suma cervantina*:

Ocho sucintos entremeses [...] bastarían para su gloria cómica. Pero algunos admiradores del siglo pasado se han afanado por aumentar su caudal ahijándole hasta diez más. Los cinco de valor más subido figuran en un tomito elegantemente prologado por Dámaso Alonso y viven al amparo de su padre putativo. El más popular —el único con algún viso de legitimidad— es el de *Los habladores*, donde el hablar constituye la verdadera acción; el más intenso y original, el de *La cárcel de Sevilla*, que con humor patibulario gemelo del de Quevedo presenta las últimas horas de un condenado a la horca. Podemos con tranquila conciencia darles de lado, concentrando la atención sobre aquellos cuya paternidad ha sido refrendada por el autor [1973: 171-172].

E. Asensio invita “con tranquila conciencia a darles de lado”, y así parece ser en lo que resta de centuria, en donde el asunto de la posible paternidad cervantina interesa únicamente a modo de inventario [Avalle-Arce, 1973; Eisenberg, 1990; Montero Reguera, 1995 y 1998]. La tónica general es la de mantener su anonimia:

Algunos investigadores del siglo pasado atribuyen a Cervantes diversos entremeses anónimos, algunos de verdadero interés, como *Los habladores*, *El hospital de los podridos*, *Los mirones* y *La cárcel de Sevilla*. Sin embargo, las pruebas aducidas fueron escasas y no llegaron a aportar un documento definitivo, por lo que la crítica actual los considera como anónimos” [Bleuca, 1979: 47].

vez de relegarlos a lo que solía constituir un capítulo aparte, teniéndolos constantemente en cuenta en el asedio polifacético que aspiraba a llevar a cabo. Por esta razón, en unas páginas de cimentación de carácter introductivo, me abstuve de volver sobre la tan dudosa como controvertida paternidad de las obras que le han sido atribuidas sin fundamento a Cervantes, como *La soberana virgen de Guadalupe*, *El hospital de los podridos* o *Los habladores*, en tanto que decidí abordar de entrada el intrincado problema de la cronología en otro terreno que el de la polémica” [2014: 123-124]”.

4.2.3 Nueva oleada de estudios sobre los entremeses atribuidos a Cervantes

Como hemos visto en los epígrafes anteriores, en el siglo XIX se identificaron como cervantinos todos los entremeses de autor desconocido que guardaban ciertas semejanzas temáticas o estilísticas con las obras de este autor; con el cambio de centuria la perspectiva varía, y todas esas atribuciones son cuestionadas y rechazadas por la crítica, que llega incluso —como es el caso de Agustín del Campo— a emplear los mismos argumentos pero en sentido contrario. Si la coincidencia de *temas* o *lenguajes* sirvió para apoyar el cervantismo de ciertos intermedios lúdicos, las diferencias en estos mismos criterios se usaron después para negarlo, y tanto una postura como la contraria están fundadas en impresiones subjetivas desprovistas de valor y credibilidad.

En los primeros años del siglo XXI, y a pesar de que sigue vigente la idea de la anonimidad, todas estas piezas siguen estando vinculadas a Cervantes. En el quinto volumen de la *Gran enciclopedia cervantina*, dedicado al entremés, se hace referencia a *La cárcel de Sevilla*, *Los habladores*, *Los mirones*, *Los refranes*, *Los romances* y *El hospital de los podridos* [2005: vol. V: 4089-4094]; y Vicente Pérez de León, encargado de escribir la entrada de los entremeses de Cervantes para el primer volumen del *Diccionario filológico de literatura española*, se manifiesta del siguiente tenor:

Quizás la paternidad del *Entremés de los romances* ha sido tradicionalmente la más polémica, por la similitud de argumento de esta obra con el *Quijote*. Así, es especialmente controvertido el asunto de su datación. Por su parte, Asensio ha estudiado en profundidad *La cárcel de Sevilla* y *El hospital de los podridos* y afirma que la atribución de ambos a Cervantes “ha de tomarse como juicio de valor”. Sin embargo, en otro lugar he defendido la posible autoría cervantina de *El hospital de los podridos* [2009: 214].

Ese “otro lugar” al que se refiere Pérez de León es su monografía *Tablas destempladas. Los entremeses de Cervantes a examen*, en donde desarrolla su teoría sobre la posibilidad de que Cervantes sea efectivamente el autor de *El hospital de los podridos*:

El hospital de los podridos es, en definitiva, un entremés en el que se reflexiona sobre un problema social en forma de arbitrio de imposible resolución en el que hay dos alusiones a los poetas, además de una expresión calcada de otras cervantinas, que cuenta con una canción final ejemplarizante, y en el que se plantea un esquema de examen sin resolución final demasiado similar a *El juez de los divorcios* para no poder, al menos, sospechar su paternidad cervantina” [2005: 129].

Maestro, al comentar el estudio de Pérez de León en un trabajo que publica en el número 25 de *Cervantes. Bulletin of the Cervantes Society of America*, defiende la coherencia metodológica de las atribuciones de ciertos entremeses “como *El hospital de los podridos*, *Los habladores*, *Los mirones* o *La cárcel de Sevilla*, etc., cuyos personajes protagonistas son individuos destemplados y excéntricos por antonomasia” [2005: 207], a pesar de que años después cuestione la validez de los estudios de autoría [2013: 75]¹²⁵. Un descrédito que a día de hoy sigue latente en ciertos sectores críticos.

Alfredo Baras Escolá, por su parte, también aborda este asunto en la magnífica edición de los entremeses de Cervantes que realiza para la Real Academia Española [2012: 196-200]. Según recoge en estas páginas, nada apunta a Cervantes ni en *Los habladores*, ni en *Los romances*, ni en *Ginetilla, ladrón*, ni cree que pueda atribuírsele *La cárcel de Sevilla* sólo por el hecho de haber estado encarcelado en ese lugar. Tampoco concede mucho crédito al resto de atribuciones, a excepción de *Melisendra* y *Durandarte y Belerma*, unidos por su andalucismo y por el orden en el que están recogidos en el código de la Colombina, y se pregunta: ¿Cabe atribuir tales innovaciones a Cervantes porque Durandarte y Belerma preceden en *Quijote*, II, 23, al retablo de Gaiferos y Melisendra?” [2012: 199]¹²⁶.

¹²⁵ En *Calipso eclipsada. El teatro de Cervantes más allá del Siglo de Oro*, leemos: “El objetivo del trabajo no será descubrir la autoría de los entremeses hasta ahora atribuidos en ocasiones a Cervantes, pues no disponemos de documentos o manuscritos hallados que lo prueben (o nieguen) positivamente [...] La literatura no es una adivinanza, ni un jeroglífico, ni el investigador ha de ser un visionario —lo que sucede con demasiada frecuencia—, ni un ideólogo —lo que es aún más frecuente—, que interpreta la obra de arte para justificar la ideología del gremio al que pertenece”. Aunque no compartimos sus opiniones, pues creemos que a día de hoy disponemos de herramientas metodológicas que permiten estudiar de manera fiable la autoría de un texto sin necesidad de disponer de un documento que la afirme o la niegue, sí nos resulta muy interesante el estudio que realiza sobre el espacio antropológico de cinco de estas piezas [2013: 75-89].

¹²⁶ Baras Escolá comete un pequeño error al decir que Rius creyó que Luis de Belmonte Bermúdez era el autor del *Entremés de Doña Justina y Calahorra*, pues en verdad lo dijo de *Los mirones*, haciendo referencia a un estudio de Menéndez Pelayo [1895-1905: I, 192].

Ignacio García Aguilar, Luis Gómez Canseco y Adrián J. Sáez, en su monografía sobre *El teatro de Miguel de Cervantes*, tampoco conceden ningún crédito a todas estas atribuciones¹²⁷:

Pese a que para todos y cada uno se han encontrado razones muy variopintas (parecidos intertextuales, temáticos, etc.) con las que ahijárselos a Cervantes, no hay cartas que permitan dar por buena esta propuesta de paternidad y este conjunto de entremeses parecen tener más de bastardos que otra cosa [2016: 86].

Aunque desde principios del siglo XX la tónica general, alentada por el descrédito de Castro, ha sido la de negar las atribuciones entremesiles decimonónicas, lo cierto es que, como hemos visto, no son pocos los que continúan apuntando tímidamente a Cervantes. Y lo hacen tímidamente porque los estudios de autoría siguen siendo vistos —y parafraseamos aquí a Eisenberg— como investigar la prostitución o algún otro tema supuestamente “indecente”. A pesar de ello, somos varios los que de un tiempo a esta parte estamos centrando nuestros esfuerzos (o tal vez *prostituyéndonos*, quién sabe) en investigar la posible autoría cervantina de todas las obras que le han sido ahijadas —también la de buena parte de los entremeses— desde las posibilidades metodológicas que ofrecen la Métrica, la Lingüística de corpus, la Estilística o, en nuestro caso, la Lingüística forense; así ha ocurrido, por ejemplo, con el *Entremés de los romances* [Ruiz Urbón, 2010, Blasco, 2013, y Rodríguez López-Vázquez, 2019], *Los habladores* [Rodríguez López-Vázquez, 2011a]¹²⁸, *Doña Justina y Calahorra* [Muñoz Ramos, 2011, y Rodríguez López-Vázquez, 2013a], *Los mirones* [Rodríguez López-Vázquez, 2011b, y Blasco, 2019] o *La cárcel de Sevilla* [Peña Muñoz, 2010; Rodríguez López-Vázquez, 2013a]. Pero ni siquiera así se ha logrado alcanzar el consenso¹²⁹.

¹²⁷ Aunque, como veremos, Sáez no se muestra tan tajante en su reciente edición de los *Entremeses* [2020], concediendo cierta credibilidad a la paternidad cervantina de *Los habladores* y *Los romances*.

¹²⁸ No incluimos aquí el trabajo de Arellano-Torres sobre *Los habladores* porque el objetivo del mismo es ofrecer “una edición fiable del texto”, si bien también incluye algunas consideraciones interesantes sobre el problema de su autoría [2018]. Según su criterio, la autoría “por el momento queda anónima”.

¹²⁹ Rodríguez López-Vázquez defiende la autoría cervantina de todos los entremeses que ha analizado hasta la fecha, a diferencia Blasco [2019a], Ruiz Urbón [2010] y Muñoz Ramos [2011].

Sin duda, Alfredo Rodríguez López-Vázquez es el que más páginas ha dedicado a la posible autoría cervantina de varios de los entremeses que le fueron prolijados en época decimonónica. Este autor establece como pilar fundamental para apoyar la teoría cervantina de *Los habladores*, *Los romances*, *Los mirones*, *Doña Justina* y *Calahorra* y *La cárcel de Sevilla* el hecho de que ciertos vocablos o secuencias léxicas de estos textos aparezcan con mayor frecuencia en el corpus indubitado de Cervantes que en el de otros autores coetáneos, usando como respaldo el *Corpus Diacrónico del Español* de la Real Academia Española (en adelante, CORDE). Sin querer desmerecer sus trabajos, que han contribuido a revalorizar los estudios de autoría como líneas de investigación objetivas y rigurosas, es importante advertir que, aunque la frecuencia de uso de ciertas unidades léxicas sea, efectivamente, una de las marcas idiosincrásicas que conforman nuestro idiolecto, la cantidad de las muestras incluidas en el CORDE puede hacer que malinterpretemos los resultados obtenidos. Por ejemplo, no resulta tan significativo que “tendido sobre” aparezca tres veces en Cervantes y ninguna en Lasso de la Vega —como indica Rodríguez López-Vázquez [2019: 236]—, si tenemos en cuenta que en el CORDE están recogidas todas las obras del primero, incluso algunas de dudosa autoría, hasta superar el millón de palabras (1.070.039), mientras que del segundo las muestras de respaldo no alcanzan las 180.000 (en concreto, 177.826). Este tipo de análisis puede resultar significativo dentro de la complementariedad metodológica a la que hacíamos referencia en el capítulo 2, pero nunca determinantes ni concluyentes para afirmar o negar autorías.

Adrián J. Saéz, en su reciente edición de los *Entremeses* de Cervantes para Cátedra, confirma este *revival* del interés por las atribuciones a Cervantes. Para él, los dos únicos entremeses que parecen tener el “ADN Cervantes” son *Los habladores* y *Los romances*, aunque no profundiza demasiado en el asunto¹³⁰. Coincidimos con él en que éste es un debate que necesita de “espacio, finura y sentido común” [2020: 30].

¹³⁰ “*Los habladores* por la estructura de desfile de figuras y la sátira de vicios, acaso como último entremés cervantino (Rodríguez López Vázquez, 2011b, 71), y el *Entremés de los romances*, que representa las aventuras de Bartolo, un loco que ha perdido la cabeza por tanto romance, y su escudero Bandurrio, como un par todavía más cómico del inicio del Quijote, que se dice pudo ser primeramente

4.3. NÓMINA DE LOS DIEZ ENTREMESES ATRIBUIDOS A CERVANTES

En las siguientes páginas daremos breve noticia de cada uno de los entremeses que alguna vez han sido prolijados a Cervantes, a pesar de la ligereza de los argumentos en los que por lo general se han sustentado tales filiaciones. Aunque no conocemos las fechas exactas de composición, es importante aclarar que todos ellos pueden circunscribirse dentro de los años en los que Cervantes estuvo centrado *en la pluma y la comedia*¹³¹:

- *Entremés de los habladores*, en prosa
- *Entremés del hospital de los podridos*, en prosa
- *Entremés de la cárcel de Sevilla*, en prosa
- *Entremés de los refranes*, en prosa
- *Entremés de los mirones*, en prosa
- *Entremés de los romances*, en verso
- *Entremés de Melisendra*, en verso
- *Entremés de Durandarte y Belerma*, en verso
- *Entremés de Doña Justina y Calahorra*, en verso
- *Entremés de Ginetilla ladrón*, en verso (perdido)

una novelita (con la primera salida, I, 1-8) luego desarrollada en toda su potencia (Baras Escolá, 1993; Rey Hazas, 2006)” [2020: 30].

¹³¹ En los primeros entremeses, siguiendo la estela loperuedesca, predomina la prosa como vehículo de expresión, pero en las primeras décadas del siglo XVII los autores empiezan a adoptar el verso, decantándose por el “endecasílabo, ya suelto con un pareado al final de cada parlamento, ya en silva con rimas frecuentes” y por el “romance” [Asensio, 1971: 63]. Teniendo esto en cuenta, es probable que la fecha de composición de los cinco entremeses en prosa sea anterior a la de los cinco en verso.

4.3.1. *Entremés de los habladores*

El entremés de los habladores narra la historia de Roldán, un hombre empobrecido y maltrecho que huye de la justicia por *hacer las maulas*. Tras interpretar erróneamente unas palabras que escucha por boca de Sarmiento, Roldán se ofrece voluntario para que éste le dé una cuchillada a cambio de dinero. La conversación entre ambos se vuelve aún más hilarante cuando Roldán empieza a utilizar las palabras que pronuncia Sarmiento para ofrecer largas disquisiciones sobre temas y asuntos varios, sin conexión aparente alguna. Sarmiento cae entonces en la cuenta de que este hombre es el único que puede hacer callar a su esposa Beatriz, “la mayor habladora que se ha visto desde que hubo mujeres en el mundo”, y decide llevarlo a su casa para que la ponga “en dos horas muda como una piedra”. Tras ganar la guerra dialéctica a Beatriz, que acaba desmayándose, Roldán es descubierto por la justicia. La pieza se cierra con varias glosas en verso sobre la irrefrenable necesidad de hablar.

Los habladores se corresponde con la estética del ‘entremés de cambio de siglo’, por lo que su composición puede fecharse a finales del Seiscientos. Se imprime por primera vez en la *Séptima parte de las comedias de Lope de Vega* (Madrid, 1617, folios 290r.-293r.), junto a *El hospital de los podridos* y *La cárcel de Sevilla*. Navarrete cita una suelta a nombre de Cervantes, impresa en Sevilla en 1624, pero que nadie ha logrado ver, y Fernández-Guerra otra impresa en Cádiz en 1646 por Francisco Juan de Velasco, también a nombre de Cervantes, que sí se conserva pero que presenta claros indicios de ser apócrifa [Madroñal Durán, 2006]¹³². Poseemos también un testimonio manuscrito de este entremés en el folio 8 del código 82-3-40 de la Biblioteca Colombina de Sevilla, antes descrito, junto a trece entremeses más (véase nota 118).

¹³² Francisco Juan de Velasco es uno de los primeros editores que se dedica a la impresión de entremeses sueltos para venderlos de manera independiente. Tras estudiar algunas de estas sueltas, que fueron agrupadas posteriormente en un volumen facticio titulado *Entremeses de Cádiz*, Madroñal Durán advierte que “da la impresión [...] que Velasco quería vender sus entremeses sueltos con el reclamo publicitario de un nombre famoso” [2006: 574], y que así ocurre en el caso de *Los habladores*, en donde “Velasco, tomando el texto de la parte de Lope de 1617, inventa el nombre de Cervantes como autor” [572].

El hecho de que Lope de Vega reniegue de las loas y entremeses que aderezan la séptima parte de sus comedias¹³³, unido al hecho de que varios autores manifiesten tener o haber visto ejemplares de *Los habladores* a nombre de Cervantes, ha provocado que éste sea el entremés más largamente atribuido al autor del *Quijote*, y también el más aceptado como tal por la crítica; incluso Cotarelo Valledor, Dámaso Alonso o Agustín del Campo, muy críticos con las atribuciones de época decimonónica, conceden a este entremés un estatus especial. Pero a pesar de que muchos especialistas defiendan el cervantismo de la pieza, lo cierto es que a día de hoy no hay datos concluyentes que permitan fijar su paternidad en esta dirección. Así lo manifiesta Madroñal Durán, que apunta además ciertos datos de peso que permiten probar la relación que existe entre *Los habladores* y la ciudad de Toledo [1993: 121-123], y su posible vinculación con el dramaturgo Gaspar de Barrionuevo:

No queremos indicar con ello que *Los habladores* pertenezcan a Barrionuevo, tan sólo pretendemos llamar la atención sobre el hecho de que algunos de esos entremeses anónimos o pseudoanónimos puedan adscribirse al buen hacer del regocijado ingenio toledano [123].

Azcune, como recuerda Baras Escolá en su edición de los entremeses cervantinos, apunta también al entorno de Lope de Vega:

Azcune [2000] aduce un dato esclarecedor: el 30 de octubre de 1600, Tomás Gracián Dantisco aprobó en el manuscrito Gálvez la comedia de Lope *Los enredos de Celauro*, fechada el 25 de enero, y con ella el “entremés de que me hizo relación Fabián de Rivera, del *Hablador*”, para ser representados. “Es evidente que el título mencionado y *Los Habladores* son el mismo” (p. 396). Entre los actores de Porres, empresario del Fénix, se contaba “para entremeses” con Fabián de Ribera, según documento de 25 de febrero de 1600 (salvo 2003: 215-216) [2012: 197].

Tan sólo Rodríguez López-Vázquez se aleja de esta tendencia de “desatribución” cervantina [2011a], si bien Arellano-Torres advierte que la lista de coincidencias léxicas o expresiones en que basa sus conclusiones no resulta muy significativa:

¹³³ En el prólogo a la decimoquinta *Parte* de sus comedias, Lope asegura que “tiene por menos mal que salgan de su casa que no de las ajenas por no las ver, como las primeras, en tal dicha, ya con loas y entremeses que él no imaginó en su vida” [Citamos por Couderc, 2009: 123].

Para Rodríguez López-Vázquez «los resultados son contundentes: de las 30 palabras rastreadas [en el CORDE] 26 de ellas aparecen en Cervantes» (2011:68). Pero palabras que se asegura no usa Quevedo aparecen por ejemplo en los poemas de *El Parnaso español*: ‘holgara’ está en poemas núms. 397, 432, 455; ‘mohino’ en los núms. 218, 359 bis, 415, 449; también usa ‘huerto’ o ‘garrote’ [...]. Los rastreos de Rodríguez López-Vázquez no parecen muy fiables” [2018: 304].

4.3.2. *Entremés del hospital de los podridos*

El argumento de este ‘entremés de figuras’, que también puede fecharse a finales del siglo XVI, gira en torno a la manía o enfermedad que tenemos todos de repudrirnos en cosas que no tienen solución, o que no nos afectan directamente. Hasta este peculiar hospital de podridos, fundado “para que se curen los heridos de esta enfermedad”, van llegando personas de todo tipo, con diversas pudriciones: Cañizares se pudre de que un hombre que trae “por lo caniculares chinelas y la espada a zurdas” pueda llegar a ser gobernador, Pedro Díaz de que los poetas digan tantos disparates, Valenzuela de que a su vecino todo le salga bien, Gálvez de que una mujer hermosa esté enamorada de un hombre calvo y con anteojos; un hombre de que su mujer tenga los ojos azules; y ésta de que su marido tenga la boca grande. El entremés concluye con un giro inesperado, por el que los jueces acaban siendo víctimas de la misma enfermedad que persiguen. Las palabras finales de Villaverde, el único que no está podrido, nos dan la clave para no caer en este mal: “No se pudra nadie / de lo que otros hacen, / [...] / dejemos a cada uno / viva en la ley que gustare, / [...] / yo no me pienso pudrir, / ni que el contento me acabe, / aunque abadejo me digan / y aunque bacallao me llamen”.

Al igual que *Los habladores*, este entremés en prosa se encuentra entre los que aderezan la *Séptima parte de las comedias* del Fénix de España (Madrid, 1617, folios 298r.-301r.) y también fue impreso como suelta, esta vez a nombre de Lope de Vega, por Francisco Juan de Velasco (Cádiz, s.f.). Fernández Guerra creyó que era de Cervantes por su “espíritu”, “genio” y “estilo” [1863: 1371], pero la crítica posterior ha cuestionado tal atribución:

Tanto *El hospital de los podridos* como *La cárcel de Sevilla*, incluidos en la *Séptima parte de las Comedias de Lope* (Madrid 1617), han sido frecuentemente atribuidos a Cervantes. Tal atribución ha de tomarse como juicio de valor, como opinión de que en la sustancia y la

forma, en la visión de la realidad, tienen un aire de *familia* con las piezas de Cervantes, de cuya pluma son dignos [Asensio: 1971: 85].

No hay ningún estudio reciente que aborde la autoría de esta pieza, aunque autores como Pérez de León [2005] sigue apuntando hacia Cervantes por su similitud con la trama y los personajes de *El juez de los divorcios*.

Diferentes detalles, tales como la existencia de expresiones casi calcadas en ambas obras cervantinas, el hecho de que los examinadores respondan a actitudes ejemplares y serias en su función de llevar a cabo el arbitrio a la práctica, además de la fijación existente contra los poetas en diferentes momentos de ambas son otros rasgos que acercan *El hospital de los podridos* a su paternidad cervantina [Pérez de León, 2005: 124]

4.3.3. *El entremés de la cárcel de Sevilla*

La Cárcel Real de Sevilla fue uno de los centros neurálgicos de la vida marginal del Siglo de Oro, del que nos ha llegado noticia a través de distintos documentos de la época, como la relación que escribió el procurador Cristóbal de Chaves, las memorias del jesuita Pedro de León, confesor de los presos, o este entremés anónimo de finales del XVI, publicado por primera vez en la *Séptima parte de las comedias de Lope de Vega* (Madrid, 1617, folios 293v-297v) y que también se conserva en el folio decimotercero del manuscrito en la Biblioteca Colombina (véase nota 118). La comicidad de toda la pieza gira en torno a la actitud lúdica y flemática de Paisano, uno de los presos, tras conocer que ha sido condenado a pena de muerte.

La vinculación de este entremés con Cervantes tiene sus orígenes en época decimonónica, y los argumentos más repetidos son el hecho de que tanto en la edición de las comedias de Lope de Vega como en el códice de la Colombina aparezca justo después de *El entremés de los habladores*, también atribuido a Cervantes, así como su cercanía temática con la *Relación de la Cárcel de Sevilla*, cuya tercera parte se disputan Chaves y Cervantes [Fernández Guerra, 1863: 1371; J. M. Asensio, 1902: 40; Castro, 1874: 8; Menéndez Pelayo, 1941: 275]. Este mismo argumento sirve a Sánchez Arjona para atribuírselo a Chaves, arguyendo además que el Cabildo de la ciudad le había pagado en 1598 por un entremés para la fiesta del Corpus:

En el libro de Caja del Ayuntamiento de Sevilla consta que se dieron este año “20 ducados á Cristóbal de Chaves, procurador de esta ciudad, que fueron por gratificación del entremés que hizo para un carro de los de la fiesta del Corpus en que se representan las grandezas de la ciudad” [...] ¿no fuera lógico suponer que si Chaves, autor de la *Relación* de donde se tomó el argumento del entremés, y con la que tiene marcados puntos de contacto, escribió algunas obras de este género, pudiese ser el autor del entremés de la cárcel de Sevilla, escrito acaso para representarse también en la fiesta del Corpus de esta ciudad? [1898: 98-99].

La autoría de Chaves también es defendida por Rodríguez Marín en su edición de *Rinconete y Cortadillo* [1905: 214], por Dámaso Alonso, no sin “cierta repugnancia”, y por Sáez [2020: 30]. E. Asensio descarta que el autor sea Cervantes, “que allí estuvo preso, pero cuya técnica y modalidades cómicas son muy distantes”, pero también que lo sea Chaves, tanto por la calidad de la pieza, muy superior a las de éste, como por su datación: “la aparición de Escarramán entre los interlocutores nos obliga a retrasar la obra por lo menos hasta 1612, diez años después de la muerte de Chaves” [1971: 90-91]. Hernández Alonso y Sanz Alonso proponen como “candidato a Agustín de Rojas:

El tipo de diálogo y la soltura de estilo del *Entremés* no desdice los del *Viaje entretenido*; en ambas piezas del lenguaje y las formas de tratamiento se adaptan a los personajes, según su nivel sociocultural, y el autor conoce bien el lenguaje coloquial popular bajo de su momento [1999: 401].

Agostini del Río, por su parte, vuelve a poner el foco en Cervantes, al advertir sus similitudes con *El rufián viudo*, *La guarda cuidadosa* y el *Quijote*. Y lo mismo hacen Peña Muñoz y Rodríguez López-Vázquez, la primera arguyendo criterios biográficos [2010] y el segundo, criterios léxicos [2013a].

4.3.4. *Entremés de refranes*

Este entremés nos presenta a Doña Sofía, una mujer sin “oficio ni beneficio” que recibe la noticia de que un tío que tenía en las Indias le ha dejado en herencia mil

ducados, que sólo podrá cobrar si “esté casada o se case”¹³⁴. Decide por ello proponer a Pedraza que finja ser su marido, sin caer en la cuenta de que finalmente será él, y no ella, quien reciba el dinero. Tras cobrar la herencia, Pedraza le dice a Doña Sofía que es “tierno de corazón”, y que por ello le entregará el dinero poco a poco para que no le dé “con los ochos y nueves”. La gracia de todo el entremés reside en que los personajes llevan al extremo la práctica de insertar refranes, modismos, proverbios, aforismos o sentencias en su habla cotidiana¹³⁵:

PEDRAZA. Quien no cree en buena madre, crea en mala madrastra. Pensé yo, señora Doña Sofía, que pescaba bogas y que tenía trapillo con dineros, en amartelar a vuesa merced; y al fin he visto que la mejor mujer, mujer: pues me deja como el carnero encantado, que fue por lana y volvió tresquilado.

DOÑA SOFÍA. Más es el ruido que las nueces, señor Pedraza. No diga vuesa merced esta boca es mía, sino punto en boca; y si no, tome las de Villadiego, y no piense que me hace los hijos caballeros: que ya está pobre; y de costal sacudido, nunca buen bodigo.

El *Entremés de refranes* es uno de los catorce que conforman el manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla (folio 107v); también está recogido en el código S-14-956 de la Biblioteca Rodríguez Moñino, junto a otros de los entremeses que nos ocupan (*Los mirones*, *Melisendra* y *Doña Justina* y *Calahorra*)¹³⁶. El primero en editarlo y atribuírselo a Cervantes, aunque luego él mismo llega a dudar de dicha

¹³⁴ No debemos confundir este entremés con el de *Los refranes del viejo celoso*, atribuido a Quevedo, y que hoy se conserva manuscrito en la biblioteca de la Hispanic Society of America (Signatura B2900). A este respecto, véase Madroñal Durán [2013].

¹³⁵ Northup llega a contabilizar ciento noventa y dos refranes [1922a]. Como se recoge en la *Gran enciclopedia cervantina*, “el estilo del entremés no era extraño ni innovador en el siglo XVI. Sus raíces se hallan en la poesía lírica de cancionero del siglo XV en una técnica que O’Kane denominó al modo germánico ‘priamel’ y que Bizzarri califica como de ‘enhebrado de refranes’, es decir, la aglutinación de paremias para componer una pieza” [2005-2009: vol. V, 4090].

¹³⁶ Según describe Madroñal Durán en el suplemento a su catálogo de entremeses de la Real Academia Española [1998], en este código del siglo XIX se recogen las siguientes piezas: (1) *Entremés del zurdo*, (2) *Entremés de los mirones*, (3) *Entremés del sacristán Soguijo*, (4) *Entremés de la villana de Getafe*, (5) *Entremés de la endemoniada fingida* y *chiste de Bacallao*, (6) *[Loa] para el entremés de Melisendra* y *Entremés de Melisendra*, (7) *Entremés del rey cachumba de Motril* y *la infanta Palancona*, (8) *Entremés de doña Justina* y *Calahorra* y (9) *Entremés de los refranes*. Todos ellos también figuran en el código 82-3-40 de la Biblioteca Colombina de Sevilla, en el que aparecen además cinco entremeses más: *El examinador Miser Palomo*, *Los Habladores*, *La cárcel de Sevilla*, *Durandarte* y *Belerma*, y *El Doctor Zarrabulleque* (véase nota 118).

atribución, es José María Asensio, en una carta dirigida a Fernández-Guerra en agosto de 1867, en la que asegura además que su opinión es compartida por Álava y Hartzenbusch [1902: 41]. No obstante, como recuerda Sbarbi en *El refranero general español*, Fernández Guerra, tras recibir la carta, “se inclinó desde luego á dudar de semejante paternidad, fundándose para ello, entre otras razones, en lo flojo de la acción”, y apuntó la posibilidad de que perteneciera “á la pluma de Quiñones de Benavente con cuyo estilo parece guardar mayor analogía” [1877: tomo VII, VI-VII].

Castro defiende la autoría cervantina por “el argumento, modo de expresión, diálogo y la facilidad en el uso de tanta multitud de refranes” [1874: 21], al igual que Menéndez Pelayo. Vidal de Valenciano, por su parte, rechaza tanto a Cervantes como a Quiñones de Benavente, y sospecha de Quevedo [1883]. No hay ningún estudio reciente —al menos no que conozcamos— que aborde la autoría de esta pieza.

4.3.5. *Entremés de los mirones*

La acción de este entremés transcurre en Sevilla, en época de Carnestolendas: varios estudiantes, pertenecientes a la cofradía de Los mirones, recorren las calles de la ciudad para observar las cosas que ocurren en esos días, e ir después a casa de su maestro para darle cuenta de ellas. A diferencia de los otros entremeses en prosa atribuidos a Cervantes, éste no se cierra con ninguna canción en verso, y muchos ven en él un *coloquio* no pensado —al menos en origen— para llevar a las tablas.

El testimonio más antiguo que poseemos de esta piececilla es el que se conserva en el manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla, al que ya tantas veces hemos hecho referencia (véase nota 118); y también está incluido en el código S-14-956 de la Biblioteca Rodríguez Moñino (véase nota 136). Castro lo edita y se lo atribuye a Cervantes porque “en el estilo se asemeja mucho al de *los perros Cipión y Berganza*” y tiene “la misma manera de presentar los pensamientos filosóficos, y la de contar las aventuras y describir las costumbres” [1874: 19]; Menéndez Pelayo, siguiendo su estela, señala primero que es de Cervantes y que “forma una trilogía con *Rinconete y Cortadillo* y el *Coloquio de los perros*” [1941: 276], pero años después se inclina a

pensar que es en verdad obra “de algún buen imitador suyo, probablemente sevillano, y que pudo ser muy bien Luis de Belmonte Bermúdez” [Rius, 1895-1905: I: 192]; y Cotarelo Valledor, por su parte, reconoce que “entre los varios entremeses atribuidos al Príncipe de los ingenios, [es] uno de los que ofrecen ciertas condiciones de probabilidad” [1915: 727].

Más recientemente, Rodríguez López-Vázquez, basándose en los resultados que arroja el CORDE para en un repertorio de treinta unidades léxicas, cree que ha sido escrito por Cervantes [2011b]; Blasco, en cambio, apunta la posibilidad de que su autor sea Salas Barbadillo, pues este entremés presenta gran identidad verbal con *El caprichoso en su gusto* y *Los mirones de la Corte*:

¿Podemos concluir con los datos que hemos visto que *Los mirones* es obra de Salas Barbadillo, y no de Cervantes? Existe una probabilidad muy alta de que, a la vista de los resultados que ofrece la estilometría, sea así. Pero yo no me atrevería a apuntarlo hasta no concluir definitivamente que *El caprichoso en su gusto* y *Los mirones de la corte* sean también obras del madrileño [2019: 166].

4.3.6. *Entremés de los romances*

Este entremés, escrito en verso entre la última década del siglo XVI y la primera del siglo XVII e impreso en la *Parte tercera* de las *Comedias de Lope de Vega y otros* (Valencia, 1611), recoge la historia de Bartolo, un labrador enloquecido de tanto leer el Romancero y que, creyéndose uno de los héroes que tanto fervor le causan, decide abandonar su casa y marcharse rumbo a otras tierras para enrolarse en la guerra y servir a su patria, llevando como escudero a su convecino Bandurrio. Si en el *Entremés de refranes* la comicidad residía en la inserción de multitud de aforismos, frases hechas y sentencias, en este caso lo que se pone en boca de los personajes son un sinfín de romances populares.

El primero en advertir su relación con la historia del *Quijote* y con Cervantes es Adolfo de Castro [1874: 129-141], y en la misma línea se manifiesta una vez más Menéndez Pelayo:

La primera reflexión que ocurre a todo el que lee este agradable juguete es su conformidad extraordinaria con los primeros capítulos del *Quijote*. No es su objeto reprender la lectura de los libros de caballerías, sino la de ciertos romances moriscos y caballerescos, inspirados o inspiradores de aquellos, asunto que se da mucho la mano con el anterior. El desarrollo de la acción es el mismo, con escasa diferencia. Para poner de manifiesto la admirable similitud que existe en los pensamientos y hasta en las palabras, pone en parangón el señor Castro largos pasajes de uno y otro. Es verdad que faltan en la figura de Bartolo muchos rasgos que a la de don Quijote distinguen, razón que nos induce a mirar la primera como el bosquejo de la segunda. La primera encarnación del pensamiento cervantino fué el rústico e ignorante Bartolo. Purificada y embellecida transformóse en don Quijote, tipo acabado de virtud e hidalguía, modelo de discreción en cuanto no se rozaba con su manía. El insignificante escudero Bandurrio vióse convertido en Sancho Panza, personificación del buen sentido y de la recta razón aplicada a las más prosaicas necesidades de la vida. En el *Entremés de los romances*, obra dignísima del aprecio y estimación de los cervantistas, vemos a las claras el desarrollo del pensamiento fundamental del *Quijote* en la mente de Cervantes [1941: 285].

A principios del siglo XX, Menéndez Pidal y Cotarelo y Mori presentaron tesis opuestas en cuanto a su datación, y a día de hoy el debate en torno a si este entremés es fuente inspiradora o herencia del *Quijote* sigue abierto. En relación con su paternidad, que ha generado menor interés por parte de la crítica, tampoco hay consenso: la gran mayoría rechaza la autoría cervantina; Northup [1922b] y Rodríguez López-Vázquez [2019] la defienden; Millé y Giménez [1930] y López Navío [1959-1960] piensan que el autor es algún enemigo de Lope de Vega, como Góngora, Salinas o el mismo Cervantes; Rey Hazas [2006 y 2007] propone como firme candidato a Gabriel Lobo Lasso de la Vega; y Pérez López [2009] y Blasco [2013], por el contrario, señalan a Liñán de Rianza, gran amigo de Lope, que pudo escribirla tal vez en colaboración.

4.3.7. *Entremés de Melisendra*

El *Entremés de Melisendra* inaugura en España el teatro burlesco, que tantos éxitos cosecha durante todo el reinado de Felipe IV; Madroñal Durán advierte que “aunque se titule así no es propiamente un entremés sino un anticipo de comedia burlesca” [2011: 70], como prueba además que en varios de los testimonios conservados lleve una “loa muy graciosa” y esté dividido en dos jornadas. La pieza recoge la historia romanceril del rescate de Melisendra, retenida en Sansueña por los moros, de manos de

su marido Gaiferos, pero en clave paródica, con diálogos disparatados y situaciones grotescas. Así se representa, por ejemplo, el momento de la liberación:

GAIFEROS. [...] Don Gaiferos soy señora.
 MELISENDRA. ¿Es posible que tal siento?
 ¡Mi regalo!
 GAIFEROS. ¡Mi contento!
 MELISENDRA. ¡Mi jarro!
 GAIFEROS. ¡Mi cantimplora!
 MELISENDRA. ¡Mi ballesta!
 GAIFEROS. ¡Mi bodoque!
 MELISENDRA. ¡Oh, cabe mío!
 GAIFEROS. ¡Oh, mi toque!
 MELISENDRA. ¡Mi ochavo!
 GAIFEROS. ¡Mi dingandux!
 MELISENDRA. ¡Oh, mi primera!
 GAIFEROS. ¡Oh, mi flux!
 MELISENDRA. ¡Mi ciruelo!
 GAIFEROS. ¡Mi albaricoque!
 Baja, señora, baja.
 MELISENDRA. Parecerseme han las piernas
 GAIFEROS. Yo os cubriré las cavernas [BC, Ms. 82-3-40, f. 59v]

Melisendra es el primero —y el único en verso— de los que cinco entremeses que aparecen en la edición valenciana de la conocida como *Primera parte de las Comedias de Lope de Vega*, en donde figura con loa y división de jornadas (Valencia, 1605)¹³⁷. Se conserva además en varios testimonios manuscritos:

¹³⁷ La edición prínceps (Zaragoza, 1604) no lleva entremeses, sólo loas y comedias. En la edición valenciana se incluyen cinco entremeses (*Melisendra*, *El padre engañado*, *El capeador*, *El doctor simple* y *Pedro Hernández y el corregidor*), seguramente con la intención de trasladar al papel el curso de la función de los corrales, y en la de Valladolid siete más (*Los alimentos*, *Los negros de San Tomé*, *El indiano*, *la cuna*, *Los ladrones engañados*, *La dama fingida* y *La endemoniada*), para subsanar la evidente asimetría entre el número de comedias y el de entremeses [Pontón y Sánchez Aguilar, 1997: 1809-1810].

- Un manuscrito de la Biblioteca Colombina, con letra cursiva de principios del siglo XVII (BC, Ms. 82-3-40, fols. 53r-62r): Aunque lleva loa, se trata de una versión reducida del impreso, sin división en jornadas.
- Una suelta manuscrita de la Biblioteca Nacional de España, con letra cursiva de principios del siglo XVII (BNE, Ms. Res/88): El entremés ocupa 10 folios. Carece de loa, pero mantiene las dos jornadas, conforme a la segunda edición impresa en Valladolid en 1609. Lleva la rúbrica de Lope de Vega y dos licencias de representación (1622 y 1623), aunque a día de hoy se descarta su autenticidad [Cuenca Muñoz, 1999; Urzáiz Tortajada, 2014].
- Un manuscrito de la Biblioteca Nacional de España, con letra cursiva de principios del siglo XVII (BNE, Ms. 4.117 fols. 131r-136v): El entremés forma parte de un volumen misceláneo que reúne poemas varios, loas teatrales y piezas satíricas.
- Un manuscrito de la Biblioteca Rodríguez Moñino, con letra del siglo XIX (BRM, Ms. S-14-956): Lleva loa.
- Un manuscrito de la Biblioteca del Instituto del teatro de Barcelona, con letra del siglo XIX (BIT, Ms. 61.595): Al igual que el manuscrito de la Biblioteca Nacional, se trata de una copia bastante fiel del entremés impreso en Valladolid (1609).

J. M. Asensio es el primero en atribuir este entremés a Cervantes, al ver en él —al igual que en *Durandarte y Belerma*— una manera de explicar el aprecio que iban teniendo en su ánimo “las invenciones caballerescas, y el punto de vista ridículo bajo el cual empezaba á considerarlas” [1902: 41]. Su tesis no tiene el respaldo de la crítica posterior, que opta por adjudicárselo a Lope de Vega, basándose en la firma que figura en uno de los testimonios conservados (BNE Res/88) (Imagen 4); pues, aunque ésta no sea auténtica, sí podría serlo la atribución. Así lo entienden Pontón y Sánchez, tal y como explican en la introducción a la edición que hacen de este entremés para el equipo PROLOPE:

Si la firma que sigue al entremés no es la de Lope, se trata cuando menos de una imitación cuidadosa, realizada por alguien que disponía de un modelo. La rúbrica es prácticamente idéntica, y también lo es la disposición de la mayoría de los trazos de las letras, aunque parecen escritas con mayor pausa de lo que es habitual en los autógrafos, donde el trazo se apresura y hay más ligaduras entre letras e incluso entre palabras [...] Sea como fuere, es preciso concluir que el manuscrito Res. 88 contiene una copia del entremés tomada de la edición de Valladolid y presentada para la concesión de las licencias oportunas bajo la atribución a Lope de Vega, atribución en la que él mismo podría haber estado implicado. [1997: 1824].

Y así lo entiende también Cuenca Muñoz, al señalar que el manuscrito, aunque no sea autógrafo, “sugiere la posibilidad de que el texto contenido en el mismo sí fuera obra del dramaturgo [1999: 173], y Percas de Ponseti, que ve claras alusiones a Lope en el capítulo de la segunda parte del *Quijote* en el que se representa la historia de Gaiferos y Melisendra [2003]. Conjeturas todas ellas que resultan arriesgadas si recordamos que el propio Lope reniega de todos los entremeses que aderezan los volúmenes de las primeras partes de sus comedias.

4.3.8. *Entremés de Durandarte y Belerma*

En este entremés en verso, al igual que en el de *Melisendra*, se parodia una historia de la tradición romanceril muy conocida por el público de la época: la de los amores de Durandarte y Belerma¹³⁸. A pesar de lo absurdo y disparatado de los diálogos, aliñados con alusiones a actividades sexuales, necesidades fisiológicas o elementos

¹³⁸ Los hermanos Cortijo Ocaña explican de este modo la prevalencia de esta historia del ciclo carolingio en el Siglo de Oro: “No podemos hacernos idea cabal de la significación de la comedia burlesca que estudiamos sin tener presente el recorrido y difusión impresa y manuscrita de los romances sobre la historia de Durandarte y Belerma durante los siglos XVI y XVII en la Península Ibérica. Romances serios, romances burlescos, glosas, ediciones en pliegos sueltos, *contrafacta* paródicos, *contrafacta* a lo divino, etc., y que abarcan desde el *Cancionero de 1511* de Hernando del Castillo, donde por primera vez se imprime un romance de este tema, hasta Góngora o Cervantes e el capítulo XXIII de su *Don Quijote*, pasando por el *Cancionero de Martín Nucio* (Ambreres, 1550) o la versión cantada para vihuela de Luis de Milán de 1535. En suma, todos estos testimonios nos documentan que el tema y el romance fueron enormemente difundidos durante todos estos siglos, lo que implica que el público de nuestra comedia burlesca tenía un referente claro (y de sobra conocido) sobre el que juzgar las modificaciones y elaboraciones del drama paródico” [2002: 76].

religiosos, la obra mantiene la coherencia de la trama original: relata el enamoramiento de los protagonistas, la muerte de Durandarte en la batalla de Roncevalles y la misión de Montesinos de llevar su corazón a Belerma, que cae desmayada tras recibir la noticia.

La pieza es una de las manuscritas en los citados códices de la Biblioteca Colombina y la de Rodríguez Moñino, aunque su extensión (1069 versos) permite adivinar que no se trata en verdad de un entremés, sino de una comedia. De hecho, es publicada como “comedia burlesca” en 1679, con el título *El amor más verdadero, Durandarte y Belerma*, a nombre de Mosén Guillén Pierres (bajo el que muchos han querido ver un pseudónimo)¹³⁹; y también en cuatro sueltas de fecha no segura¹⁴⁰. Cabe advertir que en la base de datos CLEMIT se hace referencia a una representación hecha en Pisa en 1616, aunque en realidad el entremés que allí se representó fue el de *La infanta Palancona*, y no el de *Durandarte y Belerma*¹⁴¹.

De todos los entremeses atribuidos a Cervantes, éste era, junto al perdido de *Ginetilla ladrón*, el que menos atención había recibido por parte de la crítica. Pero en los últimos años, gracias a dos proyectos de investigación del GRISO de la Universidad

¹³⁹ *El amor más verdadero, Durandarte y Belerma* se publicó dentro de las *Comedias nuevas, escogidas de los mejores ingenios de España (Parte cuarenta y cinco)* (Madrid, José Fernández de Buendía, 1679), de la que se conocen cinco versiones.

¹⁴⁰ Los datos de las cuatro sueltas son poco clarificadores, ya que de alguna de ellas ni siquiera sabemos lugar, año o impresor: suelta 1, “s.l., s.i., s.a. (fols. 1-8)”; suelta 2, “¿Sevilla?: s.i., ¿1820? (fols. 1-16)”; suelta 3, “Barcelona: Francisco Suria y Burgrda [sic], impresor calle de la Paja, ¿1790? (fols. 1-16)”; y suelta 4, “Sevilla, Joseph Antonio de Hermosilla, 172? (fols. 1-24)” [Arellano, dir., 2002: 93].

¹⁴¹ Nider documenta una representación del *Entremés de la infanta Palancona* hecha en Pisa, en 1616, por parte de la comunidad judío-portuguesa de esta ciudad, que tuvo como desenlace una denuncia inquisitorial por parte de los espectadores cristianos que asistieron a esa fiesta [2011]. Suponemos que la confusión de CLEMIT tiene su origen en que Nider hace referencia también al código de la Colombina, en el que el entremés *El rey Cachumba de Motril y la Infanta Palancona* antecede al de *Durandarte y Belerma* (ambos parodias burlescas del ciclo carolingio): “Un esempio è la citata commedia burlesca *El amor más verdadero. Durandarte y Belerma*, che, come si è anticipato, viene trascritta con il titolo di *Entremés de Durandarte y Belerma* nel codice S di seguito all’*Entremés de la Infanta Palancona*. Serralta osserva che la commedia è, fra le *burlescas*, quella con il maggior numero di riferimenti religiosi, in particolare a personaggi biblici, mentre scarseggiano quelli alle categorie e gerarchie ecclesiastiche o alla liturgia, e ipotizza che tali *impertinencias* si possano spiegare con le circostanze di creazione, ovvero con il fatto che l’autore cui l’opera viene attribuita, il non altrimenti noto Mosén Guillén Pierres, fosse un ecclesiastico e avesse scritto la commedia per rappresentarla in ambiente religioso (scuola seminario, etc.)” [2011: 175].

de Granada, uno sobre comedias burlescas y otro sobre la burla en el Siglo de Oro, la pieza ha sido rescatada del olvido: en 2002 la editaron los hermanos Cortijo Ocaña, siguiendo como texto base el manuscrito 3907 de la Biblioteca Nacional de España (fols. 274-289), y muy recientemente lo ha hecho Nider [2020], tomando como referencia el texto del códice colombino, hasta ahora inédito¹⁴².

El número de comedias burlescas de autor anónimo es porcentualmente superior al de las comedias de otros géneros, posiblemente por la delicadeza del material tratado. De hecho, los testimonios impresos a nombre de Mosén Guillén Pierres son una versión censurada (o autocensurada), en las que han sido eliminadas muchas de las referencias religiosas contenidas en el manuscrito anónimo de la Colombina.

J. M. Asensio, en la carta que dirige a Fernández-Guerra en 1867, atribuye la pieza a Cervantes [1902: 41]; Castro y Menéndez Pelayo apuestan por Mira de Amescua, aunque esa autoría “no se acepta en la actualidad” [Madroñal Durán, 1998: 137]; los hermanos Cortijo Ocaña apuntan que tras el nombre de Mosén Guillén Pierres se encuentre Jerónimo de Cáncer y Velasco o “uno de los tres ingenios «de La Habana» que se mencionan en el ms. de La muerte de Valdovinos” o Bernardo de Quirós [2002: 81]; Madroñal Durán cree que el autor puede ser el sevillano Félix Persio, Bertiso, autor del *Entremés del rey Cachumba de Motril y la infanta Palancona* [2011: 74]; y Baras Escolá vuelve a pensar en Cervantes [2012: 199].

4.3.9. *Entremés de Doña Justina y Calahorra*

Este entremés en verso es otro de los que se conserva tanto en el códice de la Biblioteca Colombina (82-3-40, folio 92) como en el de la Biblioteca Rodríguez Moñino (S-14-956). La trama gira en torno a la venganza que llevan a cabo Justina y

¹⁴² Nider da una descripción detallada de todos los testimonios conservados [2020], aunque no hace referencia al manuscrito decimonónico de la Biblioteca Rodríguez Moñino, descrito por Madroñal Durán [1998].

Clara tras descubrir que a sus dos viejos maridos, Calahorra y Matanga, “les güele mejor lo que se guisa en otras casas”.

El primero en hacer referencia a la autoría de este interludio anónimo es J. M. Asensio, pero no para atribuírselo a Cervantes sino a Quevedo. Castro, sin embargo, cree que debió escribirlo Cervantes en sus últimos años, porque “la manera de componer versos sueltos, y de empezar el diálogo, es muy propia suya” [1874: 20]. A día de hoy, salvo Rodríguez López-Vázquez, nadie parece apoyar las tesis de Castro y, de un tiempo a esta parte, se está poniendo el acento en las similitudes que hay entre este entremés y varios pasajes de *La socarrona Olalla*, de Quiñones de Benavente, o *Los refranes del viejo celoso*, también atribuido al contador toledano [Muñoz, 2011: 305-306; Madroñal Durán, 2013: 172].

4.3.10. *Entremés de Ginetilla ladrón*

La nómina de los diez entremeses atribuidos a Cervantes se cierra con *Ginetilla ladrón*, un entremés perdido y del que muy poco se sabe. Es J. M. Asensio el que da cuenta de él en la primera de las dos cartas que dirige a José de Palacio Vitery y Mariano Pardo de Figueroa en 1874.

Guardaba nuestro docto Álava un cuaderno manuscrito de diferentes letras, todas, al parecer, del siglo XVII, que, entre otras piezas menores, contenía dos ó tres *Entremeses*. Uno disparatado se intitulaba *El poeta*; otro se llamaba *Ginetilla, ladrón*; y me lo mostró y leyó mil veces, porque, en su concepto, era obra *descarriada y sin el nombre de su dueño, perteneciente á Cervantes*, en la cual se vislumbraba algo que quería parecerse al embrión del gobierno de Sancho en la Insula Barataria. Hace muchos años que vi el manuscrito, y solamente recuerdo que *Ginetilla* se fingía corregidor de un pueblo, y sus compañeros iban por fiscal, escribano y alguaciles, y daban algunas providencias, como podían esperarse de tal gente [1874: 557].

Asensio, aunque cuestiona la paternidad cervantina de esta pieza, recuerda que los dos argumentos en los que basaba Álava su atribución a Cervantes son la cercanía del nombre del protagonista con el de Ginés de Pasamonte y la semejanza entre un pasaje

de este entremés y otro de *La elección de los alcaldes de Daganzo*¹⁴³. Mientras no encontremos el manuscrito al que hacen referencia ambos cervantistas —y que según E. Asensio “debe parar con los papeles de Álava en el Seminario de Vitoria” [1971: 99]— nada podemos decir en relación a su autoría.

¹⁴³ El fragmento que transcribe Asensio dice así:

- GIN. ¿Qué es lo que más se usa en vuestro oficio?
- BOTIC. Señor, de la geringa el ejercicio.
- GIN. Gran oficial seréis, que es peregrina
 y general salud la melecina.
 Llégose á mi una vez cierto harriero,
 que avía perdido el pobre cuatro mulos
 pidiéndome remedio para hallarlos;
 y yo le aconsejé que al mismo instante
 se enflautase una buena melecina.
 Así lo hiço, y en saliendo al campo
 para hacer de su cuerpo purgatorio
 halló los muslos, y esto es muy notorio [1874: 558].



Imagen 2. Portada de la *Séptima parte de las comedias de Lope de Vega* (Madrid, 1617), en donde se imprimen por primera vez el *Entremés de los habladores*, el *Entremés de la cárcel de Sevilla* y el *Entremés del hospital de los podridos*. Fuente: Biblioteca Digital Hispánica.

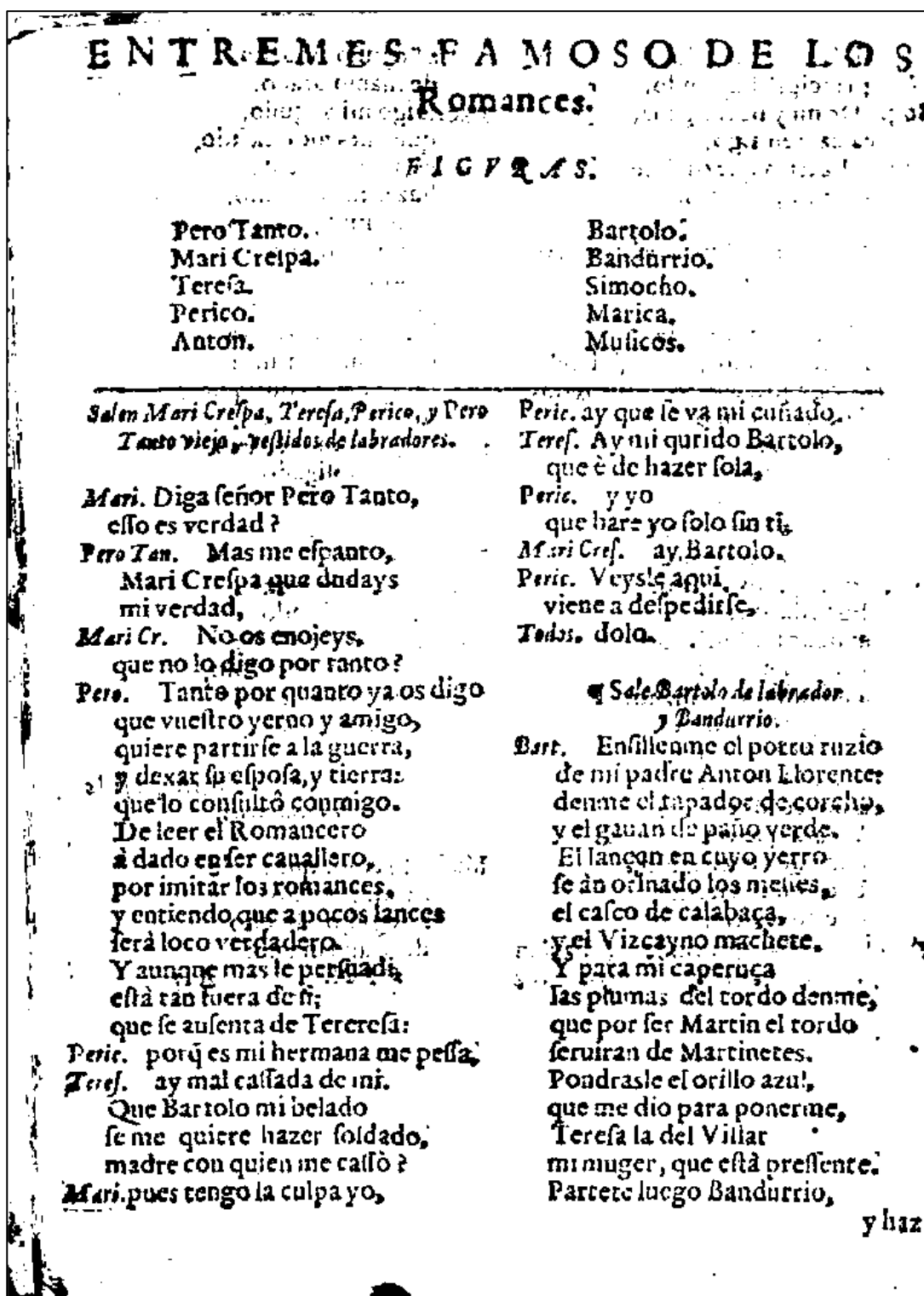


Imagen 3. Primer folio del *Entremés famoso de los romances*, tomado de la *Tercera parte de las comedias de Lope de Vega* (Barcelona, 1612) Fuente: Biblioteca Nacional de España.

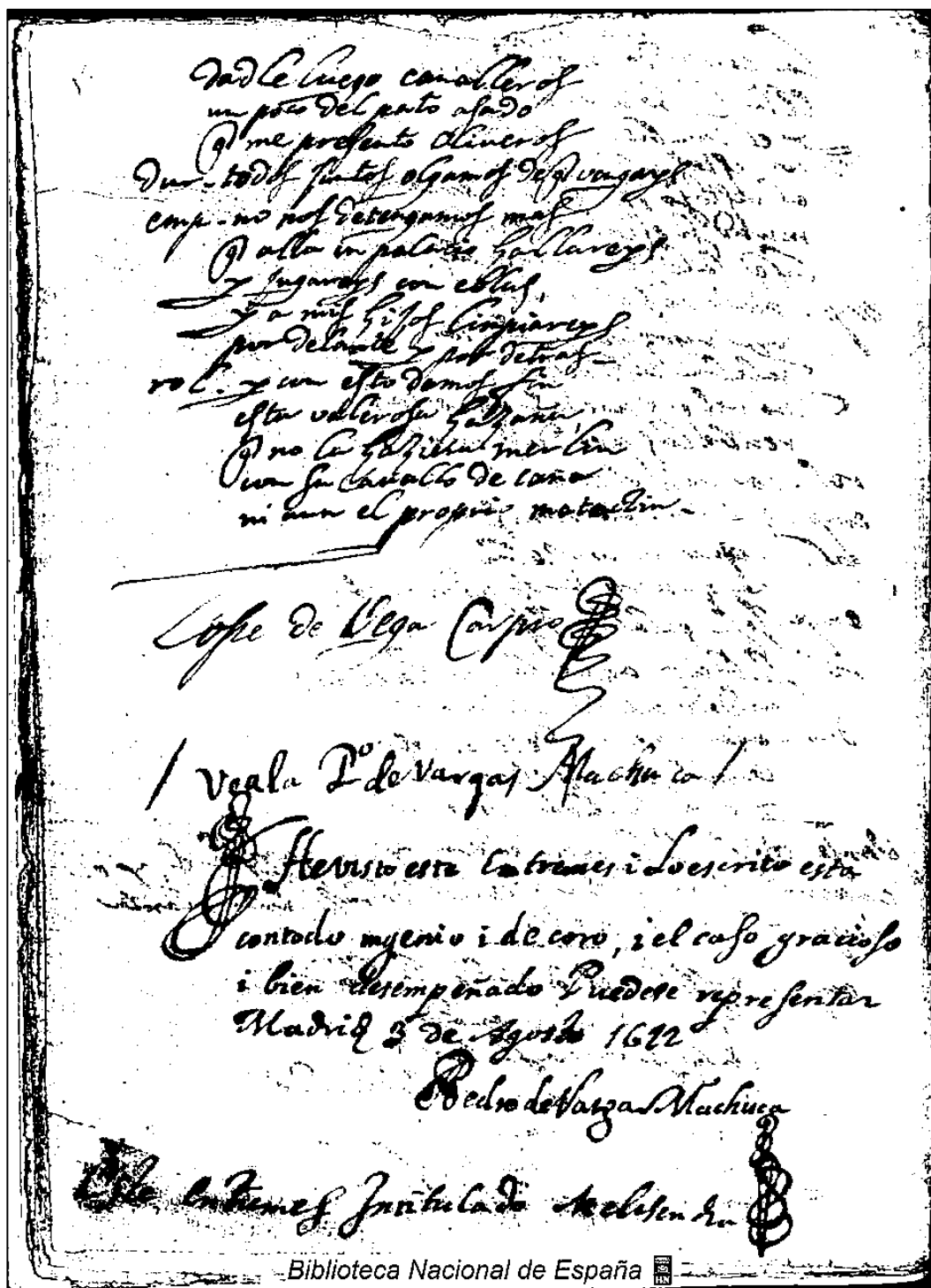


Imagen 4. Último folio del *Entremés de Melisendra*, tomado del manuscrito Res/88 de la Biblioteca Nacional de España, en el que figura la rúbrica —no auténtica— de Lope de Vega. Fuente: Biblioteca Digital Hispánica.

Entremes de los habladores

Vn procurador	y muerriada
Sarmiento	Vn algaazil
Roldan	un miano
Doña Beatri	Cochete

Salen procurador y sarmiento y detras Roldan en abito de cura y paday catillas.

Sarm. Sumere por procurador y uay dantos aducientos ducados y do y patabra a vmd que ayan que meos para que aya en los fustgares que fueran la cuchillada de vmd tantos puntos.

Proc. Vmd a hecho como cauallero en darsela y como erudiano en pagarselo y lleuo el dinero contento de que me se canse y exenmede.

Rol. aca allu es vmd procurador.

Proc. que es lo que manda vmd.

Rol. que dinero es?

Proc. Damiseste cauallero para pagar la parte a que me diouna cuchillada de doce puntos.

Rol. y quanto es el dinero?

Proc. Diezientos ducados.

Rol. vago vmd con diez.

Proc. Diez es a diez vmd.

Rol. de cauallero.

Sarm. Fami gente en bre.

Rol. a vmd digo?

Sarm. y que es lo que manda.

Rol. albrase vmd que no hablar patabra.

Sarm. y que me dizeis.

Rol. Sentencia yo soy un pordiosero del go que me uisto en honra ter go neceidad y a vmds que vmd adado a un hombre de auctos ducados a quina habido una cuchillada y por vmd tiene de by se dadas vengo a que vmd.

Imagen 5. Primer folio del Entremés de los habladores, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

ep morder con hallar
 Ambido mure en or
 yomena y en rigor
 aun quidi gasa mugel
 por no me dar de comer
 Vite y casa hablada
 Vanse todos dando se bay a lo que se da fin

—Entremés famoso de la cárcel
de Sevilla

Garay	Varecruano
Sotapo	Frullilla
Payzano	Del rana
Alcayde	Un proculador
copilla pizaro	dos murel
Baragan.	

Suenan dentro ruydo de grillos cadel y prios
 y diensin salira fuer

Gar.	Abra qui Alcay de que nos io manochincha
Sola.	Abra qui lo Alcay de que nos io manochincha
Pay.	Saque nos amhar. Alcay de

Salen Garay Sotapo y Payzano con grillos
 en lot y jet y quitarras -

Gar.	Coado sea Dios que vos dios de rito
Sola.	Coado sea Dios que vos co el nubi era
Pay.	Coado sea Dios que vos co el sen pite no
Sola.	Se vos mios todos con qui tarros que us eno

Imagen 6. Primer folio del *Entremés famosos de la cárcel de Sevilla*, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

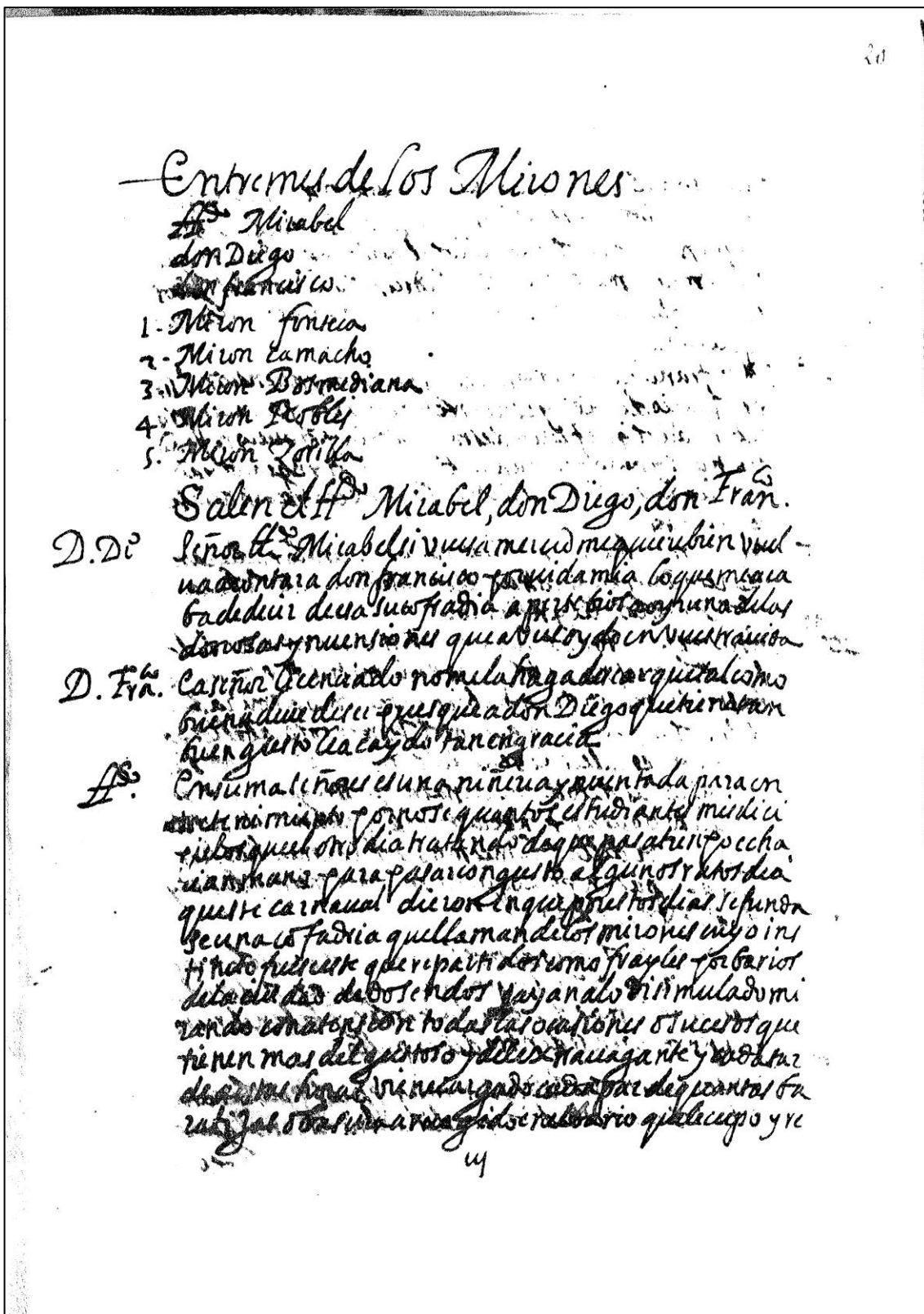


Imagen 7. Primer folio del *Entremés de los mirones*, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

Famoso entremés de Melisendra

Salen vestidos al traçe de los fidalgo de blucas
 y alomuy y piao con ruyos de la nobleza quando
 cuellos grandes de papel espadas que lo abia y al
 guinas de palo. Eora muy vieja se camina que
 seguira el pu en media dilla. Gay por una parte
 ro duan abra de alto gran folio y de la otra parte del
 son bruo hecha del mismo. Hicieron y lusnada en la
 oracion capucitas y son bruo de las parras un
 follado de una calca y parras y un cuerno de una
 vnacabina al cuello de la. Des parte turna en
 gala y un foro de opus in celo y muy y tra. Mili
 sendra en brua en un os cabitros blancos y un os
 patos y por las un palio de una folada. Vella
 las varas de caña y de otras de un modo de la

Un muchacho que se halaba
 don Rodon
 Duandante
 vn pintero de palacio
 oliuros
 Galalon
 don gajeros

Un melico ofe han
 con yndor laia
 baldauin
 Melisendra
 don marcos
 vnacabina

Mucha.

Loa

Senado hume y prudente
 aqui al bra un chormet
 que dilo mira la gente
 de la cabeza al b. gij
 pare que es de repente
 solo se uen a castaya
 para hacer en mo lino
 ados que se candecusar
 y que se paso el papino
 no ay si no de i. mollar

Imagen 8. Primer folio del Famoso entremés de Melisendra, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

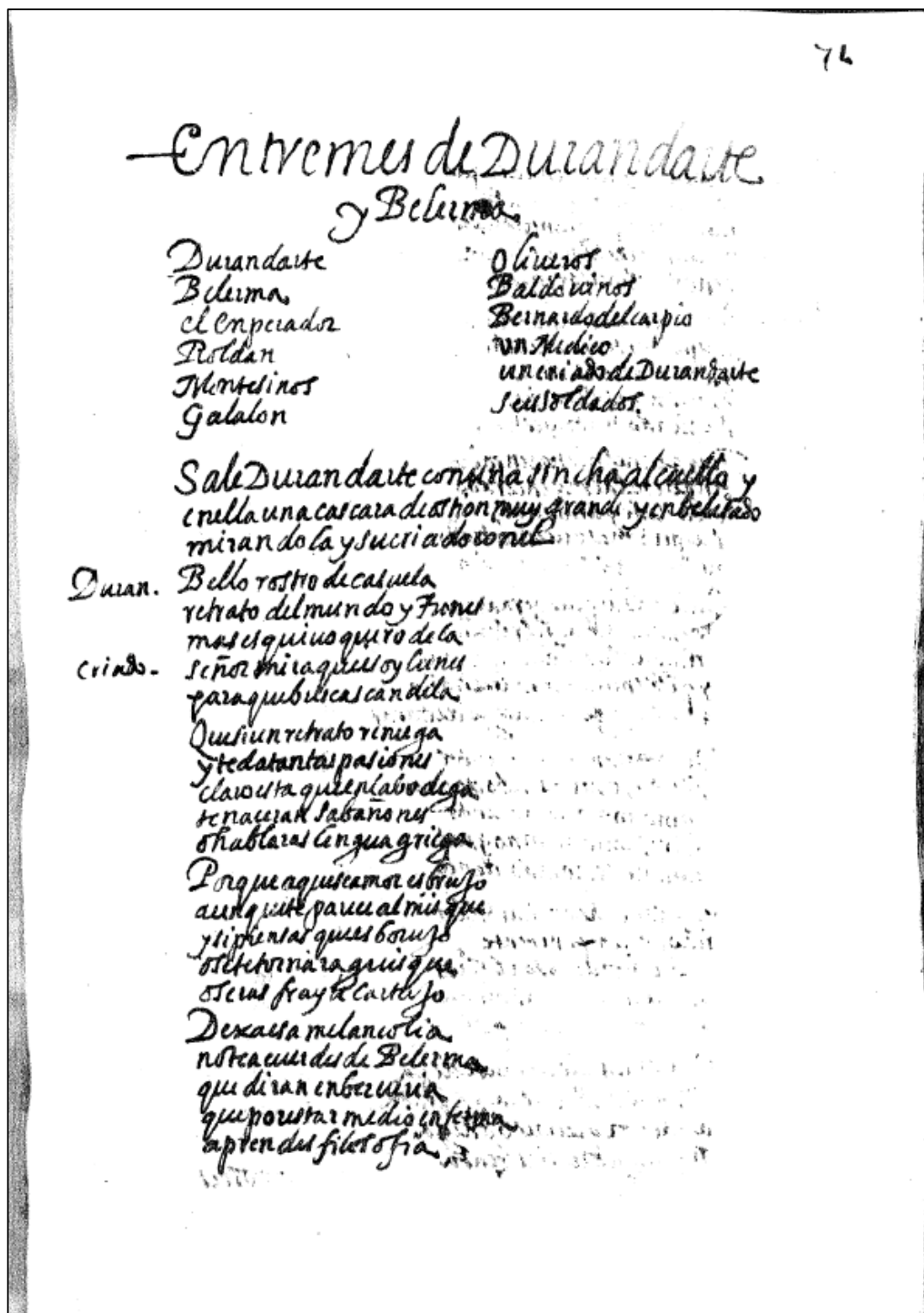


Imagen 9. Primer folio del *Entremés de Durandarte y Belerma*, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

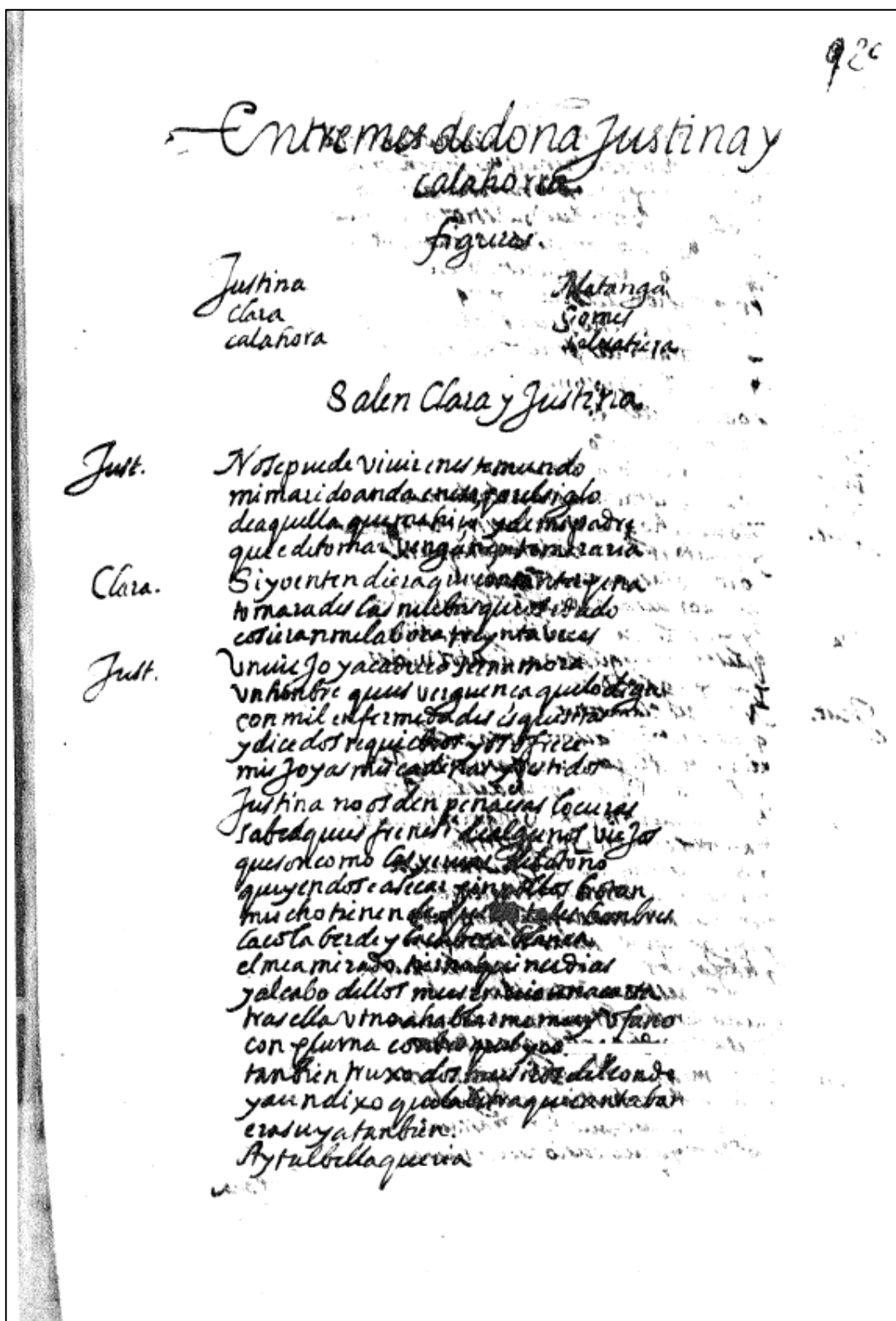


Imagen 10. Primer folio del *Entremés de Doña Justina y Calahorra*, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

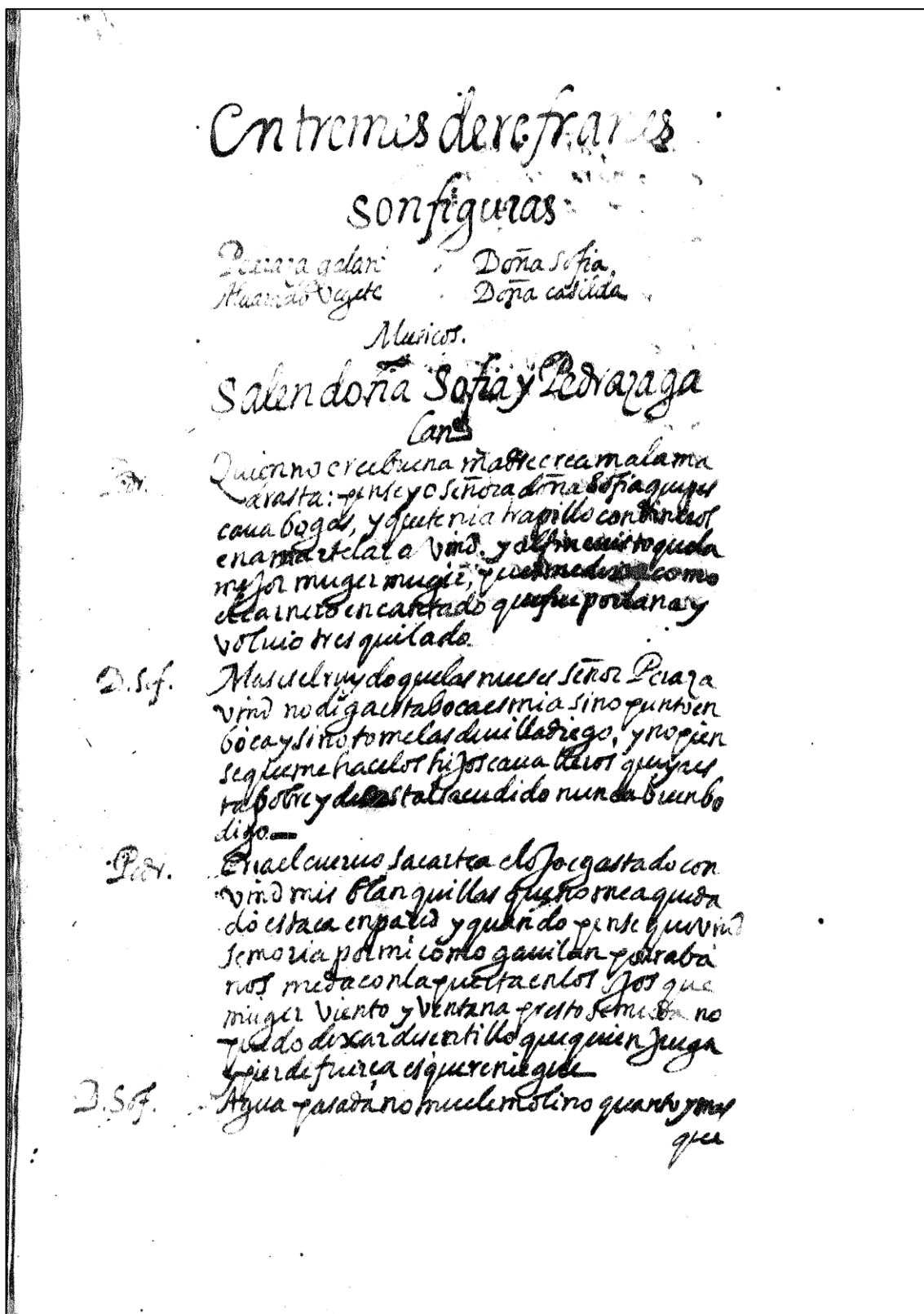


Imagen 11. Primer folio del *Entremés de los refranes*, tomado del manuscrito 82-3-40 de la Biblioteca Colombina de Sevilla. Fuente: Biblioteca Colombina.

CAPÍTULO 5

APROXIMACIÓN AL CASO DE ESTUDIO

“Como cervantinos se han considerado, por ejemplo, algunos entremeses de paternidad ciertamente dudosa que hay que revisar, pues no hay nada definitivo al respecto” (José Montero Reguera).

5.1. OBJETO DEL ESTUDIO

El presente proyecto de tesis doctoral surgió con el objeto de investigar, desde la metodología desarrollada por la Lingüística forense, si los diez entremeses que habían sido atribuidos a Cervantes en época decimonónica eran suyos o no, y con qué grado de probabilidad. Como éramos conscientes de la imposibilidad de realizar un análisis exhaustivo de las nueve piezas conservadas, optamos por acotar nuestro estudio a dos de ellas: *La cárcel de Sevilla* y *Los romances*. Esta selección no fue aleatoria, sino que estuvo condicionada esencialmente por dos motivos: en primer lugar, el interés por analizar un entremés en prosa y otro en verso (motivo lingüístico) y, en segundo lugar, el deseo de averiguar quién estaba tras la firma de estas dos obras tan vinculadas a Cervantes (motivo literario), pues recordemos que el primero de los entremeses dubitados guarda relación temática con la *Relación de la cárcel de Sevilla*, cuya tercera parte se ha atribuido también a Cervantes, y el segundo con la primera salida de don Quijote.

5.2. MARCO METODOLÓGICO

Tal y como postula la Lingüística forense, para poder determinar si un entremés dado ha sido escrito o no por Cervantes (y con qué grado de probabilidad) debemos llevar a cabo una comparación entre la obra dubitada y las indubitadas de este autor, haciendo especial hincapié en la comparación con las obras del mismo género textual (es decir, los ocho entremeses publicados en 1615); asimismo, debemos confrontar todos estos textos con otros textos no cervantinos, lo más cercanos posible en fecha y género al dubitado, pues sólo así podremos establecer si las posibles coincidencias pueden entenderse como marcas propias e idiosincrásicas de autoría o si, por el contrario, responden a usos comunes de la lengua de la época o del género literario. La investigación, siguiendo el principio de certeza, valorará las evidencias lingüísticas desde un punto de vista cualitativo y cuantitativo.

El complejo proceso de creación artística y la larga cadena de eslabones que conforman la transmisión textual en el Siglo de Oro español dificultan no sólo la identificación de los autores reales, sino también la conservación fidedigna de las obras, lo que nos lleva a tener que desconfiar de “las letras de los textos heredados y las autorías que en ellos constan” [Vega, 2002: 1292]¹⁴⁴. Estas circunstancias se agravan aún más en el caso de los entremeses, por el carácter secundario que tienen respecto de la obra principal, y, sobre todo, por su naturaleza mixta y su brevedad: tal y como explica Asensio en su famoso *Itinerario*, el entremés es un género inestable, de corta extensión, fabricado a veces con remiendos cómicos y que se mueve “entre la historieta y la revista, la fantasía y el cuadro de costumbres” [1971: 40]. Por este motivo, a los inconvenientes generales de los estudios de atribución de autoría —ya comentados en el capítulo 2— hay que sumar toda una serie de limitaciones metodológicas derivadas de las especiales características de estos textos dubitados. Además, el hecho de que los textos conservados no sean manuscritos autógrafos, imposibilita poder utilizar variantes

¹⁴⁴ Como advertíamos en el apartado 1.3., los textos teatrales del Siglo de Oro son productos abiertos, condicionados por su carácter mercantil, y presumiblemente modificados a manos de los *autores* de comedias, los censores, los copistas, los impresores, etc.

relacionadas con las grafías o los signos de puntuación, a pesar de la operatividad de las mismas.

5.3. DELIMITACIÓN Y DESCRIPCIÓN DEL CORPUS DE ANÁLISIS

5.3.1 Delimitación del corpus de análisis y tratamiento de los textos

El primer paso para llevar a cabo un estudio sobre la posible autoría cervantina de los entremeses dubitados es el de establecer un corpus de análisis para la comparación. Asignaremos a cada texto un tejuelo o etiqueta identificativa, formada por tres unidades:

- la primera hace referencia al autor, distinguiendo entre autoría dubitada (Dub) o indubitada, en cuyo caso indicamos el nombre abreviado (“Cerv” para Cervantes, “Av” para Francisco de Ávila, “Barr” para Gaspar de Barrionuevo y Carrión, “Belm” para Luis de Belmonte Bermúdez, “Cast” para Alonso de Castillo de Solórzano, “Hurt” para Antonio Hurtado de Mendoza, “Quev” para Francisco de Quevedo, “Quin” para Luis Quiñones de Benavente, “Salas” para Alonso Jerónimo de Salas Barbadillo, “Vel” para Luis Vélez de Guevara, “Chav” para Cristóbal de Chaves y “An” para los anónimos, a los que no catalogamos aquí como dubitados porque no son textos susceptibles de ser cervantinos).
- la segunda, al género o subgénero literario, clasificando los textos en novelas (N), novelas ejemplares (NE), comedias (C), relaciones (R) o entremeses, en cuyo caso indicamos además si es en prosa (EP) o en verso (EV);
- y la tercera, al título de la obra, del que tomamos como referencia la primera palabra con significado léxico.

Así, por ejemplo, el entremés en verso de Cervantes titulado de *La elección de los alcaldes de Daganzo* tendría la etiqueta “Cerv_EV_eleccion”. Advertimos que no utilizamos ni la letra ñ, ni tildes, para evitar posibles problemas derivados del procesamiento de este tipo de caracteres.

5.3.1.1. *Corpus dubitado: los diez entremeses alguna vez atribuidos a Cervantes*

El corpus dubitado —es decir, aquel cuya autoría se cuestiona— está integrado por todos los entremeses que alguna vez han sido atribuidos a Cervantes (Tabla 7). Como el entremés de *Ginetilla ladrón* se ha perdido, el corpus lo conforman realmente los nueve restantes: *Los habladores*, *El hospital de los podridos*, *La cárcel de Sevilla*, *Los refranes* y *Los mirones*, en prosa; y *Melisendra*, *Los romances*, *Doña Justina* y *Calahorra* y *Durandarte y Belerma*, en verso.

AUTOR	TÍTULO	ETIQUETA
¿Miguel de Cervantes?	<i>Los habladores</i>	Dub_EP_habladores
	<i>El hospital de los podridos</i>	Dub_EP_hospital
	<i>La cárcel de Sevilla</i>	Dub_EP_carcel
	<i>Los mirones</i>	Dub_EP_mirones
	<i>Los refranes</i>	Dub_EP_refranes
	<i>Melisendra</i>	Dub_EV_melisendra
	<i>Los romances</i>	Dub_EV_romances
	<i>Doña Justina y Calahorra</i>	Dub_EV_justina
	<i>Durandarte y Belerma</i>	Dub_EV_durandarte
	<i>Ginetilla ladrón</i> (perdido)	Dub_EV_ginetilla

Tabla 7. Corpus integrado por los entremeses atribuidos a Cervantes (textos dubitados)

5.3.1.2. *Corpus indubitado del candidato a autor: las obras de Cervantes*

El corpus indubitado está integrado por las obras de Cervantes de autoría cierta, entre las que no incluimos ni *El cerco de Numancia*, ni *El trato de Argel*, por las circunstancias atributivas ya comentadas (Tabla 8). De todas ellas, las que mejor nos permitirán establecer la comparación de ciertos parámetros serán los ocho entremeses publicados en 1615, por su coincidencia genérica con las piezas dubitadas: *El juez de los divorcios* (prosa), *El rufián viudo* (prosa), *La elección de los alcaldes de Daganzo* (verso), *La guarda cuidadosa* (prosa), *El vizcaíno fingido* (prosa), *El retablo de las maravillas* (prosa), *La cueva de Salamanca* (prosa) y *El viejo celoso* (verso).

APROXIMACIÓN AL CASO DE ESTUDIO

AUTOR	TÍTULO	ETIQUETA
Miguel de Cervantes	<i>La Galatea</i>	Cerv_N_Galatea
	<i>Don Quijote de la Mancha (I)</i>	Cerv_N_quijote1
	<i>La gitanilla</i>	Cerv_NE_gitanilla
	<i>El amante liberal</i>	Cerv_NE_amante
	<i>Rinconete y Cortadillo</i>	Cerv_NE_rinconete
	<i>La española inglesa</i>	Cerv_NE_española
	<i>El licenciado vidriera</i>	Cerv_NE_licenciado
	<i>La fuerza de la sangre</i>	Cerv_NE_fuerza
	<i>El celoso extremeño</i>	Cerv_NE_celoso
	<i>La ilustre fregona</i>	Cerv_NE_ilustre
	<i>Las dos doncellas</i>	Cerv_NE_doncellas
	<i>La señora Cornelia</i>	Cerv_NE_señora
	<i>El casamiento engañoso</i>	Cerv_NE_casamiento
	<i>El coloquio de los perros</i>	Cerv_NE_coloquio
	<i>Don Quijote de la Mancha (II)</i>	Cerv_N_quijote2
	<i>Viaje del Parnaso</i>	Cerv_PN_viaje
	<i>El gallardo español</i>	Cerv_C_gallardo
	<i>La casa de los celos y selvas de Ardenia</i>	Cerv_C_casa
	<i>Los baños de Argel</i>	Cerv_C_baños
	<i>El rufián dichoso</i>	Cerv_C_rufian
	<i>La gran sultana</i>	Cerv_C_sultana
	<i>El laberinto de amor</i>	Cerv_C_laberinto
	<i>La entretenida</i>	Cerv_C_entretenida
	<i>Pedro de Urdemalas</i>	Cerv_C_pedro
	<i>El juez de los divorcios</i>	Cerv_EP_juez
	<i>El rufián viudo llamado Trampagos</i>	Cerv_EV_rufian
	<i>La elección de los alcaldes de Daganzo</i>	Cerv_EV_eleccion
	<i>La guarda cuidadosa</i>	Cerv_EP_guarda
	<i>El vizcaíno fingido</i>	Cerv_EP_vizcaino
	<i>El retablo de las maravillas</i>	Cerv_EP_retablo
	<i>La cueva de Salamanca</i>	Cerv_EP_cueva
	<i>El viejo celoso</i>	Cerv_EP_viejo
<i>Los trabajos de Persiles y Segismunda</i>	Cerv_N_trabajos	

Tabla 8. Corpus integrado por las obras de Cervantes (Textos indubitados)

Además, para facilitar ciertos análisis, crearemos un texto con todas las obras indubitadas de Cervantes (Cerv_total), otro con los ocho entremeses (Cerv), otro con los seis entremeses en prosa (Cerv_EP) y otro con los dos entremeses en verso (Cerv_EV).

5.3.1.3. *Corpus de respaldo: entremeses de época cervantina*

El corpus de respaldo lo conforman 30 entremeses de finales del XVI y principios del XVII, tanto en prosa como en verso (Tabla 9): un entremés de Gaspar de Barrionuevo y Carrión (*El triunfo de los coches*)¹⁴⁵, cuatro de Alonso Jerónimo de Salas Barbadillo (*El buscaoficios*, *Los mirones de la Corte*, *El caprichoso en su gusto* y *El comisario contra los malos gustos*), tres de Francisco de Quevedo (*El marido Pantasma*, *El niño* y *Peralvillo de Madrid* y *La vieja Mutañones*), dos de Francisco de Ávila (*Los invencibles hechos de don Quijote de la Mancha* y *El mortero y chistes del sacristán*), dos de Alonso de Castillo de Solórzano (*El barbador* y *El comisario de figuras*), dos de Luis de Belmonte Bermúdez (*Sierra Morena de las mujeres* y *Una rana hace ciento*), tres de Antonio Hurtado de Mendoza (*Getafe*, *El examinador Miser Palomo* y *Miser Palomo, el médico del espíritu*), cuatro de Luis Quiñones de Benavente (*El ventero*, *Los sacristanes*, *El barbero* y *El borracho*), tres de Luis Vélez de Guevara (*La burla más sazónada*, *Los atarantados* y *Antonia y Perales*) y seis anónimos, pero que no son sospechosos de haber sido escritos por Miguel de Cervantes (*Pedro Hernández* y *el Corregidor*, *El estrólogo borracho*, *Mazalquiví*¹⁴⁶, *La mómola*, *El padre engañado* y *El doctor simple*)¹⁴⁷.

¹⁴⁵ No debemos confundir a Gaspar de Barrionuevo y Carrión con Gabriel de Barrionuevo, como hiciera en su día La Barrera [Madroñal, 1993].

¹⁴⁶ Recordemos lo que Cotarelo Valledor decía de las atribuciones de entremeses a Cervantes: “es muy probable que algunos de los anónimos de fines del siglo XVI, como el famoso de *Mazalquiví*, sean suyos” [1915: 70].

¹⁴⁷ Nos hemos visto obligados a tener que contar con entremeses anónimos porque hay pocos entremeses en prosa de autor conocido. No los hemos catalogado como dubitados, sino como anónimos, porque el objeto de este trabajo no está relacionado con su autoría.

APROXIMACIÓN AL CASO DE ESTUDIO

AUTOR	TÍTULO	ETIQUETA
Anónimo	<i>Pedro Hernández y el Corregidor</i>	Anon_EP_pedro
	<i>El estrólogo borracho</i>	Anon_EP_estrolago
	<i>Mazalquiví</i>	Anon_EP_mazalquivi
	<i>La mánola</i>	Anon_EP_mamola
	<i>El padre engañado</i>	Anon_EP_padre
	<i>El doctor simple</i>	Anon_EP_doctor
Gaspar de Barrionuevo y Carrión	<i>El triunfo de los coches</i>	Barr_EP_triunfo
Alonso Jerónimo de Salas Barbadillo	<i>El buscaoficios</i>	Salas_EP_buscaoficios
	<i>Los mirones de la Corte</i>	Salas_EP_mirones
	<i>El caprichoso en su gusto</i>	Salas_EV_caprichoso
	<i>El comisario contra los malos gustos</i>	Salas_EV_comisario
Francisco de Quevedo	<i>El marido Pantasma</i>	Quev_EV_marido
	<i>El niño y Peralvillo de Madrid</i>	Quev_EV_niño
	<i>La vieja Mutañones</i>	Quev_EP_vieja
Francisco de Ávila	<i>Los invencibles hechos de don Quijote de la Mancha</i>	Av_EV_quijote
	<i>El mortero y chistes del sacristán</i>	Av_EV_mortero
Alonso de Castillo de Solórzano	<i>El barbador</i>	Cast_EV_barbador
	<i>El comisario de figuras</i>	Cast_EV_comisario
Luis de Belmonte Bermúdez	<i>Sierra Morena de las mujeres</i>	Belm_EV_sierra
	<i>Una rana hace ciento</i>	Belm_EV_rana
Antonio Hurtado de Mendoza	<i>Getafe</i>	Hurt_EV_getafe
	<i>El examinador Miser Palomo</i>	Hurt_EV_miser1
	<i>Miser Palomo, el médico del espíritu</i>	Hurt_EV_miser2
Luis Quiñones de Benavente	<i>El ventero</i>	Quin_EV_ventero
	<i>Los sacristanes</i>	Quin_EV_sacristanes
	<i>El barbero</i>	Quin_EV_barbero
	<i>El borracho</i>	Quin_EV_borracho
Luis Vélez de Guevara	<i>La burla más sazónada</i>	Vel_EV_burla
	<i>Los atarantados</i>	Vel_EV_atarantados
	<i>Antonia y Perales</i>	Vel_EV_antonia

Tabla 9. Corpus integrado por los entremeses de respaldo (Textos indubitados)

5.3.1.4. Otros textos de respaldo

También utilizaremos como corpus de respaldo los textos contenidos en el CORDE, circunscribiéndonos cronológicamente al periodo temporal comprendido entre 1547 (año del nacimiento de Cervantes) y 1617 (año de la publicación de *Los trabajos de Persiles y Segismunda*, que vio la luz, póstumamente, un año después del fallecimiento del autor). Utilizaremos este recurso con la cautela que requiere, pues en el CORDE están recogidas todas las obras de Cervantes, Quevedo, Góngora o Lope, pero no todas las de los autores considerados ‘menores’¹⁴⁸. Asimismo:

- para el estudio del entremés de *La cárcel de Sevilla*, utilizaremos las tres partes de la relación a la que está temáticamente ligado: la *Relación de las cosas de la cárcel de Sevilla y su trato* y el *Ensanche de la segunda parte de las cosas que pasan en la cárcel*, ambas de Cristóbal de Chaves (Chav_R_carcel1 y Chav_R_carcel2), y la *Tercera parte de las cosas de la cárcel de Sevilla*, cuya autoría se disputan el mismo Chaves y Cervantes (Anon_R_carcel3);
- y para el estudio del entremés de *Los romances*, utilizaremos un texto conformado por los seis primeros capítulos de *El Quijote* de 1605 (Cerv_N_quijote1_1.6).

5.3.1.5. Localización y tratamiento de los textos

Todos los textos de Cervantes han sido tomados de las ediciones de Florencia Sevilla Arroyo que figuran en la Biblioteca Virtual Miguel de Cervantes. Los entremeses, además de en este mismo recurso, se han extraído de la *Flor de entremeses y sainetes de diferentes autores* (Madrid, 1657) [Menéndez Pelayo, ed., 1903], de la *Colección de entremeses* de Emilio Cotarelo y Mori [1911], del *Itinerario del entremés*

¹⁴⁸ Esta desigualdad en el número total de palabras por autor hace que podamos usar este recurso para testar la fuerza idiosincrásica de una determinada secuencia de palabras en época cervantina, pero no tanto para comparar su frecuencia de uso entre dos autores (salvo que calculáramos las frecuencias relativas).

de Eugenio Asensio [1971] y de algunas ediciones específicas [Arellano y García Valdés, 1997 y Nider, 2020]. Por su parte, las tres partes de la *Relación de la cárcel de Sevilla* se han extraído del trabajo de Hernández Alonso y Sanz Alonso sobre *Germanía y sociedad en los Siglos de Oro* [1999]¹⁴⁹.

Para facilitar el tratamiento automatizado de los textos y posibilitar su comparación, los hemos convertido a formato UTF-8 (*8-bit Unicode Transformation Format*) y hemos modernizado su ortografía, pero manteniendo las variantes gráficas que implican una variante fonética (como *agora* o *escura*), las contracciones (*desto*, *deso*, *dello*, *daquel*, *esotro*...), las desinencias verbales enclíticas (del tipo *cogióla* o *dejáronles*) y las formas verbales arcaicas en *-alle*, *-elle* e *-illa* (*satisfacelle*, *vencelle*, *decilla*...). Asimismo, hemos optado por eliminar los títulos y, en el caso de las obras dramáticas, la lista de figuras y el nombre del personaje que antecede a cada intervención; sí hemos mantenido, en cambio, las acotaciones, que en el caso de Cervantes resultan idiosincrásicas. El corpus de trabajo puede consultarse en el siguiente repositorio de *GitHub*: <https://github.com/Cruizurbon/Entremeses-SdO>.

5.3.2. Descripción de las muestras textuales

Junto a la más que probable contaminación textual y/o autorial derivada de las circunstancias de transmisión del entremés barroco, la mayor limitación metodológica de nuestro estudio está relacionada con el tamaño de las muestras.

¹⁴⁹ Sin duda, uno de los mayores inconvenientes de los investigadores que trabajamos con herramientas de análisis textual en el ámbito hispánico es la dificultad de poder acceder a textos en formato XML/TEI, lo que nos obliga a tener que emplear un tiempo considerable al procesamiento de los mismos. No podemos estar más de acuerdo con las palabras de Calvo Tello: “En mi opinión una de las principales dificultades para trabajar en diferentes métodos de Humanidades Digitales con textos en español es precisamente la falta de textos literarios disponibles en formatos aceptables. Un ejemplo paradigmático de esto, aunque no el único, es la Biblioteca Virtual Miguel de Cervantes, un proyecto que ha editado miles de textos en formato XML/TEI, lo que merece un enorme reconocimiento, pero que no solo no lo publica en ese formato, sino que tampoco lo facilita a investigadores que lo solicitan. Esto hace que cada investigador tenga que encargarse de conseguir un texto digno a partir de las obras troceadas en diferentes páginas HTML en formato hoy en día obsoleto. Por supuesto comenzar el trabajo desde ese HTML siempre es mejor que tener que comenzar desde el escaneo del libro” [2016: 146-147].

CAPÍTULO 5

OBRA	EXTENSIÓN	OBRA	EXTENSIÓN
Dub_EP_habladores	2 574 palabras	Av_EV_quijote	2 584 palabras
Dub_EP_hospital	2 492 palabras	Av_EV_mortero	2 268 palabras
Dub_EP_carcel	3 240 palabras	Barr_EP_triunfo	4 701 palabras
Dub_EP_mirones	9 220 palabras	Belm_EV_sierra	1 012 palabras
Dub_EP_refranes	2 384 palabras	Belm_EV_rana	871 palabras
Dub_EV_melisendra	3 019 palabras	Cast_EV_barbador	1 294 palabras
Dub_EV_romances	2 243 palabras	Cast_EV_comisario	1 422 palabras
Dub_EV_justina	1 936 palabras	Hurt_EV_getafe	1 602 palabras
Dub_EV_durandarte	5 293 palabras	Hurt_EV_miser1	2 353 palabras
Dub_EV_ginetilla	(Perdido)	Hurt_EV_miser2	2 108 palabras
Cerv_EP_juez	3 004 palabras	Quev_EV_marido	1 748 palabras
Cerv_EV_rufian	2 748 palabras	Quev_EV_niño	1 390 palabras
Cerv_EV_eleccion	2 386 palabras	Quev_EP_vieja	2 102 palabras
Cerv_EP_guarda	3 684 palabras	Quin_EV_ventero	1 437 palabras
Cerv_EP_vizcaino	4 418 palabras	Quin_EV_sacristanes	1 824 palabras
Cerv_EP_retablo	3 300 palabras	Quin_EV_barbero	986 palabras
Cerv_EP_cueva	3 568 palabras	Quin_EV_borracho	1 859 palabras
Cerv_EP_viejo	3 790 palabras	Salas_EP_buscaoficios	3 941 palabras
Anon_EP_pedro	2 271 palabras	Salas_EP_mirones	2 254 palabras
Anon_EP_estrolago	1 497 palabras	Salas_EV_caprichoso	3 510 palabras
Anon_EP_mazalquivi	1 510 palabras	Salas_EV_comisario	2 729 palabras
Anon_EP_mamola	1 988 palabras	Vel_EV_burla	1 144 palabras
Anon_EP_padre	2 691 palabras	Vel_EV_atarantados	1 440 palabras
Anon_EP_doctor	2 103 palabras	Vel_EV_antonia	1 094 palabras

Tabla 10. Extensión de todos los entremeses que conforman el corpus de análisis (Nº total de palabras)

La extensión de los entremeses que conforman el corpus de análisis es de [986-9220] palabras (Tabla 10), algo que complica enormemente nuestra labor, pues, como advertimos en el capítulo 2, la precisión de los análisis cuantitativos decrece significativamente cuando la longitud del texto es inferior a las 1000 palabras [Hirst y Feiguina, 2007].

También es importante advertir que dos de los entremeses dubitados tienen una extensión considerablemente mayor a la propia de un entremés, lo que nos lleva a replantearnos su identidad genérica. Así ocurre con *El entremés de Durandarte y Belerma*, de 5293 palabras, que fue publicado en 1679 como “comedia famosa y burlesca” de Mosén Guillén Pierres, bajo el título *El amor más verdadero, Durandarte y Belerma*; y con *El entremés de los mirones*, de 9220 palabras, estructurado como un coloquio en prosa más cercano a una novela ejemplar (como *Rinconete y Cortadillo* o el *Coloquio de los perros*) que a un entremés, pues, de hecho, carece de algunos elementos básicos de este género (triplica la extensión habitual, no se cierra con ninguna canción en verso, no hay agilidad en los diálogos, etc.).

Junto a la contaminación (autorial y textual) y la extensión, otra de las limitaciones a tener en cuenta estará relacionada con la temática, los personajes y los espacios de la ficción, ya que debemos ser conscientes de que la refundición total o parcial de unas piezas en otras fue una tendencia constante en toda la historia del teatro aurisecular español; además, el costumbrismo literario de los entremeses, que tienden a caricaturizar ciertos usos y costumbres de la época, potencia esta reutilización de tramas y argumentos.

5.3.2.1. *La ridiculización del matrimonio y el adulterio*

En el género entremesil, el matrimonio es abordado desde múltiples y muy variadas perspectivas, pero siempre en tono satírico-burlesco, pues el principal objetivo es provocar la risa. En *La guarda cuidadosa*, *El triunfo de los coches* o *Los atarantados* se trata el tema de los acuerdos matrimoniales, en los que prima el dinero por encima del amor. En *El marido fantasma* se muestra la animadversión al casamiento a través del personaje de Lobón, quien primero desaconseja a su amigo Muñoz que se case, pero que después, tras quedar viudo, le aconseja que lo haga, “que todo puede pasarse / por ver ir en procesión, / kiriada de los niños, / la mujer que nos cansó”. En *El juez de los divorcios* se plantea la problemática de cuatro matrimonios que quieren romper su unión por distintas razones (insatisfacción sexual, desavenencias de tipo económico, celos desmedidos...), pero que terminan cantando “que vale el peor concierto / más que el

divorcio mejor”. Y en *La cueva de Salamanca*, *El viejo celoso*, *El mortero y chistes del sacristán*, o *Doña Justina y Calahorra* se aborda el tema del adulterio; aunque generalmente son las mujeres jóvenes las que engañan a sus maridos viejos e impotentes (como en los antiguos *fabliaux* medievales), en algunos casos ocurre a la inversa, como en el entremés atribuido a Cervantes, en donde son Calahorra y Matanga los que intentan traicionar a Doña Justina y Doña Clara.

5.3.2.2. *El motivo del disfraz y otros engaños*

El disfraz, uno de los *lazzi* característicos de los *Pasos* de Lope de Rueda, aparece en muchos de los entremeses de finales del XVI y principios del XVII. El entremés anónimo de *El doctor simple* presenta a un criado que, vistiéndose con las ropas de su amo, se disfraza de médico y prescribe insólitas recetas a una serie de pacientes que desfilan ante él; el motivo, lejos de ser original, aparecía ya en antiguas comedias clásicas y en ciertos *fabliaux* medievales y está presente en otras muchas obras dramáticas de la época, como el *Paso primero del médico simple* y *Coladilla paje y el doctor Valverde*, atribuido a Lope de Rueda, el entremés de *El doctor [Borrego]*, atribuido a Rojas Zorrilla, o las comedias *El acero de Madrid*, de Lope de Vega, y *La fingida Arcadia*, de Tirso de Molina [Ruiz Urbón, 2007]. También encontramos la treta del disfraz en *El padre engañado*, en donde el galán viste de mujer al criado del viejo para poder raptar a su hija Isabelilla, o en *Doña Justina y Calahorra*, en donde los dos vejetes aceptan disfrazarse de mujeres para poder consumir su infidelidad¹⁵⁰.

Pero el disfraz no es el único engaño característico de los entremeses. En el entremés de *La mómola*, Sofía hace creer a tres galanes que ha tenido un hijo suyo para sacarles el dinero; en *El barbador*, Piruétano y Pescaño fingen tener poderes mágicos para hacer crecer el pelo y la barba a los lampiños; en *El retablo de las maravillas*,

¹⁵⁰ Mención aparte merece *Una rana hace ciento*, de Belmonte, en donde Juan Rana aparece vestido como este animal, con “sayo y capirote verdinegro”. En este caso, el disfraz, lejos de formar parte de una treta argumental, sirve para motivar el lucimiento personal de este mítico actor, que aparece en escena cantando y bailando.

Chanfalla y Chirinos simulan una representación teatral ante todo un pueblo; en *Los refranes*, Sofía dice estar casada con Pedraza para poder cobrar una herencia de mil ducados; en *La burla más sazónada*, el estudiante Tabaco burla a un alguacil, imitando la voz de la dama a la que está persiguiendo; en *El mortero y chistes del sacristán*, Gigorro lleva a cabo la clásica burla del intercambio entre el mortero y el manteo, que ya estaba presente en uno de los cuentos del *Decamerón*; y un intercambio entre objetos vemos también en *El vizcaíno fingido*, en donde Solórzano y Quiñones engañan a Cristina, una prostituta sevillana muy famosa por su taimería, con un cambiazo entre dos cadena de oro, una verdadera y otra falsa, en un juego que presenta muchos puntos en común con uno de los capítulos del *Guzmán de Alfarache*.

5.2.2.3. *La cuestión de la limpieza de sangre*

Otro leitmotiv recurrente en los entremeses barrocos —y en toda la literatura aurisecular— es el de la limpieza de sangre y la honorabilidad del linaje, que se representa a través de chistes antijudáicos y antimoriscos y de toda una serie de tópicos sobre los conversos, relacionados, principalmente, con la repugnancia por ciertos alimentos [Asensio, 1971: 150]. En *La elección de los alcaldes de Daganzo*, Cervantes nos presenta a cuatro labradores que quieren ser alcaldes y que reivindicán, ante Panduro, Alonso Algarroba, Pesuña y Pedro Estornudo, su condición de cristianos viejos (un argumento que luego retomará Quiñones de Benavente en *Los alcaldes encontrados*). En *El retablo de las maravillas*, Chirinos y Chanfalla hacen creer al gobernador y a las autoridades del pueblo que aquellos vecinos que tengan “alguna raza de confeso, o no sea habido y procreado de sus padres de legítimo matrimonio” no podrán ver las maravillas que suceden dentro del retablo. El tema de la calidad del cristiano viejo también se recoge en *El juez de los divorcios*, a través del personaje de Ganapán, y en *El mortero y chistes del sacristán*, donde el vejete Pero Díaz dice haber reñido “con ese barberillo / porque se ha demandado de la lengua / en vituperio de mi honrosa estirpe, / publicando en la plaza que soy moro”.

5.2.2.4. *Presencia de elementos costumbristas*

Los entremeses del Siglo de Oro son un buen reflejo de las costumbres de la época (fiestas, diversiones, juegos, etc.).

En varios de los entremeses de nuestro corpus se aborda la moda de los coches de caballos y, en concreto, la pasión desmedida que sienten las mujeres por ellos; así aparece en *El triunfo de los coches*, de Barrionuevo, en *El comisario contra los malos gustos* y *El buscaoficios*, de Salas Barbadillo, en *Los atarantados*, de Vélez de Guevara, en *El marido pantasma*, de Quevedo, o en *El vizcaíno fingido*, de Cervantes, en donde Brígida siente que se le “arranca el alma” al enterarse “que quitaban los coches”¹⁵¹. La presentación caricaturesca de la obsesión femenina por los coches es un motivo recurrente en toda la dramaturgia del Siglo de Oro español; lo vemos también, sólo por citar algún otro caso, en la comedia *La llave de la honra*, atribuida a Lope, en el *Entremés de los coches*, de Benavente, en el *Entremés de la condesa*, de Juan de Alarcón, o en la comedia *Mañanas de abril y mayo*, de Calderón, en donde se menciona a una mujer que resucita al sentirse dentro de un coche [Fernández Oblanca, 1991; Briosos Santos, 1996; Urzáiz Tortajada, 1997].

Otro de los temas recurrentes es el gusto desmedido por el romancero, el refranero o las novelas de caballerías, que sabemos tuvo su correlato en la realidad. Los entremeses de *Durandarte y Belerma* y *Melisendra* parodian conocidos romances del ciclo carolingio, que fueron enormemente populares en los siglos XV, XVI y XVII; *Los romances* nos presenta a un labrador que ha perdido el juicio por leer el Romancero; y el don Quijote del entremés de Ávila, al igual que el de la novela cervantina, quiere ser armado caballero y liberar “a Dulcinea del Toboso / del castillo encantado, donde asiste / en poder de gigantes y de leones”.

¹⁵¹ El contrapunto a todas estas mujeres obsesionadas con los coches lo encontramos en la villana Francisca, del entremés de *Getafe*, que no sucumbe antes los requiebros de Don Lucas y sus promesas de coches, joyas, vestidos y tocados.

También está presente en varios de los entremeses de nuestro corpus el tema de las salteadoras de la calle Mayor de Madrid. Éste es el motivo central de *Sierra Morena de las mujeres*, en donde cuatro bandoleras asaltan a Garañón y Paris¹⁵²; y a él se hace referencia también en los entremeses de Quevedo *La vieja Mutañones* y *El niño y Peralvillo de Madrid*.

Pero sin duda, dos de los entremeses en los que mejor se da cuenta de ciertos usos de la época son *Los mirones* y *Los mirones de la corte*. En estas dos piezas se hace referencia a la figura del mirón, es decir, aquél que explora el fluir de la vida cotidiana por las calles y plazas de las ciudades (en estos casos, Sevilla y Madrid, respectivamente).

5.2.2.5. Personajes: tipos tradicionales y figuras

Los entremesistas barrocos recurren constantemente a unos mismos personajes (ya sean tipos tradicionales o modernas figuras), que se encuentran a caballo entre dos espacios: el de la realidad cotidiana y el de la literatura y el folclore.

En las obras que representan nuestro corpus encontramos mujeres interesadas (*La mómola*, *El triunfo de los coches...*), malmariadas (*El juez de los divorcios*, *El viejo celoso...*), vejetes (*El padre engañado*, *El borracho*, *El viejo celoso*, *Doña Justina y Calahorra*, *Los refranes*, *El barbero...*), criados bobos (*El padre engañado*, *La mómola*, *El doctor simple...*), galanes (*El padre engañado*, *El borracho*, *Los refranes...*), sacristanes (*La mómola*, *La cueva de Salamanca*, *La guarda cuidadosa*, *Los sacristanes*, *Los atarantados...*), soldados (*El borracho*, *La guarda cuidadosa*), valientes (*Antonia y Perales*, *El examinador miser Palomo...*), boticarios (*El triunfo de los coches*, *El caprichoso en su gusto...*), barberos (*La cueva de Salamanca*, *El barbero...*), alcaldes

¹⁵² Garañón pide a Paris que le acompañe a la calle Mayor, “que, según infinitos pareceres, / Sierra Morena es ya de las mujeres, porque en ella saltean”. La comparación entre Sierra Morena y la calle Mayor de Madrid por los asaltos aparece también en *Todo es ventura*, de Juan Ruiz de Alarcón, en donde leemos: “Yo, señor, / salí a la calle Mayor, / Sierramorena en Madrid, / pues allí roban a tantos / mil damas ricos despojos, / llevando armas en los ojos, / y máscaras en los mantos” [Citamos por Cortijo Ocaña y Cortijo Ocaña, 2006: 64].

rurales (*El retablo de las maravillas, La elección de los alcaldes de Daganzo...*), estudiantes (*La cueva de Salamanca, Los mirones, La burla más sazónada...*), astrólogos (*El estrólogo borracho, La cueva de Salamanca...*), viejas (*La vieja Mutañones, El caprichoso en su gusto...*), rufianes o hampones (*El rufián viudo llamado Trampagos, La cárcel de Sevilla...*), médicos (*El doctor simple*), mesoneros y venteros (*Getafe, El ventero, Los invencibles hechos de don Quijote de la Mancha...*), etc.; y algunos han querido ver en los personajes de Lucas y Francisca del entremés de *Getafe* a los antecedentes del petimetre y la maja dieciochescos [Asensio: 1971: 121].

Además, son varios los entremeses de nuestro corpus que responden a lo que se ha definido como “desfile de tipos o figuras”, en donde distintos personajes destemplados van pasando ante una figura que ejerce como juez o examinador (*El hospital de los podridos, El juez de los divorcios, El buscaoficios, El comisario de figuras, El comisario contra los malos gustos y El examinador miser Palomo*).

5.2.2.6. Espacios de la ficción.

La villa y Corte de Madrid es uno de los espacios más representados, como vemos, por ejemplo, en *El triunfo de los coches, El caprichoso en su gusto, Sierra Morena de las mujeres, El niño y Peralvillo de Madrid* o *Los mirones de la Corte*; junto a Madrid, encontramos ciudades como Sevilla (*La cárcel de Sevilla* o *Los mirones*) o Salamanca (*La cueva de Salamanca*). Otros entremeses encuadran la acción en aldeas (*La elección de los alcaldes de Daganzo, El retablo de las maravillas, Los romances...*) o en las ventas y mesones que unen unas ciudades con otras, como ocurre en *Getafe*.

Estos son sólo algunos ejemplos de las redes temático-argumentales y las concomitancias de personajes y espacios que configuran el espectro entremesil de los siglos XVI y XVII y que podrían afectar, en cierta medida, a los análisis cuantitativos basados en la frecuencia de palabras o el nivel de similitud textual.

5.4. OPERATIVIDAD DEL CORPUS DE ANÁLISIS

Tras haber descrito el corpus de trabajo y poner sobre la mesa sus posibles limitaciones metodológicas, el siguiente paso será el de realizar una serie de análisis de tipo cuantitativo que nos permitan testar su operatividad. Las cuatro comprobaciones que llevaremos a cabo serán de tipo léxico y ortográfico:

- Análisis de la distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*)
- Análisis del porcentaje de similitud textual (*similarity*)
- Análisis de la longitud de secuencias de palabras idénticas (*verbatim*)
- Análisis de la frecuencia de n-gramas de 5 caracteres (*character 5-grams*)

No se trata tanto de testar los parámetros que utilizaremos, sino la validez de los textos seleccionados, pues somos conscientes, como ya hemos señalado en varias ocasiones, que muchos de ellos pueden haber sufrido contaminación autorial o textual. Además, al trabajar con textos relativamente breves, todo proceso de contaminación tendrá más peso en el resultado global¹⁵³.

5.4.1. Análisis de la distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*)

El paquete Stylo, desarrollado en el año 2013, es una de las herramientas más empleadas en Estilometría [Eder, Rybicki y Kestemont, 2016]. La medida mejor acogida hasta la fecha es la del algoritmo Delta, propuesto por John Burrows, que parte

¹⁵³ Nos preocupa especialmente este asunto, ya que la extensión media de un entremés son las 2500 palabras. Hernández Lorenzo, en su estudio sobre la poesía de Fernando de Herrera, excluye del corpus a los autores que no alcanzan las 1800 palabras [2019a: 317]. Por su parte, Campión Larrumbe y Cuéllar consideran que los análisis estilométricos empiezan a ofrecer “resultados fiables a partir de textos con 2000-2500 palabras” [2021: 62].

de la idea de que la variación en las frecuencias de las palabras más empleadas en un texto (*tokens*) permiten reconocer su autoría: “the mean of the absolute differences between the z-scores for a set of word-variables in a given text-group and the z-scores for the same set of word-variables in a target text” [Burrows, 2002: 271].

El algoritmo de Burrows (Classic Delta) ha sido actualizado más tarde por Hoover (Hoover Delta) [2004b], Argamon (Delta Lineal) [2008], Smith y Aldridge (Cosine Delta) [2011] y Maciej Eder (Eder’s Delta) [2013b], que propone una modificación pensada específicamente para las lenguas flexivas¹⁵⁴. Aunque la mayor parte de los estudios con Delta se han hecho con textos en lengua inglesa, alemana, francesa, polaca, húngara o latina, también han probado su efectividad para el español Rißler-Pipka [2016], Wrisley [2016], Calvo Tello [2016], Fradejas Rueda [2016a y 2019b], Blasco [2016, 2019a y 2019b], de la Rosa y Suárez [2016], Rojas Castro [2017], Soler y Calvo Tello [2019], García-Reidy [2019], Hernández Lorenzo [2019a], Vega García-Luengos [2021] y Campión Larumbe y Cuellar [2021]¹⁵⁵.

¹⁵⁴ Para una explicación más detallada de todos los algoritmos, véase Evert, Proisl, Jannidis, Reger, Pielström, Schöch y Vitt [2017].

¹⁵⁵ Los estudios empíricos han demostrado que la eficacia de los algoritmos varía no sólo de unas lenguas a otras, sino también en función del tipo de texto [Rybicki y Eder, 2011]. En español, se han empleado diferentes algoritmos. Blasco, en su trabajo sobre Avellaneda, combina Classic Delta con Eder’s Delta. Fradejas Rueda elige Classic Delta para realizar diferentes análisis estilométricos de textos medievales [2019]. García-Reidy, sin embargo, opta por Cosine Delta para estudiar la autoría de la comedia *Siempre ayuda a la verdad* [2019]. En su estudio de los textos poéticos de Fernando Herrera, Hernández Lorenzo, tras testar diferentes medidas y parámetros, concluye: “La mejor combinación es la aplicación de la Delta de Eder sobre las 4000 palabras más frecuentes con una media de éxito o aciertos superior al 85%, mientras que la segunda es la aplicación de Cosine Delta sobre las 5000 palabras más frecuentes con una media de éxito superior al 84%, y la tercera vuelve a ser con la Delta de Eder, pero esta vez con las 5000 palabras más frecuentes. La Delta de Burrows se muestra menos eficiente que las distancias anteriores con nuestro corpus, pero sobre las 3000 palabras más frecuentes también obtiene una alta media de éxito, superior al 80%, encontrándose en el quinto puesto de las mejores combinaciones de parámetros” [2019a: 314-315]. En el proyecto *ETSO. Estilometría aplicada al Teatro del Siglo de Oro*, están usando el algoritmo Delta propuesto por Burrows (Classic Delta); véanse, por ejemplo, los trabajos publicados recientemente por Vega García-Luengos [2021] o Campión Larumbe y Cuellar [2021] en el número 3 de *Talía. Revista de estudios teatrales*. Nosotros, al igual que Soler y Tello [2019], optamos por Eder’s Delta, que, al menos para nuestro corpus de entremeses indubitados, arroja los resultados más fiables.

Delta trabaja del siguiente modo: tokeniza los textos del corpus, calcula diferentes datos de frecuencia de los *tokens*, los estandariza mediante democratización léxica (*z-scores*) y compara los resultados de la suma de cada palabra en cada pareja de textos (*delta-scores*) para agruparlos jerárquicamente [Calvo Tello, 2016]; aunque Stylo permite acceder a los archivos de cada proceso (lista de palabras, tabla de frecuencias, valores delta para cada pareja de textos, etc.), también ofrece los resultados en forma de gráfico, como dendrogramas (*Cluster Analysis*) o árboles de consenso (*Bootstrap Consensus tree*), lo que resulta ciertamente útil para su comprensión en el ámbito de las Humanidades digitales¹⁵⁶. Delta, al igual que otros métodos estilométricos, basa sus resultados en la comparación de un número *n* textos; en consecuencia, si añadimos o quitamos uno o varios de esos textos, los nodos de ramas pueden variar¹⁵⁷.

En este primer apartado vamos a realizar una serie de análisis contrastivos basados en la distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*), seleccionando el idioma español, el algoritmo Delta de Eder (*Eder's Delta*) y la visualización en clúster (*Cluster Analysis*). Para poder interpretar correctamente los resultados ofrecidos por *Cluster Analysis*, es importante advertir que el 0 indica la identidad total y que, por tanto, cuanto más alejada del 0 se produzca la asociación entre dos textos dados, menos significativa será dicha asociación (o, dicho de otro modo, las asociaciones serán más sólidas y fiables cuanto más a la derecha se produzcan). Asimismo, es importante tener en cuenta que, si hay varias subramas dentro de una misma rama, las que están situadas en el centro serán más similares entre sí que las dos de los extremos (la de arriba y abajo).

¹⁵⁶ El análisis de conglomerados (*Cluster analysis*) es una técnica estadística multivariante cuya finalidad es dividir un conjunto de objetos en grupos. Para evitar la arbitrariedad del investigador en la selección de las MFW (*cherry picking*), Eder introdujo el *bootstrap* para crear árboles de consenso que ofrezcan la agrupación más estable, es decir, aquella que permanece a lo largo de un porcentaje mínimo de iteraciones (*consensus value*): “The procedure introduced above aims to help eschew the problem of arbitrariness. In a very large number of iterations, the variables needed to construct a dendrogram were chosen randomly, and a virtual dendrogram for each iteration was generated. Next, these numerous virtual dendrograms were combined into a single compact consensus tree” [Eder, 2013b].

¹⁵⁷ Por eso, aunque son análisis muy fiables, están limitados por el llamado “*nearest neighbor problem*”. Esto será lo que ocurra con Barrionuevo, del que sólo tenemos un texto indubitado (*El triunfo de los coches*), que se acercará, por defecto, a su “vecino más cercano”).

CAPÍTULO 5

Tomamos en primer lugar el corpus de todos los entremeses indubitados, considerando las 800 y las 1000 palabras más frecuentes (MFW) (Gráficos 1 y 2)¹⁵⁸.

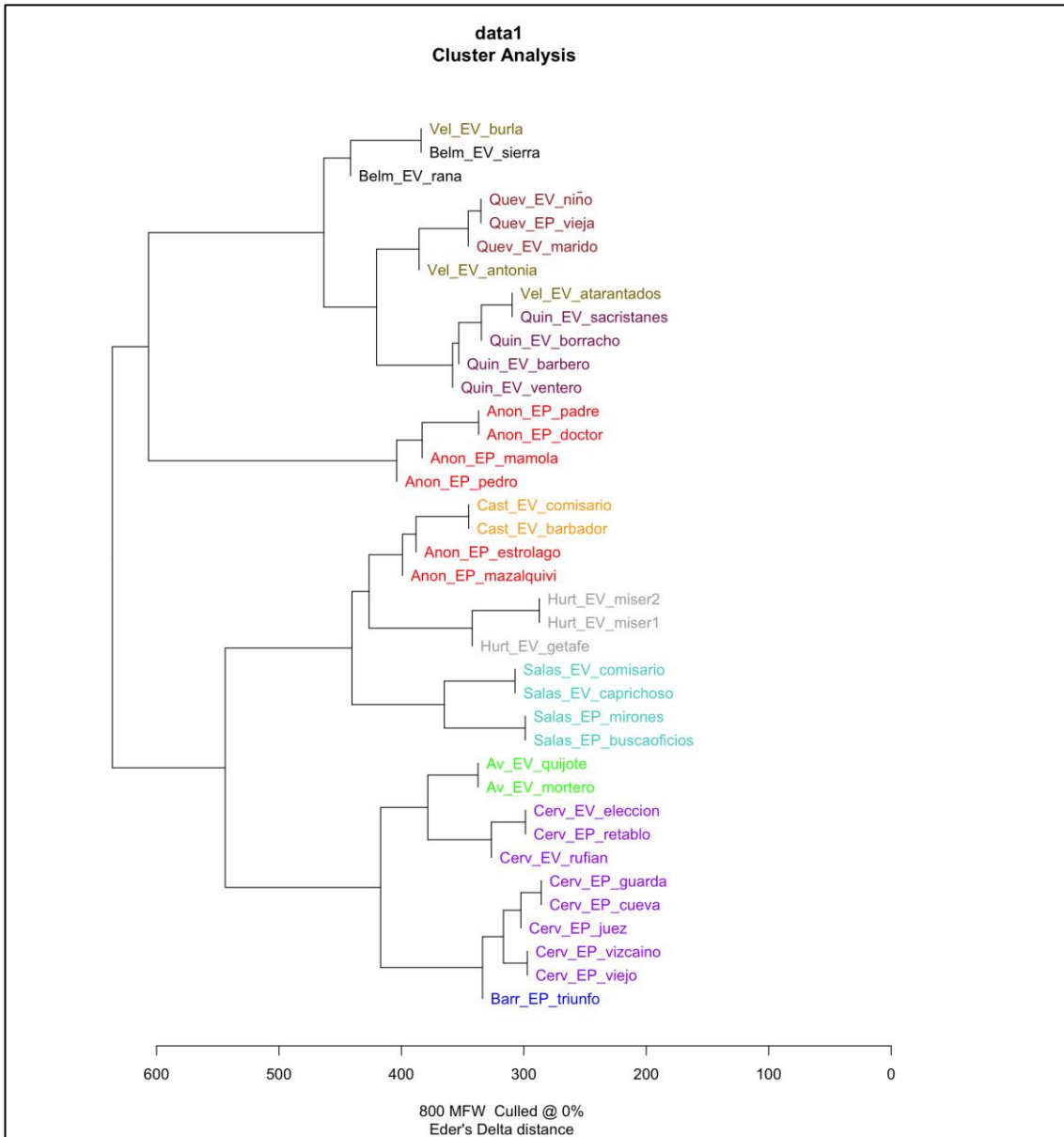


Gráfico 1. Clúster basado en la frecuencia de las palabras más empleadas en el total de entremeses indubitados (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 800. Culled 0%)

¹⁵⁸ En todos los análisis con Delta marcamos la opción “0% culled”, ya que no consideramos necesario que las palabras más frecuentes aparezcan en un número determinado de textos.

APROXIMACIÓN AL CASO DE ESTUDIO

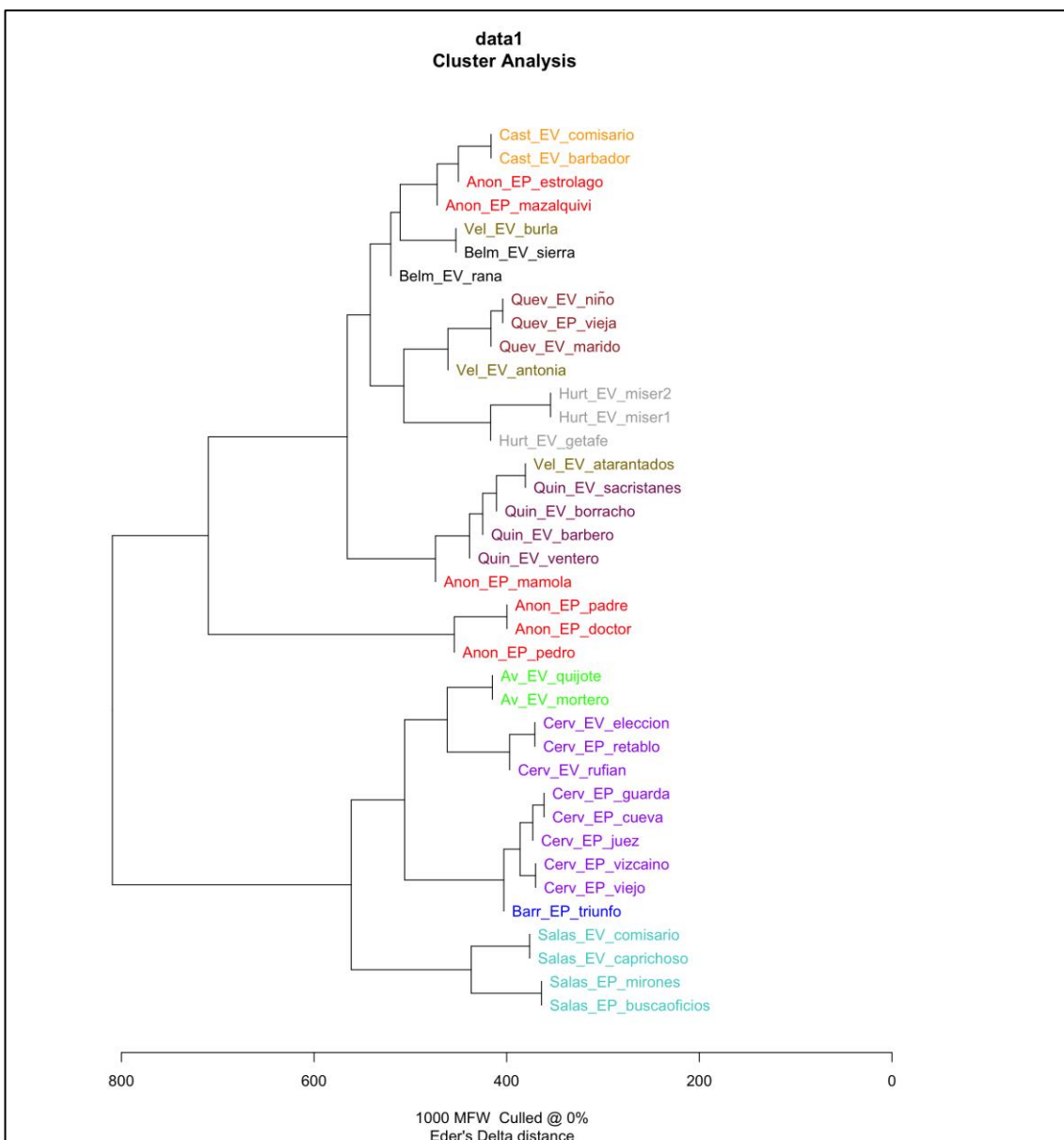


Gráfico 2. Clúster basado en la frecuencia de las palabras más empleadas en el total de entremeses indubitados (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Cullled 0%)

Aunque entraremos a valorar los resultados más adelante, es necesario señalar que no se aprecian grandes diferencias entre las 800 y las 1000 MFW, a excepción de las relaciones que se establecen entre los seis entremeses anónimos. En el Gráfico 1, cuatro de ellos se agrupan en una misma subrama (*El padre engañado*, *El doctor simple*, *La mámola* y *Pedro Hernández*) mientras que en el Gráfico 2 sólo lo hacen tres de ellos (*La mámola* pasa en este análisis a formar parte del grupo de Quiñones de Benavente).

CAPÍTULO 5

En los dos casos, *El padre engañado* y *El doctor simple* se asocian directamente (¿son del mismo autor o simplemente son los “vecinos más cercanos”, es decir, los dos más parecidos dentro de este corpus?).

Como el objetivo ahora es el de testar la operatividad del corpus de entremeses indubitados, repetiremos los dos análisis sólo con las obras de autor conocido (Gráficos 3 y 4):

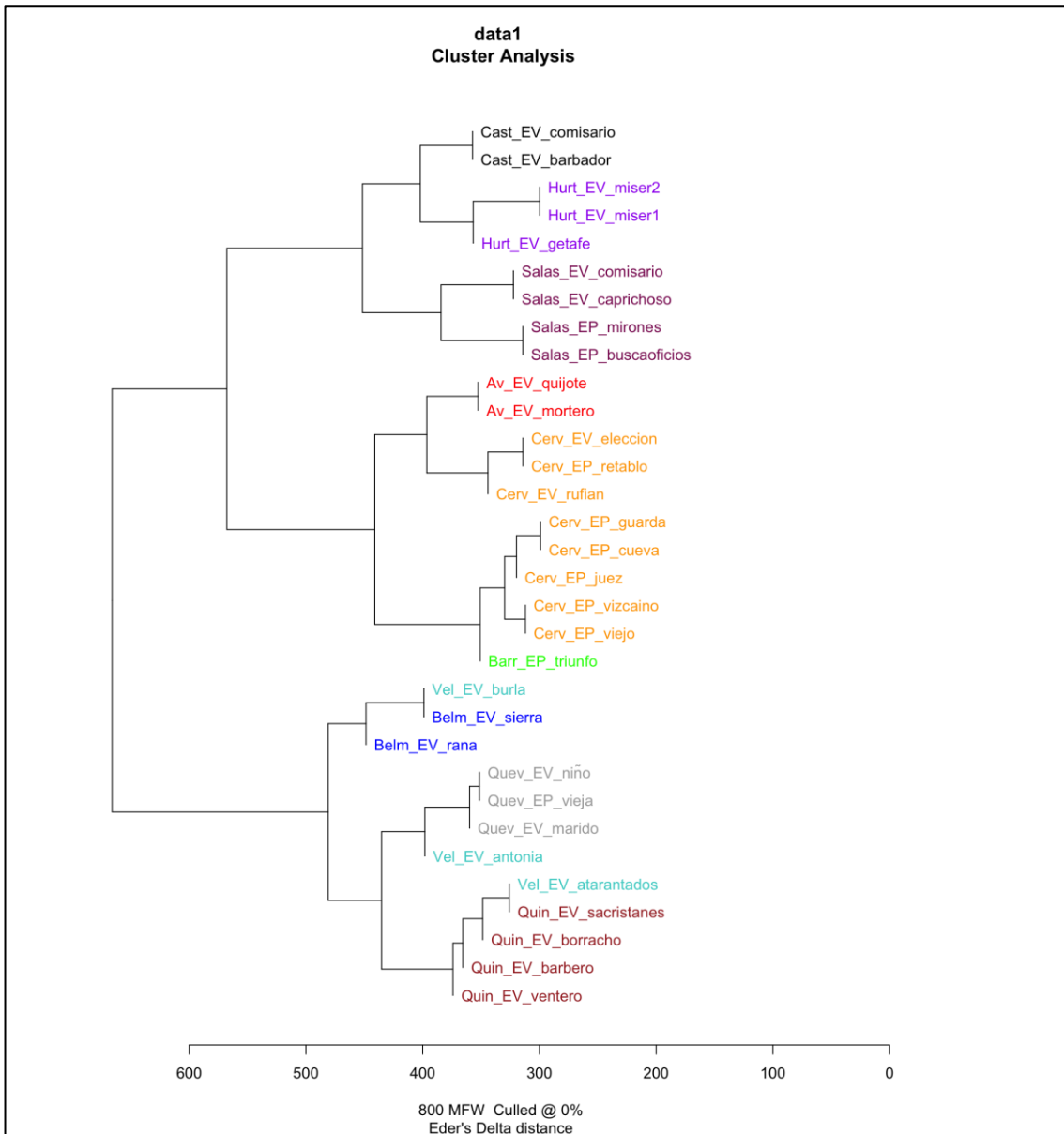


Gráfico 3. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses indubitados de autor conocido (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 800. Culled 0%)

APROXIMACIÓN AL CASO DE ESTUDIO

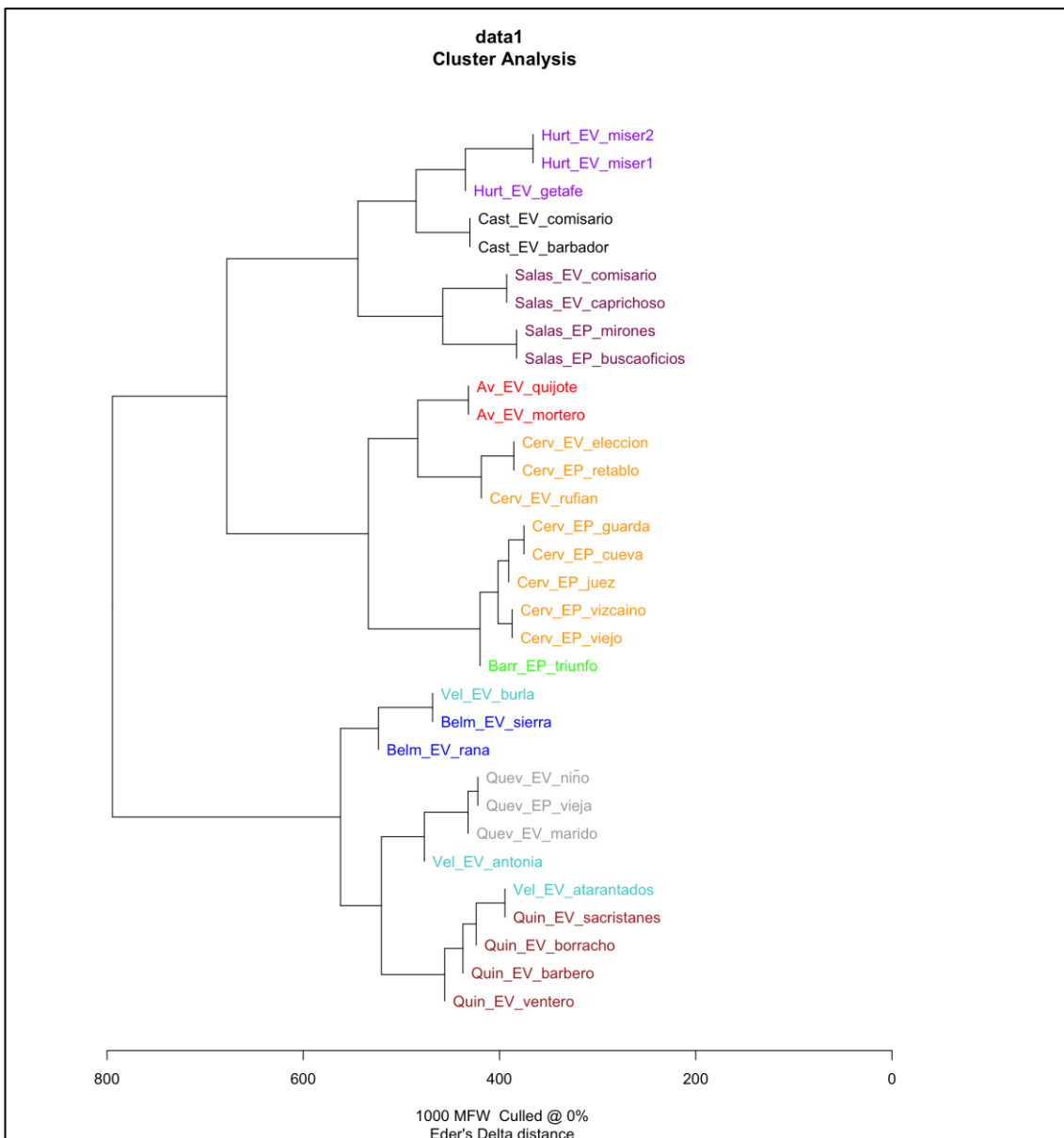


Gráfico 4. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses indubitados de autor conocido (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

Los gráficos nos llevan a las siguientes consideraciones:

- a) A pesar de la brevedad de las muestras analizadas, el algoritmo Eder's Delta presenta un resultado de agrupación altamente fiable (tanto con 800 como con 1000 MFW). En un primer nivel de agrupación, se distribuyen las obras en dos grandes ramas: por un lado, las de Hurtado de Mendoza, Castillo de Solórzano, Salas Barbadillo, Ávila, Cervantes y Barrionuevo; y, por otro, las de Vélez de

Guevara, Belmonte, Quevedo y Quiñones. En los siguientes niveles de agrupación, todos los textos de autor conocido se han identificado con otros textos del mismo autor, a excepción de los tres entremeses de Vélez de Guevara, que no se asocian nunca directamente entre sí. Este resultado confiere al análisis de la frecuencia relativa de palabras un porcentaje de acierto del 100% en primer nivel y superior al 90% en el resto de niveles (29 de 32 asociaciones).

- b) No podemos profundizar en las causas que ocasionan que los tres entremeses de Vélez de Guevara no se asocien entre sí. Es cierto que son tres de los textos más breves (1 094, 1 144 y 1 440 palabras) y que, como ya hemos explicado, cuanto menor es la extensión, menor es también la fiabilidad de los resultados obtenidos; pero también es cierto que otros textos de similar extensión sí se han asociado correctamente (véanse, por ejemplo, *El barbero* de Quiñones, con 986 palabras, los dos textos de Belmonte, con 871 y 1 012 palabras, y los dos textos de Castillo de Solórzano, con 1 294 y 1 422). Las causas, posiblemente, estén relacionadas con la baja calidad de las muestras (contaminación autorial y/o textual) o con las prácticas pseudoplagiarias de Vélez¹⁵⁹.
- c) Todos los entremeses de Cervantes parten de un troco común, si bien se advierten dos ramas diferenciadas: una compuesta por todos los entremeses en prosa, a excepción de *El retablo de las maravillas*, y otra compuesta por este entremés y los dos en verso. Aunque necesitaríamos realizar un estudio mucho más exhaustivo para establecer la posible causa de esta diferenciación, nos aventuramos a presuponer que está relacionada con dos épocas o momentos en la redacción de los entremeses¹⁶⁰.

¹⁵⁹ *La burla más sazonada*, *Los atarantados* y *Antonia y Perales* fueron publicados con el nombre de Vélez de Guevara en la *Flor de entremeses y sainetes de diferentes autores* (Madrid, 1657). Menéndez Pelayo es el primero en desconfiar de los “nombres ilustres” que figuran en esta colección [Menéndez Pelayo, ed., 1903: VII] y Urzáiz Tortajada, por su parte, nos recuerda que Vélez se valió frecuentemente de títulos, argumentos y versos de autores ajenos [1997].

¹⁶⁰ Son muchos los autores que han expuesto su teoría sobre la fecha de composición de los ocho entremeses publicados en 1615, aunque ninguno lo ha hecho basándose en análisis estilométricos. La gran

- d) *El triunfo de los coches*, de Barrionuevo, no se asocia directamente con ningún entremés (algo esperado, pues es el único entremés conservado de este autor), pero muestra cierta proximidad —es decir, cierta identificación léxica— con una de las dos subramas cervantinas (la integrada por *La guarda cuidadosa*, *La cueva de Salamanca*, *El juez de los divorcios*, *El vizcaíno fingido* y *El viejo celoso*).
- e) Los dos entremeses de Francisco de Ávila se asocian entre sí, aunque guardan también cierta cercanía léxica con la subrama cervantina formada por *La elección de los alcaldes de Daganzo*, *El retablo de las maravillas* y *El rufián viudo llamado Trampagos*.
- f) Parece haber mayor relación de proximidad entre los textos escritos en verso, por un lado, y los escritos en prosa, por otro; así ocurre en los entremeses de Salas Barbadillo y en la mayoría de los de Cervantes (pero no en los de Quevedo).

A pesar de las dificultades metodológicas de estudiar la autoría de unas piezas literarias tan breves y volátiles como son los entremeses barrocos, el análisis basado en el estudio de la variación de las palabras más frecuentes muestra una fiabilidad del 100% en un primer nivel y de más del 90% en los siguientes niveles, sin variaciones significativas entre las 800 y las 1000 MFW. Ahora bien, deberemos estudiar —y lo haremos más adelante—, si esa similitud léxica entre los entremeses de Cervantes y los de Francisco de Ávila y Gaspar de Barrionuevo y Carrión se da también en otro tipo de variables (como, por ejemplo, en la frecuencia relativa de las clases de palabras o en el estudio de las etiquetas morfosintácticas).

mayoría cree que los dos entremeses en verso fueron los de composición más tardía, en consonancia con la evolución general del género; otros, sin embargo, como Cotarelo Valledor, se decantan por el orden en que están impresos. Para una revisión de este aspecto, véase el apartado de “cronología” de la introducción de Baras Escolá a su edición de los entremeses de Cervantes [2012: 170-176].

5.4.2. Análisis del porcentaje de similitud textual (*similarity*)

El programa CopyCatch Gold es una herramienta muy utilizada para detectar casos de plagio lingüístico, ya que permite calcular los porcentajes de similitud textual, de vocabulario compartido (una sola vez y más de una vez) y de vocabulario único, así como el número de frases coincidentes (o cuasi coincidentes) entre dos textos dados. Para obtener un resultado preciso, debemos aplicar una *stopword list* que bloquee la identificación de todas aquellas palabras vacías o de función, es decir, aquéllas sin significado léxico¹⁶¹. Es importante advertir que usamos siempre una *stopword list* adaptada al castellano del Siglo de Oro, en la que incluimos formas arcaicas ya en desuso, como, por ejemplo, *dél, deso, daquesto, aqueso, esotro, ansí, ansina, maguer, vuestra merced, vusted, vuesasted, cascuno*, etc.

Turell, basándose en resultados empíricos obtenidos para calcular el nivel de similitud textual entre varios textos sobre un mismo tema, advierte que “es bastante normal que el parecido oscile entre el 30 y el 50%” [2005: 281]; más recientemente, Garayzabal Heinze, Queralt Estévez y Reigosa Riveiros precisan un umbral del 40% para textos de “temática similar” [2019: 119]. Aunque estos porcentajes indican el umbral de similitud a partir del cual podemos establecer la existencia de plagio entre dos textos que versan sobre un mismo tema, el nivel de similitud textual también resulta relevante para detectar casos de autoría (pues la autoría y el plagio son, en realidad, dos caras de una misma moneda).

En las siguientes páginas, recogemos los resultados obtenidos tras analizar los niveles de similitud textual existentes entre los pares formados por dos entremeses cervantinos (representados en color azul) y los formados por un entremés de Cervantes y otro de un autor distinto (en color rojo).

¹⁶¹ Lo cierto es que el debate sobre qué clases de palabras transmiten significado léxico y aquéllas en que éste es débil o incluso nulo sigue creando controversia entre los gramáticos

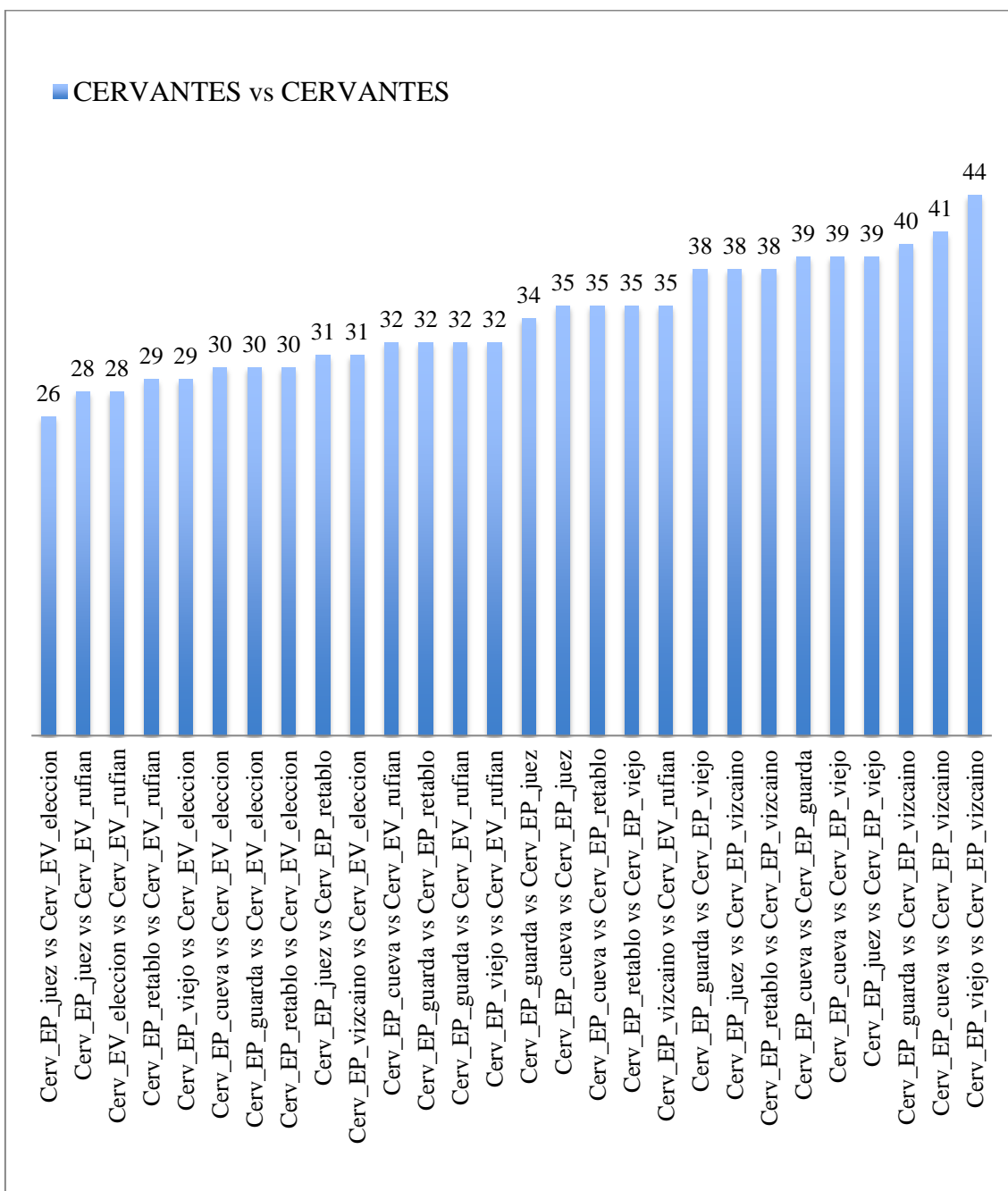


Gráfico 5. Nivel de similitud textual entre los pares formados por dos entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

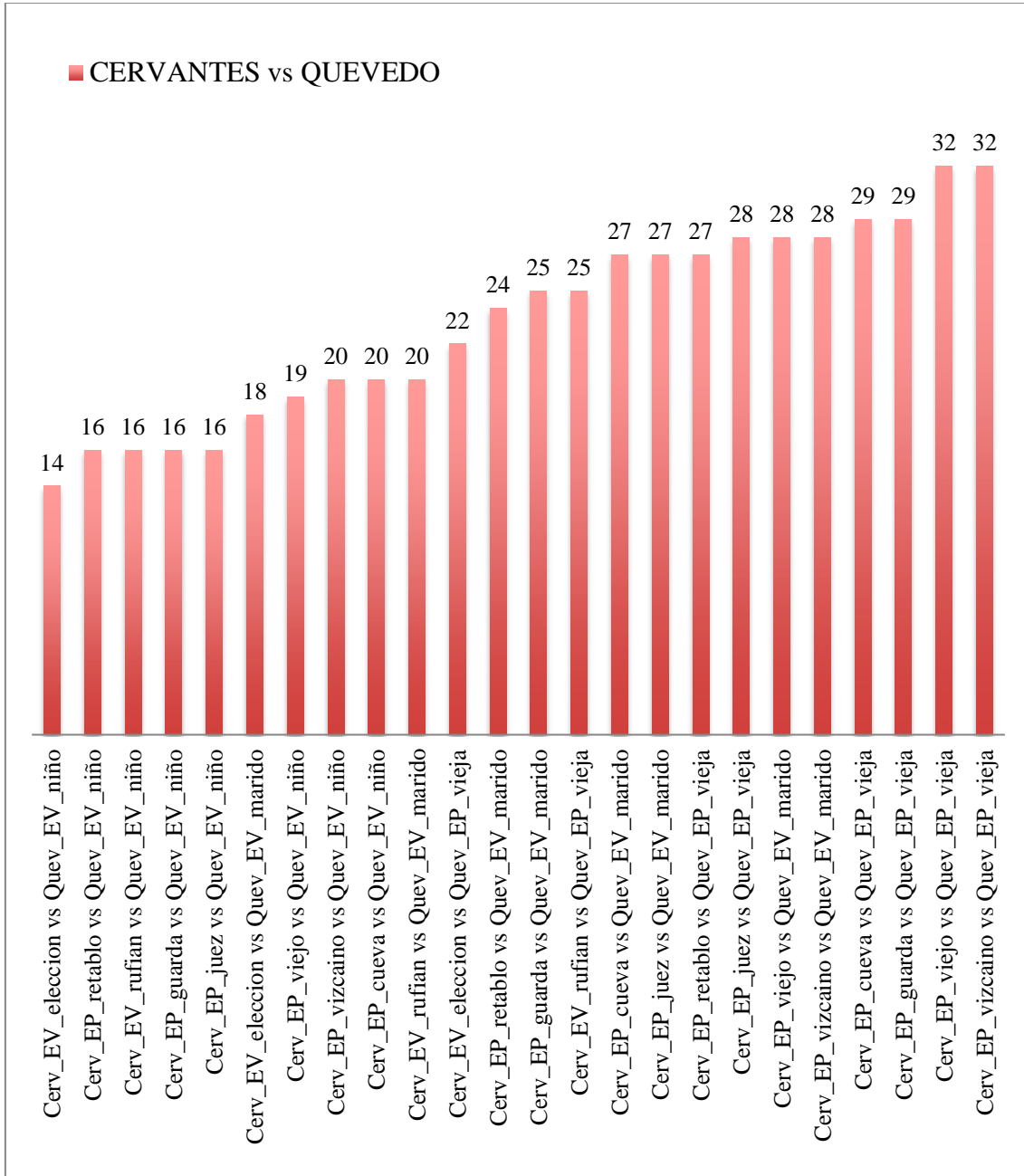


Gráfico 6. Nivel de similitud textual entre los pares formados por un entremés de Cervantes y otro de Quevedo (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

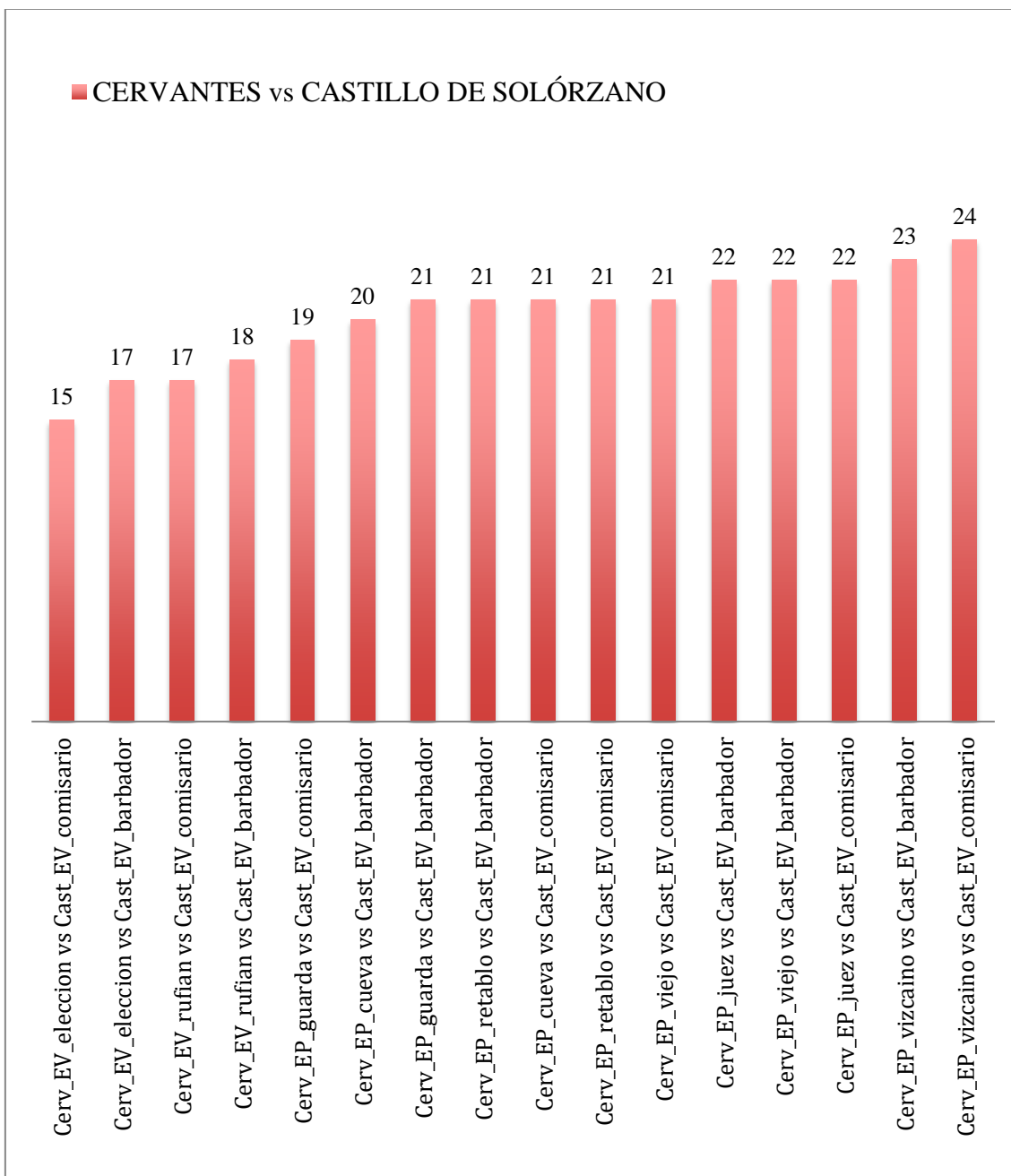


Gráfico 7. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Castillo de Solórzano (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

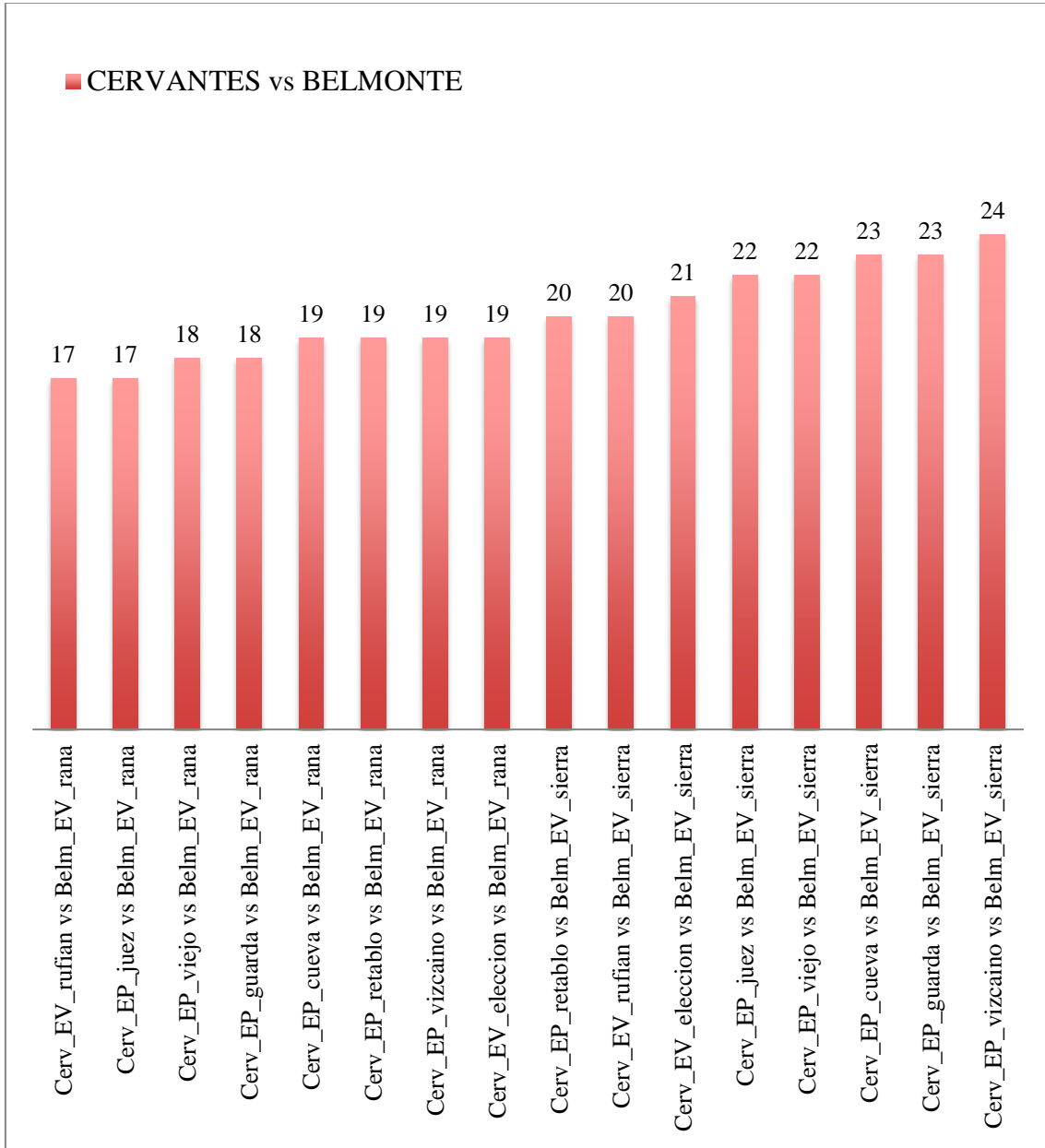


Gráfico 8. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Belmonte (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

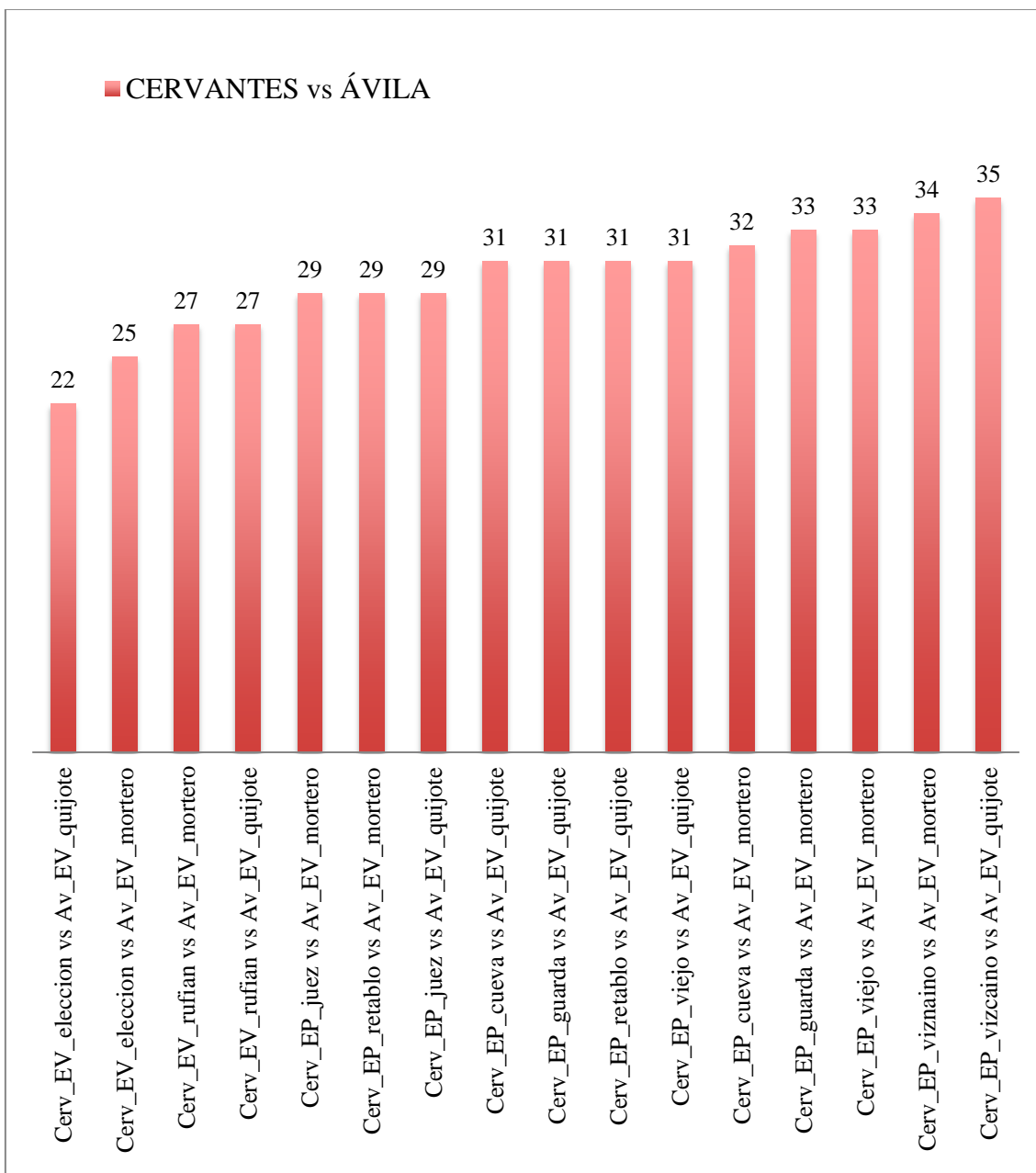


Gráfico 9. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Ávila (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

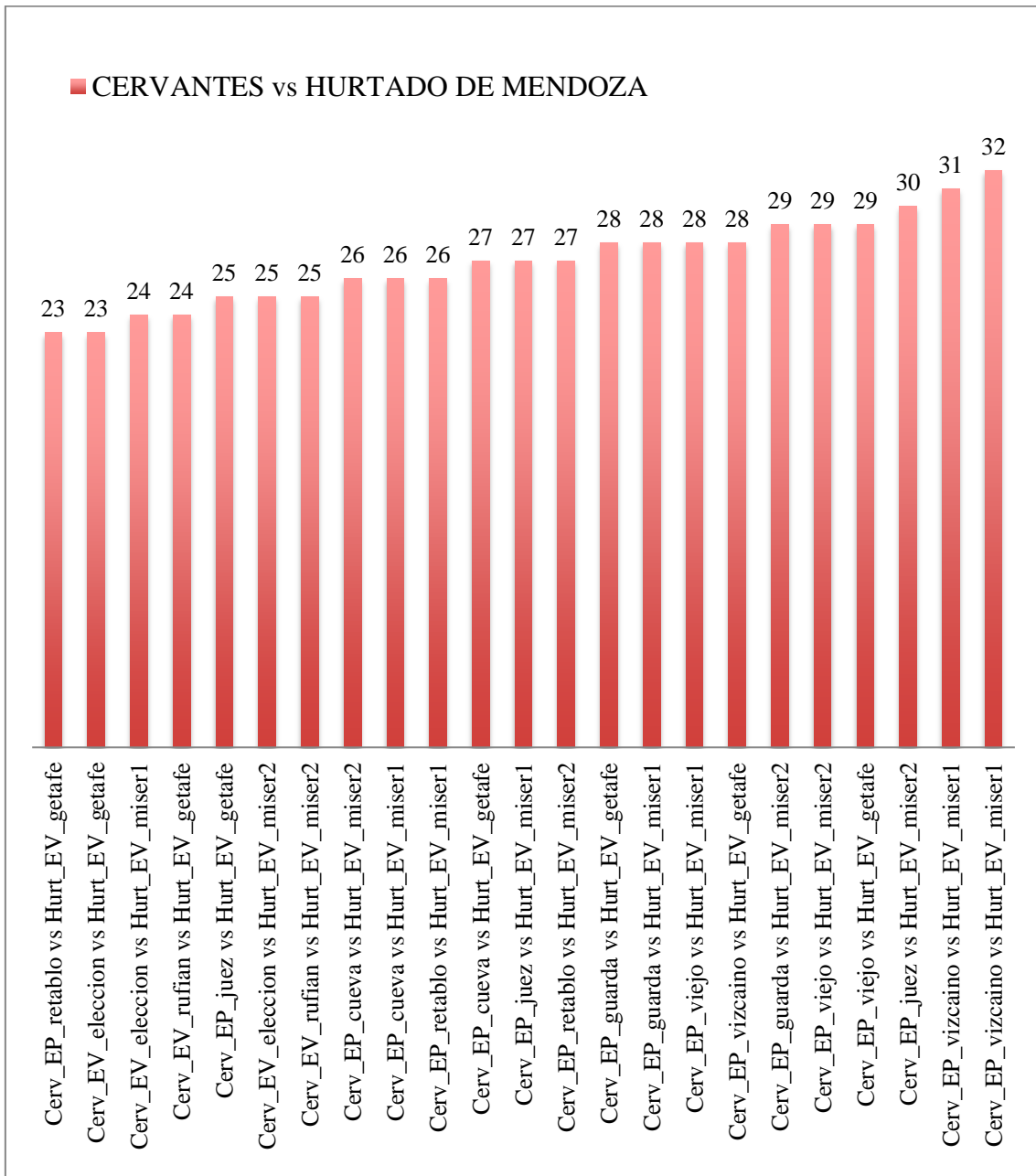


Gráfico 10. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Hurtado de Mendoza (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

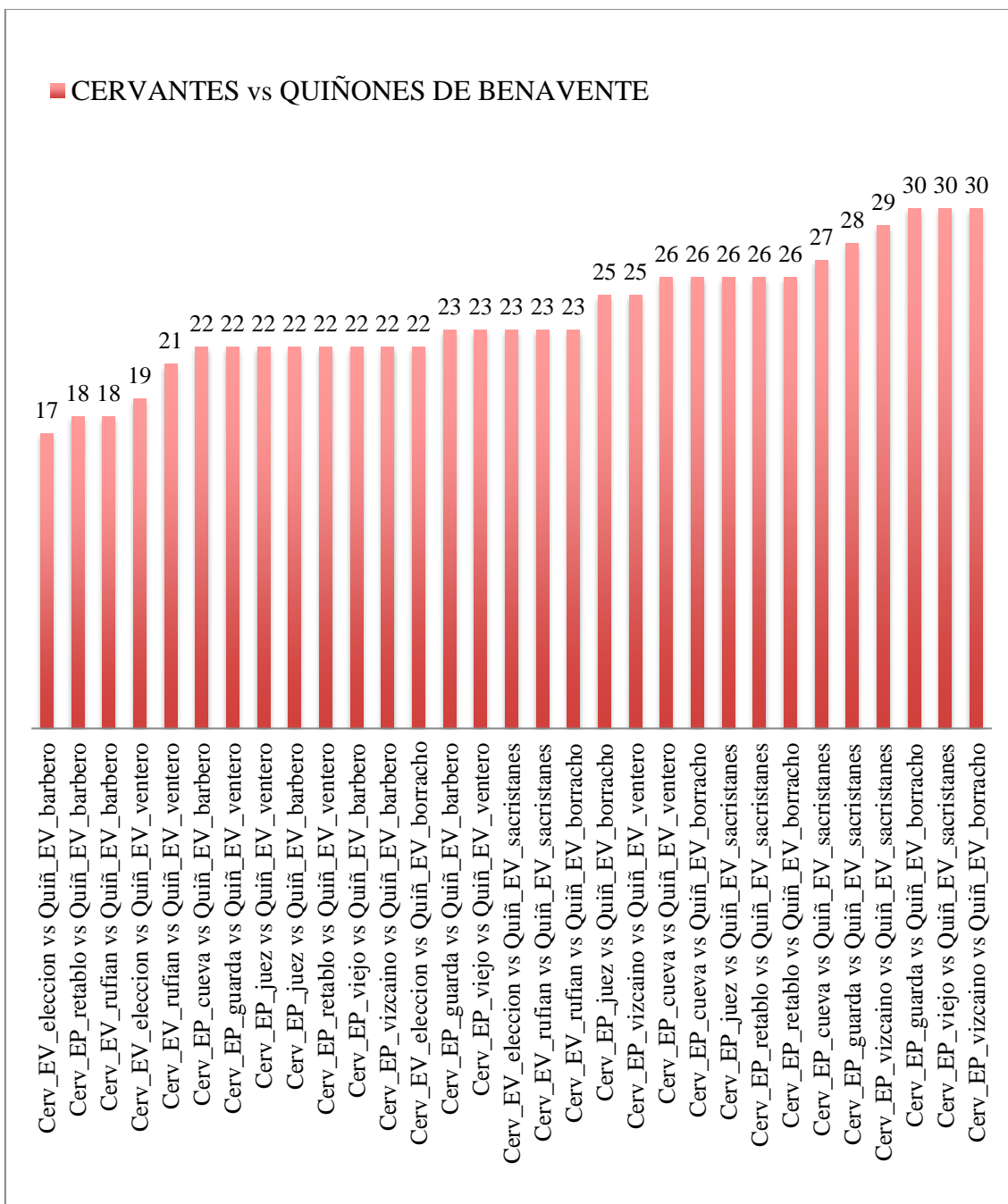


Gráfico 11. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Quiñones de Benavente (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

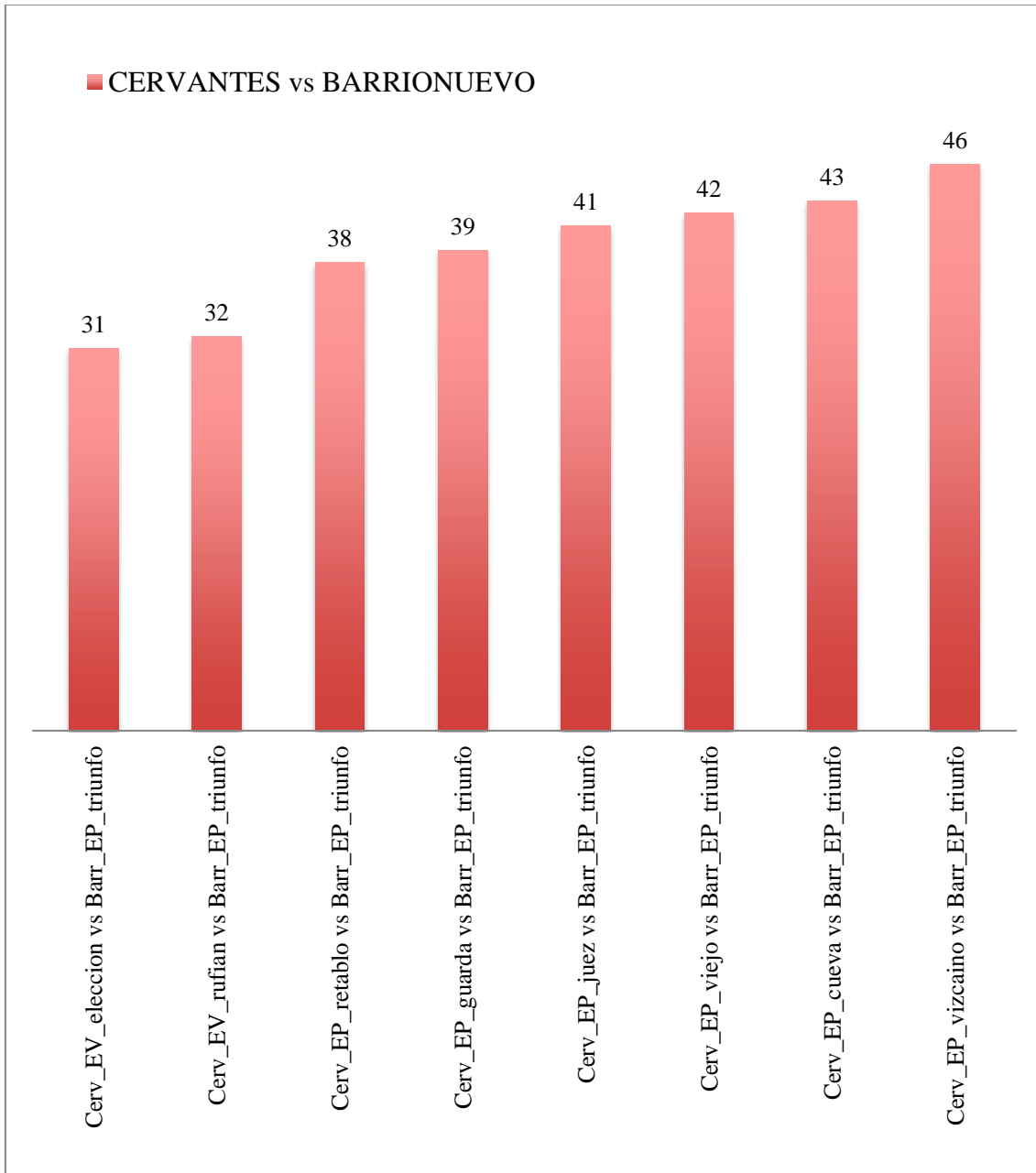


Gráfico 12. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Barrionuevo (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

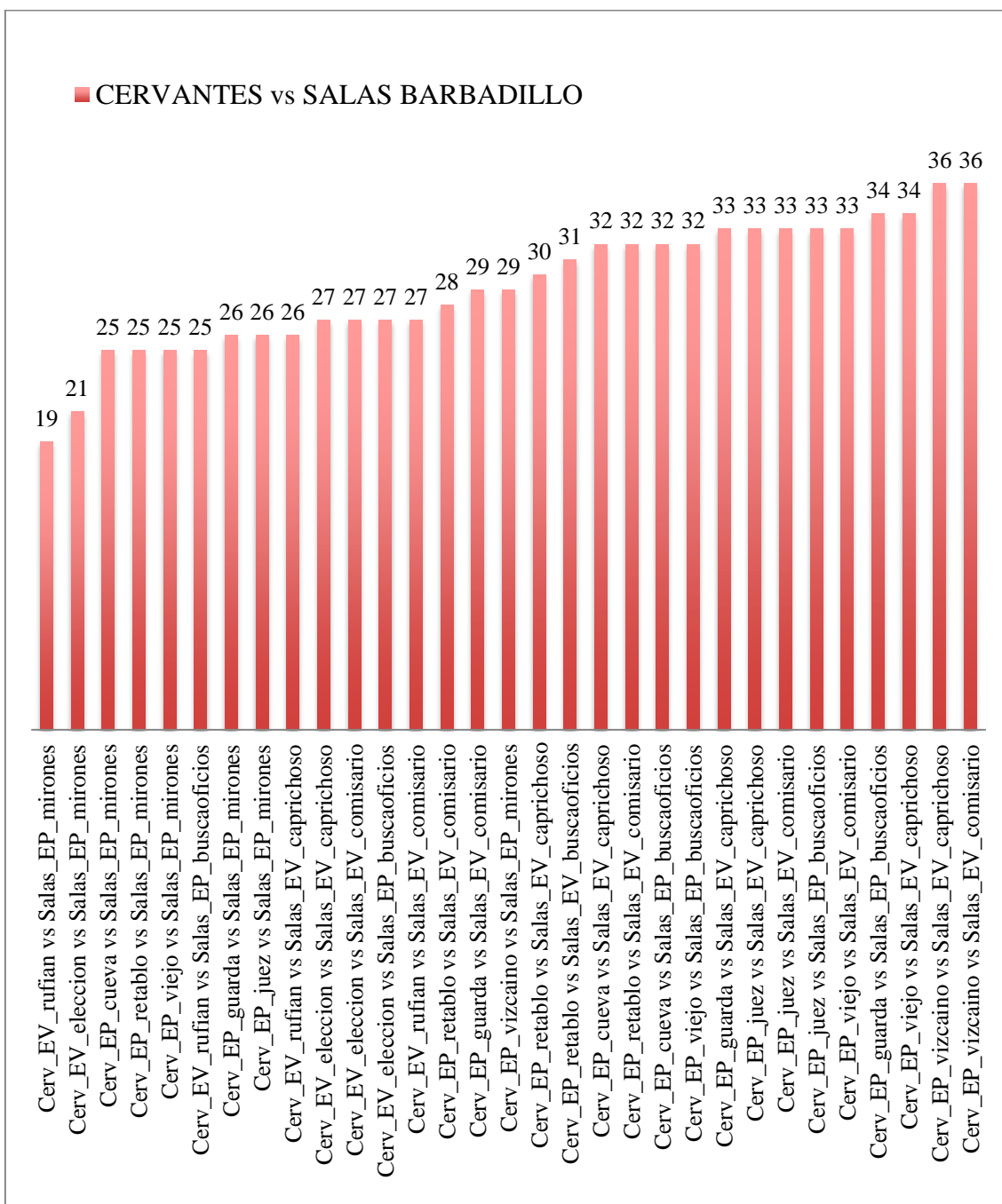


Gráfico 13. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Salas Barbadillo (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

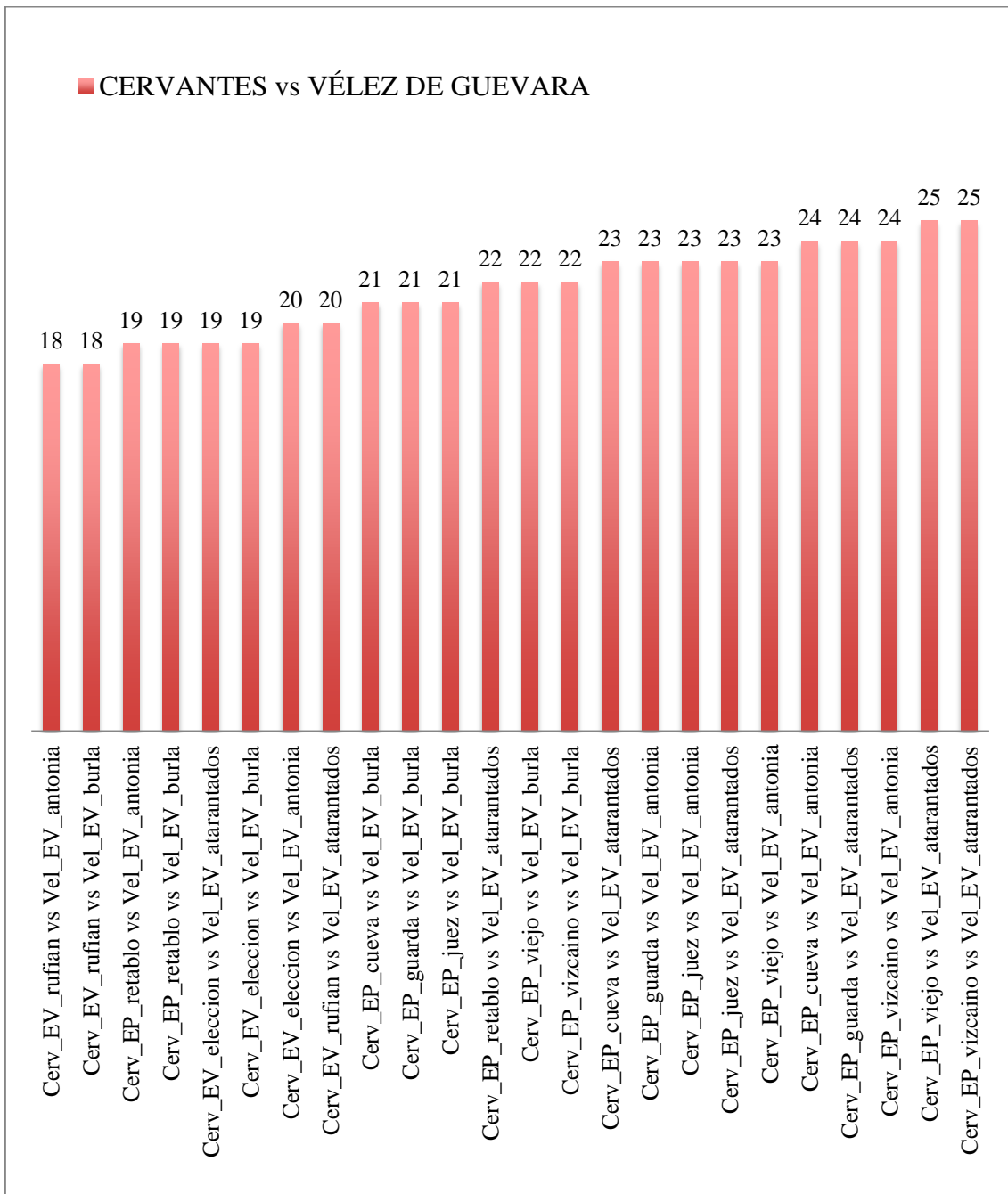


Gráfico 14. Nivel de similitud textual de los pares formados por un entremés de Cervantes y otro de Vélez de Guevara (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

El Gráfico 5 permite advertir que el nivel de similitud textual de los entremeses de Cervantes es del [26 , 44] %, superior al que mantiene con los entremeses del resto de autores, a excepción de Barrionuevo. Basándonos en estos intervalos, podemos establecer tres niveles de proximidad:

- a) En el nivel más bajo se encontrarían Castillo de Solórzano [15 , 24] %, Belmonte [17 , 24] % y Vélez de Guevara [18 , 25] %. En estos tres autores, el mayor porcentaje de similitud textual entre un entremés suyo y otro de Cervantes es todavía inferior al menor porcentaje de similitud textual entre dos entremeses cervantinos (26%).
- b) En un nivel medio, pero aún bastante alejados de los porcentajes cervantinos, encontramos a Quiñones de Benavente [17 , 30] %, Quevedo [14 , 32] %, Ávila [22 , 35] %, Hurtado de Mendoza [23 , 32] % y Salas Barbadillo [19, 36] %. En ninguno de estos autores, el porcentaje más bajo entre un entremés suyo y otro de Cervantes es mayor que el más bajo entre dos entremeses de Cervantes (26%), ni el porcentaje más alto entre un entremés suyo y otro de Cervantes es mayor al más alto entre dos entremeses de Cervantes (44%).
- c) En el nivel más alto si sitúa Barrionuevo [31 , 46] %. El entremés de este autor presenta unos porcentajes de similitud textual con el teatro breve cervantino muy parecidos a los que presentan los entremeses de Cervantes entre sí, superando el controvertido umbral del 40% y superando, incluso, el porcentaje más alto de Cervantes (44%) en hasta dos puntos (46%).

De hecho, como recogemos en el Gráfico 15, sólo hay siete pares textuales (de los 220 analizados) que igualan o superan ese umbral del 40%, lo que induce a pensar en una dependencia textual que ya no viene condicionada por el tema de los textos, sino por la autoría común o la copia. Marcamos, en azul, los casos en que esos pares están formados por dos entremeses de Cervantes y, en rojo, los casos formados por un entremés de Cervantes y otro de otro autor:

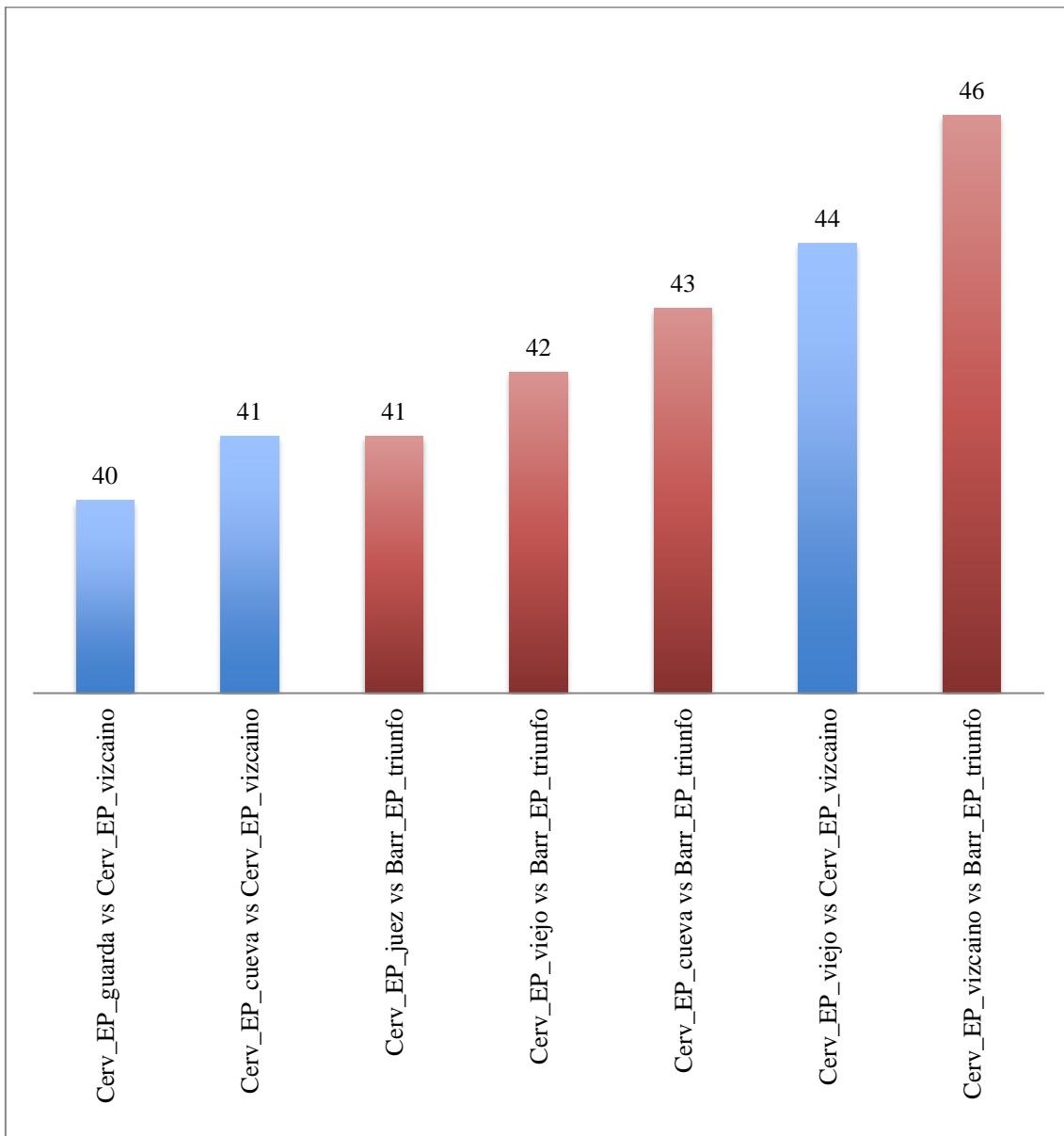


Gráfico 15. Pares textuales analizados que igualan o superan el umbral del 40% de similitud textual. (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

La similitud entre *El vizcaíno fingido* y *La guarda cuidadosa* (40%), *La cueva de Salamanca* (41%) y *El viejo celoso* (44%) era esperable y explicable, pues nos encontramos ante tres pares textuales de un mismo autor; sin embargo, no era esperable la similitud textual que se establece entre *El triunfo de los coches* y *El juez de los divorcios* (41%), *El viejo celoso* (42%), *La cueva de Salamanca* (42%) y *El vizcaíno fingido* (46%).

5.4.3. Análisis de secuencias de palabras idénticas (*verbatim*)

El número de palabras iguales y en el mismo orden que pueden producir dos escritores distintos se reduce drásticamente en función de la longitud de la frase [Coulthard, 2004]. En español se tiene como referencia el estudio realizado por Marquina y Queralt [2014], en el que se concluye que es poco probable que dos personas distintas produzcan de forma independiente secuencias idénticas de 5 o más palabras, siempre y cuando dichas secuencias no sean una expresión común de la lengua, compartida por multitud de hablantes.

En todos los entremeses analizados, hay un larguísimo listado de *verbatim* de dos ítems, que son compartidos por la mayoría de autores (“por qué”, “para que”, “las fiestas”, “la causa”, “los hombres”, “mi casa”, “para bailar”, “¡Vive Dios!”, etc.

Son también muy frecuentes los *verbatim* de 3 ítems, que pueden aparecer en la obra de un mismo autor, como es el caso de Quevedo con “aparécese lleno de” (Quev_EV_marido, Quev_EV_niño y Quev_EV_niño) o de Salas Barbadillo con “me admiro que” (Salas_EV_caprichoso y Salas_EV_comisario), “que se pierda” (Salas_EV_caprichoso y Salas_EV_comisario), “vive el cielo” (Salas_EV_caprichoso y Salas_EV_comisario) o “qué os parece” (Salas_EV_caprichoso y Salas_EV_comisario), pero que, por lo general, son compartidos por distintos autores: “de la vista” (Belm_EV_rana y Salas_EP_mirones), “de mis ojos” (Quev_EV_niño y Vez_EV_atarantados), “luego al momento” (Av_EV_quijote, Av_EV_mortero y Barr_EP_triunfo), “de mi casa” (Belm_EV_rana y Quev_EP_vieja), “de mi pluma” (Quev_EV_niño y Salas_EV_caprichoso), “en los corrales” (Belm_EV_rana y Vez_EV_burla), “¿hay tal disparate?”, (Cerv_EP_guarda y Av_EV_mortero), “para que digan” (Quin_EV_sacristanes y Vélez Vez_EV_antonia), “de una vez” (Cast_EV_barbador y Salas_EP_mirones), “Dios os haga” (Av_EV_quijote y Belm_EV_rana), “cuerpo de Dios” (Cerv_EP_retablo y Quin_EV_barbero y Quin_EV_ventero), “váyase con Dios” (Cerv_EP_cueva, Cerv_EP_guarda y Barr_EP_triunfo), “un hombre que” (Quev_EV_marido y Vez_EV_burla), “pernil de tocino” (Cerv_EP_juez y Av_EV_quijote), etc.

Menor es el número de *verbatim* de 4 ítems¹⁶². En estos casos seguimos encontrando secuencias compartidas por distintos autores, como “ay, desdichada de mí” (Cerv_EP_guarda y Vel_EV_burla), “¿qué voces son éstas?” (Cerv_EP_vizcaino y Barr_EP_triunfo)¹⁶³, “bueno, por vida mía” (Cerv_EV_rufian, Salas_EP_mirones y Barr_EP_triunfo), “no falta más que” (Cerv_EP_retablo y Av_EV_quijote), “¿qué hemos de hacer?” (Salas_EV_comisario y Vel_EV_atarantados), “¿qué te ha / habemos hecho?” (Av_EV_mortero y Vel_EV_atarantados) o “pero, ¿qué es esto?” (Cerv_EP_juez y Quin_EV_sacristanes), pero también hay ya un número significativamente alto de secuencias de 4 ítems iguales compartidas en obras de un mismo autor; así, en Ávila encontramos “a manos de un” (Av_EV_mortero y Av_EV_quijote) o “que soy peor que” (Av_EV_mortero y Av_EV_quijote)¹⁶⁴; en Cervantes, “Dios que va de” (Cerv_EV_rufian y Cerv_EV_eleccion), “no es / será posible que” (Cerv_EP_viejo, Cerv_EP_vizcaino y Cerv_EP_viejo), “aquí como de molde” (Cerv_EP_vizcaino y Cerv_EV_rufian) o “yo así lo creo” (Cerv_EP_cueva y Cerv_EP_viejo); y en Hurtado de Mendoza, “voto a Cristo que” (Hurt_EV_getafe y Hurt_EV_miser2) y “vase y sale él” (Hurt_EV_miser 1 y Hurt_EV_miser 2).

Pero como señalábamos antes, lo verdaderamente relevante es detectar los *verbatim* de 5 o más ítems (que extraemos tras su rastreo en la pestaña “Sentence” de CopyCatch Gold) y comprobar si su uso es idiosincrásico o no, mediante su búsqueda en un corpus de la época (CORDE) entre los años 1547 y 1617¹⁶⁵.

¹⁶² No hemos considerado los casos en los que hay una alteración del orden de las palabras, pues en esos casos no podemos hablar propiamente de *verbatim*. Por ejemplo: “hallo por mi cuenta” (Salas_EP_buscaoficios) frente a “yo por mi cuenta hallo” (Salas_EP_buscaoficios), “yo he de perder el juicio” (Quin_EV_barbero) frente a “el juicio he de perder” (Cast_EV_comisario) o “ya no lo puedo sufrir” (Cerv_EP_cueva) frente a “no lo puedo ya sufrir” (Quin_EV_sacristanes).

¹⁶³ En Hurtado de Mendoza encontramos esta secuencia, pero en otro orden: “¿qué voces éstas son?” (Hurt_EV_miser2). Como hemos explicado, lo realmente relevante en un caso de *verbatim* es el mantenimiento exacto del orden de la cadena de palabras.

¹⁶⁴ En este caso la similitud va más allá: “que soy peor que Judas si me enojo” (Av_EV_mortero) y “que soy peor que el diablo si me enojo” (Av_EV_quijote).

¹⁶⁵ Utilizamos comodines (“?” y “*”) para los casos en los que pueda haber flexión verbal o vacilación gráfica.

APROXIMACIÓN AL CASO DE ESTUDIO

VERBATIM	ÍTEMS	TEXTOS	EN EL CORDE
a pagar de mi dinero	5	Cast_EV_comisario Vel_EV_antonia	2
otra vez torno a decir	5	Cerv_EP_vizcaino Cerv_EP_viejo	1 (1 Cerv)
pueda / puedo ir a la mano	5	Cerv_EV_eleccion Vel_EV_atarantados	15 (2 Cerv)
que me parece a mí	5	Cerv_EV_eleccion Salas_EV_comisario	11
que es lo mismo que	5	Salas_EV_caprichoso Vel_EV_antonia	160
qué es lo que queréis	5	Cerv_EP_guarda Cerv_EP_retablo	4
quién es aquí el señor	5	Cerv_EP_retablo Cast_EV_barbador	2 (1 Cerv)
vuesa merced, por su vida	5	Cerv_EP_vizcaino Cerv_EP_retablo	0
quién te mete a ti en	6	Cerv_EV_rufian Cerv_EV_eleccion	4 (2 Cerv)
beso a vuestas mercedes las manos	6	Cerv_EP_retablo Cerv_EP_viejo ¹⁶⁶	0 ¹⁶⁷
llevar / llevarán un pan como unas nueces	6	Vel_EV_atarantados Vel_EV_burla	1
que se me arranca el alma	6	Cerv_EP_vizcaino Cerv_EP_viejo	1 (1 Cerv)
eso haré yo de muy buena gana, y	8	Cerv_EP_vizcaino Cerv_EP_juez	0
pero, ¿por qué me lo pregunta vuesa merced?	8	Cerv_EP_vizcaino Cerv_EP_guarda	0
yo desde luego te doy mi bendición para que, con la de dios y con ella, salgas bien de todas las aventuras cortesanias	23	Salas_EP_mirones Salas_EP_buscaoficios	0

Tabla 11. Estudio de los *verbatim* de 5 o más ítems entre los entremeses indubitados de autor conocido

¹⁶⁶ En Salas Barbadillo encontramos esta misma secuencia, pero en otro orden: “Beso las manos a vuestas mercedes” (Salas_EP_buscaoficios).

¹⁶⁷ El *verbatim* “beso a vuestas mercedes las manos” sólo aparece en Cervantes, pero que en *La pícara Justina* encontramos “besa a vuestas mercedes las manos”. Como ya hemos señalado antes, lo verdaderamente significativo es que las palabras de la secuencia se den en un mismo orden.

Tal y como se desprende de la tabla anterior, los *verbatim* de 5 ítems, aunque significativos, no son casi nunca una marca propia de autoría; sin embargo, todos los casos de más de 5 ítems, sí lo son, a excepción de la frase hecha “llevar / llevarán un pan como unas nueces”¹⁶⁸.

5.4.4. Análisis de la distribución de la frecuencia de n-gramas de 5 caracteres (*character 5-grams*)

Otra variante utilizada en los estudios de atribución de autoría es la frecuencia de n-gramas de grafemas o de caracteres, aunque lo cierto es que su fiabilidad no siempre alcanza los mismos niveles de identificación en español que en inglés¹⁶⁹. Tal y como determinamos en un estudio de hace unos años, “la fiabilidad de los perfiles de n-gramas de grafemas [en español] crece conforme aumenta el número de grafemas analizados (23% para 2, 30% para 3, 36% para 4 y 50% para 5), algo que contrasta con lo que ocurría en lengua inglesa, en donde el perfil más fiable era 2-gram profile (79%) y el menos 9-gram profile (18%) [Blasco y Ruiz Urbón, 2009: 43].

De un tiempo a esta parte se tiende a trabajar con caracteres en lugar de grafemas, pues son capaces de capturar matices a nivel léxico, sintáctico y estructural [Houvardas y Stamatatos, 2006]. En los dos siguientes gráficos mostramos los dendrogramas resultantes de aplicar Stylo al estudio de la distribución de la frecuencia de los n-gramas de 5 caracteres en el corpus de indubitados de autor conocido, primero con 800 y

¹⁶⁸ En el resultado del número de concordancias, no contabilizamos los casos en los que CORDE remite a alguno de los entremeses de la columna “TEXTOS”, pues lo que tratamos de probar es el grado de rareza de ese *verbatim*. Indicamos el número total de concordancias y, dado el caso, entre paréntesis, el número de ellas que pertenecen a alguno de los autores de los textos en los que hemos detectado esos casos de *verbatim*. Por ejemplo, en el caso de “otra vez torno a decir”, CORDE sólo detecta una concordancia en esas fechas (además de la del *El viejo celoso* y *El vizcaíno fingido*), y esa concordancia es también de Cervantes (*La ilustre fregona*).

¹⁶⁹ Recordemos que, siguiendo a Grieve [2007], usamos *grafemas* para referirnos sólo a las letras del alfabeto y *carácter* para cualquier tipo de unidad gráfica (letras, dígitos, espacios en blanco y signos de puntuación). Es importante advertir que en el concepto de grafema obviamos, por lo ya explicado, los signos de puntuación.

después con 1000 MFW (0% culled), seleccionado el idioma español, el algoritmo mejorado de Eder (Eder's Delta) y la visualización en clúster (Cluster Analysis).

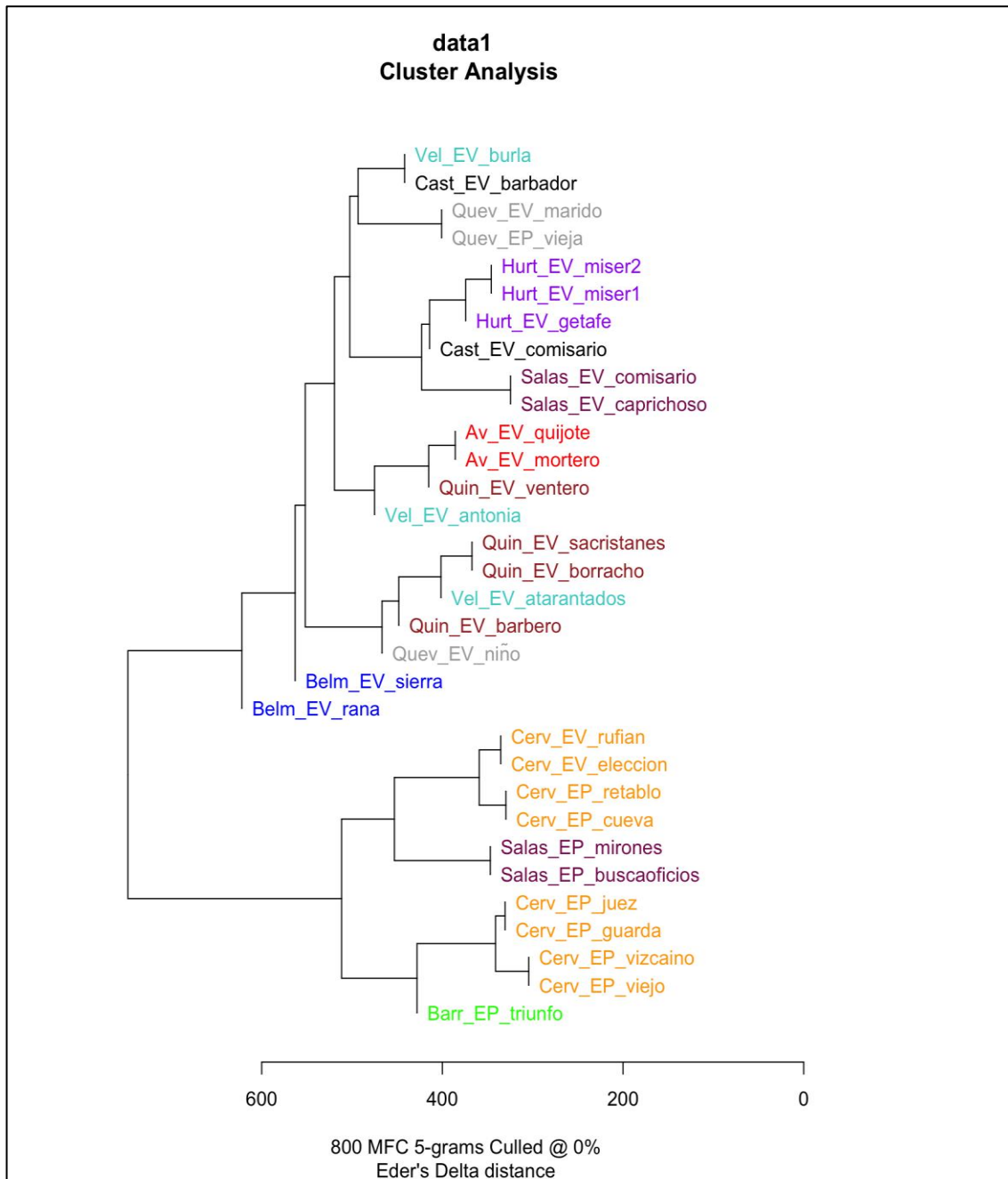


Gráfico 16. Clúster de los entremeses indubitados de autor conocido basado en la frecuencia de 5-gramas de caracteres (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFC 800 5-grams. Culled 0%)

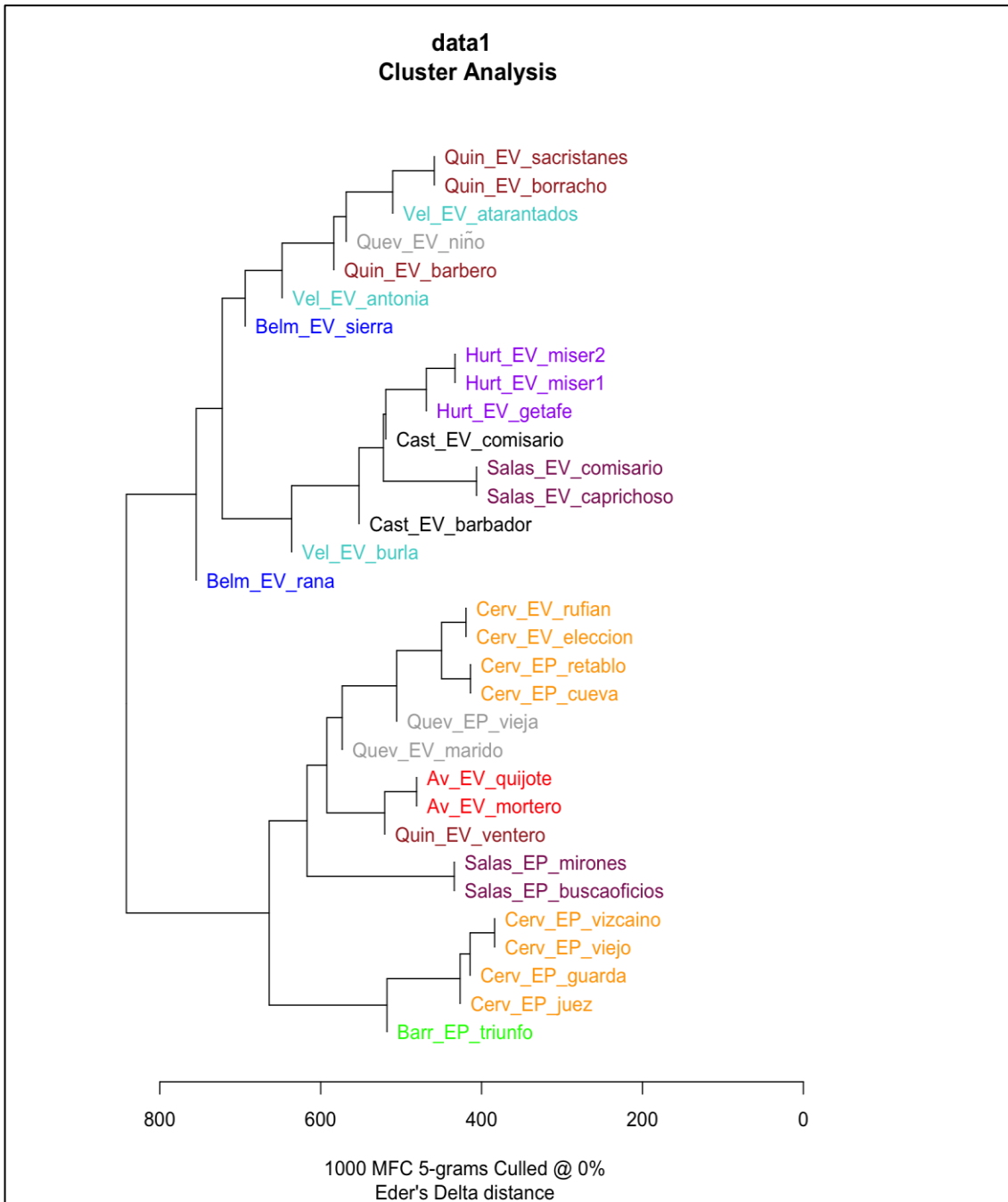


Gráfico 17. Clúster de los entremeses indubitados de autor conocido basado en la frecuencia de 5-ngramas de caracteres (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFC 1000 5-grams. Culled 0%)

Dejando a un lado el caso de Vélez de Guevara, el resultado vuelve a ser altamente satisfactorio: todas las obras que en el último nivel se asocian directamente con otra, son del mismo autor. Sorprende, sin embargo, que los cuatro entremeses de

Salas Barbadillo se distribuyan por pares en dos ramas distintas (prosa *vs* verso) o que alguno de los entremeses de Quevedo, Quiñones de Benavente o Castillo de Solórzano parezcan estar más próximos a los de otros autores que a los de su verdadero autor. Por eso, aunque con un porcentaje de efectividad algo más bajo que con palabras (*tokens*), el estudio de la frecuencia de 5-gramas de caracteres también parece resultar efectiva en español para discriminar autorías.

Estos primeros análisis realizados, de tipo ortográfico y léxico, nos permiten validar el corpus indubitado, aunque con pequeñas observaciones:

- Hay una similitud ortográfica y léxica significativamente alta entre *El triunfo de los coches* y varios de los entremeses cervantinos. El hecho de no poseer otro entremés indubitado de Gaspar de Barrionuevo nos imposibilita saber si la cercanía intraescritor (entre dos entremeses de Barrionuevo) sería mayor que la que detectamos interescriptor (entre el entremés de Barrionuevo y varios de los entremeses de Cervantes) (*nearest neighbor problem*)¹⁷⁰.
- Los dos entremeses de Francisco de Ávila se aproximan a varios de los indubitados cervantinos desde el punto de vista léxico, pero no desde el ortográfico. En todo caso, la distancia interescriptor es mayor que la distancia intraescritor.
- En ninguna de las pruebas realizadas los entremeses de Luis Vélez de Guevara se comportan como si fueran del mismo autor, lo que nos lleva a pensar en algún tipo de contaminación textual y/o autorial¹⁷¹.

¹⁷⁰ Tal y como explicamos en el capítulo 4, Madroñal atribuye a Barrionuevo *Los habladores*, incluido en nuestro corpus como dubitado de Cervantes, por su estrecha relación con Toledo [1993: 121-123]. Otros le han atribuido *El toquero*, que Madroñal, “basándose en el lenguaje y el estilo, no cree suyo” [Huerta Calvo, 2008: 214].

¹⁷¹ El *verbatim* de 6 ítems que comparten *Los atarantados* y *La burla más sazónada* (“llevar / llevarán un pan como unas nueces”) es significativo, pues el CORDE sólo identifica el uso de esta secuencia de palabras en otro texto de la época (*Sucesos de Sevilla de 1592 a 1604*, de Francisco de Ariño). En el *Vocabulario* de Correas aparece “dar un pan como unas nueces” con el significado de dar “palos, golpes y pesadumbres” [1924].

5.5. HIPÓTESIS DE TRABAJO

Aunque el objetivo de nuestro proyecto de investigación sólo aborda el estudio de autoría de dos de los diez entremeses atribuidos a Cervantes, consideramos importante plantear aquí una serie de hipótesis de trabajo sobre las nueve piezas conservadas, que puedan invitar a otros siglodoristas (o a nosotros mismos en un futuro) a continuar con esta labor:

El entremés X parece ser / no parece ser de Cervantes

En nuestro caso particular, nos centraremos únicamente en los resultados obtenidos para *La cárcel de Sevilla* (prosa) y *Los romances* (verso), cuyo estudio completo de autoría desarrollaremos en los capítulos 6 y 7, respectivamente, siguiendo las hipótesis de trabajo aquí consignadas.

Un primer análisis de agrupación basado en el estudio de la distribución de las palabras más frecuentes (1000 MFW)¹⁷², cuya efectividad ha sido demostrada en el apartado precedente, permite advertir:

- que los dos entremeses dubitados que más se aproximan al comportamiento de los ocho cervantinos son *La cárcel de Sevilla* y *Los mirones*: el primero se sitúa junto a *La elección de los alcaldes de Daganzo*, *El retablo de las maravillas* y *El rufián viudo*; el segundo, junto a *La guarda cuidadosa*, *La cueva de Salamanca*, *El juez de los divorcios*, *El vizcaíno fingido* y *El viejo celoso*.
- que los tres entremeses dubitados más alejados de los entremeses de Cervantes son *Melisendra*, *Durandarte* y *Belerma* y *Los romances*, que aparecen en otra rama.

¹⁷² Como los análisis realizados anteriormente no revelan diferencias significativas entre las 800 y las 1000 palabras más frecuentes, optamos por utilizar de ahora en adelante el parámetro 1000 MFW.

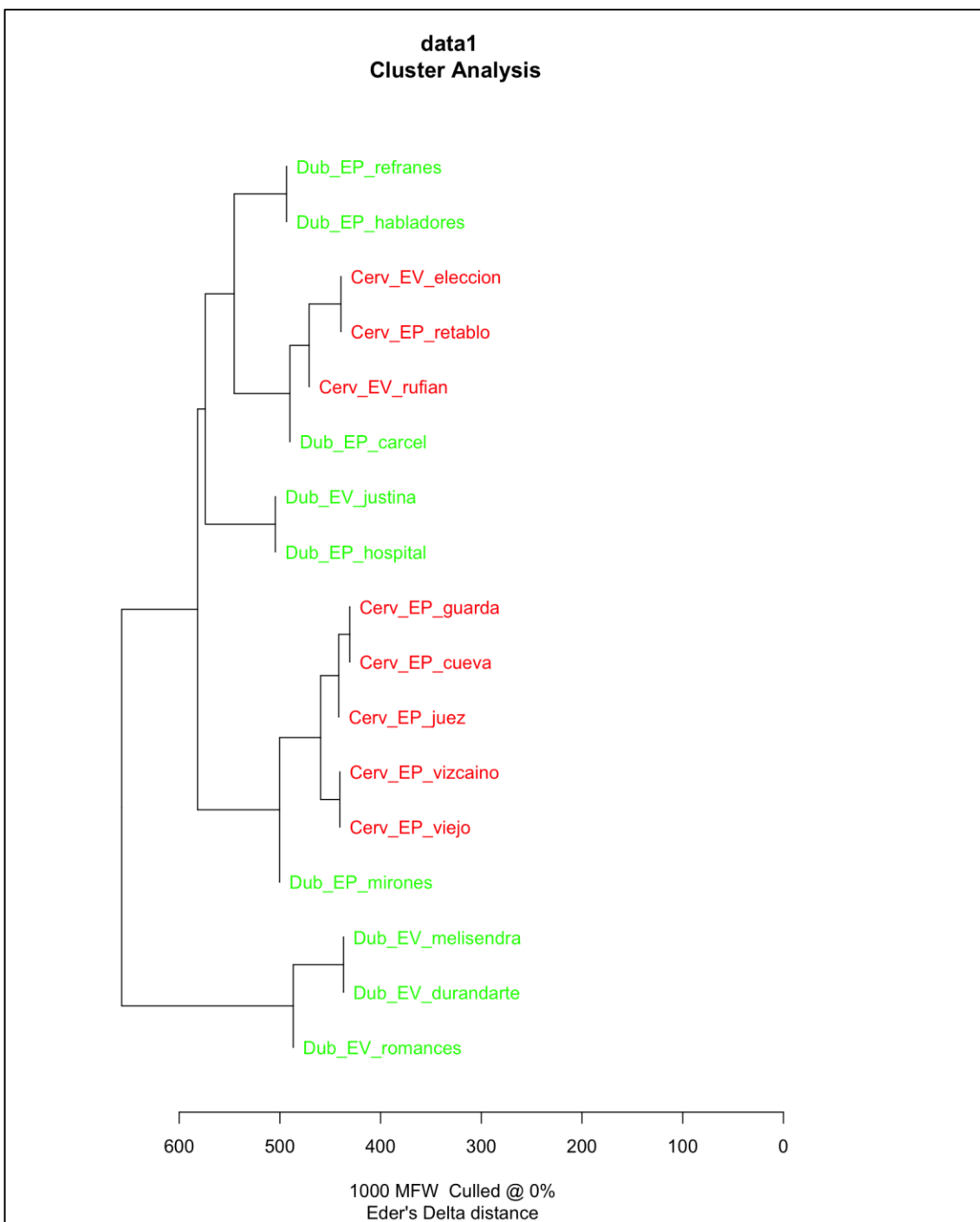


Gráfico 18. Clúster de agrupación basado en la frecuencia de las palabras más empleadas en los entremeses dubitados e indubitados de Cervantes (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

Esas fortalezas y debilidades asociativas también se aprecian en el siguiente árbol de consenso, realizado mediante *Bootstrap* para mostrar las relaciones más estables entre 800 y 1000 MFW, con un *consensus strength* de 0,5¹⁷³:

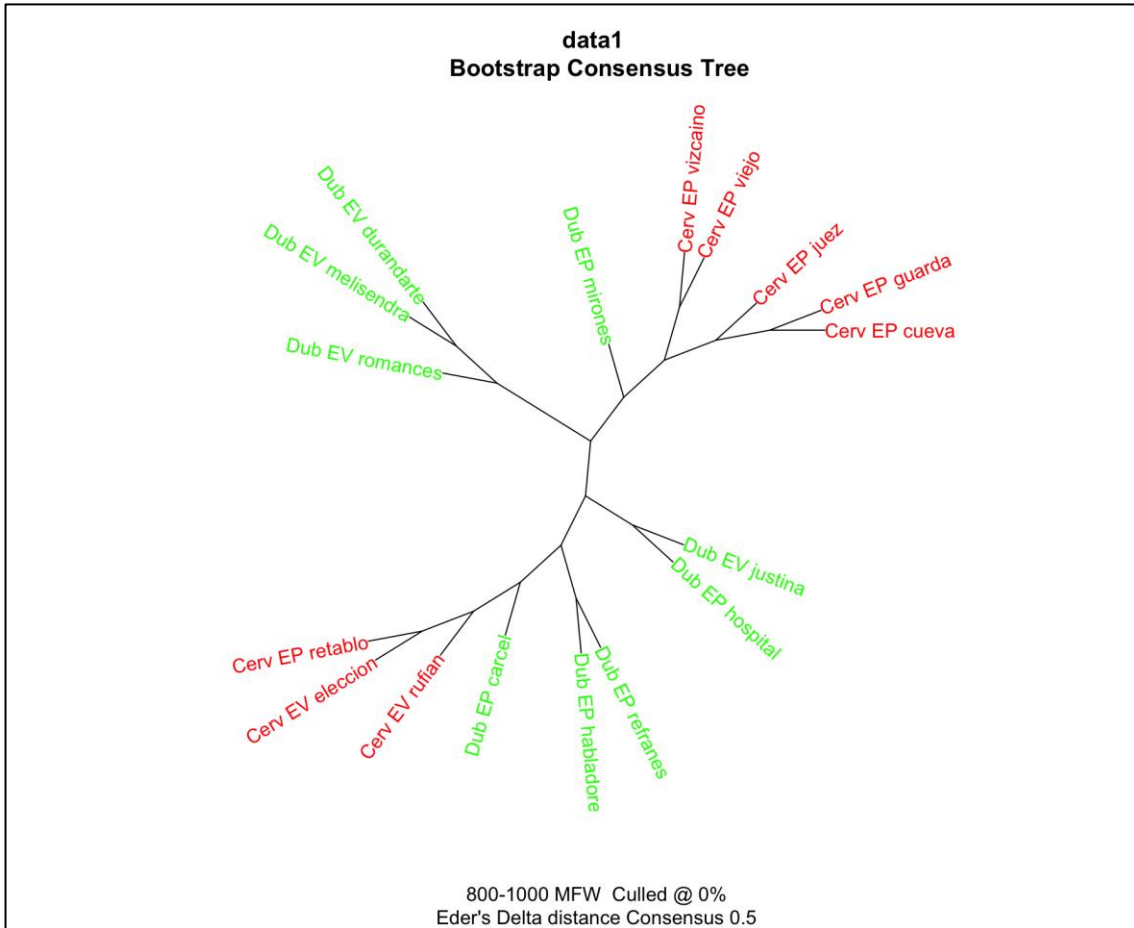


Gráfico 19. Árbol de consenso de los entremeses dubitados e indubitados de Cervantes (RStudio. Stylo. Bootstrap Consensus Tree. Consensus strength 0.5. Eder's Delta. MFW 800-1000. Culled 0%)

Algo parecido ocurre al realizar el análisis de componentes principales (PCA), una técnica estadística de síntesis de la información que consiste en describir un

¹⁷³ Como explicar Eder, los *Bootstrap Consensus Tree* permiten evitar el problema de la arbitrariedad en la selección de las MFW, ya que el árbol generado muestra las relaciones más estables en un gran número de iteraciones [2013 y 2017].

conjunto de datos en términos de nuevas variables (o componentes) no correlacionadas, con el objetivo es reducir la dimensionalidad de dicho conjunto [Jolliffe, 2002: 2]¹⁷⁴:

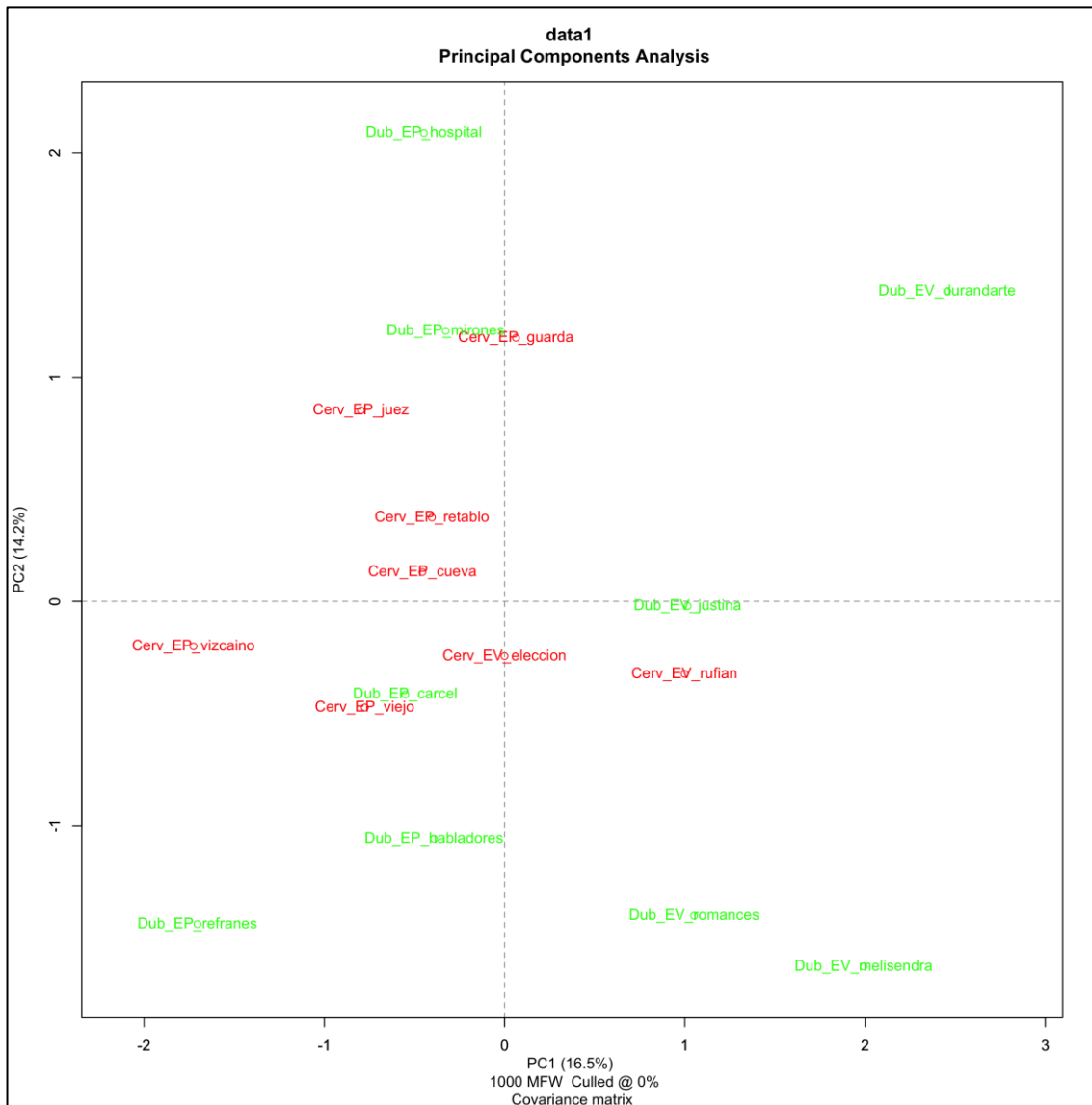


Gráfico 20. Gráfica de puntuaciones del PCA de las 1000 palabras más frecuentes en los entremeses dubitados e indubitados de Cervantes (RStudio. PCA. Eder's Delta. 1000 MFW. Culled 0%)

¹⁷⁴ El análisis de componentes principales (PCA), al igual que la distancia Delta (*Delta distance*), ha resultado muy efectivo en los estudios de atribución de autoría. De la Rosa y Suárez lo han aplicado al caso del *Lazarillo* [2016].

Los dos primeros componentes principales (PC1 y PC2) cubren respectivamente el 16,5% y el 14% de la varianza total. Llama la atención que primer componente (eje horizontal) muestra una clara diferenciación entre los entremeses en verso (en la mitad derecha) y los entremeses en prosa (en la mitad izquierda), a excepción de *La guarda cuidadosa*, que se sitúa en el lado derecho de PC1, aunque ciertamente próximo al lado izquierdo¹⁷⁵. *La cárcel de Sevilla* y *Los mirones* aparecen casi superpuestos a dos obras cervantinas; de estas dos aproximaciones, la verdaderamente significativa es la de *La cárcel de Sevilla* con *El viejo celoso*, pues la de *Los mirones* con *La guarda cuidadosa* se produce en cuadrantes distintos. Los cuatro entremeses dubitados más alejados de los puntos de distribución de los indubitados cervantinos son *El hospital de los podridos*, *Durandarte* y *Belerma*, *Los habladores*, *Los romances*, *Los refranes* y *Melisendra*; todos ellos, representados en los extremos, son los menos definidos por las características del grupo.

Aunque estos primeros análisis ya resultan relevantes, creemos conveniente estudiar el comportamiento de la distribución de la frecuencia de las palabras más empleadas (*Delta distance*), del porcentaje de similitud textual (*similarity*) y de la longitud de secuencias de palabras idénticas (*verbatim*) en cada uno de estos textos, a fin de poder establecer de manera más sólida nuestras hipótesis de trabajo. No realizamos aquí el análisis de la frecuencia de 5-gramas de caracteres (*character 5-grams*) por no ser tan fiable como las otras tres variables (algo en lo que influye, sin duda, el hecho de que la puntuación sea responsabilidad del editor moderno).

¹⁷⁵ El primer componente discrimina entre entremeses escritos en prosa y entremeses escritos en verso, aunque esta diferenciación podría corresponderse en realidad con la datación de los mismos, ya que los entremeses en prosa son anteriores en el tiempo a los entremeses en verso. Recientemente hemos validado el uso del PCA para mostrar la evolución estilística de un autor a lo largo del tiempo, tomando como referencia los *Episodios nacionales* de Benito Pérez Galdós o varias obras de Miguel Delibes [Celma Valero y Ruiz Urbón, 2021]. Sería interesante, y tal vez lo hagamos en un futuro próximo, proponer una nueva teoría de la datación de los entremeses de Cervantes teniendo en cuenta, entre otros factores, el análisis de componentes principales.

APROXIMACIÓN AL CASO DE ESTUDIO

Comenzamos por el *Entremés de los habladores* (Dub_EP_habladores). El estudio de la variación de la frecuencia en las 1000 MFW lo sitúa en la misma subrama en la que se encuentran tres de los entremeses indubitados de Cervantes y los dos de Ávila, aunque asociado más directamente con los del segundo (Gráfico 21):

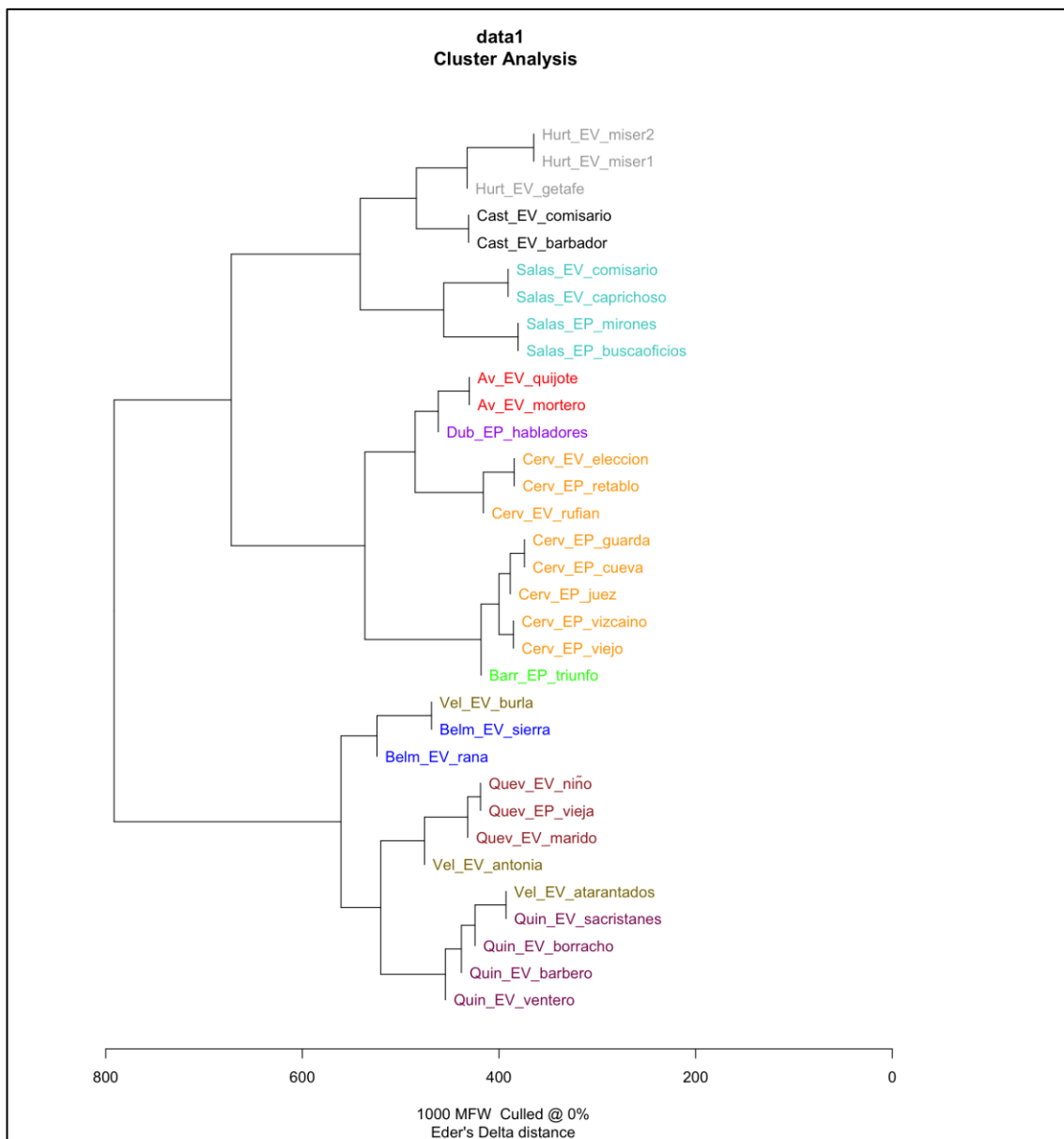


Gráfico 21. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Los habladores* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

Este entremés presenta un intervalo de similitud textual con los de Cervantes del [24 , 36] %, lo que lo sitúa en un nivel medio de proximidad (Gráfico 22).

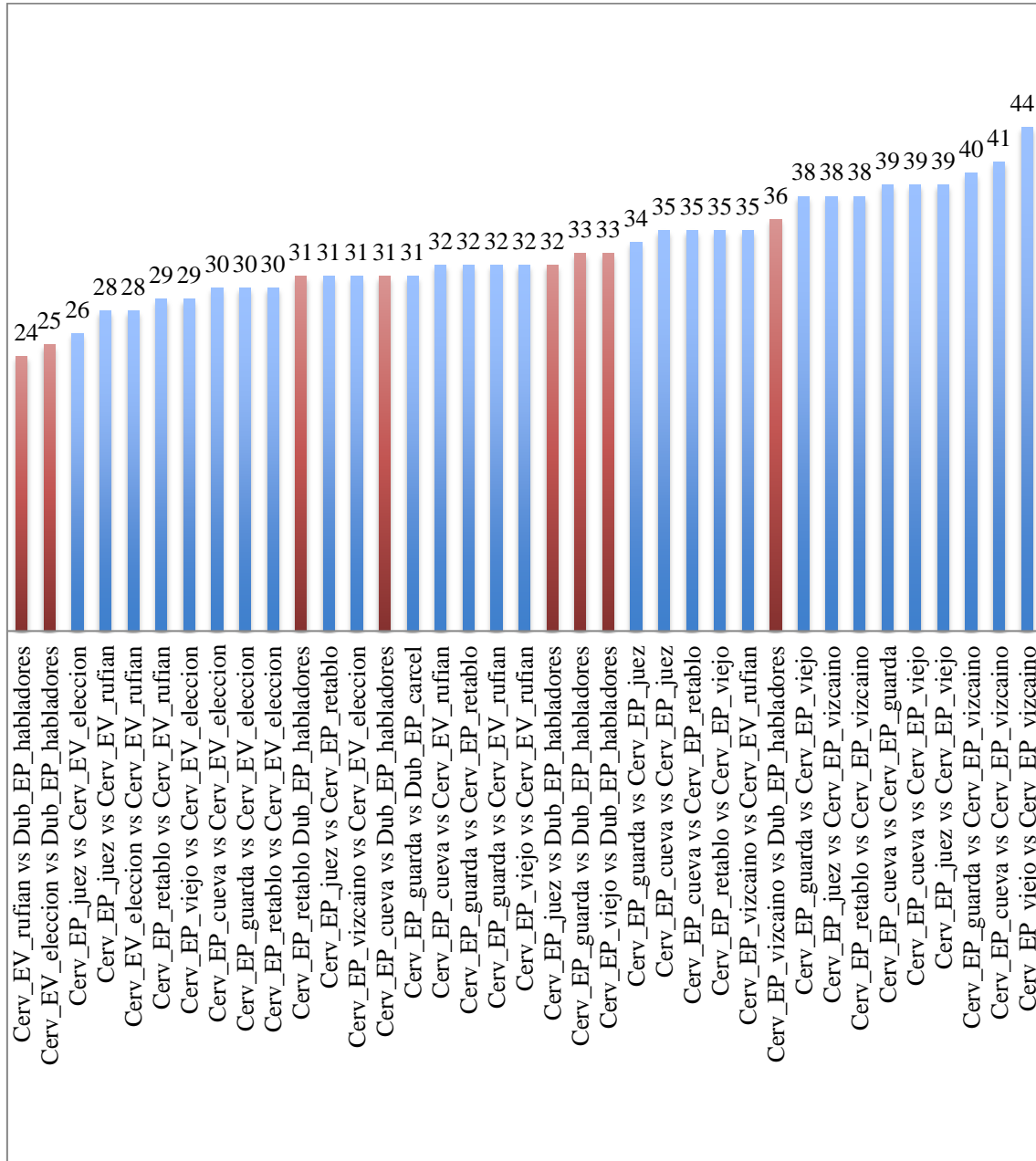


Gráfico 22. Nivel de similitud textual entre *Los habladores* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

No hemos detectado ningún caso de *verbatim* de más de 5 o ítems entre *Los habladores* y el resto de entremeses indubitados de Cervantes.

APROXIMACIÓN AL CASO DE ESTUDIO

Pasamos ahora a realizar estos mismos análisis en *El hospital de los podridos* (Dub_EP_podridos). El estudio con Delta lo sitúa en la misma subrama en la que se encuentran cinco de los ocho entremeses de Cervantes y el único de Barrionuevo, aunque sin asociarse directamente con ellos (Gráfico 23). El análisis con CopyCatch Gold (Gráfico 24) revela que mantiene una similitud textual con los de Cervantes del [27 , 36] %, sin alcanzar en ningún caso el umbral del 40% :

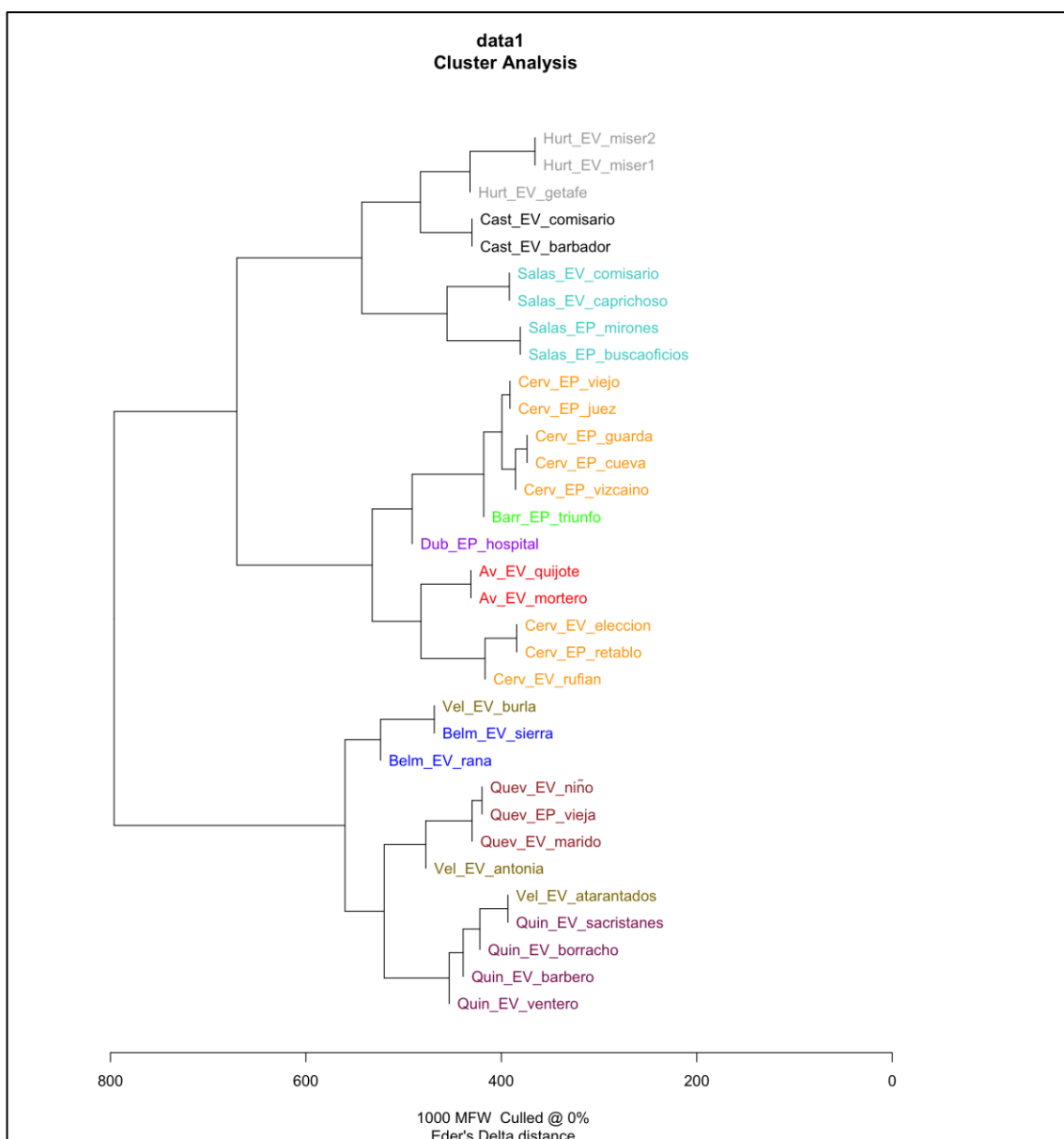


Gráfico 23. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *El hospital de los podridos* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

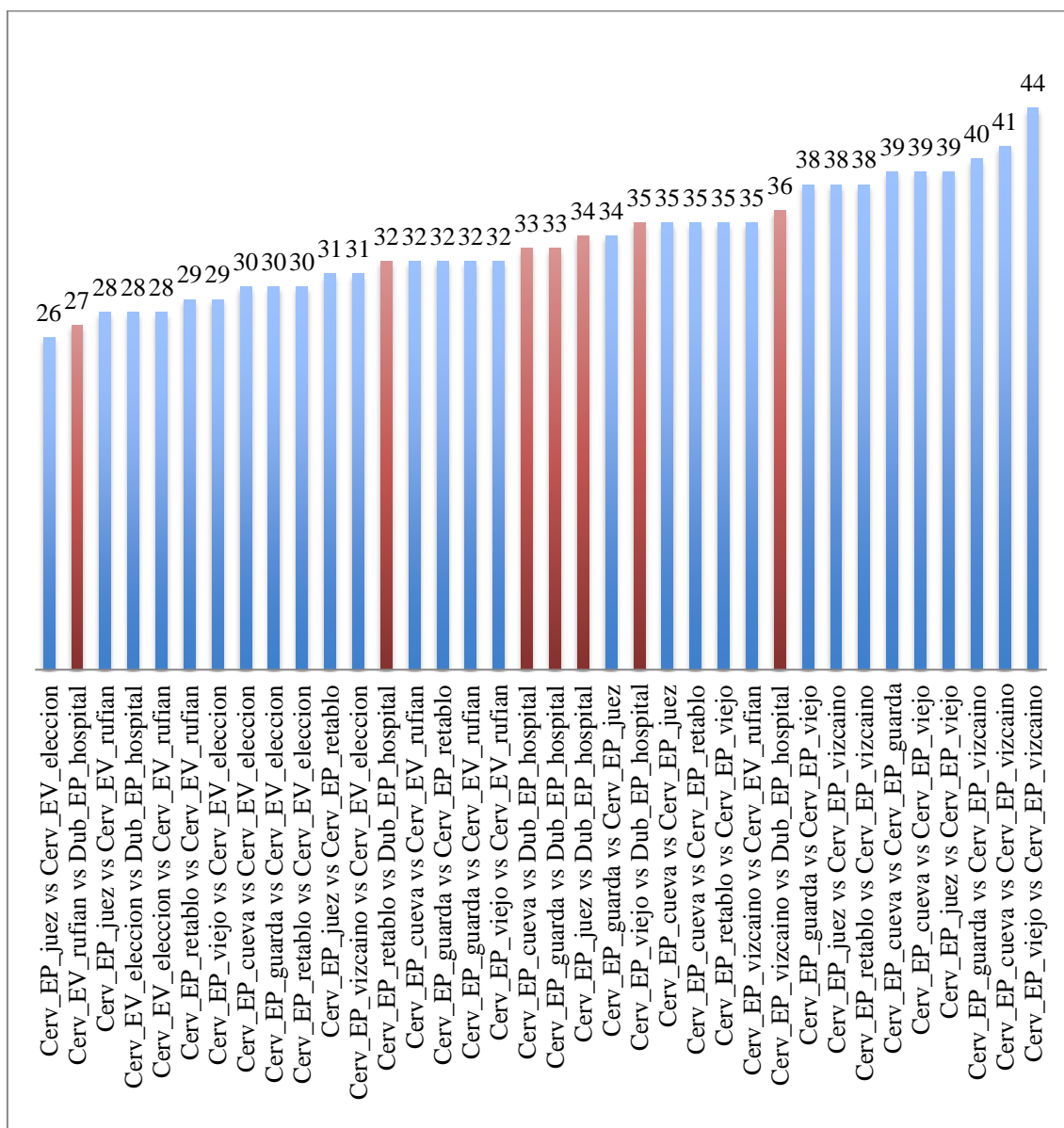


Gráfico 24. Nivel de similitud textual entre *El hospital de los podridos* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

No hemos encontrado ningún caso de *verbatim* de más de 5 ítems entre *El hospital de los podridos* y el resto de entremeses cervantinos ¹⁷⁶.

¹⁷⁶ En *El hospital de los podridos* encontramos “¿por qué lo pregunta vuesa merced?” y en dos de los entremeses de Cervantes aparece esa secuencia pero con el objeto indirecto explícito: “¿por qué me lo pregunta vuesa merced?”(Cerv_EP_guarda y Cerv_EP_vizcaino). No es, por tanto, un caso de *verbatim*.

APROXIMACIÓN AL CASO DE ESTUDIO

Analizamos, en tercer lugar, el *Entremés de la cárcel de Sevilla* (Dub_EP_carcel). El análisis con Delta lo sitúa en la misma subrama en la que se encuentran tres de los entremeses indubitados de Cervantes; además, su inclusión debilita la asociación que se producía entre éstos y los entremeses indubitados de Ávila (Gráfico 25):

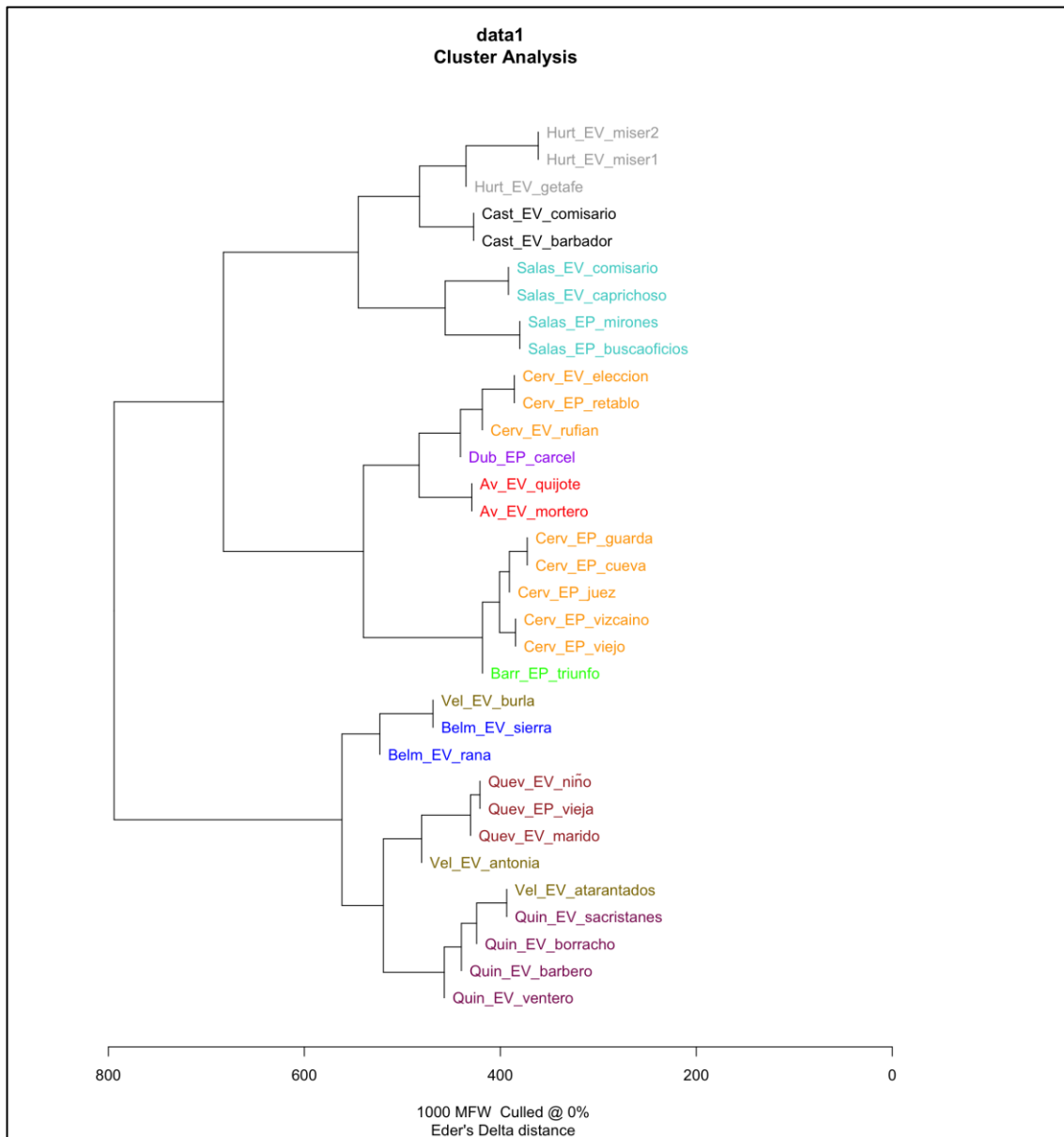


Gráfico 25. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *La cárcel de Sevilla* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

Su intervalo de similitud textual con los indubitados de Cervantes es del [25 , 34] %, por lo que se situaría en un nivel medio de proximidad (Gráfico 26):

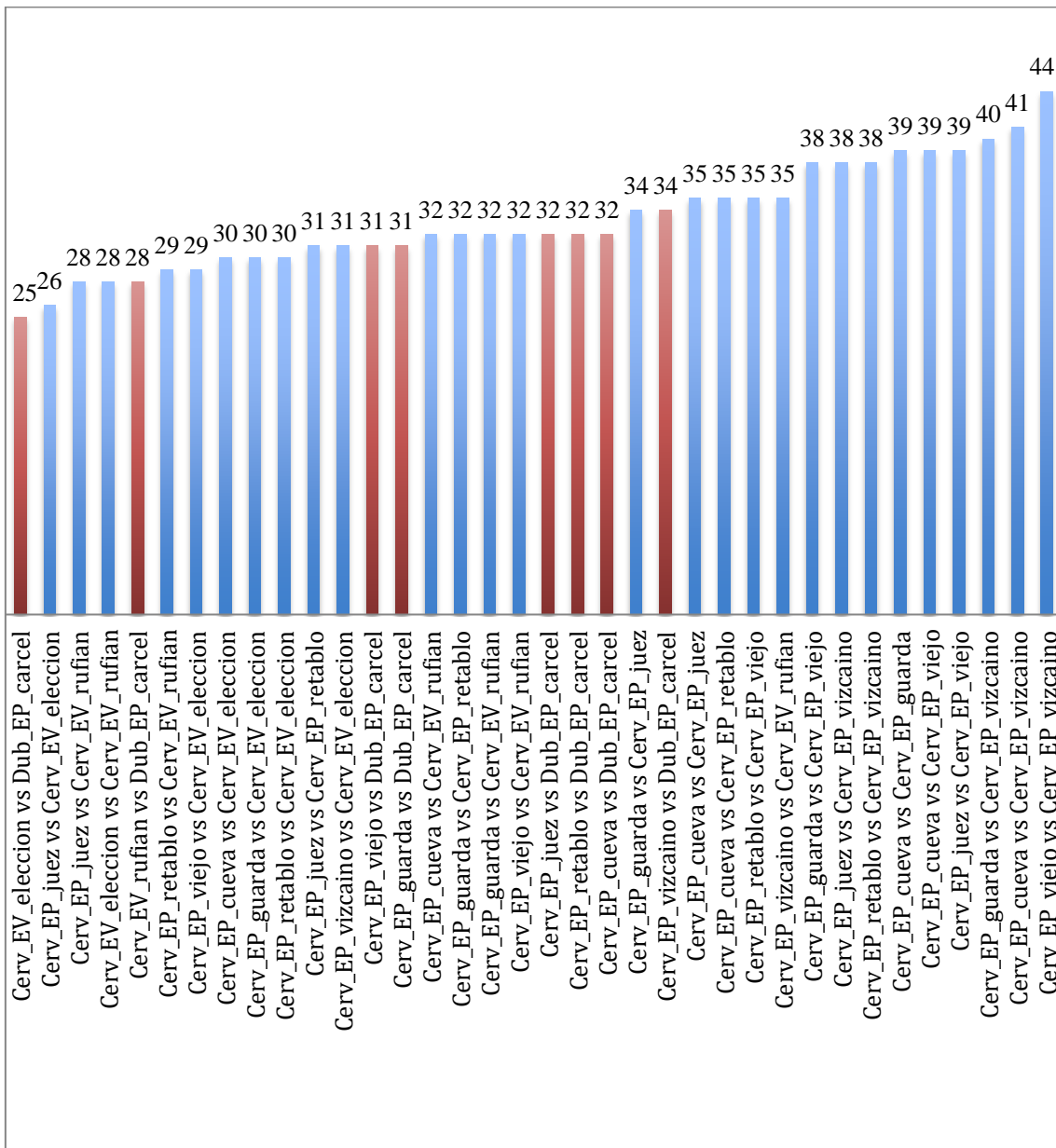


Gráfico 26. Nivel de similitud textual entre *La cárcel de Sevilla* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

La cárcel de Sevilla y los entremeses de Cervantes presentan un *verbatim* de 7 ítems (“eso haré yo de muy buena gana”). El CORDE confirma que no es exclusivo de Cervantes, pero sí es muy característico de su escritura, pues aparece dos veces en el *Quijote* de 1605, tres en el *Quijote* de 1615 y una en *Rinconete y Cortadillo*:

APROXIMACIÓN AL CASO DE ESTUDIO

VERBATIM	ÍTEMS	TEXTOS	CORDE (1547-1617)
eso haré yo de muy buena gana	7	Cerv_EP_juez Cerv_EP_vizcaino Dub_EP_carcel	8 (6 Cerv)

Tabla 12. Casos de *verbatim* de más de 5 ítems entre Dub_EP_carcel y los entremeses de Cervantes

Pasamos ahora al *Entremés de los refranes* (Dub_EP_refranes):

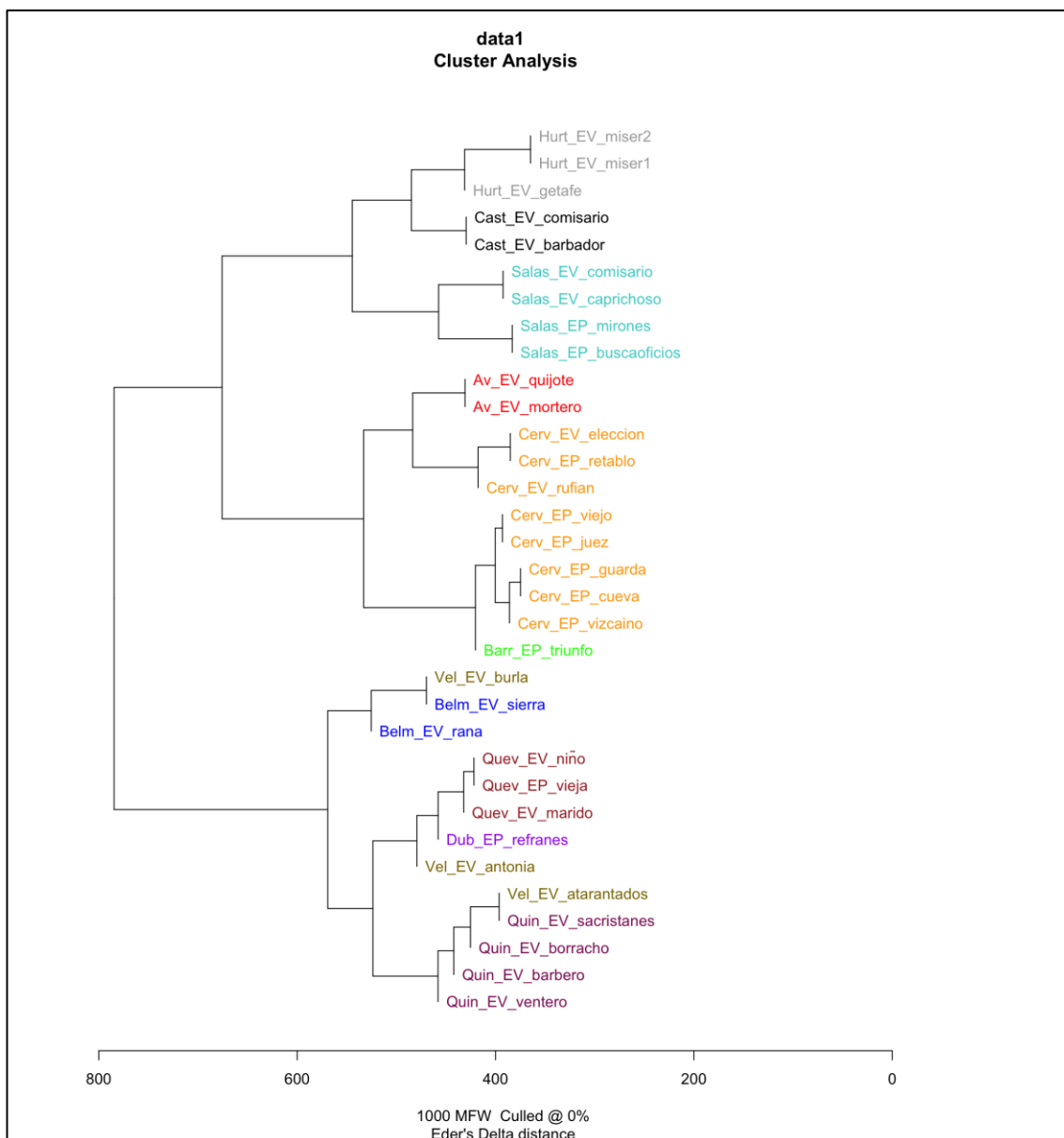


Gráfico 27. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Los refranes* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

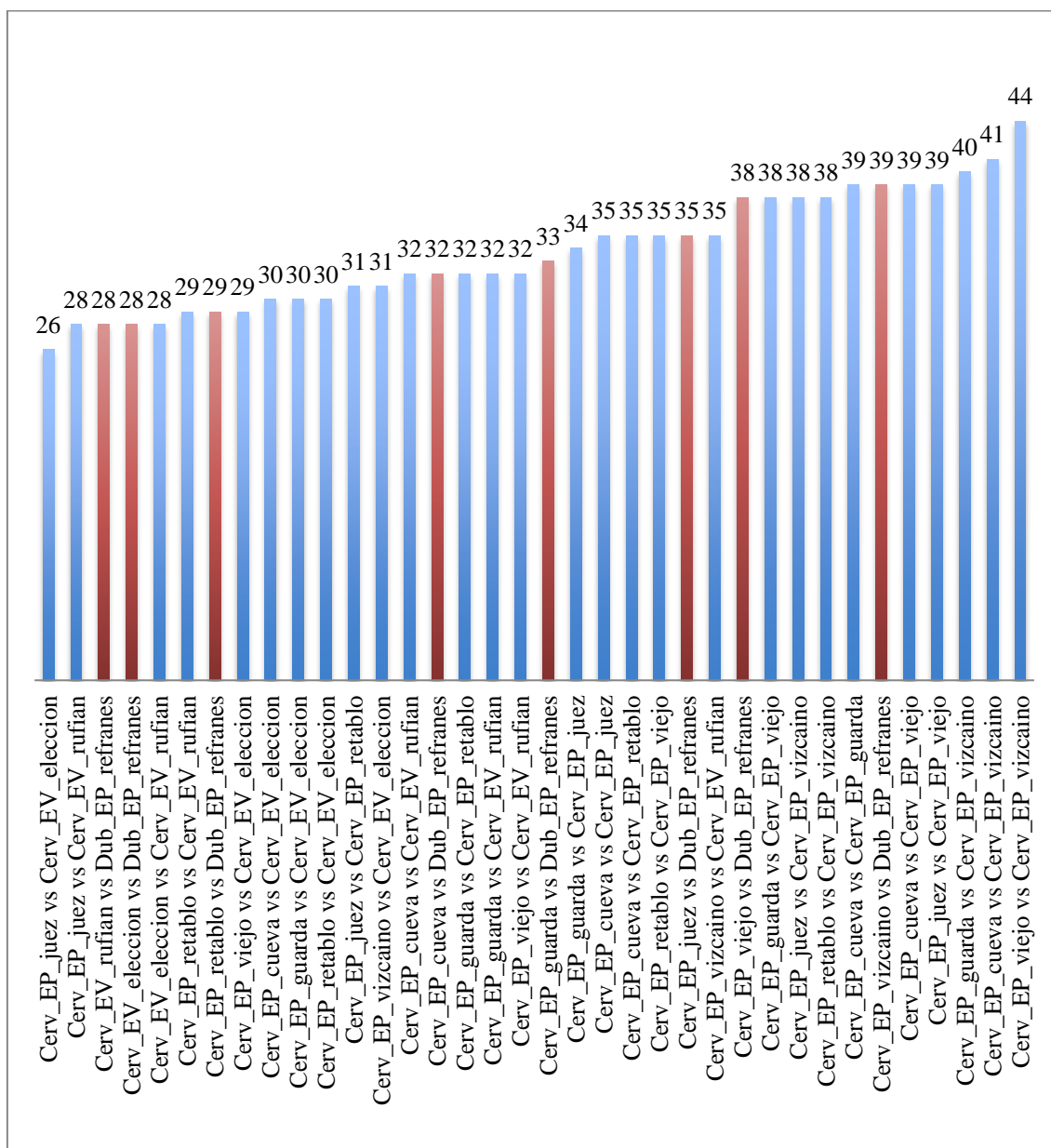


Gráfico 28. Nivel de similitud textual entre *Los refranes* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

El estudio de la variación de la frecuencia léxica en las 1000 MFW aleja a *Los refranes* de los indubitados cervantinos (Gráfico 27), pero no tanto su similitud textual, que aunque nunca alcanza el 40%, sí está dentro de la horquilla de los de Cervantes [28 , 39] % (Gráfico 28). Tal como mostramos en la Tabla 13, hemos localizado un caso de *verbatim* de 11 ítems entre este entremés dubitado y *El juez de los divorcios*; aunque CORDE no recoge ningún otro caso en esas fechas, se trata de un conocido refrán.

APROXIMACIÓN AL CASO DE ESTUDIO

VERBATIM	ÍTEMS	TEXTOS	CORDE (1547-1617)
riñas de por San Juan son paz para todo el año	11	Cerv_EP_juez Dub_EP_refranes	0

Tabla 13. Casos de *verbatim* de más de 5 ítems entre Dub_EP_refranes y los entremeses de Cervantes

Analizamos, en quinto lugar, el caso de *Los mirones* (Dub_EP_mirones).

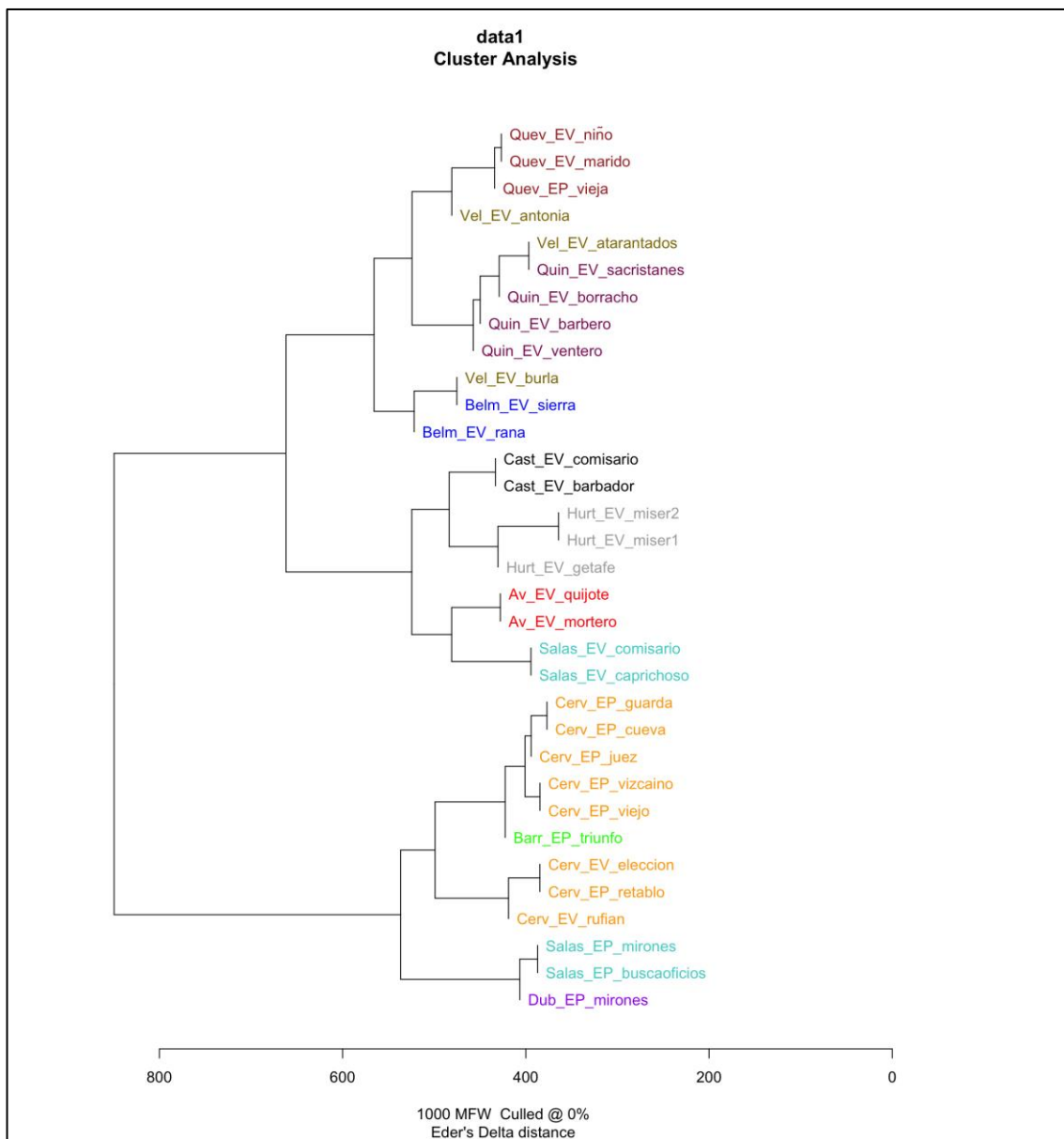


Gráfico 29. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Los mirones* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

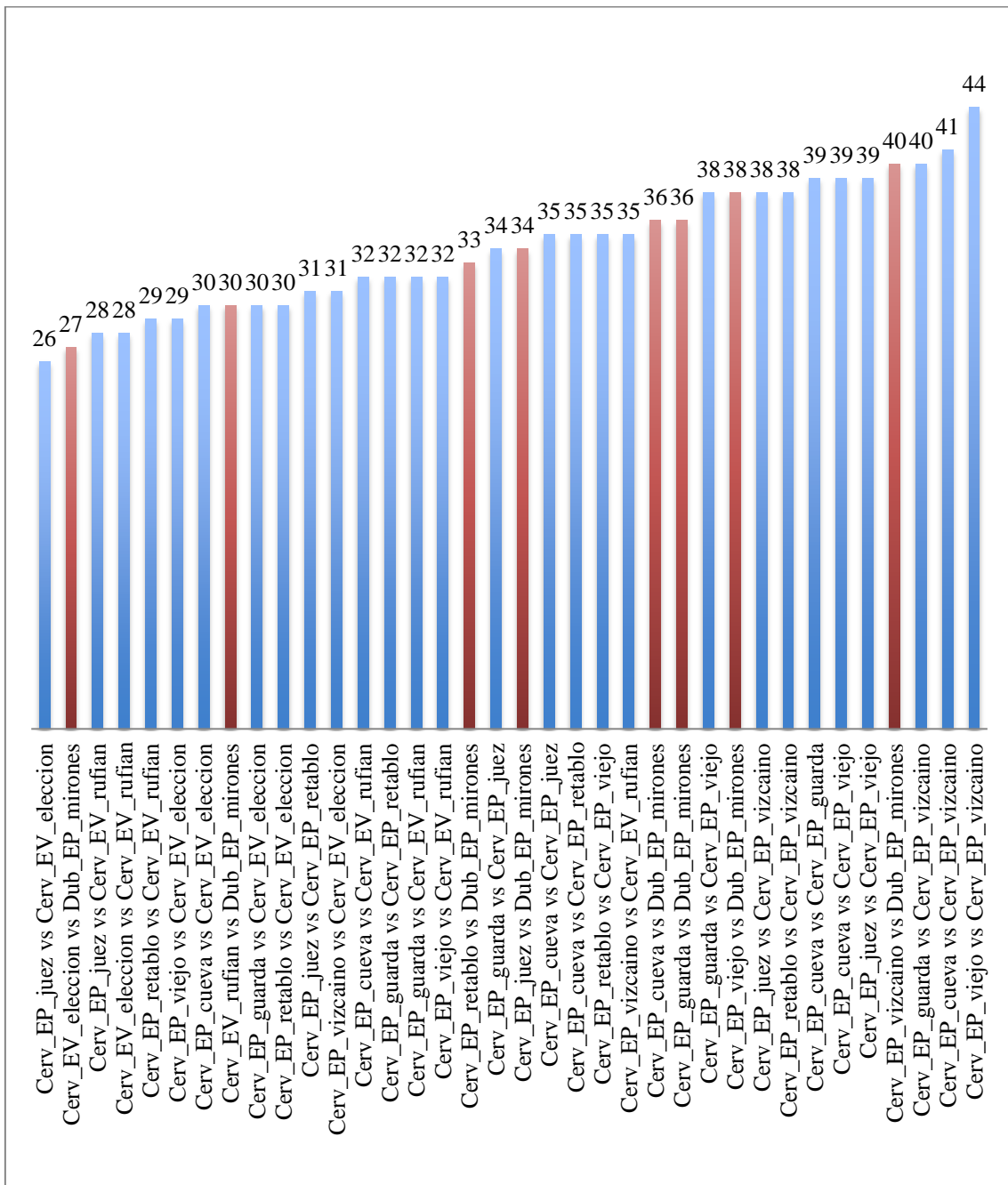


Gráfico 30. Nivel de similitud textual entre *Los mirones* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

El algoritmo Eder's Delta lo identifica con los dos entremeses en prosa de Salas Barbadillo (Gráfico 29) y el análisis con CopyCatch Gold revela que hay un par textual formado por este entremés y uno cervantino que alcanza una similitud textual del 40%

APROXIMACIÓN AL CASO DE ESTUDIO

(Cerv_EP_vizcaino vs Dub_EP_mirones) (Gráfico 30). No hemos localizado ningún caso de *verbatim* de más de 5 ítems entre *Los mirones* y el resto de entremeses de Cervantes.

Pasamos al *Entremés de los romances* (Dub_EV_romances):

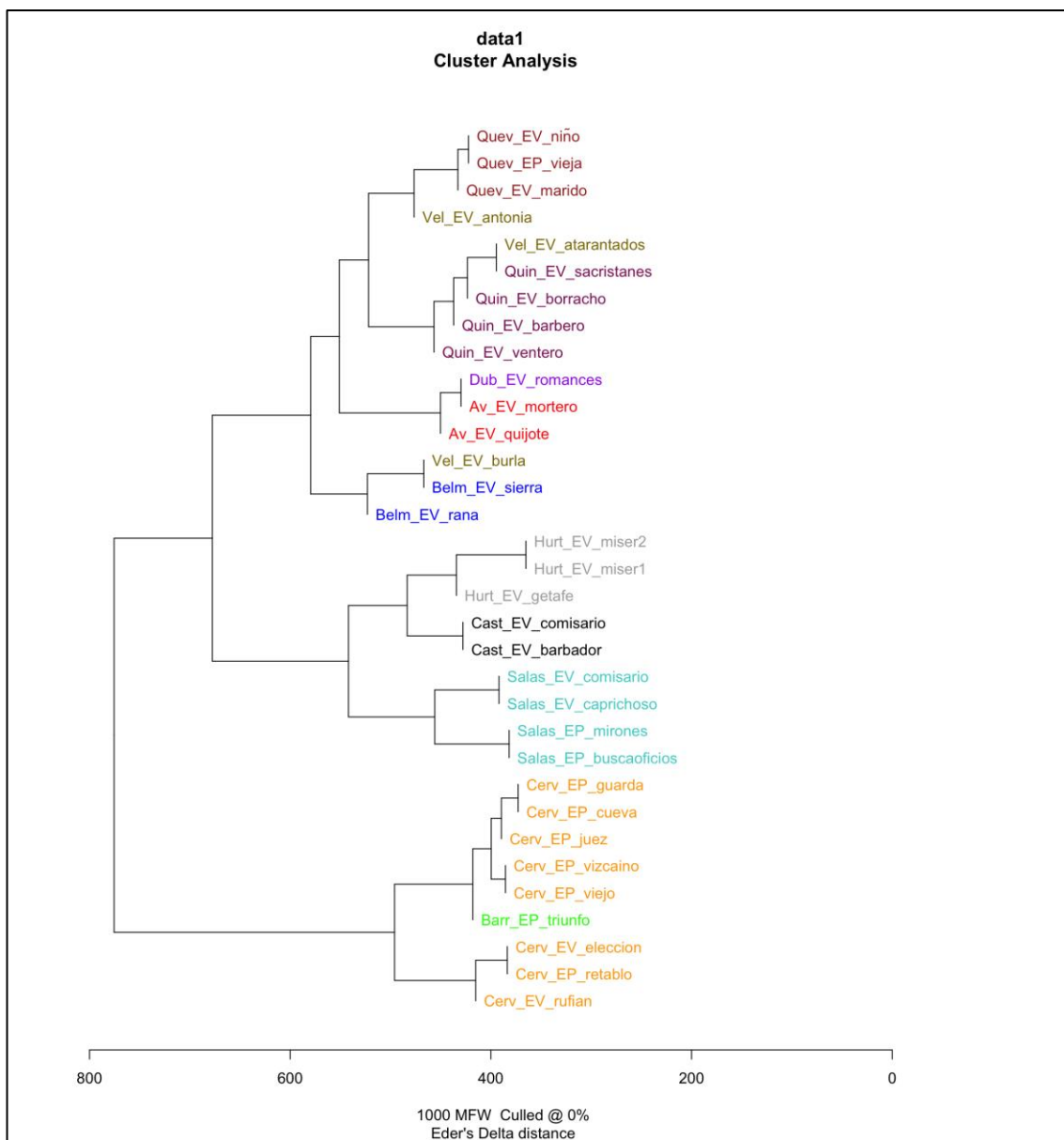


Gráfico 31. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Los romances* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

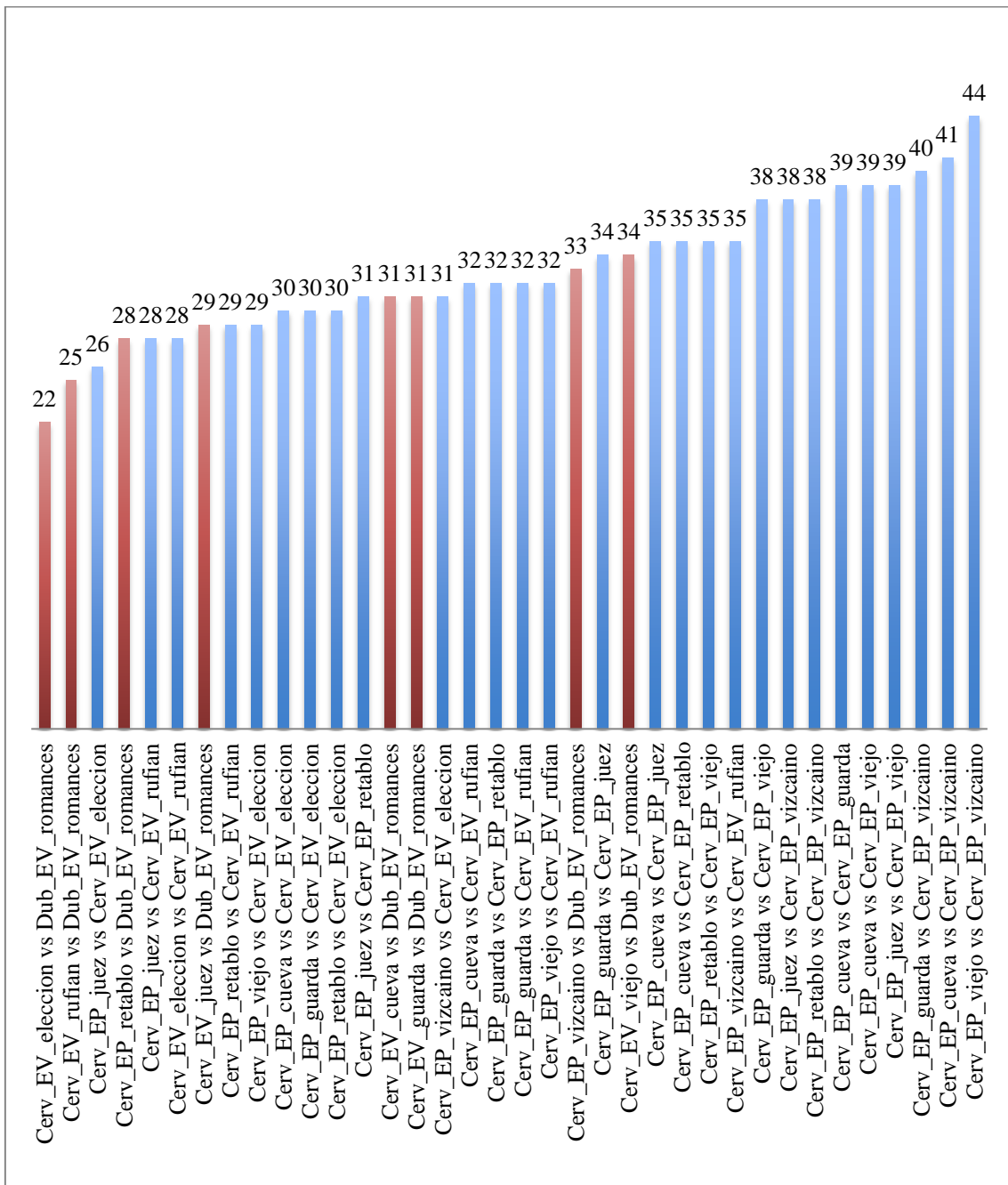


Gráfico 32. Nivel de similitud textual entre *Los romances* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

Los romances no sólo no se asocia con los indubitados cervantinos, sino que, además, su inclusión en el análisis con Delta aleja las obras de Ávila de las de Cervantes (Gráfico 31). Su intervalo de similitud textual con los entremeses de Cervantes es poco significativo (Gráfico 32) y no hay entre ellos casos de *verbatim* de más de 5 ítems.

Continuamos con el *Entremés de Melisendra* (Dub_EV_melisendra):

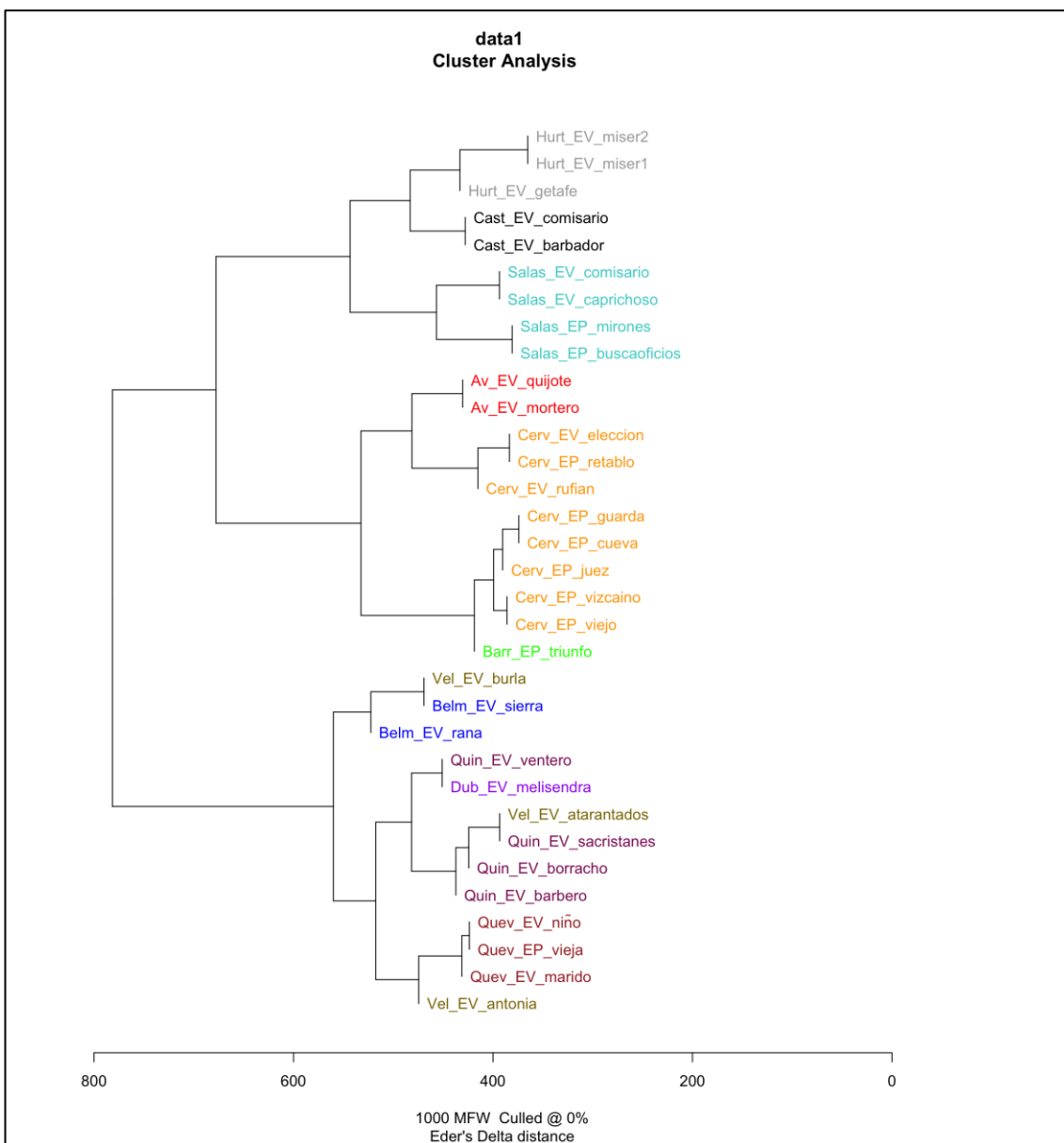


Gráfico 33. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Melisendra* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

El comportamiento léxico de *Melisendra* no se asemeja al de Cervantes, tanto en la distancia Delta (Gráfico 33) como en el análisis de similitud textual (Gráfico 34).

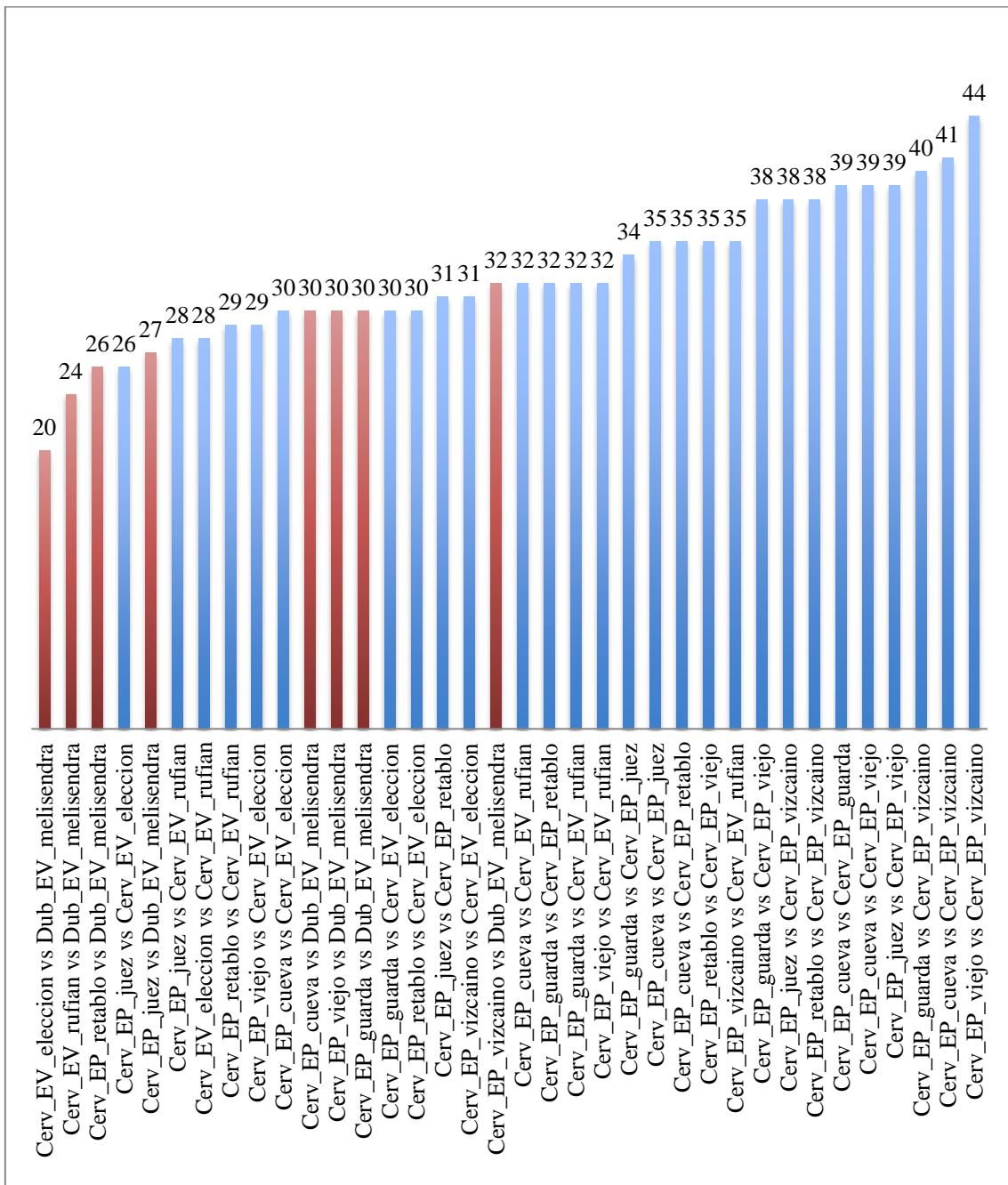


Gráfico 34. Nivel de similitud textual entre *Melisendra* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

Además, no hay casos de *verbatim* de más de 5 ítems entre *Melisendra* y el resto de entremeses cervantinos.

APROXIMACIÓN AL CASO DE ESTUDIO

Veamos qué ocurre con los análisis del *Entremés de Durandarte y Belerma* (Dub_EV_durandarte). La distancia Delta lo asocia directamente con *El rufián viudo*, que, recordemos, es el entremés de Cervantes que muestra un comportamiento más diferenciado del resto en esta variante (Gráfico 35).

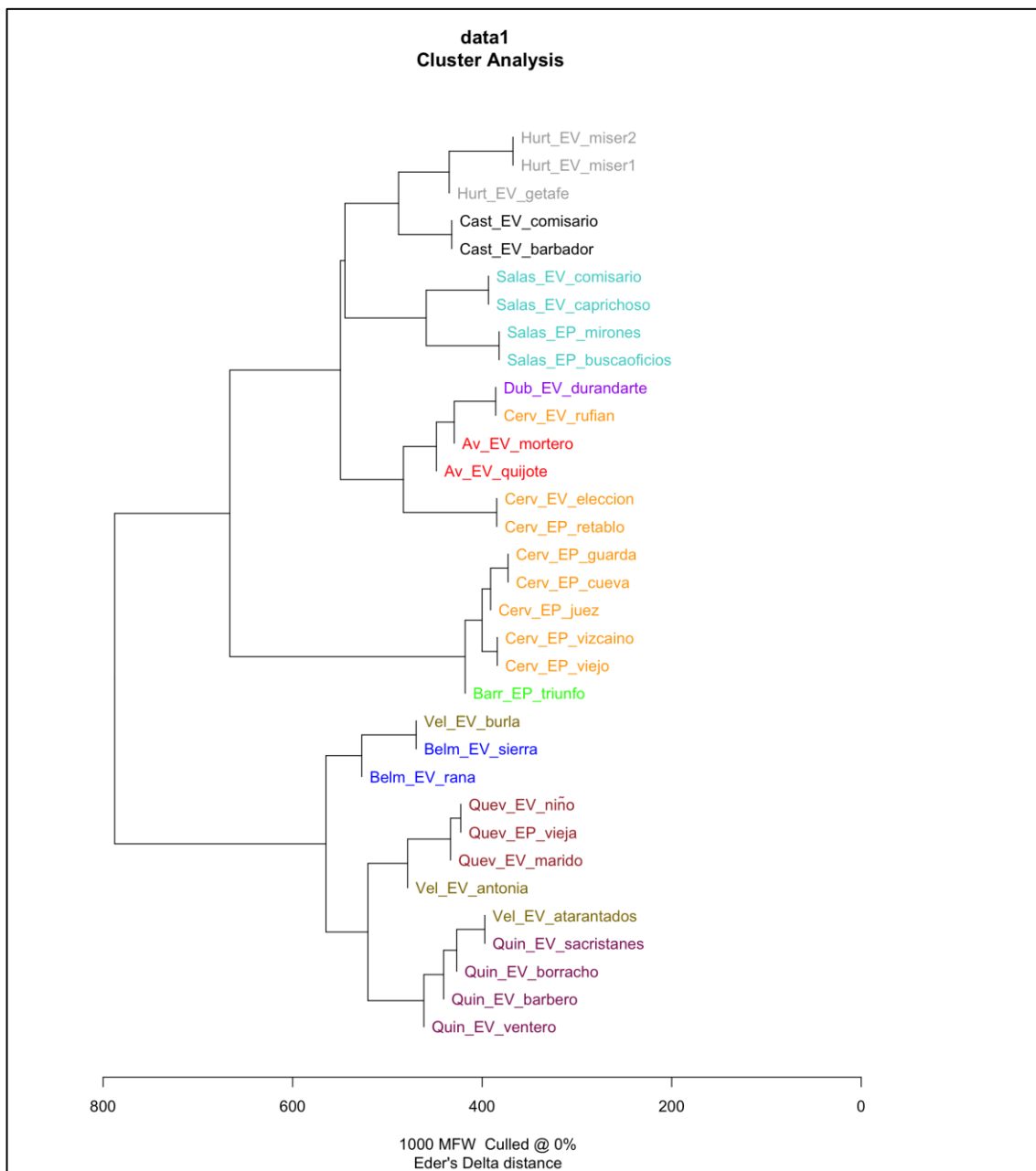


Gráfico 35. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Durandarte* y *Belerma* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

Sus porcentajes de similitud textual con los entremeses cervantinos son bajos, pues en ningún caso alcanzan el 30% (Gráfico 36):

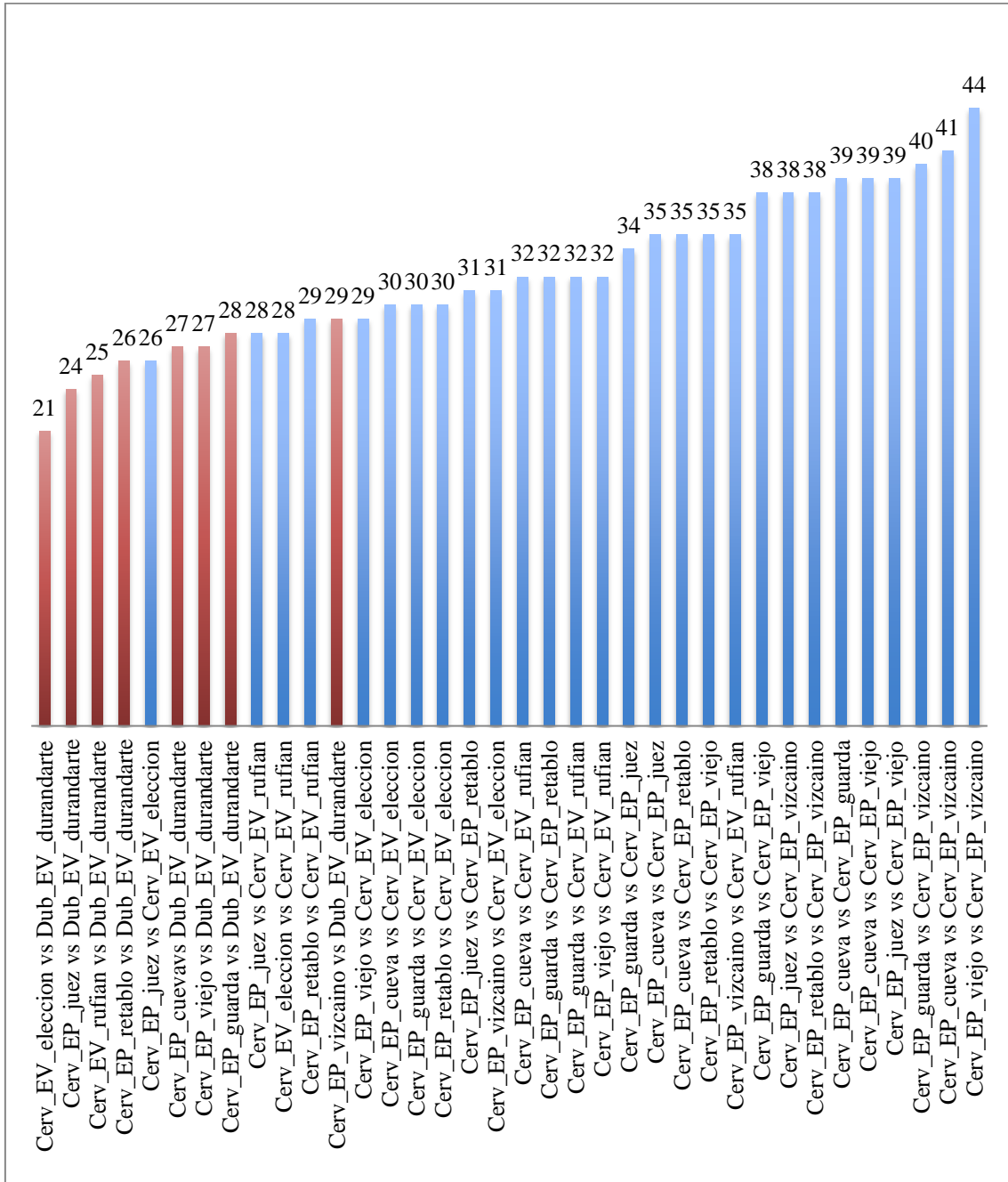


Gráfico 36. Nivel de similitud textual entre *Durandarte* y *Belerma* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

APROXIMACIÓN AL CASO DE ESTUDIO

No hay casos de *verbatim* de más de 5 ítems entre *Durandrate* y *Belerma* y el resto de entremeses indubitados de Cervantes.

Analizamos en último lugar el *Entremés de Doña Justina y Calahorra* (Dub_EV_justina):

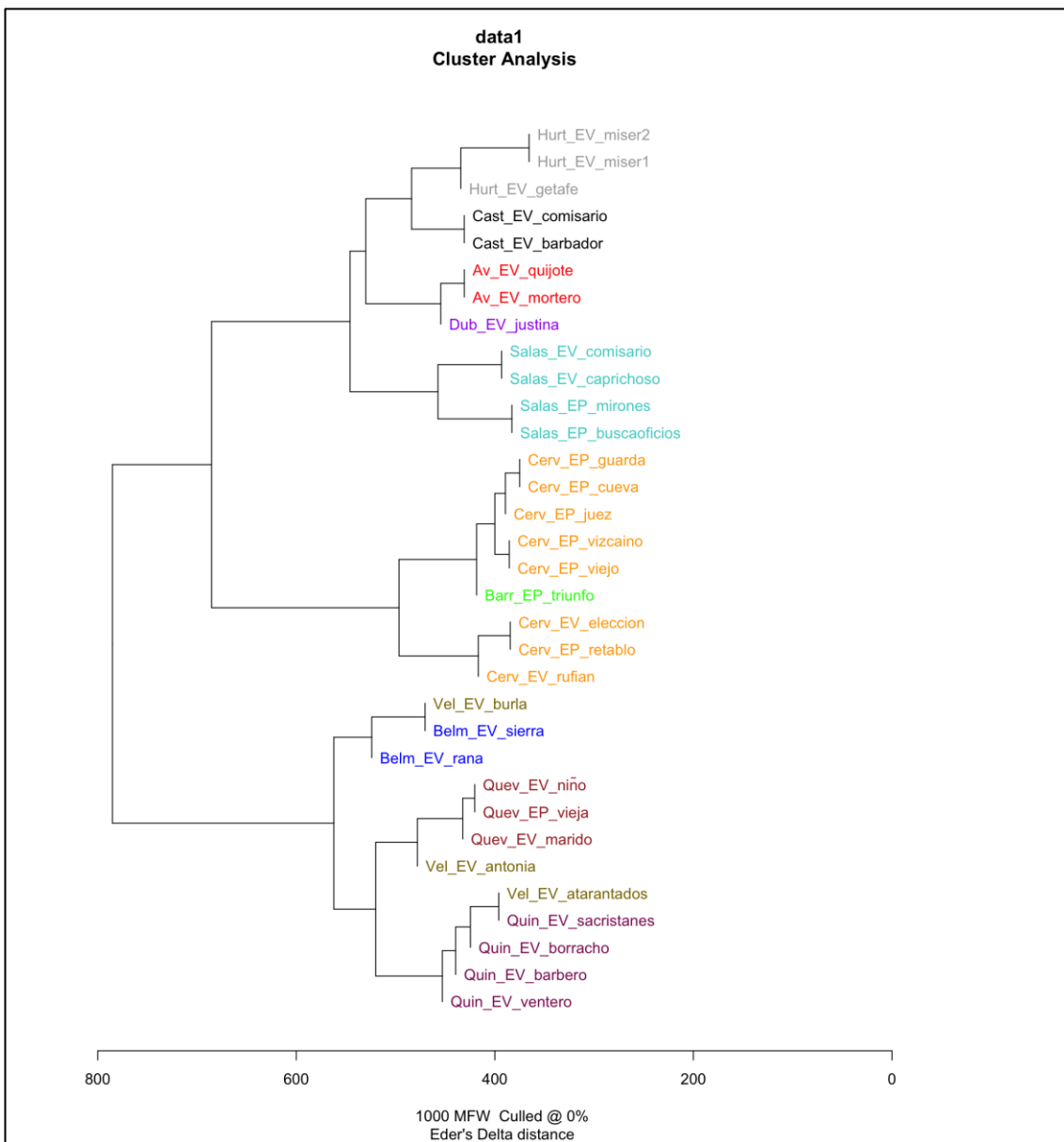


Gráfico 37. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de Cervantes y *Doña Justina* y *Calahorra* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000)

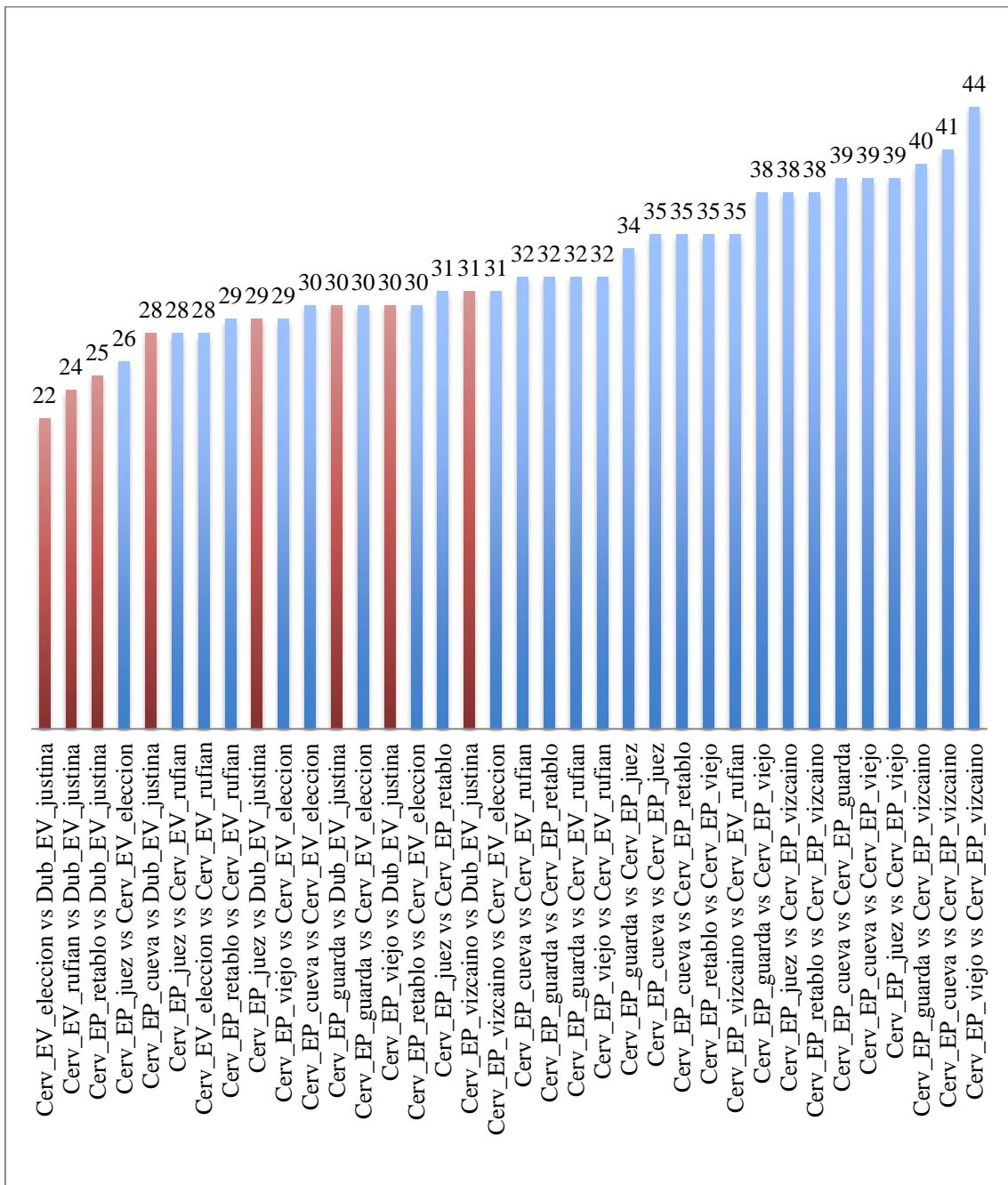


Gráfico 38. Nivel de similitud textual entre *Doña Justina* y *Calahorra* y los entremeses de Cervantes (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

El comportamiento léxico de *Durandarte* y *Belerma* no se asemeja al de Cervantes (Gráficos 37 y 38). Tampoco hay casos de *verbatim* de más de 5 ítems entre *Durandarte* y *Belerma* y el resto de entremeses de Cervantes.

A la luz de los resultados obtenidos (*PCA, Delta distance, similarity y verbatim*), podemos establecer las siguientes hipótesis de trabajo para cada uno de los nueve entremeses conservados:

- *La cárcel de Sevilla* parece ser de Cervantes.
- *Los habladores, El hospital de los podridos, Los refranes, Los romances, Melisendra y Doña Justina y Calahorra* no parecen ser de Cervantes.
- No podemos establecer hipótesis concluyentes sobre *Los mirones y Durandarte y Belerma*, pues, en realidad, no son entremeses (o al menos no lo son tal y como nos han llegado).

Se trata, como hemos explicado antes, de meras hipótesis de trabajo, que deberán ser oportunamente validadas mediante estudios de atribución de autoría rigurosos y exhaustivos, que contemplen otras muchas variantes de análisis (morfológicas, sintácticas, etc.). Así lo haremos nosotros con *La cárcel de Sevilla* y *Los romances*, en los capítulos 6 y 7, respectivamente.

5.6. ESCALA DE OPINIÓN VERBAL PARA EL CASO DE ESTUDIO

Como ya hemos explicado en algún otro lugar, los estudios de atribución de autoría literaria deben seguir los mismos preceptos que propugna la Lingüística forense, entre los que se encuentra el principio de probabilidad, que establece que la acumulación de datos o evidencias permite inferir grados de probabilidad en relación a los hechos analizados. Por eso, una de las aportaciones de nuestra investigación es la de incorporar este principio a los estudios de autoría en el campo literario, desterrando para siempre las conclusiones categóricas (sí / no) o las que, sin llegar a serlo, (como “probablemente es” o “parece que no es”) no se circunscriben a una escala de opinión verbal que permita cuantificar ese grado de probabilidad.

CAPÍTULO 5

La escala de opinión verbal que proponemos a continuación —inspirada en las de Coulthard [Coulthard, Johnson y Wright, 2017: 197 y 19] y McMenamín [2002: 125], que traducíamos al español en el capítulo 2 (Tablas 4 y 5)— nos permitirá encuadrar las conclusiones alcanzadas para cada una de las hipótesis posibles (SÍ / NO) dentro de seis grados de probabilidad, así como, dado el caso, declarar la imposibilidad de establecer una conclusión:

ESCALA DE OPINIÓN VERBAL

Se puede determinar que Miguel de Cervantes Saavedra SÍ es el autor del entremés X, con un grado de probabilidad	(5) MUY ALTO (4) BASTANTE ALTO (3) ALTO (2) BAJO (1) BASTANTE BAJO (0) MUY BAJO
Se puede determinar que Miguel de Cervantes Saavedra NO es el autor del entremés X, con un grado de probabilidad	(0) MUY ALTO (-1) BASTANTE ALTO (-2) ALTO (-3) BAJO (-4) BASTANTE BAJO (-5) MUY BAJO
No se puede determinar que Miguel de Cervantes Saavedra SÍ o NO es el autor del entremés X	NO CONCLUYENTE

Tabla 14. Escala de opinión verbal propuesta para los estudios de autoría

CAPÍTULO 6

ESTUDIO DE AUTORÍA DEL *ENTREMÉS* *DE LA CÁRCEL DE SEVILLA* (PROSA)

“Y ha habido hombre que estando jugando a los naipes le han notificado sentencia de muerte y que se confiese, y ha respondido que le dejen ver su suerte; y tornándole a decir que mire que le notifican aquello, ha respondido a el escribano que haga su oficio y no pase de ahí: “Mire que me enojaré”. Otros, que muy en su juicio responden a el escribano cuando les hace semejantes notificaciones: “¿Quién dio esta sentencia?” Y diciéndole que el alcalde de la Justicia, o el teniente, ha respondido: “Puédelo hacer como juez; pero sea él tan honrado, que con una espada en la mano salga a reñir conmigo, Y veremos quién mata quién”. Y saliendo el escribano, santiguándose de desemejante disparate y atrevimiento, torna a la baraja a decir: “Digo mi parte”. Y por qué algunos bachilleres presos le aconsejan que antes que se vaya al escribano diga que apela, dice a voces: “¡Ah señor! A él digo: ponga que apelo treinta veces” (Cristóbal de Chaves, *Segunda parte de las cosas que pasan en la cárcel de Sevilla*, 1592).

6.1. ANTECEDENTES Y ESTADO DE LA CUESTIÓN

La cárcel de Sevilla es el único de los entremeses dubitados que, al menos según las primeras comprobaciones, tiene visos de ser cervantino. Apareció publicado en la *Séptima parte de las comedias de Lope de Vega* (Madrid, 1617) y está copiado también en el folio decimotercero de un códice del siglo XVII que se conserva en la Biblioteca Colombina, al que ya tantas veces hemos hecho referencia (BC, Ms. 82-3-40). Además, gracias a un apógrafo dieciochesco de la comedia *La viuda, casada y doncella* de Lope

de Vega (BNE, Ms. 22 423), tenemos noticia de su licencia de representación a cargo de Tomás Gracián Dantisco en el año 1600:

Si el entremés de *La cárcel de Sevilla* se representare en ésta (o en otra) comedia, el sentenciado a ahorcar no saque cruz (con la cual después esgrime como daga) ni otra imagen; y en las letanías que los otros presos le dicen cuando le dan el pésame, no nombre a Nuestro Señor y María Santísima, ni ningún santo; sino el cantar sea de cosas ridículas, sin el tono de letanías, procurando siempre que en el lenguaje no se diga deshonestidad cubierta o descubiertamente, principalmente cuando las mujeres hablan con el procurador [...] En Madrid a 15 de octubre de 1600 [Citamos por la ficha de este entremés en CLEMIT <<http://buscador.clemit.es/indice.php>>].

Este documento resulta de vital importancia, pues nos permite situar su fecha de composición en los últimos ocho años del siglo XVI, mucho antes de lo que la crítica tradicional ha defendido¹⁷⁷.

¹⁷⁷ El entremés hubo de ser escrito entre 1592, año en que Cristóbal de Chaves finaliza la segunda parte de la *Relación de la cárcel de Sevilla* —de la que, como puede verse en la cita con la que abrimos el capítulo, se toma el tema central de la pieza— y el 15 de octubre de 1600, fecha en la que Tomás Gracián Dantisco rubrica su licencia de representación. La presencia de Escarramán ha hecho que muchos sitúen su fecha de composición con posterioridad a los años 1610-1612 [Asensio, 1971: 90-91], pero eso pasa por presuponer que este personaje literario nace con el romance de Quevedo “Ya está metido en la trena / tu querido Escarramán”, que circulaba de boca en boca a partir de esa fecha. Hay otros, sin embargo, que consideran que Escarramán ya estaba en el imaginario común mucho antes, pues hubo de ser un famoso ladrón sevillano condenado a galeras a finales del XVI [Cotarelo y Mori, 1911]. Para Di Pinto, que ha estudiado la tradición escarramanesca en el teatro del Siglo de Oro, no hay datos objetivos sobre la existencia de Escarramán como persona real, pues, además, todos los archivos de la Audiencia sevillana en los que poder documentar su rastro se han quemado; no obstante, especula con la posibilidad de que Cervantes pudiera conocerlo en la cárcel sevillana: “El nombre de Escarramán siempre ha estado vinculado en romances y jácaras al de otro malhechor, Maladros, cuyo testamento en el *Romance del Testamento de Maladros*, presenta al final la fecha del 27 de mayo de 1570 y está firmado en la cárcel de Sevilla. ¿Cabe suponer que alrededor de esa fecha Escarramán no estuviera muy lejos de su compañero Maladros? ¿O precisamente que él estaba libre? Es incluso posible que escarramán compartiera ‘estancia’ con Cervantes en la cárcel de Sevilla. De hecho, Cervantes hace hablar a Escarramán en *El rufián viudo* de una manera muy especial: ‘Cobré mi libertad y ya soy mío. / Hice voto y promesa inviolable / de no mudar de ropa ni de carga / hasta colgar la de los muros santos / de una devota ermita, que en mi tierra / llaman de San Millán de la Cogolla’; es harto curioso lo que sugiere Cervantes, pudiera ser que, de haber sido compañeros de prisión, tuviera ‘información privilegiada’ y por lo tanto supiera que escarramán era del norte de España, riojano, aunque ‘ejerciera’ en Sevilla” [2005: 22-23]. Pedrosa, en la reseña que hace al monográfico de Di Pinto, señala otro dato interesante: el propio nombre (o apodo) de “ESCARRAMÁN” es lingüísticamente cercano a “SCARAMouche” o “SCARAMuccia” [2006: 326]. Sea como fuere, lo cierto es que la licencia de representación de Tomás Gracián Dantisco permite acotar la fecha de composición a los años 1592-1600.

La Cárcel Real de Sevilla no sólo tiene relevancia histórica, sino también literaria. En aquel edificio de la actual calle Sierpes estuvieron presos Mateo Alemán o Miguel de Cervantes, quien parece que esbozó allí los primeros capítulos de su inmortal *Quijote*; y son muchos los que se animaron a poner por escrito lo que ocurría entre aquellas cuatro paredes, desde los ya citados escritores hasta Alonso de Morgado (*Historia de Sevilla*), Agustín de Rojas Villandrando (*El viaje entretenido*), Cristóbal de Chaves (*Relación de las cosas de la cárcel de Sevilla y su trato y Ensanche de la segunda parte de las cosas que pasan en la cárcel*), el padre Pedro de León (*Compendio de las cosas tocantes al ministerio de las cárceles*) o los autores anónimos de la *Tercera parte de las cosas de la cárcel de Sevilla* y del *Entremés de la cárcel de Sevilla*, cuya atribución a Cervantes estudiamos en este capítulo.

Desde que Aureliano Fernández Guerra considerara que el *Entremés de la cárcel de Sevilla* era de Cervantes [1863: 1371], fueron muchos los que recogieron su testigo, basándose en su proximidad temática con la tercera parte de la *Relación* y su cercanía editorial con *El entremés de los habladores* (a quien sigue tanto en la versión impresa como en la manuscrita), sin importarles que no hubiese ninguna certeza de la autoría cervantina de esas dos obras [Asensio, 1902; Castro, 1874; Menéndez Pelayo, 1941]¹⁷⁸.

Poco después, esta teoría cayó en descrédito. Emilio Cotarelo y Mori defendió que, al no haber pruebas, “tal atribución debe tenerse por no hecha” [1911: LXV], al igual que Armando Cotarelo Valledor [1915], Adolfo Bonilla y San Martín [1916] y Dámaso Alonso [1936]. Agustín del Campo dijo no ver en ella nada de Cervantes, a pesar de que retratara, al igual que *El rufián viudo*, los bajos fondos de la sociedad [1948: 45]. Y tampoco apoyaron la atribución cervantina ni Eugenio Asensio [1970 y 1071], ni José Manuel Blecua [1973], ni César Hernández Alonso y Beatriz Sanz

¹⁷⁸ Otros, como Sánchez-Arjona [1989: 99] o Rodríguez Marín [1905: 202-208], apuntaron la autoría de Cristóbal de Chaves.

Alonso [1999]¹⁷⁹, ni Canavaggio [2003], ni Alfredo Baras Escolá [2012], ni Adrián J. Sáez [2020].

Aunque la tónica general desde principios del siglo XX ha sido la de desvincularlo del autor del *Quijote*, Alfredo Rodríguez López-Vázquez volvió a defender hace unos años la paternidad cervantina [2013a], con un método que, en nuestra opinión, y como ya hemos comentado en algún otro lugar, no puede considerarse determinante:

extracción de un repertorio léxico caracterizado por la baja frecuencia de uso en el CORDE, cotejo entre un elenco de autores posibles y verificación ulterior del resultado a partir de la obra de un autor que no haya sido incluido en el elenco inicial [2013a: 11].

De este modo, tras analizar un total de 12 vocablos o sintagmas en el CORDE (siete característicos del léxico de germanía y otros cinco no argóticos)¹⁸⁰, concluye que “es de Cervantes sin la menor duda”, ya que “su estancia en la en la cárcel sevillana y la coincidencia del léxico del hampa con una notable cantidad de obras cervantinas lo deja bastante claro” [2013a: 14]; una opinión controvertida si tenemos en cuenta que en el CORDE están recogidas todas las obras de Cervantes, pero no otras muchas en las que se refleja ese vocabulario de germanía al que se refiere. Vamos a ejemplificar esta limitación metodológica con uno de los términos que propone:

Guzpátaro/s. El glosario de Juan Hidalgo registra ‘guzpatero’, con el valor de ‘agujero’ y ‘guzpatatero’ con el valor de ‘Ladrón que horada las paredes’, lo que llamamos un butronero. El CORDE, además de la cita de este entremés sólo registra su uso en dos autores, Mateo Alemán y Miguel de Cervantes, en *Rinconete* y *Cortadillo*, donde, además, se glosa en el texto el significado del vocablo: “guzpátaros –que son agujeros–”. Dado que de esta novela ejemplar ya circulan copias manuscritas en 1605, parece claro que Cervantes no necesitaba aguardar a la publicación del *Romancero* de Juan Hidalgo para estar informado, por experiencia propia, del significado del término [2013a: 12].

¹⁷⁹ Estos dos autores, como explicamos en el apartado 4.3.3., creen que el “presunto autor” es Agustín de Rojas, autor del *Viaje entretenido* [1999: 402].

¹⁸⁰ Este “micro-repertorio”, como él mismo lo define, se compone de las siguientes unidades léxicas: “guzpátaro/s”, “trainel/es”, “trato airado”, “la de ganchos”, “aguilucho/s”, “leva/s”, “murcio”, “calles acostumbradas”, “gente desalmada”, “remediadore/s”, “más de la marca”, “borracherías”. Rodríguez López-Vázquez no analiza en ningún caso secuencias de 5 o más ítems, a pesar de ser las más significativas para discriminar autorías.

Al sólo tener en cuenta el CORDE, Rodríguez López Vázquez presume que *Rinconete y Cortadillo* es la primera obra en la que se utiliza “guzpátaro/s”, pues ésta circulaba manuscrita antes de la publicación de los *Romances de germanía de varios autores, con el vocabulario al cabo por orden del a. b. c. para declaración de sus términos y lengua* de Juan Hidalgo, en 1609. Ahora bien, lo cierto es que este vocablo, propio de la germanía, aparece ya en otros testimonios anteriores, como el romance anónimo “De Toledo sale el jaque” (no recogido en el CORDE) o la propia *Relación de la cárcel de Sevilla* de Chaves, en la que se inspira este entremés (testimonio que sí que está recogido en el CORDE, pero que Rodríguez López-Vázquez no detecta al aparecer con la variante gráfica “guspátaro”)¹⁸¹. También aparece en los capítulos del *Compendio* que el padre Pedro de León dedica a la cárcel de Sevilla, y que debió redactar en torno a 1616.

El método de Rodríguez López Vázquez puede arrojar ciertas pistas, pero nunca puede ser válido por sí sólo para discriminar autorías¹⁸². En primer lugar, porque una atribución rigurosa debe contemplar diferentes variables de análisis, de distintos niveles de la lengua, y responder al principio de certeza, combinando métodos cualitativos y cuantitativos; en segundo lugar, porque el CORDE es un repositorio limitado en el que no están representados todos los autores por igual; y, en tercer lugar, porque al no estar unificados los criterios de transcripción de los textos, puede ocurrir, como en el caso de *guzpátaro*, que no se localicen ciertas formas de uso.

Tampoco nos parece acertado el método utilizado por Peña Muñoz, basado exclusivamente en la biografía del autor del *Quijote*:

¹⁸¹ Además, según afirma el propio Chaves al final de la segunda parte de su *Relación*, tenía un vocabulario de germanía escrito de su puño y letra, por lo que es más que probable que este término también estuviera recogido en él. Esta afirmación lleva a Hernández Alonso y Sanz Alonso a presuponer que bajo el pseudónimo de Juan Hidalgo se esconde en realidad el procurador de la Audiencia sevillana [1999: 219, 304 y 399-400].

¹⁸² Este tipo de análisis son un claro ejemplo de la falacia de prueba incompleta, conocida en inglés como *cherry picking* (Eder, 2013). Esta falacia hace referencia a la elección (a veces inconsciente) de los resultados más “acordes” con la hipótesis.

Resumiendo, extraemos de lo dicho como conclusión que el entremés se puede atribuir con justicia a Cervantes a partir de su propia experiencia de la cárcel de Sevilla entre 1597 y 1598. Autobiografía y autoría están entrelazadas [2010].

Como veremos en el siguiente apartado, que un autor pueda estar en un determinado lugar y conocer a unas determinadas personas o unos determinados documentos pueden servir para proponerlo como candidato, pero nunca para afirmar su autoría. Por todos estos motivos, consideramos que aún no se ha hecho un estudio de atribución de autoría que permita establecer, de manera fiable, qué probabilidad existe de que Cervantes sea o no el autor de este entremés.

6.2. ¿PUDO CERVANTES ESCRIBIR ESTE ENTREMÉS?

En un estudio de lo que tradicionalmente se ha llamado “atribución de autoría” [Turell, 2011], una de las preguntas que debemos hacernos antes de realizar el análisis lingüístico comparado es si realmente esa persona pudo escribir ese texto. *El entremés de la cárcel de Sevilla* permite inferir lo siguiente de su autor: que hubo de ser un escritor en activo en la última década del siglo XVI, con experiencia teatral, que conocía el lenguaje de germanía y que tuvo acceso directo a las dos partes de la *Relación* de Cristóbal de Chaves (escritas entre 1591 y 1592), así como a alguno de los romances que luego verán la luz en la obra de Juan Hidalgo (1609)¹⁸³.

Cervantes cumple todos esos requisitos. Autor teatral desde la década de los 80, sabemos que intentó retomar la escritura dramática en los 90, como queda patente en el contrato que firma con Rodrigo Osorio en 1592. También sabemos que entre 1587 y 1600 o 1601 estuvo vinculado profesionalmente a la provincia de Sevilla, donde ejerció como comisario de abastos y recaudador de impuestos, y que en 1597 pasó unos meses

¹⁸³ “El seguimiento fiel del texto de la *Relación de la cárcel de Sevilla* y el casi calco del mencionado romance nos dice claramente que quien escribió este entremés, cuanto menos, tenía delante los textos de la *Cárcel* y del romance; y conocía bien el mundo de la germanesca” [Hernández Alonso y Sanz Alonso, 1999: 399].

encarcelado en la capital sevillana, donde pudo conocer a Cristóbal de Chaves, pues era “el procurador de la Audiencia más solicitado y que a mayor número de presos defendía” [Hernández Alonso y Sanz Alonso, 1999: 219]. Además, queda patente su conocimiento del mundo del hampa sevillano y del léxico de germanía en muchas de las obras que compuso, como *El rufián dichoso*, *Pedro de Urdemalas*, *Rinconete y Cortadillo*, *La ilustre fregona*, *El rufián viudo llamado Trampagos*, *El coloquio de los perros* o el *Quijote*, a través del personaje de Ginés de Pasamonte.

Como vemos, la autobiografía posibilita pensar en Cervantes, pero también en los otros dos autores propuestos por la crítica: Cristóbal de Chaves y Agustín de Rojas Villandrando. El primero conoció todo lo referente a la Real Cárcel de Sevilla y al léxico de los hampones gracias a su trabajo como procurador, como se refleja en las dos partes de su *Relación* y en el vocabulario de Juan Hidalgo, que algunos le atribuyen (véase nota 181); asimismo, sabemos que recibió 20 ducados por un entremés que escribió para las representaciones sevillanas del Corpus de 1598 [Sánchez-Arjona, 1898: 98]. El segundo, al que en Sevilla se le conocía como “el caballero del Milagro”, fue ladrón y formó parte del hampa andaluza durante un tiempo; además, participó en varias compañías de renombre (como las de Antonio de Villegas o Nicolás de los Ríos) y dejó cuenta de sus conocimientos teatrales en *El viaje entretenido*, publicado en 1603 [Sánchez-Arjona, 1898: 105; Hernández Alonso y Sanz Alonso, 1999: 400-401]. Ante estos hechos, el estilo idiolectal se convierte en la única prueba fiable para determinar autorías.

6.3. ANÁLISIS DE AUTORÍA BASADO EN LA COMPARACIÓN FORENSE DE TEXTOS

Siguiendo el protocolo metodológico de los estudios de autoría de la Lingüística forense, realizaremos el análisis comparado de una serie de variables de todos los niveles de la lengua, desde el fónico/gráfico hasta el textual o discursivo, a fin de determinar si el *Entremés de la cárcel de Sevilla* es o no de Cervantes (y con que grado de probabilidad). Aunque nos centraremos esencialmente en la comparación entre la

pieza dubitada y las indubitadas cervantinas, en ciertas ocasiones utilizaremos, como hemos explicado ya, textos de otros autores, a fin de poder determinar si las posibles coincidencias son marcas propias e idiosincrásicas de autoría o si, por el contrario, responden a usos comunes de la lengua de la época o del género literario¹⁸⁴. Los análisis que vamos a realizar son los siguientes¹⁸⁵:

- Tema, argumento, personajes, segmentación y estructura
- Análisis cualitativo de marcas idiosincrásicas
- Distribución de la frecuencia de los 5-gramas de caracteres más empleados en los textos (*character 5-grams*)
- Distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*)
- Análisis de componentes principales (PCA) aplicado a la *Delta distance*
- Clasificación supervisada por aprendizaje automático (Classify)
- Porcentaje de similitud textual (*similarity*) y porcentajes de vocabulario compartido una sola vez (*Once only in both*), de vocabulario compartido más de una vez (*Shared > 1*) y de vocabulario único y exclusivo (*Only in*)
- Clasificación secuencial de aprendizaje automático (Rolling classify)
- Verificación de autoría con General Imposters (GI)
- Longitud de secuencias de palabras idénticas (*verbatim*)
- Coocurrencias y redes léxicas
- Estudio de las palabras de función: orden de frecuencia, frecuencia relativa y PCA

¹⁸⁴ Para todo lo referente a la delimitación del corpus de análisis y el tratamiento de los textos, véase el apartado 5.3.1.

¹⁸⁵ No son, desde luego, todas las variables analizables, pero sí las más fiables para este caso en concreto (y, en nuestra opinión, suficientes para poder establecer conclusiones robustas sobre la posible autoría cervantina del *Entremés de la cárcel de Sevilla*).

- Ley de Zipf
- Función tf-idf
- Frecuencia de clases de palabras y bigramas y trigramas de secuencias morfosintácticas mediante etiquetadores gramaticales automáticos (POS tagging)

No hemos contemplado aquí el análisis de riqueza léxica, ni de la ratio *type / token* de palabras de contenido, pues no alcanza en lengua española la fiabilidad de otros parámetros. Cabe destacar, no obstante, que los usaremos en el estudio de autoría del *Entremés de los romances*, en donde la extensión del texto original apenas supera las 1000 palabras y en donde, además, deberemos prescindir del análisis de bigramas y trigramas de clases de palabras, ya que éstos pueden estar condicionados por las exigencias del metro y la rima. Tampoco utilizaremos ninguna variable relacionada con la puntuación, como la frecuencia relativa de los principales signos sobre el total de caracteres o el promedio de palabras por frase, a pesar de su operatividad [Blasco y Ruiz Urbón, 2009: 42; De la Rosa y Suárez, 2016: 410], ya que, como hemos explicado anteriormente, los signos de puntuación que aparecen en los textos conservados son marcas personales de los copistas, impresores o editores, pero no de sus autores primigenios¹⁸⁶.

Asimismo, es importante recordar que la similitud o disimilitud de un solo rasgo o factor no servirá en modo alguno para afirmar o negar una atribución, sino que se precisa de la acumulación de varios rasgos o factores (cualitativos y cuantitativos) para poder establecer, con mayor o menor grado de probabilidad, un comportamiento análogo o dispar y, en consecuencia, sostener o rechazar dicha autoría.

¹⁸⁶ Sólo podríamos usar estas variables si se conservaran los manuscritos autógrafos.

6.3.1. Tema, argumento, personajes, segmentación y estructura

6.3.1.1. Tema central y argumento

El entremés en prosa de *La cárcel de Sevilla* muestra, a través de una serie de personajes absolutamente hilarantes, el mundo de la germanesca y del hampa sevillano. Tras un comienzo cantado, la acción del entremés se centra en Paisano, uno de los presos, que recibe su sentencia de muerte mientras está jugando a las cartas con Barragán, Solapo y Garay. Paisano, tras decirle al escribano que apela “treinta veces”, le pide al alcaide que le lleve a la enfermería para ponerle el hábito de la Caridad, exigiéndole que no esté viejo y apolillado, como lo estaba el del último ahorcado, pues él quiere “salir como hombre honrado”; también pide que le digan letanías y que avisen a su *amiga* Beltrana, “a ver si tiene remedio esta desgracia”. Cuando los presos empiezan a loar al condenado, irrumpen en escena Beltrana y Torbellina, que entre lloros y sollozos se despiden de Paisano, ya vestido de condenado. Paisano quiere hacer a Beltrana heredera de los “bienes, muebles y raíces” de su calabozo y *acomodarla* con Solapo, pero pronto descubre que ésta ya ha dado su “palabra a otro”. Los presos, entre los que se encuentra Escarramán, cantan letanías a Paisano, que, justo cuando va a comenzar su procesión por la cárcel, recibe la noticia de que han aceptado su apelación. Paisano pide que “se regocijen y alegren” y todos cantan y bailan.

Como han estudiado Hernández Alonso y Sanz Alonso, la letrilla cantada con la que se abre el entremés se corresponde con un romance de germanía del siglo XVI titulado “Alta mar esquiva”, aunque con ligeras variantes¹⁸⁷, y la canción que lo cierra “recuerda muy de cerca a uno de los *Romances de germanía* de Juan Hidalgo” [1999: 403]. Pero más allá de la inclusión de romances y canciones conocidas y reconocibles por el público, como era lo habitual en el género, este entremés guarda muchas

¹⁸⁷ “Alta mar esquiva” apareció publicado por primera vez en el *Quinto cuaderno de varios romances* (Valencia, 1598); se trata de una endecha de 66 versos, dividida en cinco estrofas desiguales de 14, 14, 10, 10 y 18 versos, puesta en boca de un único prisionero. La versión del entremés es más breve, pues consta de 52 versos divididos en tres estrofas, de 24, 12 y 16 versos, cada una de las cuales es entonada por un personaje distinto [Jones, 1976: 224].

similitudes temático-argumentales con las dos primeras partes de la *Relación de la cárcel de Sevilla*, de las que toma no sólo el motivo central, sino también el nombre de alguno de los personajes y ciertas expresiones y situaciones:

DUB_EP_CARCEL	CHAV_R_CARCEL
<p>Hablan las personas siguientes: GARAY SOLAPO</p>	<p>Llámanse de ordinario los que sirven de limpiar y lo demás coplilla, venturilla, trapaña y mojarrilla, cambalosos y jamones; y los valientes a quien se acude con el provecho el Paisano, Barragán, Maladros, Pecho de acero, Garay, y otros nombres que acuden al oficio y ánimo dellos (Chav_R_carcel1).</p>
<p>PAISANO ALCAIDE COPLILLA, pícaro</p>	<p>A esto estaba otro junto a el que había de morir, a quien preguntó paso que si le había cargado Barragán en decir que la justicia le podía dar pesadumbre (Chav_R_carcel1).</p>
<p>BARRAGÁN UN ESCRIBANO TORBELLINA BELTRANA UN PROCURADOR</p>	<p>Dijo otro: “una muerte había vuesacerced de morir: ¡Bienaventurado el que muere por la justicia! De la señora Beltrana no lleve vuesacerced cuidado; que aquí quedo yo, y nadie la dará pesadumbre” (Chav_R_carcel1).</p>
<p>DOS MÚSICOS</p>	<p>He sabido que mientras cumplo el tiempo de galeras te has acomodado con el Paisano, hombre desflorado, a quien los demás no sólo no respetan, pero aún le quitan lo que tú le das (Chav_R_carcel2).</p>
<p>GARAY. Abre aquí, alcaide, que nos comen chinchas.</p>	<p>Y hay tres pícaros en cada uno, que los dos limpian las paredes de las chinchas, raen el suelo, espulgan las mantas, vacían los servicios; y el otro enciende las luces; y si es verano, hacen aire toda la noche a los germanes (Chav_R_carcel1).</p> <p>hirviendo de piojos y chinchas, que hacían nido en él como si fuera pared y sudando de calor, por el aposento en que estaba (Chav_R_carcel1).</p>
<p>ALCAIDE. [...] Por vida del rey, que he de pasar alguno a la otra cárcel, o que ha de dormir en el cepo.</p>	<p>Y demás del castigo que llevan del sotaalcaide los que quebrantan esto, el mayor que sienten es que luego dé petición a los señores alcaldes diciendo e informando que son incorregibles y que para la quietud de la cárcel conviene pasarlos a la del Audiencia o de Hermandad, porque luego se provee. Y lo sienten mucho porque, en pasándolo, luego es preso nuevo en la otra y no habla palabra hasta que sea antiguo, por manera que con esto pierde la antigüedad (Chav_R_carcel2).</p>

ALCAIDE. Paisano, aquí os vienen a notificar una sentencia; pésame que es de muerte.

ESCRIBANO. Oíd, hermano, lo que os quiero notificar.

PAISANO. Barahe voacé, y quite esos encuentros.

ESCRIBANO. ¿Oye lo que le digo, hermano?

PAISANO. Aguarde voacé, que más me va en esto que en esotro.

Y ha habido hombre que estando jugando a los naipes le han notificado sentencia de muerte y que se confiese, y ha respondido que le dejen ver su suerte; y tornándole le a decir que mire que le notifican aquello, ha respondido a el escribano que haga su oficio y no pase de ahí: “Mire que me enojaré” (Chav_R_carcel2).

PAISANO. ¿Quién dio esta sentencia?

ESCRIBANO. El juez que entiende de vuestra causa.

PAISANO. Puédelo hacer, que es mi juez. Mas dígame voacé que sea tan honrado que nos veamos en el campo solos, él con su fallo y yo con una espada de siete palmos; veamos quién mata. Estos juecicos en tiniendo un hombre embanastado como besugo, luego le fallan, como espada de la maesa: “¡Fallo que debo de condenar, y condeno, que sea sacado por las calles acostumbradas en un asno de albarda...” ¡Que todo lo diga! ¡Válgate el diablo sentencia de pepitoria! ¿No es mejor decir que muera este hombre y ahorrar de tanta guarnición?

Otros, que muy en su juicio responden a el escribano cuando les hace semejantes notificaciones: “¿Quién dio esta sentencia?” Y diciéndole que el alcalde de la Justicia, o el teniente, ha respondido: “Puédelo hacer como juez; pero sea él tan honrado, que con una espada en la mano salga a reñir conmigo, Y veremos quién mata quién”. Y saliendo el escribano santiguándose de desemejante disparate y atrevimiento, torna a la baraja a decir: “Digo mi parte”. Y por qué algunos bachilleres presos le aconsejan que antes que se vaya al escribano diga que apela, dice a voces: “¡Ah señor! A él digo: ponga que apelo treinta veces”. Y diciendo el escribano que para quién apela, responde: “Apelo para Dios y ¿qué se yo? Digo que apelo para esos señores padres de la Audiencia”, diciendo por los alcaides. Y luego queda diciendo él y sus camaradas, por el escribano: “¡Mire con que venía el señor escribano! Vaya con Dios, que ahí se remediará; que no queremos esa sentencia, ni sabemos qué es, ni la oímos” (Chav_R_carcel2).

ESCRIBANO. ¡Por Dios, que estoy por ponello así, visto tanta desvergüenza!

ALCAIDE. Váyase vuesa merced, señor escribano, y no haga caso desta gente desalmada.

GARAY. Señor Paisano, llámele voacé, y dígame que apela.

PAISANO. A él digo: ¡ah seor escribano!, venga acá voacé.

ESCRIBANO. ¿Qué queréis, hermano?

¿Cómo se va voacé después que queda un hombre cargado hasta las entrañas? Ponga ahí voacé que apelo treinta veces.

PAISANO. Con una basta. ¿Y para quién diremos que apeláis?

ESCRIBANO. Apelo para Dios; que si yo apelo para esos señores padres de la audiencia, remediadores de los fallos, pienso que no tendré ningún remedio.

ALCAIDE. Hermano, esto va muy de rota; el escribano me ha notificado que os suba a la enfermería, y que os ponga el hábito de la Caridad.

Y habiéndolos puesto en la enfermería, lugar común para todos los que han de morir, porque allí los ponen junto a un altar y dos bancos, donde se sientan juntos dos padres confesores y los visitan tres días que dura la confesión y comunión, conforme a el estilo tan piadoso que esta ciudad tiene (Chav_R_carcel2).

PAISANO. Pues, señor alcaide, voacé me haga merced de que no se me ponga el hábito de la Caridad que sacó el ahorcado del otro día, que estaba viejo y apollado, y no me lo he de poner por ninguna cosa: que ya que haya de salir, quiero salir como hombre honrado y no hecho un pícaro; que antes me quedaré en la cárcel.

Acabada la cena, entró la persona a cuyo cargo está poner los hábitos blancos de la Caridad. Y acertó a cabelle al postrero un hábito no tan bueno ni tan a gusto como él quisiera; y habiendo lo mirado se lo quiso quitar, jurando a Dios de no llevarle, si no le daban otro; y dándole una caperuza vieja, la echó por ahí diciendo que votaba a Dios si no le daban otra, de no llevarla, que bastaba que llevaba el hábito (Chav_R_carcel2).

(Salen Torbellina y Beltrana, mujeres de la casa, con mantos doblados y mandiles blancos, y su Procurador con ellas).

[...]

BELTRANA. ¡Ay, hermana! ¿Qué es esto? ¡Jesús, que me muero! (Desmábase).

TORBELLINA. Téngala, seor procurador; mire que se ha desmayado.

[...]

(Sale el Paisano, vestido de ahorcado, y una cruz en la mano, y el Alcaide con él).

[...]

BELTRANA. ¡Ay, sentenciado de mis ojos! ¿Qué es esto?

[...]

BELTRANA. ¡Ay, sentenciado de mi ánima y de mi vida! (Llora).

Cuando se sabe en la mancebía o en la casa de la mujer que tiene por amiga el que ha de morir, viene acompañada de otras semejantes a la prisión, puesta de duelo; y a voces, como si fuera su marido, dice: “¡Afuera! no me detenga a nadie. ¿Dónde está el sentenciado de mi anima?” Y antes de llegar al aposento se desmaya en los brazos de veinte bergantes, que unos dicen que no la dejen entrar, y otros que sí (Chav_R_carcel1).

PAISANO. Beltrana, no me digas nada. El alma te encargo, pues el cuerpo te ha servido en tantas ocasiones; y una de tus amigas (no lo hagas tú por el escándalo que puede haber), cuando estuviere ahorcado, me limpiará el rostro, porque no quede feo como otros probetos. Y me traerás un cuello almidonado y más de la marca, y abierto con bolo, y puntas y todo negocio; que quiero ver antes que deste mundo vaya quien hace esta denunciación.

Y ha acaecido el que ha de morir decir a su amiga: “Leona, encárgote el alma, pues el cuerpo te ha servido en todas las ocasiones. Conciértate con el verdugo que no me quite la camisa y calzón: y una destas Señoras, cuando esté colgado, me limpie apriesa, porque no quede feo como otros pobretes”. A esto da voces ella diciendo: “¡Hasta la muerte es limpio y pulido mi bien!” (Chav_R_carcel1).

BELTRANA. Aun hasta en la muerte fue limpio mi amor; yo apostaré que no ha habido mejor ahorcado en el mundo.

(Salgan todos los que pudieren, en orden de figurillas, con velas encendidas en las manos, y cantando las ledanías)

Y cuando hay hombre de quien hacer justicia van todos los presos con su cera cantando las letanías hasta el lugar donde está recogido el que ha de morir (Chav_R_carcel1).

CAPÍTULO 6

<p>GARAY. Hable el seor Barragán, que es más honrado y más antiguo.</p> <p>BARRAGÁN. Yo no haré: hable el seor Solapo.</p> <p>SOLAPO. Así me vea en aquella calle con libertad, que no diga palabra: hable el seor Cuatro.</p> <p>CUATRO. El Cuatro no lo hará: hable el seor Garay.</p> <p>GARAY. Garay no lo hará: no hay qué decir.</p>	<p>Pararon con esto, y por curiosidad me puse a ver el pésame y despedimiento: y los delanteros tuvieron comedimiento de los unos a los otros quién hablaría primero (Chav_R_carcel1).</p>
<p>SOLAPO. Por no perder la costumbre antigua que se tiene con los presos honrados, digo así, que en estos luctos echará de ver voacé lo que lo sienten sus camaradas. Plega a Dios lo seamos en el cielo. Y mal haya el diablo, que dos sentencias tengo de muerte; ¿por qué no vino la otra, para acompañar a voacé?</p>	<p>Habló a otro y dijo: “En estos lutos verá vuesaerced lo que sienten sus camaradas. ¡Pluguiera a Dios que lo fuéramos en el cielo! Que una sentencia me falta; y mal haya el diablo porque la otra no viene hoy, por acompañar a vuesaerced” (Chav_R_carcel1).</p>
<p>BARRAGÁN. So Paisano, consuéllese voacé con que la justicia lo hace; que otro no podía con voacé en el mundo. Y esta puede dar pesadumbre a voacé y a todo el mundo. Voacé déjelos; que... no digo nada.</p>	<p>Y dando la mano a Barragán dijo: “Cosuélese vuesaerced, señor Fulano, con que la justicia que hace y otro no pudiera en el mundo con vuesaerced, y esta puede darle pesadumbre. Y vuesaerced es honrado, y morirá como honrado” (Chav_R_carcel1).</p>
<p>PAISANO. Ninguno en socolor de amigo piense cargarme en este despedimiento. Quiero saber si es cargo lo que dijo el seor Barragán en decirme que la justicia me puede dar pesadumbre.</p>	<p>A esto estaba otro junto a el que había de morir, a quien preguntó paso que si le había cargado Barragán en decir que la justicia le podía dar pesadumbre (Chav_R_carcel1).</p>
<p>PAISANO. Esa vaya en aumento. Y pues que toma a cargo lo de los testigos, me hará merced voacé de cortar al uno las orejas y al otro las narices, y a los demás borrajales las caras con una daga; y con esto iré contento para la otra vida.</p>	<p>Otros le prometen, si Dios les da libertad, de matar a la otra parte, porque lo siguió y no quiso perdonar. Otros le certifican que harán otro tanto al que lo sopló. Otros que los testigos que dijeron contraer harán lo mesmo (Chav_R_carcel2).</p>
<p>ESCARRAMÁN. Voacé tenga la muerte como ha tenido la vida, pues ninguno se la hizo que no se la pagase.</p> <p>[...]</p> <p>ESCARRAMÁN. Y si al bajar lloren las personas, no las vuelva el rostro ni sea predicador en el sitio desta desgracia, que es hijo de vecino de Sevilla y no ha de mostrar punto de cobardía.</p>	<p>Replicó otro enlutado: “Vuesaerced tenga la muerte como ha tenido la vida, pues ninguno se la ha hecho que no se la pagase; y lleve buen ánimo; y cuando saliere, si lloren las presas no les vuelva el rostro; ni sea predicador en el sitio desta desgracia, pues es hijo de Sevilla, y no ha de mostrar punto de cobardía” (Chav_R_carcel1).</p>
<p>ESCARRAMÁN. Y al verdugo que apretó tanto las cuerdas a voacé, que le hizo decir lo que no había hecho, si yo salgo... No digo nada.</p>	<p>Y al verdugo, porque le dio el tormento muy grave estando pagado, por lo cual confesó lo que no hizo: por manera que en su muerte le traiga la memoria tantas muertes como he dicho, que parece que son cochinos que quieren asesinar (Chav_R_carcel2).</p>

Tabla 15. Comparación entre el contenido del *Entremés de la Cárcel de Sevilla* y las dos primeras partes de la *Relación de Chaves*

Como se desprende de la Tabla 15, el autor del entremés dubitado no sólo hubo de haber leído la *Relación* de Chaves, sino que también había de tenerla delante al escribir su obra, pues no pueden explicarse de otro modo los *verbatim* que se aprecian entre uno y otro texto, salvo que el entremés también fuera obra de Chaves (un aspecto en el que profundizaremos más adelante). Esta práctica —la de la reescritura creativa— era habitual en Cervantes, pero también en otros muchos dramaturgos del momento, por lo que no aporta ninguna prueba respecto de su autoría¹⁸⁸.

6.3.1.2. Personajes

La *dramatis personae* del entremés dubitado, tanto en la versión impresa como en la manuscrita, hace referencia a doce personajes (Garay, Solapo, Paisano, Alcaide, Coplilla, Barragán, un escribano, Torbellina, Beltrana, un procurador y dos músicos), aunque en el texto dialogado participan dos más (Escarramán y Cuatro), por lo que el número real asciende a catorce: dos mujeres y doce hombres (dos de ellos, los músicos). En los entremeses de Cervantes encontramos un número y distribución similar, ya que, como ha estudiado Agostini, los personajes fluctúan entre siete y catorce, con dos o tres mujeres como máximo [Agostini, 1964: 275].

La gran mayoría de los personajes que aparecen en *La cárcel de Sevilla* pertenecen a los bajos fondos (presos y prostitutas), como era habitual en el género¹⁸⁹. Al igual que en cuatro de los entremeses cervantinos (*El vizcaíno fingido*, *La cueva de Salamanca*, *La guarda cuidadosa* y *El viejo celoso*), el foco de la acción se centra en un único personaje; en este caso, Paisano, un rufián al estilo de Trampagos, que ya no es sólo el tipo fanfarrón y pretencioso del teatro de Lope de Rueda, sino un valentón con humor satírico y habla propia.

¹⁸⁸ Como recuerda Asensio en su Itinerario, “cualquier tipo literario podrá traspasar materiales al entremés, desde la fantasía satírica hasta la literatura documental” [1971: 32].

¹⁸⁹ Recordemos que el propio Lope de Vega, en su *Arte nuevo*, explicaba que la acción del entremés era “entre plebeya gente, / porque entremés de rey jamás se ha visto”.

6.3.1.3. Segmentación y estructura externa e interna

Basándonos en la clasificación que hace Asensio de los ocho entremeses de Cervantes [1970: 16-41], *La cárcel de Sevilla* sería un entremés de acción y ambiente, al estilo de *El rufián viudo*, *La guarda cuidadosa* o *El retablo de las maravillas*, en donde el desenlace de la acción viene ralentizado por la inclusión de varios personajes no funcionales, como Escarramán, que sirven para evocar el mundo del hampa y la germanía. Este tipo de entremeses están a caballo entre los de pura acción (*El vizcaíno fingido*, *La cueva de Salamanca* y *El viejo celoso*) y los estáticos o de revista de personajes (*El juez de los divorcios* y *La elección de los alcaldes de Daganzo*).

La segmentación de este entremés es bastante sencilla. Como vemos en la siguiente tabla, la pieza consta de una única secuencia, integrada por cuatro unidades de farsa: la primera es puramente ambiental y las tres últimas se corresponden con la presentación, el nudo y el desenlace.

1 secuencia → 4 unidades de farsa	<ol style="list-style-type: none"> (1) Rebelión de los presos y conversación del Alcaide con Barragán y Paisano sobre lo ocurrido. Incluye una letrilla cantada sobre la vida en prisión. (2) Comunicación de la sentencia de muerte a Paisano, su solicitud de apelación y la petición de sus voluntades finales. Esta unidad concluye con la salida de Paisano, que se dirige a la enfermería, acompañado por el alcaide, para ponerse el hábito de la caridad. (3) Despedida de las prostitutas Beltrana y Torbellina y letanías de los presos a Paisano, entre los que se incluye Escarramán. (4) Comunicación a Paisano de la aceptación de su apelación y celebración cantada y bailada.
-----------------------------------	--

Tabla 16. Segmentación del *Entremés de la Cárcel de Sevilla*

Al estudiar la estructuración, podemos centrarnos en su forma externa, es decir, en las partes que lo conforman (comienzos, movimientos escénicos, acotaciones, lugares de acción y desenlaces), o en su forma interna, que guarda relación con el elemento que

da unidad al conjunto de piezas de un autor, o, al menos, a las de un mismo género literario¹⁹⁰ [Baras Escolá, ed., 2012: 176].

Encontramos ciertas similitudes entre la estructura externa de *La cárcel de Sevilla* y la de alguno de los entremeses de Cervantes, cuyos componentes han sido estudiados minuciosamente por Agostini [1964b: 504-516]:

- El entremés comienza con el escenario vacío y, desde “adentro”, se escuchan las protestas de Garay, Solapo y Paisano, que pronto salen a escena para tañer y cantar una letrilla¹⁹¹. Aunque ninguno de los ocho entremeses de Cervantes comienza con voces desde el interior, sí aparece este recurso en *La cueva de Salamanca*, en donde se escucha la voz del estudiante “como de muy lejos” antes de aparecer en escena, y en *El viejo celoso*, cuando llega la Justicia.
- En relación con el movimiento escénico, hemos contabilizado un total de trece entradas y salidas de personajes, lo que deja patente el dinamismo de la pieza (dinamismo propio de todos los entremeses de Cervantes, a excepción de *El juez de los divorcios* y *La elección de los alcaldes de Daganzo*, que, recordemos, son los dos entremeses de revista de personajes)¹⁹². El autor utiliza la técnica del zigzagueo, que es la misma que aparece en *El rufián viudo* y *La guarda cuidadosa*, para terminar con la acumulación de todos los personajes en escena. Es importante advertir, en favor de la teatralidad de la pieza, que el escenario no queda vacío en ningún momento.

¹⁹⁰ Pensemos, por ejemplo, en la dedicatoria del *Quijote* de 1605, que es una reelaboración de la que antepuso Fernando de Herrera a las obras de Garcilaso (con algunos fragmentos, además, del prólogo de Medina), o en *La española inglesa*, en donde, según Zimic, se establece un diálogo intertextual con el *Amadís de Gaula*. Son muchos, además, los que han establecido que el *Quijote* o el *Persiles* son un compendio de textos de muy diversa índole, desde novelas de caballerías o pastoriles hasta libros de viajes y tratados de cartografía.

¹⁹¹ Las quejas y súplicas que dan comienzo a la obra recuerdan, aunque muy sucintamente, a las que pronuncia Doña Lorenza al comienzo del *Entremés del viejo celoso*. Mientras Garay, Solapo y Paisano lamentan la prisión física, la mujer del viejo celoso se queja de su prisión matrimonial.

¹⁹² Hay un pequeño error en la parte final, tanto en la versión impresa como en la manuscrita, pues se anuncia la entrada del Alcaide cuando no se ha especificado antes su salida.

- La gran mayoría de las acotaciones sirven para marcar la salida y entrada de los personajes (“sale(n)”, “va(n)se”, etc.) o sus acciones (“háblale al oído”, “da una castañeta”, “pónense a jugar”, “lee”, “desmáyase”, “grita”, “tañen, cantan y bailan”). En alguna otra ocasión, como vemos también en Cervantes, informan sobre el estado anímico de los personajes (“llora”), su indumentaria (“con grillos en los pies”, “con la ropilla de Solapo”, “con mantos doblados y mantillos blancos”, “vestido de ahorcado”) o los objetos que portan (“guitarras”, “almohadas y cuchillos”, “una cruz en la mano”, “velas encendidas en las manos”)¹⁹³. En contadas ocasiones, las acotaciones no tienen valor espectacular, algo que también sucede excepcionalmente en los entremeses cervantinos: una de ellas sirve para catalogar al personaje (“sale Coplilla, pícaro”) y otra adelanta una información que aún no se ha explicado en el texto (“sale Garay, con la ropilla de Solapo, que se la ha ganado”).
- Todo el entremés transcurre en un mismo lugar (el interior de la prisión sevillana), como hace Cervantes en cuatro de sus ocho entremeses: *El juez de los divorcios*, *El rufián viudo*, *La guarda cuidadosa* y *La entretenida*¹⁹⁴.
- El entremés finaliza con música, canto y baile, como es habitual en Cervantes y en todos los entremesistas del momento:

Cantos o música o ambos habrá en *El juez*, *El vizcaíno*, *La guarda*, *La cueva*, *El retablo* y *El viejo celoso*. Pero donde realmente se baila y se canta es en la *Elección de los alcaldes*, en *El rufián viudo* y en el entremés de *La entretenida* [Agostini, 1964b: 488].

¹⁹³ En dos de las acotaciones del entremés dubitado se especifica que el personaje lleva el objeto “en la mano”; así aparece también en el entremés de *La guarda cuidadosa* (“con unas chinelas pequeñas nuevas en la mano”) o en las comedias de *El rufián dichoso* (“trae el rosario en la mano” y “con un crucifijo en la mano”), *La gran sultana* (“una velilla de cera encendida en la mano” y “con un rosario en la mano”), *El laberinto de amor* (“en la mano un manto”) y *La entretenida* (“trae un papel en la mano”, “con un billete en la mano”).

¹⁹⁴ Este aspecto guarda relación con la unidad de lugar, que veremos más adelante, pues Cervantes sólo vacía los escenarios cuando debe cambiar el lugar de acción (*El vizcaíno fingido*, *El retablo de las maravillas*, *La cueva de Salamanca* y *El viejo celoso*) [Agostini, 1964b: 514].

Centrándonos ahora en la estructura interna, que estaría relacionada con la acción que da unidad al conjunto, coincidimos con Pérez de León en que “existen dificultades para encontrar un elemento unitario que aborde los entremeses cervantinos como obras que se relacionan entre sí” [2005: 93]. Para él, finalmente, este elemento cohesivo sería la teatralización de unos personajes desequilibrados, desviados por su naturaleza, edad o condición, en relación con las teorías médico-filosóficas de Huarte de San Juan:

En los ochos entremeses se utilizan protagonistas destemplados en situaciones límites que la máquina social del resto de personajes pretende convertir en un estado más o menos “templado” o armonioso hacia la conclusión de estas obras. Sin embargo, a riesgo de que el entremés pierda verosimilitud, nunca se presentan individuos completamente equilibrados o finales absolutamente resueltos [2005: 109].

La explicación de Pérez de León es perfectamente aplicable a *La cárcel de Sevilla*, en donde tenemos un protagonista absolutamente destemplado (Paisano), que se enfrenta a una situación límite (pena de muerte) que no queda del todo resuelta (se acepta la apelación, pero no hay absolución). La máquina social estaría representada por el alcaide, el escribano y el procurador.

Aunque el análisis de contenido permite pensar que *La cárcel de Sevilla* pueda ser de Cervantes, no permite tampoco descartar a otros muchos autores del momento, ya que los temas, argumentos, estructuras y personajes del género seguían —a pesar de no haber una preceptiva al respecto— unas características bien definidas y marcadas. Por este motivo, serán los siguientes análisis, relacionados con el idiolecto del texto, los que permitan establecer conclusiones fiables sobre su autoría.

6.3.2. Análisis cualitativo de marcas idiosincrásicas

En este apartado, vamos a realizar un análisis cualitativo de la presencia o ausencia en el texto dubitado de marcas características del idiolecto de los entremeses cervantinos. Para no caer en el *cherry picking*, hemos seleccionado esas marcas —que contemplan distintos niveles del lenguaje— teniendo en cuenta los trabajos de Agostini [1964b], Rosenblat [1978], Gutiérrez Cuadrado [1988], Ariza Viguera, [1998] y

Maestro [1998] y Baras Escolá [ed., 2012]: presencia del superlativo en *-ísimo*, casos de leísmo, utilización de adverbios en *-mente*, uso del diminutivo en *-ico*, empleo de reduplicaciones, anáforas y repeticiones diseminadas, uso frecuente de figuras de dicción, creación de palabras nuevas, utilización de expresiones latinas, predilección por los sinónimos coordinados, utilización pragmática de los diálogos de enunciados interrogativos y adaptación del lenguaje al nivel sociocultural de los personajes.

En ningún caso, la presencia o ausencia de una de estas marcas en *La cárcel de Sevilla* servirá para afirmar o negar la autoría cervantina. Tan sólo la acumulación de un alto número de ellas permitirá reforzar o debilitar nuestra hipótesis de trabajo.

6.3.2.1. Presencia del superlativo en *-ísimo*

Los superlativos indican el mayor grado que puede alcanzar una determinada cualidad, ya sea respecto a su relación con los otros (relativos) o respecto a sí mismos (absolutos). Las construcciones en modo superlativo son muy habituales en los textos cervantinos, sobre todo las del absoluto analítico en *-ísimo*, lo que las convierte en un rasgo muy idiosincrásico de la lengua de este autor, pues, a pesar de que desde el siglo XVI su uso se va haciendo cada vez más frecuente por influencia del latín y del italiano, “en la época del *Quijote* se sentía todavía como literario, y quizá a veces como afectado” [Rosenblat, 1978: 190].

Gutiérrez Cuadrado, al estudiar los rasgos generales de la lengua del *Quijote*, advierte el empleo relativamente frecuente de este superlativo en Cervantes, a pesar de que aún era “considerado un rasgo culto en el siglo XVI” [1998]. Y Rosenblat señala, además, que su aplicación a sustantivos “seguramente debía ser cómico” [1978: 190].

Al igual que ocurre en el *Quijote*, los entremeses de Cervantes presentan cierta predilección (aunque no muy alta) a utilizar la forma analítica del adjetivo en grado superlativo, que no aparece, sin embargo, en *La cárcel de Sevilla*.

SUPERLATIVO ABSOLUTO EN -ÍSIMO ¹⁹⁵		
	Casos	N total
Cerv_EP_juez	sanísimo; *infinitísimos; *bastantísimamente	3
Cerv_EP_guarda	bonísima	1
Cerv_EP_vizcaino	bonísima (2); *discretísimamente	3
Cerv_EP_retablo	dignísimo; valentísimo; grandísimos	3
Cerv_EP_cueva	grandísimo	1
Cerv_EP_viejo	bonísimo	1
Cerv_EV_rufian	clarísima	1
Cerv_EV_eleccion	benditísimo; rarísima	2
Dud_EP_carcel	∅	0

Tabla 17. Estudio del superlativo absoluto *-ísimo*

6.3.2.2. Casos de leísmo

En los entremeses cervantinos, como en muchas otras obras de Cervantes, encontramos numerosos casos de leísmo de persona (“le lleve”, “le honre”) y de cosa (“me le prestaban”, “me le querían cubrir”). Esta marca, habitual en autores del centro peninsular, también se da en *La cárcel de Sevilla*: “había de meter un escrito, y agora le mete, agora le saca”, “ha de soterrar alguno algún puñal que no se le saque del cuerpo otro que Dios” o “¡qué a mano le tenéis, ladrón!”.

6.3.3.3. Utilización de adverbios en *-mente*.

La formación de adverbios modales con el sufijo *-mente*, muy fructífera en la lengua de Cervantes, aparece en casi la totalidad de sus entremeses y, también, en *La cárcel de Sevilla*:

¹⁹⁵ El asterisco señala los casos en los que el sufijo *-ísimo* no se ha aplicado a adjetivos.

ADVERBIOS EN <i>-MENTE</i>		
	Casos	N total
Cerv_EP_juez	encarecidamente, rectamente, bastantísimamente	3
Cerv_EP_guarda	altamente	1
Cerv_EP_vizcaino	verdaderamente, discretísimamente, ligeramente	3
Cerv_EP_retablo	principalmente, hondamente, especialmente, finalmente, verdaderamente,	5
Cerv_EP_cueva	∅	0
Cerv_EP_viejo	forzosamente, solamente	2
Cerv_EV_rufian	piadosamente	1
Cerv_EV_eleccion	bravamente, solamente	2
Dub_EP_carcel	públicamente, naturalmente	2

Tabla 18. Estudio del adverbio en *-mente*

Los dos adverbios en *-mente* que aparecen en *La cárcel de Sevilla* son “públicamente” y “naturalmente”, cuyo uso es muy habitual en Cervantes (13 y 12 veces, respectivamente), pero también en otros autores del momento¹⁹⁶.

6.3.2.4. *Uso del diminutivo en -ico*

Otro uso característico de los entremeses cervantinos es el uso frecuente del sufijo diminutivo *-ico* con valor afectivo y despectivo, muy por encima de lo habitual en esa época [Ariza Viguera, 1998]. Como vemos en la siguiente tabla, este sufijo también aparece en una ocasión —con valor despectivo— en *La cárcel de Sevilla*:

¹⁹⁶ Tal y como refleja CORDE (1547-1617), “públicamente” aparece 13 veces en Cervantes, 9 en Lope de Vega, 8 en Quevedo y ninguna en Góngora; por su parte, “naturalmente” lo hace 12 veces en Cervantes, 23 en Lope, 8 en Quevedo y ninguna en Góngora.

SUFIJO DIMINUTIVO EN -ICO		
	Casos	N total
Cerv_EP_juez	bonica, Perico, cedacico	3
Cerv_EP_guarda	Cristinica (7), añicos,	8
Cerv_EP_vizcaino	polvico, menudico, borrico	3
Cerv_EP_retablo	morenico,	1
Cerv_EP_cueva	Cristinica (5), demonicos	6
Cerv_EP_viejo	Cristinica (9), Lorencica (3), tortolica	13
Cerv_EV_rufian	levadicas, melindricos	2
Cerv_EV_eleccion	polvico (2)	2
Dub_EP_carcel	juececicos	1

Tabla 19. Estudio del sufijo diminutivo -ico

6.3.2.5. Empleo de reduplicaciones, anáforas y repeticiones diseminadas

La repetición literal de palabras o grupos de palabras es un recurso habitual en los entremeses cervantinos que, lejos de responder al descuido de su autor, se conforma como una estrategia de acrecentar la expresividad de estos elementos. Encontramos multitud de ejemplos de reduplicación, tanto de modos adverbiales del tipo “no, no” (Cerv_EP_juez, Cerv_EP_cueva, Cerv_EP_viejo), “al cabo, al cabo” (Cerv_EP_juez), “luego, luego” (Cerv_EP_guarda, Cerv_EV_rufian), como de nombres, verbos o interjecciones: “mire, mire”, “callad, callad” (Cerv_EP_juez); “mujeres, mujeres”, “celos, celos”, “duelos, duelos”, “señora, señora” (Cerv_EP_guarda); “señora Cristina, señora Cristina” (Cerv_EP_vizcaino); “Dios te libre, Dios te libre”, “échense todos, échense todos”, “cánsala, cánsala”, “Húcho ho, húcho ho, húcho ho!” (Cerv_EP_retablo); “demonico, demonico” (Cerv_EP_cueva), “vivo, vivo”, “señor compadre, señor compadre”, “vecinas, vecinas”, “pux, pux, pux”, (Cerv_EP_viejo); “la justicia, la justicia”, “viva, viva” (Cerv_EV_rufian); “miente, miente” (Cerv_EV_eleccion), etc. Especial valor estilístico tiene la reiteración con intensificación final, como “vueltas y más vueltas” (Cerv_EP_retablo), “viejo y reviejo y más que viejo” (Cerv_EP_viejo),

CAPÍTULO 6

“divorcio, divorcio, y más divorcio, y otras mil veces divorcio” (Cerv_EP_juez) y “venturón, venturón y mil veces venturón” (Cerv_EP_vizcaino).

También encontramos reduplicaciones en los parlamentos de los personajes, que empiezan su intervención repitiendo lo mismo (o casi lo mismo) que ha dicho el personaje anterior, a modo de anáfora. Así ocurre, por ejemplo, en *El retablo de las maravillas*:

CHANFALLA. Séanme testigos que me amenaza el Alcalde.
CHIRINOS. Séanme testigos que dice el Alcalde que, lo que manda S. M., lo manda el sabio Tontonelo.

Y también repeticiones diseminadas, en donde lo reiterado se convierte en un hilo conductor, operando a modo de estribillo:

CAPACHO. Basta: *de ex illis es*.
GOBERNADOR. *De ex illis es. De ex illis es*.
JUAN. Dellos es, dellos el señor Furrier, dellos es.
FURRIER. ¡Soy de la mala puta que los parió; y, por Dios vivo, que si echo mano a la espada, que los haga salir por las ventanas, que no por la puerta!
CAPACHO. Basta: *de ex illis es*.
BENITO. Basta: dellos es, pues no vee nada.

En *La cárcel de Sevilla* también aparece la reduplicación de palabras (como “¡hola, hola!”), en hasta dos ocasiones, o “de venta en venta”) y las repeticiones cuasi literales de oraciones enteras puestas en boca de distintos personajes, que, como ocurre en Cervantes, tienen por objeto dotar de ritmo y sonoridad al texto. La obra se abre con este recurso,

GARAY. Abre aquí, alcaide, que nos comen chinchas
SOLAPO. Abra aquí, so alcaide que nos comen garrapatas

y un poco después encontramos,

GARAY. Loado sea Dios, que veo el cielo de Cristo.
SOLAPO. Loado sea Dios, que veo el nubífero.
PAISANO. Loado sea Dios, que veo el Sempiterno.

Las repeticiones diseminadas también son frecuentísimas en *La cárcel de Sevilla*. Así ocurre, por ejemplo, en este pasaje,

- ALCAIDE. ¿Qué ruido es este? Por vida de el rey, que he de pasar alguno a la otra cárcel, o que ha de dormir en el cepo.
- BARRAGÁN. Cuando usted haga pasar alguno a la otra cárcel, hay aquí hombres que no se les da esta. (*Da una castañeta*).
- PAISANO. Cuando usted haga pasar alguno a la otra cárcel, hay aquí alguno que no se le dará nada;

y también en el momento en el que los presos tienen que loar al condenado, pasándose la palabra de uno a otro,

- GARAY. Hable el seor Barragán, que es más honrado y más antiguo.
- BARRAGÁN. Yo no haré: hable el seor Solapo.
- SOLAPO. Así me vea en aquella calle con libertad, que no diga palabra: hable el seor Cuatro.
- CUATRO. El Cuatro no lo hará: hable el seor Garay.

6.3.2.6. *Uso frecuente de figuras de dicción*

En los entremeses indubitados de Cervantes son frecuentes las figuras de dicción, muy relacionadas con ese gusto por la repetición de vocablos que veíamos en el punto anterior; así, por ejemplo, encontramos la duplicación de palabras con variación de género gramatical (“el diablo o la diablo”, en Cerv_EP_vizcaino, o “no hay gitanos ni gitanas”, en Cerv_EV_eleccion), el uso pleonástico de verbos que llevan un complemento de su misma raíz (“oídos con que oír”, en Cerv_EP_juez), la sucesión de variantes flexivas de un mismo verbo (“busco en esta calle lo que hallo, y tú buscas y no hayas” o “haber tenido y tener respeto a las órdenes que tiene”, en Cerv_EP_guarda; “todos digan como digo”, “doy lo dicho por no dicho” o “siéntome, y me siento”, en Cerv_EV_eleccion), la utilización de vocablos con igual significante pero diferente significado (“no me haría vuesa merced una merced” en Cerv_EP_guarda), la proximidad de palabras con un mismo lexema (“otro gallo cantará en su gallinero” y “estaba un pregonero pregonando”, en Cerv_EP_vizcaino; “la encina da bellotas; el pero, peras; la parra, uvas y el honrado, honra”, “del mar movible la inmóvil roca”, en

Cerv_EV_rufian; “demasiadas demasías”, en Cerv_EV_eleccion; “baile el bailarín” en Cerv_EP_viejo) o de palabras con sonidos afines (“hombre honrado”, en Cerv_EP_retablo; “no solamente canta sino encanta”, en Cerv_EV_eleccion), etc. Dentro de todos estos juegos de palabras, cabe hacer mención especial a una fórmula muy idiosincrásica cervantina, que consiste en la utilización de dos elementos iguales en coordinación, pero el segundo de ellos intensificado por el uso de un prefijo de repetición (generalmente *re-*): “pulirme y repulirme”, “tocar y retocar”, “me siga y me persiga” y “ha pasado y aún repasado”, en Cerv_EP_vizcaino; “cenara y recenara”, en Cerv_EP_cueva; “viejo y reviejo y más que viejo” y “revueltas y vueltas” en Cerv_EP_viejo; “sudando y trasudando” en Cerv_EV_rufian; “vivan y revivan”, etc.

Aunque en *La cárcel de Sevilla* no encontramos esta última fórmula enfática, sí aparecen el resto de juegos señalados:

- disyuntiva de género: “*testigo* o *testiga*”;
- usos pleonásticos: “echaldas fuera” o “mal morir”;
- sucesión de variantes flexivas de un mismo verbo: “pelea y peleará”, “debo condenar y condeno”, “decirme las ledanías que se suelen decir”, “haya de salir, quiero salir”;
- utilización de vocablos con igual significante pero diferente significado: “voacé me hará merced” (en sentido estricto no es el mismo significante porque aparece “voacé”, es decir, la forma sincopada de “vuestra merced”, pero nos resulta significativo porque es el mismo juego de palabras que aparece en *La guarda cuidadosa*);
- proximidad de palabras con un mismo lexema: “si yo apelo para esos señores padres de la audiencia, remediadores de los fallos, pienso que no tendré ningún remedio”, “no hay hombre cargado... que yo también he sido carga de muladar”;
- utilización de palabras cercanas con sonidos afines: “pasar alguno algún puñal”, “hombre honrado” (que aparece en dos ocasiones y que es utilizado por Cervantes en *El retablo de las maravillas*);

- chistes y juegos de palabras: “me ha traído engañada, diciendo que había de meter un escrito; y agora le mete, agora le saca”.

6.3.2.7. Creación de nuevas palabras

Cervantes utiliza procedimientos de composición, derivación y parasíntesis para crear palabras nuevas, cargadas de comicidad.

▪ Por composición. La formación de palabras nuevas a partir de la unión de dos ya existentes es muy habitual en Cervantes; a nadie se le escapa la gran presencia que tienen en el *Quijote* palabras como *baciyelmo* (I, 44), *barbitaheño*, *barbilucio* (II, 1), *barbiponiente* (II, 1) o *toquiblanca* (II, 48). Como vemos en el siguiente cuadro, la creación léxica por composición aparece en cuatro de sus ocho entremeses (aunque no siempre se trata de palabras inventadas por él) y, también, en *La cárcel de Sevilla*, con un claro efecto humorístico:

PALABRAS CREADAS COMPOSICIÓN ¹⁹⁷		
	Casos	N total
Cerv_EP_juez	*raspahilando, *rostrituerta	2
Cerv_EP_guarda	∅	0
Cerv_EP_vizcaino	∅	0
Cerv_EP_retablo	∅	0
Cerv_EP_cueva	sacridiablo	1
Cerv_EP_viejo	*pelinegro	1
Cerv_EV_rufian	comevivos, mondanispolas, *gallipavos, Maribobales	4
Cerv_EV_eleccion	sorbehuevos, catavinos	2
Dub_EP_carcel	perniquebrados, *boquirrubio ¹⁹⁸	2

Tabla 20. Estudio de las palabras formadas por composición

¹⁹⁷ El asterisco señala los casos en los que la palabra tiene presencia en otros autores del CORDE (1547-1617).

¹⁹⁸ No incluimos *soplavivos* porque aparece dentro del romance “Alta mar esquivá”.

▪ Por derivación y parasíntesis. La formación de palabras nuevas mediante prefijación o sufijación aparece con frecuencia en el *Quijote* con la finalidad de provocar la carcajada: *venteril* (I, 3), *andantesca* (I, 7), *semidoncellas* (I, 43), *sotaermitaño* (II, 24), *antojuna* (II, 48), etc. En alguno de sus entremeses también está presente este recurso cómico: *automedones* se forma al añadir el sufijo *-es* al nombre del personaje de Homero llamado Automedón (Cerv_EP_cueva), *escarramanesco* al sumar el sufijo *-esco* al nombre del famoso Escarramán (Cerv_EP_guarda), *atontoneleada* y *atontoneleados* al añadir simultáneamente el prefijo *a-* y el sufijo *-ado* al nombre de Tontonelo (Cerv_EP_retablo) y *sotasacristán* al anteponer al nombre común *sacristán* el prefijo *sota-* (Cerv_EP_guarda). En *La cárcel de Sevilla*, sin embargo, no se da ningún caso de neologismo por derivación o parasíntesis¹⁹⁹.

6.3.2.8. Inclusión de expresiones latinas

Otro de los recursos habituales en el teatro breve cervantino es el de poner *latinicos* en boca de sus personajes. Aunque generalmente se utiliza para caracterizar el habla de jueces (Cerv_EP_juez), sacristanes (Cerv_EP_guarda) o bachilleres (Cerv_EV_eleccion), también lo encontramos por boca de pícaros y rufianes: en *El retablo de las maravillas* es Chirinos (pícaro) la que pronuncia la expresión “ante omnia”, que el alcalde Benito Repollo, en su analfabetismo, interpreta como “Antonia”; y en *El rufián viudo*, son Juan Claros (rufián), Trampagos y la Mostrenca (prostituta) los que usan las expresiones “sicut erat in principio”, “rapio rapis” y “sine numero”²⁰⁰.

Como vemos en la Tabla 21, la utilización de expresiones o palabras latinas aparece en seis de los ocho entremeses cervantinos, pero no en el texto dubitado:

¹⁹⁹ Está la palabra *corchatesca* (*corchete* + *esca*) para referirse a los corchetes, un tipo de alguaciles, pero se da dentro del romance “Alta mar esquiva”, por lo que no es atribuible al autor del texto.

²⁰⁰ “Aunque pudiera parecer inconcebible, en germanía nos encontramos con una serie de latinismos ‘crudos’, no debidos a la cultura de los escritores cultos que compusieron diversas obras con léxico germanesco, como hemos visto, sino a la presencia de unas pocas fórmulas fijas, muchas de ellas comunes del habla popular [Hernández Alonso y Sanz Alonso, 1999: 194].

EXPRESIONES LATINAS		
	Casos	N total
Cerv_EP_juez	<i>quia nullam invenio causam</i>	1
Cerv_EP_guarda	<i>tu dixisti</i>	1
Cerv_EP_vizcaino	∅	0
Cerv_EP_retablo	<i>ante omnia (2), in corbona, de ex illis (3)</i>	6
Cerv_EP_cueva	∅	0
Cerv_EP_viejo	<i>amicus usque ad aras, usque ad portam</i>	2
Cerv_EV_rufian	<i>sicut erat in principio, rapio rapis, sine numero</i>	3
Cerv_EV_eleccion	<i>redeamus ad rem, sufficit</i>	2
Dub_EP_carcel	∅	0

Tabla 21. Estudio de la inclusión de expresiones latinas

6.3.2.9. Predilección por los sinónimos coordinados

La acumulación coordinada de sinónimos, generalmente bimembre, es habitual en la lengua de Cervantes [Rosenblat, 1978: 123] y los ejemplos en sus entremeses son innumerables: “hospitalera ni enfermera”, “muela ni diente”, “vaivenes y empujones” y “castas y honestas” en Cerv_EP_juez; “mudables y antojadizas” y “joyas y preseas” en Cerv_EP_guarda; “juramentos y maldiciones” en Cerv_EP_vizcaino; “farándula y carátula” en Cerv_EP_retablo; “penas y dolores” y “origen y principio” en Cerv_EP_cueva; “ungüentos y medicinas” y “rejas y celosías” en Cerv_EP_viejo; “sus dientes y sus muelas” en Cerv_EV_rufian; y “borrascas y tormentas” en Cerv_EV_eleccion; entre otros muchos. Este tipo de sinonimia bimembre también aparece en numerosas ocasiones en *La cárcel de Sevilla*: “quietos y sosegados”, “viejo y apolilado”, “borrachería y barahúnda”, “huérfana y sola”, “regocijen y alegren” o “en la calle y en la libertad”.

6.3.2.10. Utilización pragmática de los diálogos de enunciados interrogativos

Algo sobre lo que llamaba la atención Maestro al estudiar la construcción dialógica de los entremeses de Cervantes es la alta presencia de diálogos fundamentados en enunciados y discursos interrogativos.

Los personajes cervantinos saben muy bien que toda pregunta exige por parte de su interlocutor no sólo una información, sino una información formalizada y modalizada por el interlocutario a quien se dirige la pregunta, y cuyo modelo de enunciación impone siempre la dirección en que debe realizarse la respuesta. Todo ello explica que la información no pueda existir como experiencia aislada e independiente. La información siempre es resultado de un discurso intercambiado, compartido. Como sabemos, la comunicación es una actividad cooperativa, y su éxito depende de la intensidad y la eficacia de esta cooperación [1998: 601].

El diálogo de enunciados interrogativos tiene consecuencias pragmáticas, pues sirve para definir las relaciones interpersonales que se establecen entre los sujetos dramáticos, siendo generalmente el débil el que quiere saber (pregunta) y el dominante el que tiene la información (respuesta). De este modo, el papel de los interlocutores dentro de la interrogación dialógica ayuda al autor a definirlos como personajes. A través de la dialogía interrogativa advertimos que el pretendiente más inseguro de *La guardia cuidadosa* es el soldado y no el sacristán, que recurre a la apelación interrogativa de manera recurrente (Cerv_EP_guarda); también somos conscientes de la superioridad de Chirinos y Chanfalla sobre el resto de convecinos, que les plantean preguntas constantemente (Cerv_EP_retablo); deducimos la posición enfrentada que mantienen Lorenza y Cristina ante el tema del adulterio en sus preguntas retóricas (Cerv_EP_viejo); nos percatamos de que las preguntas que realizan Leonarda y Cristina al estudiante tienen la funcionalidad estratégica de descubrir si están ante un hombre discreto que guarde el secreto de sus adúlteras intenciones (Cerv_EP_viejo), presenciamos cómo Trampagos revela al rufián la información sobre la muerte desde su propia subjetividad (Cerv_EV_rufian), etc.

En *La cárcel de Sevilla*, el diálogo de enunciados interrogativos cumple funciones pragmáticas muy similares a las cervantinas. En el primer enfrentamiento entre los presos y el alcaide, es este último el que lanza distintas preguntas a Barragán y Paisano para quejarse de su actitud, las más de las veces sin respuesta, lo que pone de manifiesto

su poca autoridad dentro de la cárcel (“¿qué ruido es éste?”, “¿qué es esto, Barragán?”, “¿ya tomáis vos las mañanas de Paisano”, “¿y decís eso?”). Cuando Paisano recibe la sentencia de muerte, es él el que va dirigiendo de manera estratégica el diálogo con el escribano, a través de una serie de preguntas concretas con las que desea obtener la máxima información posible (“¿qué buenas pascuas nos viene a notificar?”, “¿quién dio esta sentencia?”, “¿no es mejor decir que muera este hombre y ahorrar de tanta guarnición?”) y provocar su compasión (“¿cómo se va voacé después que queda un hombre cargado hasta las entrañas?”); sólo en la parte final de la conversación, el escribano formula una pregunta, ante la incredulidad que le genera la apelación de Paisano (“¿Y para quién diremos que apeláis?”). Y en el momento en el que Paisano ve llegar a Torbellina y Beltrana, formula una serie de preguntas jocosas con finalidad exclamativa (“¿quién me ha traído aquí estas ayudas a costa de mal morir?”, “¿quién me ha traído aquí estos teatrillos infernales?”).

6.3.2.11. *Adaptación del lenguaje al nivel sociocultural de los personajes*

Cervantes caracteriza a sus personajes a través del lenguaje. Como han estudiado muchos especialistas —entre ellos, Agostini [1964b]— éstos parecen hablar como lo harían “en la vida corriente, de acuerdo con su modo de ser y su categoría social”:

Así, el soldado de *El juez*, de relativa fecundidad verbal, comienza con frases cortas para proseguir con un período largo con superabundancia de la conjunción “y”. El juez es sobrio. Al genio chancero del soldado de *La guarda* corresponde un estilo figurado, ampuloso a trechos, lleno de humor siempre; el vizcaíno habla con las transposiciones correspondientes al tipo; el gobernador, con dignidad; el lenguaje del titiritero es tan gráfico y tan vivo que parecería milagro que no vieran los aldeanos verdaderamente las maravillas del famoso retablo; el estudiante cuenta su historia en un estilo cortado, pero en la parodia del conjuro emplea una octava de arte mayor, y tras el tono solemne y las palabras de doble intención vuelve a la lengua corriente y graciosa con sus “demonicos”, “sacridiablo”, etc.; la alcahueta usa un lenguaje sobrio y directo al hablar con Lorenza, que cambia en uno de período largo al dirigirse al vejete, y lleno de equívocos al final del entremés; los rufianes emplean giros de germanía; las mujeres hablan con la viveza propia del sexo y de la edad [1964b: 519].

En *La cárcel de Sevilla*, al igual que ocurre en Cervantes, no todos los personajes hablan del mismo modo:

▪ El habla de los presos es la propia de los hampones. Encontramos términos y expresiones como *murcio* ('ladrón'), *trainel* ('criado de un rufián'), *guzpátaro* ('agujero'), *lanudo* ('cobarde'), *corchete* ('oficial de la justicia'), *ventoso* ('soplón'), *retablo* ('rostro'), *águila* y *aguilucho* ('ladrón' y 'ladrón joven'), *bueyes* ('baraja'), *leiva* ('marca en los naipes'), *la de ganchos* ('daga'), *trato airado* ('vida rufanesca'), *borrajar* ('herir con un arma blanca'), *acomodar* ('poner a una prostituta bajo la protección de un rufián') o *vendimiar la vida* ('matar'), habituales en la germanía y, la gran mayoría, presentes en Cervantes. También encontramos otras características propias de esta jerga y del nivel sociocultural de sus representantes, como la utilización metafórica de términos relacionados con la marina (*borrascas*, *naufragios*), la anteposición del pronombre persona de primera persona en el sujeto ("yo y el seor Paisano"), el uso de juramentos y fórmulas blasfematorias ("voto a Cristo"), la aspiración del sonido /x/ ("barahe") o el empleo de formas apocopadas, como *seor* y *sor* por *señor* y *voacé* y *voacedes* por *vuesa merced* y *vuestras mercedes* [Hernández Alonso y Sanz Alonso, 1999].

▪ En contraposición, el habla del escribano y del procurador es normativa y carente de todos estos giros lingüísticos. Usan *señor* o *hermano / hermana* cuando hablan con los presos, pero *vuesa merced* cuando lo hacen con el Alcaide. Es reseñable la lectura de la sentencia por parte del escribano, que presenta rasgos propios del léxico jurídico, también muy habitual en Cervantes ("fallo", "debo condenar y condeno", "do", "so pena", etc.).

▪ El habla del alcaide —al igual que su actitud— está a medio camino entre unos y otros. Por lo general no utiliza los términos de la germanía (que sí entiende a la perfección), pero emplea en una ocasión *navío* con el valor metafórico de 'cuerpo', posiblemente porque, como él mismo reconoce, también ha sido "carga de muladar". En cuanto a las fórmulas de tratamiento, emplea *hermano* o *vos* para referirse a los presos y *vuesa merced* cuando dialoga con el escribano²⁰¹.

²⁰¹ En época cervantina empezó a generalizarse el empleo de la fórmula de tratamiento *vuestra merced*, "que por desgaste, a través de una serie de variantes, se fue transformando en *usted*" [Rosenblat,

Todas estas particularidades son idénticas a las que usa Cervantes para caracterizar el habla de los rufianes que aparecen en sus textos. En *Rinconete y Cortadillo* encontramos “murcio”, “guzpátaros” y “trainel”; en *La gitanilla*, “aguilucho”; en *El rufián dichoso*, “trato airado” o “la de ganchos”; en *El rufián viudo*, “hampesco”, etc. Estos personajes, además, utilizan las formas apocopadas *voacé / voacedes* y *sor y seor*.

El análisis cualitativo de marcas idiosincrásicas del texto dubitado parece corresponderse con el comportamiento cervantino. Encontramos casos de leísmo, tanto de persona como de cosa, adverbios en *–mente*, un diminutivo en *–ico* con valor despectivo (*juececicos*), ejemplos de reduplicaciones, anáforas y repeticiones diseminadas, varias figuras de dicción, palabras creadas por composición, acumulaciones coordinadas de sinónimos, diálogos de enunciados interrogativos y un lenguaje adaptado al nivel sociocultural de los personajes. Tan sólo hay disimilitud en el empleo de superlativos en *–ísimo* (cuya presencia en los entremeses de Cervantes oscila entre 1 y 3 repeticiones) y en la utilización de expresiones latinas (que no se da en todos los entremeses de Cervantes).

6.3.3. Distribución de la frecuencia de los 5-gramas de caracteres más empleados en los textos (*character 5-grams*)

Como vimos en el capítulo anterior, la aplicación del algoritmo Eder’s Delta a los 1000 5-gramas de caracteres más frecuentes también funcionaba relativamente bien para el corpus indubitado de autor conocido (Gráficos 16 y 17) y, por eso, creemos conveniente repetir el análisis incluyendo ahora *La cárcel de Sevilla* (Gráfico 39).

1978:180]. En todos sus entremeses aparece la variante *vuesa merced / vuestas mercedes*, aunque con una presencia muy superior en los seis en prosa (128 veces) que en los dos en verso (3 veces), que alternan esta forma con “sus mercedes” (2 veces). El uso de “voacé” aparece en varias ocasiones en Cerv_EV_rufian para caracterizar el habla de los rufianes.

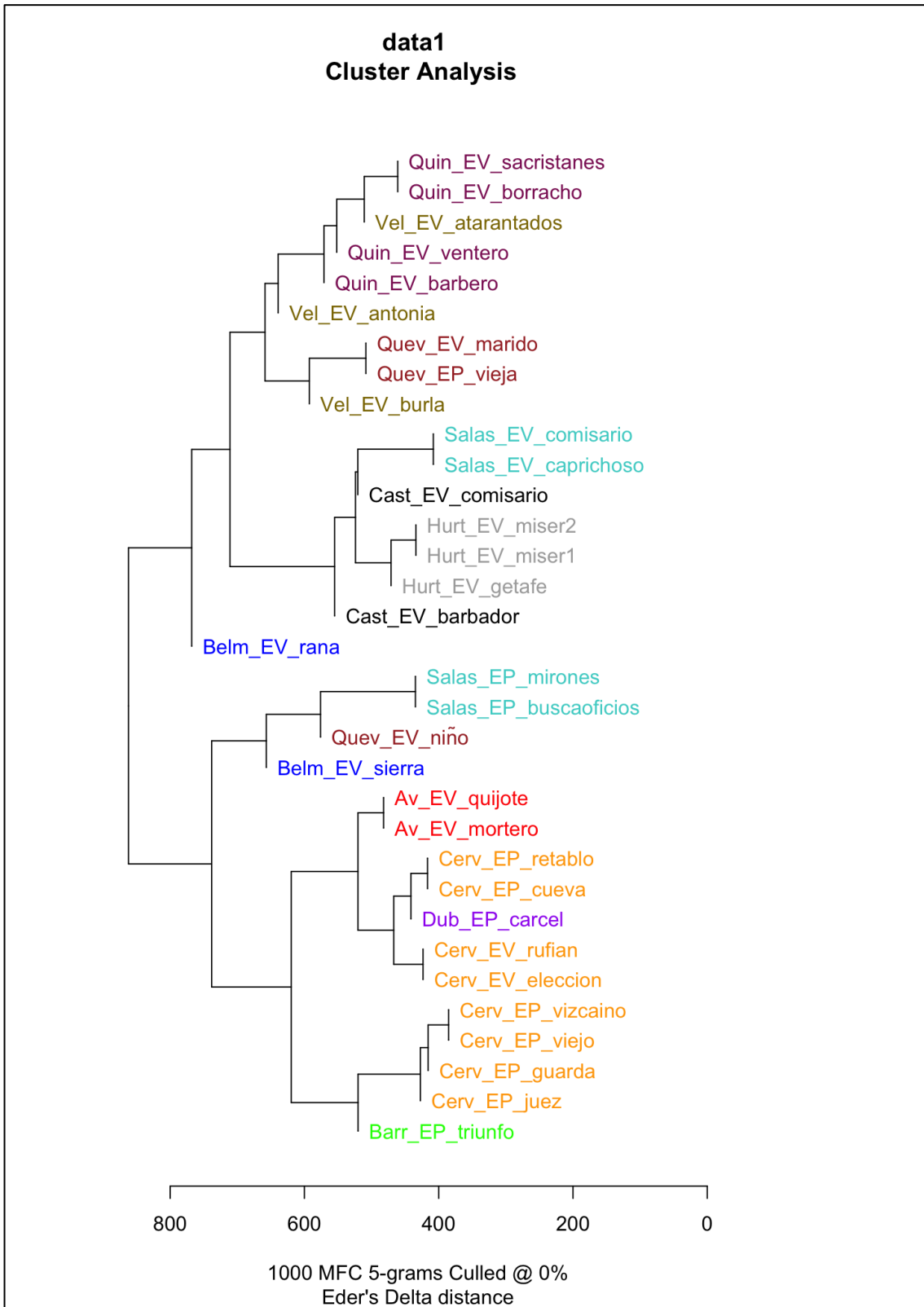


Gráfico 39. Clúster basado en la frecuencia de 5-gramas de caracteres para *La cárcel de Sevilla* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFC 1000 5-grams. Culled 0%)

Como vemos en el dendrograma anterior, el entremés dubitado presenta un comportamiento ortográfico similar a *El rufián viudo*, *La elección de los alcaldes de Daganzo* y, en mayor medida, *El retablo de las maravillas* y *La cueva de Salamanca*. Aunque se sitúa en la misma rama en la que están también los entremeses de Ávila (*Los invencibles hechos de don Quijote de la Mancha* y *El mortero y chistes de sacristán*) y Barrionuevo (*El triunfo de los coches*), las asociaciones con Cervantes son más sólidas, pues se producen en el centro y más a la derecha.

6.3.4. Distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*)

Como vimos en el capítulo 4, el análisis de la frecuencia de las 1000 MFW con Stylo revela que *La cárcel de Sevilla* tiene una distribución pareja a la de los indubitados cervantinos, especialmente a *El rufián viudo*, *La elección de los alcaldes de Daganzo* y *El retablo de las maravillas* (Gráfico 25). Para profundizar en este rasgo, hemos realizado dos nuevos análisis, manteniendo los mismos criterios:

- Si tomamos como referencia el corpus total de entremeses indubitados en prosa —los dieciséis de autor conocido y los seis anónimos—, la identificación de *La cárcel de Sevilla* con Cervantes sigue siendo evidente, pues el texto dubitado se sitúa en la rama en la que se agrupan todos los entremeses en prosa del autor del *Quijote*, en la que también está, una vez más, el entremés de Barrionuevo (Gráfico 40).
- Si realizamos el mismo análisis incluyendo también los textos de la *Relación* —tanto las dos partes de Chaves (Chav_R_carcel1 y Chav_R_carcel2) como la continuación anónima (Anon_R_carcel3)— vemos que el resultado apenas varía, a pesar de la cercanía léxica que se presupone a todos los textos que hablan de la prisión sevillana. Los dos textos de Chaves se asocian indiscutiblemente entre sí, al ser los dos de todo el corpus en los que más a la derecha se produce la asociación; la parte tercera se sitúa en la misma rama, pero bastante alejada de éstos. El entremés dubitado de *La cárcel de Sevilla* se

mantiene en la rama de los entremeses de Cervantes y Barrionuevo, asociado directamente con *El retablo de las maravillas* (Gráfico 41).

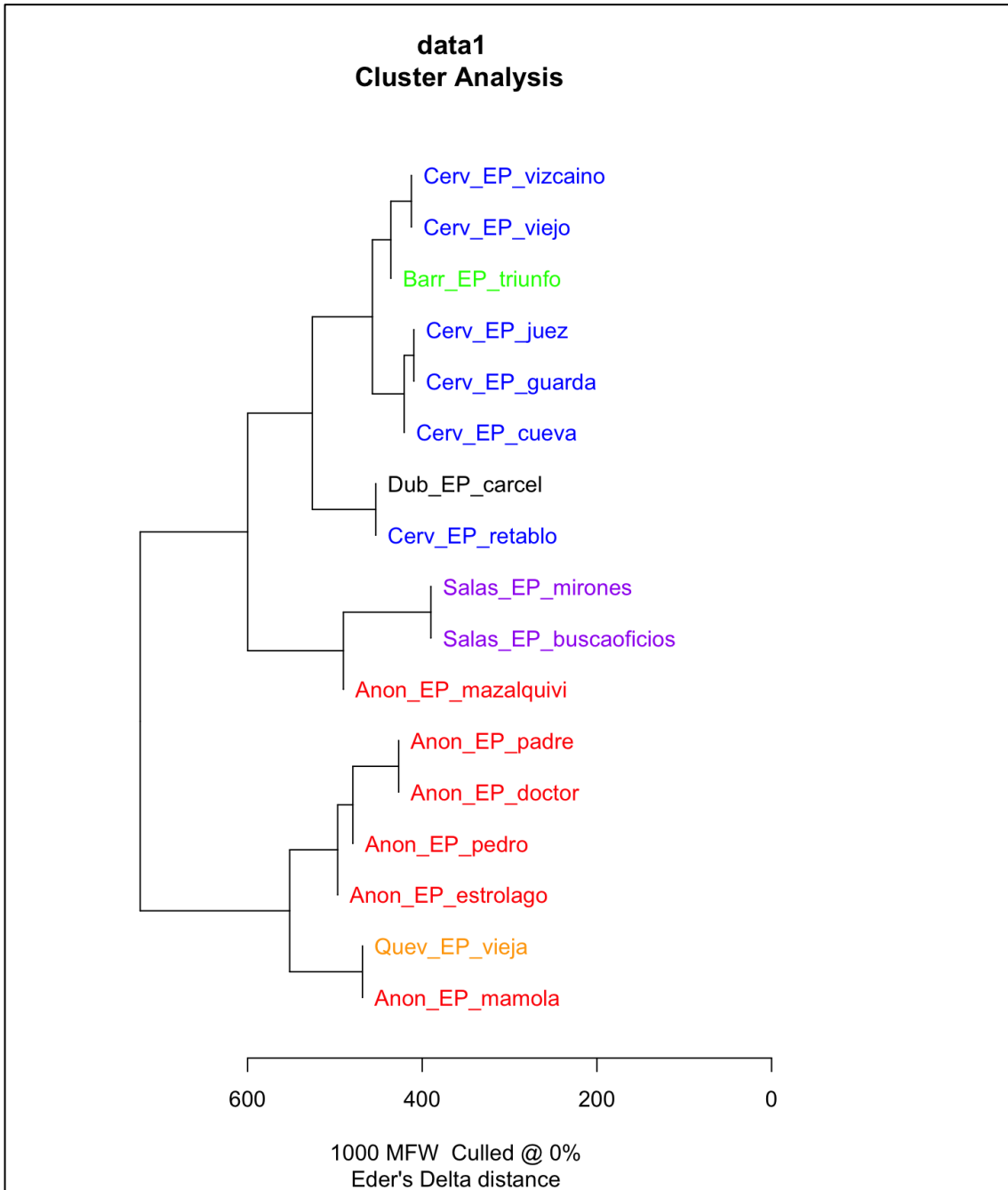


Gráfico 40. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses indubitados en prosa y *La cárcel de Sevilla* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

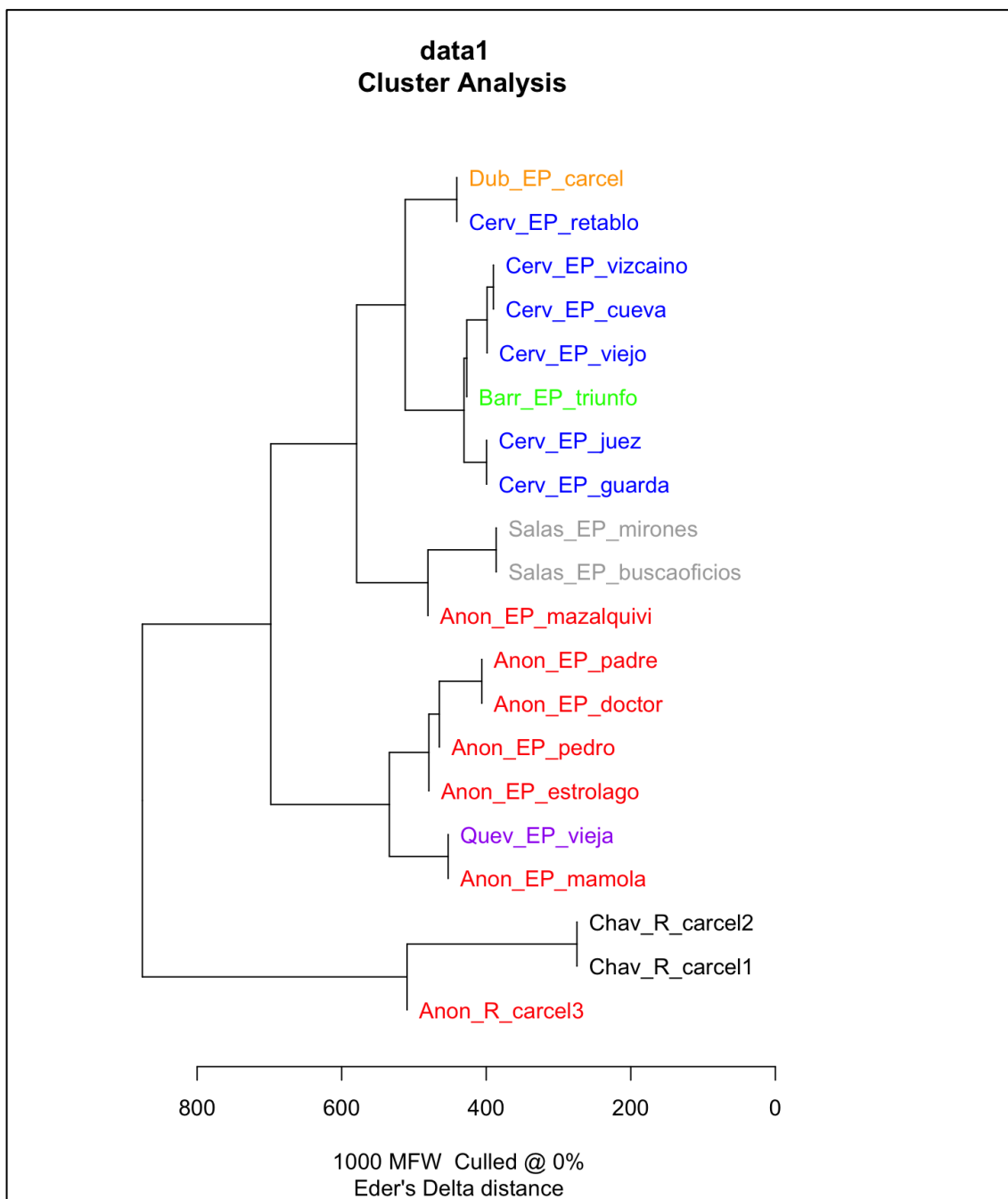


Gráfico 41. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses indubitados en prosa, *La cárcel de Sevilla* y las tres partes de la *Relación* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

Todos estos gráficos (25, 40 y 41) demuestran la similitud que existe entre la distribución de la frecuencia léxica de *La cárcel de Sevilla* y varios de los entremeses de Cervantes.

Esta similitud también se pone de manifiesto en el árbol de consenso que mostramos a continuación, en donde *La cárcel de Sevilla* se emparenta con *El retablo de las maravillas*, *El rufián viudo* y *La elección de los alcaldes de Daganzo*:

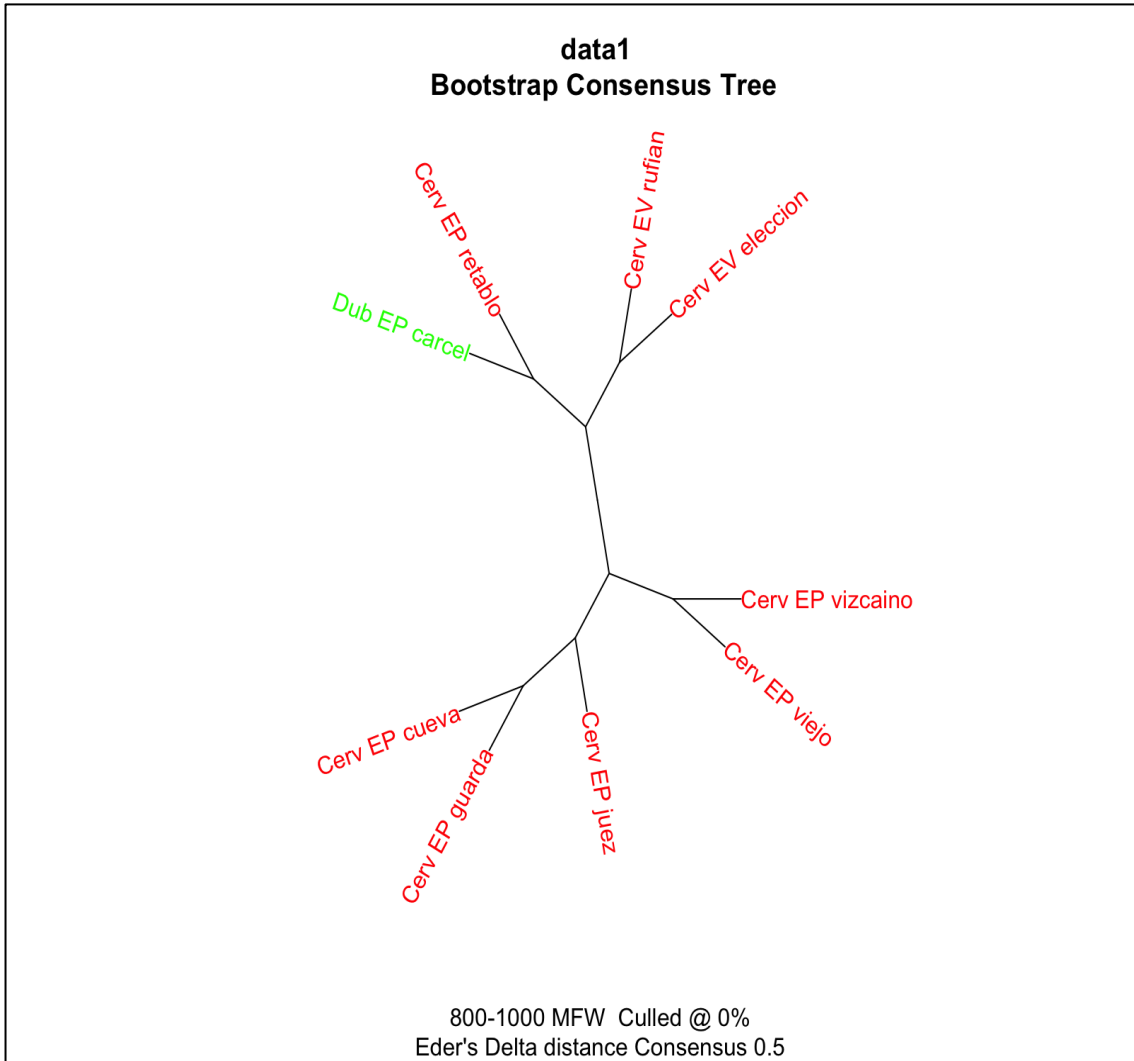


Gráfico 42. Árbol de consenso basado en la frecuencia de las palabras más empleadas (RStudio. Stylo. Bootstrap Consensus Tree. 0.5 consensus strength. Eder's Delta. MFW 800-1000)

En esta variable, *El retablo de las maravillas* está más próximo a *La cárcel de Sevilla* que a cualquiera de los otros siete entremeses de Cervantes.

6.3.5. Análisis de componentes principales (PCA) aplicado a la *Delta distance*

Como avanzamos en el capítulo anterior, el análisis de componentes principales (PCA) es una de las técnicas estadísticas más utilizadas para resumir patrones comunes de variación entre variables, lo que permite entender la relación que existe entre ellas. Ha tenido aplicaciones fructíferas en el campo del reconocimiento facial o la compresión de imágenes, así como en la identificación de la autoría de textos escritos, también en lengua española [De la Rosa y Suárez, 2016].

The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set [Jolliffe, 2002: 2].

Si la visualización del análisis del PCA de las 1000 palabras más frecuentes aplicado al conjunto de entremeses dubitados e indubitados cervantinos permitía advertir la identificación entre *La cárcel de Sevilla* y *El viejo celoso* (Gráfico 20), en este apartado vamos a ver qué ocurre si tenemos en consideración sólo el texto dubitado objeto de estudio²⁰².

Como vemos en esta nueva gráfica de puntuaciones (Gráfico 43), los dos primeros componentes (PC1 y PC2) presentan una varianza acumulada del 37,9% (PC1 cubre el 21% de la varianza total y PC2 el 16,9%). Al igual que ocurría en el Gráfico 20, el PC1 (eje horizontal) muestra una tendencia general a diferenciar entre entremeses en prosa (izquierda) y entremeses en verso (derecha), salvo contadas excepciones, que se sitúan en el lado opuesto, aunque ciertamente próximas a la línea divisoria (Cerv_EP_retablo y Dub_EP_carcel). *La cárcel de Sevilla* se sitúa muy cerca de la intersección, es decir, que sus componentes principales se identifican con los del grupo mayoritario. De hecho, este entremés guarda mayor sintonía con el conjunto de entremeses cervantinos que *El rufián viudo* (Cerv_EV_rufian), que aparece en el extremo derecho del cuarto cuadrante.

²⁰² No realizamos una PCA del total de indubitados ya que, como hemos explicado, este tipo de análisis busca reducir la dimensionalidad del conjunto para crear conglomerados. Por eso, es más significativo ver si el texto dubitado forma parte o no del conglomerado integrado por los ocho entremeses de Cervantes.

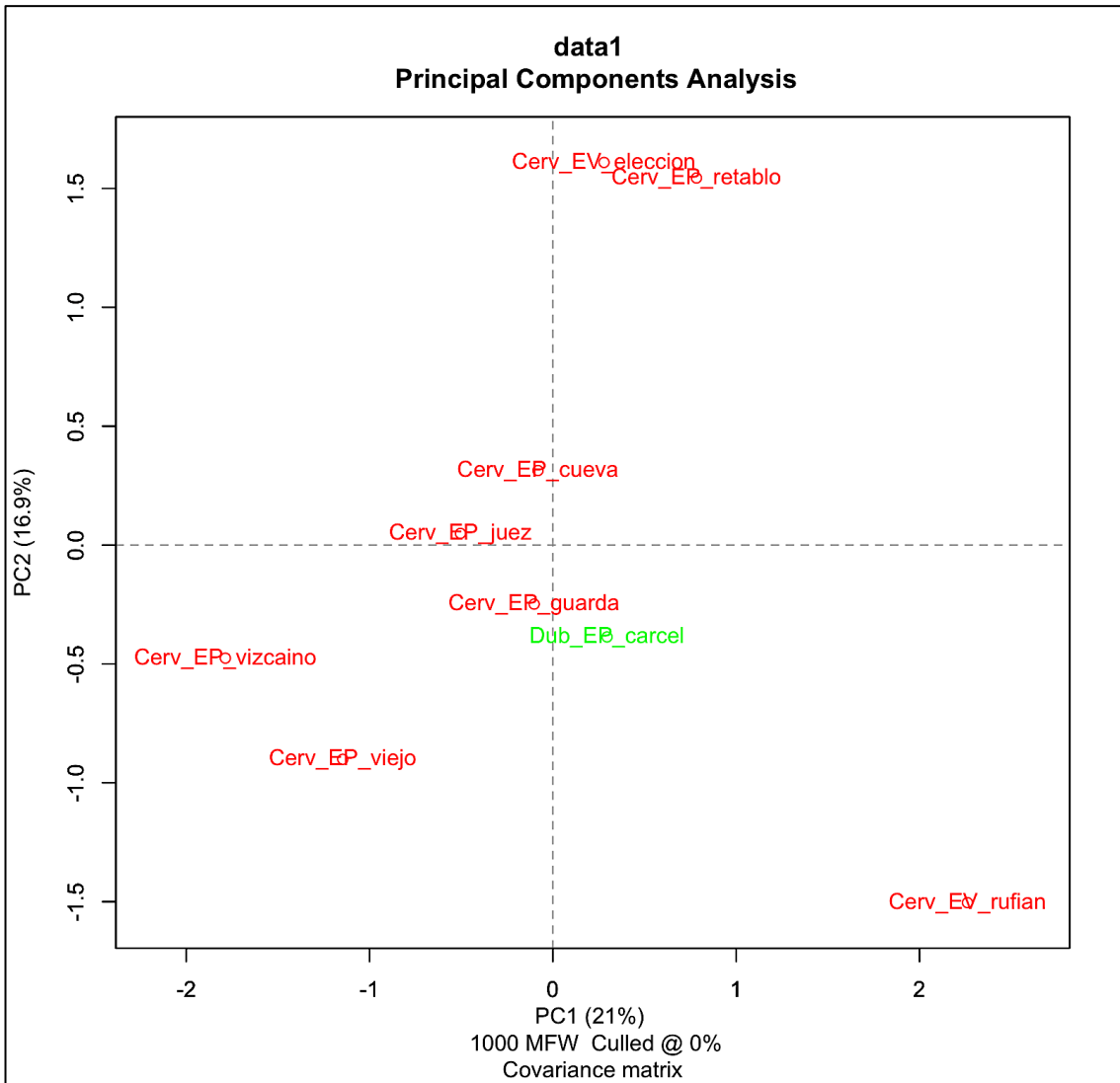


Gráfico 43. Gráfica de puntuaciones del PCA de las 1000 palabras más frecuentes en *La cárcel de Sevilla* y los entremeses indubitados (RStudio. PCA. 1000 MFW. Culled 0%)

6.3.6. Clasificación supervisada por aprendizaje automático (Classify)

Otra de las funcionalidades del paquete Stylo, de la que no hemos hablado hasta ahora, es Classify, que está basada en una serie de métodos de aprendizaje automático especialmente concebidos para su aplicación a la estilística computacional: Delta, k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), Naive Bayes y Nearest Shrunken Centroids (NSC) [Jockers y Witten, 2010; Eder, Rybicki y Kestemont, 2016].

Nosotros vamos a aplicarlo al algoritmo Delta, ya antes explicado [Burrows, 2002]. El procedimiento es sencillo: Classify extrae una tabla de frecuencias de los textos que incluimos en un corpus de entrenamiento (*primary set*), aprende los parámetros caracterizadores de cada texto, compara esas frecuencias con las de cada uno de los textos que conforman el corpus de prueba (*secondary set*) y propone asociaciones.

Nuestro corpus de entrenamiento estará conformado por una serie de textos indubitados de Cervantes (Cerv_EP_vizcaino y Cerv_EP_retablo), Ávila (Av_EV_mortero), Salas Barbadillo (Salas_EP_mirones), Quiñones de Benavente (Quin_EV_borracho), Hurtado de Mendoza (Hurt_EV_getafe), Quevedo (Quev_EV_marido) y Chaves (Chav_R_carcel1) y nuestro corpus de prueba por el texto dubitado (Dub_EP_carcel) y otros tantos indubitados de los mismos autores (Cerv_EP_viejo y Cerv_EP_guarda, Av_EV_quijote, Salas_EV_caprichoso, Quin_EV_barbero, Hurt_EV_miser1, Quev_EV_vieja y Chav_R_carcel2); de este modo, los textos indubitados del corpus de prueba nos servirán para confirmar la fiabilidad de los resultados obtenidos para el texto dubitado.

En la siguiente tabla mostramos los resultados obtenidos tras aplicar Classify al estudio de la distribución de la frecuencia léxica, seleccionando el algoritmo Delta y un intervalo de 100 a 1000 palabras más frecuentes (con un incremento de 100 en 100):

RESULTADOS			
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
100 MFW	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Quev_EP_vieja	-->	Cerv_EP_retablo
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino

CAPÍTULO 6

200 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Quev_EV_vieja	-->	Cerv_EP_retablo
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino
300 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Quev_EV_vieja	-->	Cerv_EP_retablo
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino
400 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino
500 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino

ESTUDIO DE AUTORÍA DEL ENTREMÉS DE LA CÁRCEL DE SEVILLA

600 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino
700 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino
800 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino
900 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino

	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Chav_R_carcel2	-->	Chav_R_carcel1
	Dub_EP_carcel	-->	Cerv_EP_retablo
1000 MFW	Cerv_EP_guarda	-->	Cerv_EP_vizcaino
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Quev_EV_vieja	-->	Quev_EV_marido
	Cerv_EP_viejo	-->	Cerv_EP_vizcaino

Tabla 22. Clasificación supervisada por aprendizaje automático de Delta para Dub_EP_carcel (RStudio. Classify. Eder's Delta. MFW 100 - 1000)

La función Classify identifica correctamente todos los textos indubitados del corpus de prueba a partir de las 700 palabras más frecuentes. Centrándonos en los porcentajes de éxito atributivo, calculados en base a las asociaciones entre indubitados, podemos hablar del 62,5% en las 100, 200 y 300 MFW (5 de 8), del 75% en las 400 y 500 MFW (6 de 8), del 87,5% en las 600 MFW (7 de 8) y del 100% en las 700, 800, 900 y 1000 MFW (8 de 8). En todos los casos, desde las 100 a las 1000 MFW, el *Entremés de la cárcel de Sevilla* se asocia directamente con *El retablo de las maravillas*, de Cervantes, y nunca con Chaves, a pesar de ser su fuente documental.

6.3.7. Porcentaje de similitud textual (*similarity*) y porcentajes de vocabulario compartido una sola vez (*Once only in both*), de vocabulario compartido más de una vez (*Shared > 1*) y de vocabulario único y exclusivo (*Only in*)

En los análisis realizados en el capítulo anterior para establecer una hipótesis de trabajo con respecto a *La cárcel de Sevilla*, calculamos el nivel de similitud textual que mantenía con cada uno de los entremeses cervantinos (Gráfico 26). A pesar de que ningún par textual superaba ese 40% considerado como significativo para empezar a sospechar de la existencia de un idiolecto común, consignamos la hipótesis de la autoría cervantina por el comportamiento que mantenía en el resto de variantes analizadas (PCA, Eder's Delta y *verbatim*).

A fin de ahondar en este análisis, mostramos en el siguiente gráfico el nivel de similitud textual que mantiene este entremés con todos los indubitados en prosa, incluidos los anónimos. Cabe recordar que estos resultados se obtienen tras aplicar una *stopword list* de palabras funcionales del Siglo de Oro (véase apartado 5.4.2.):

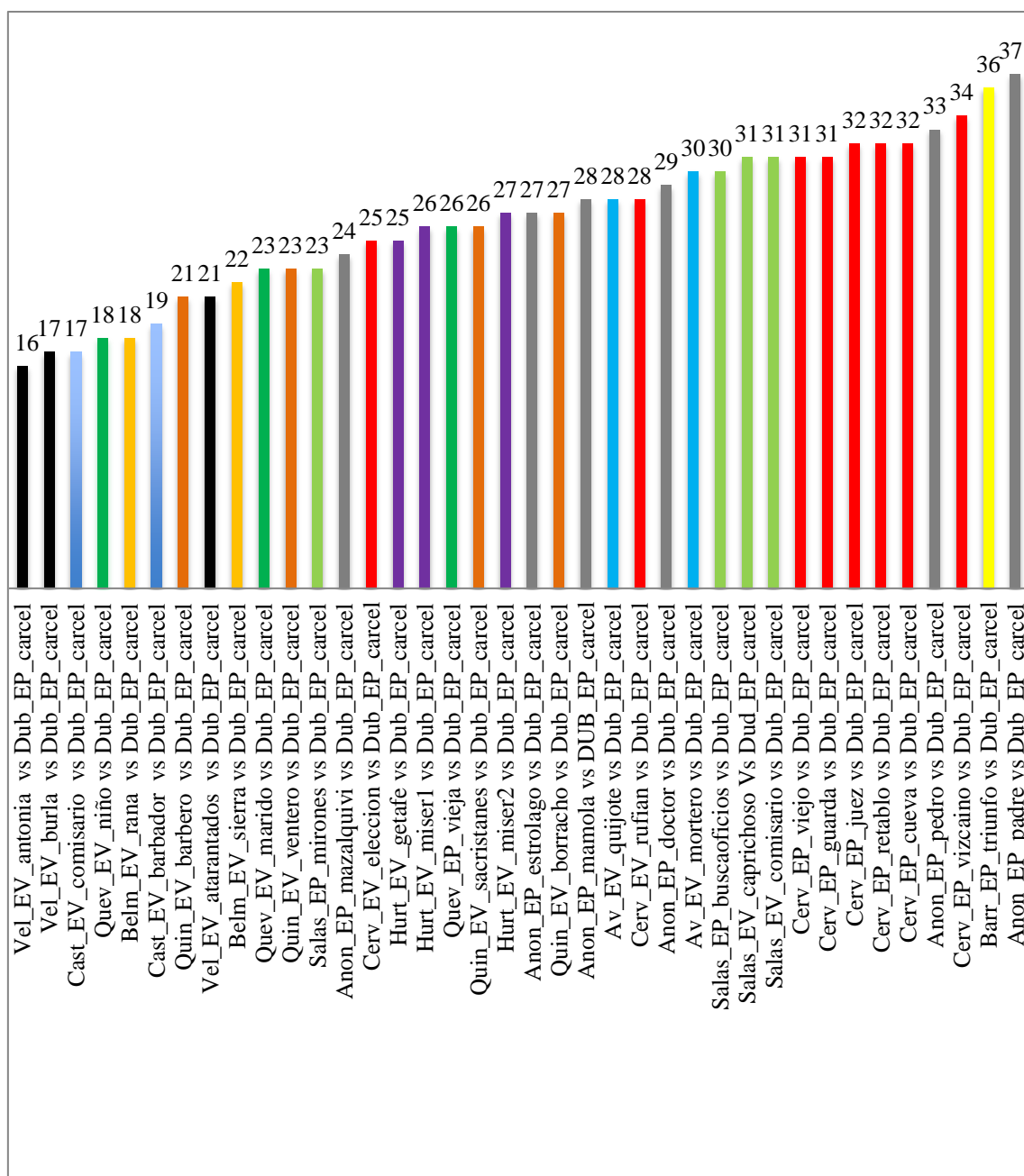


Gráfico 44. Nivel de similitud textual entre *La cárcel de Sevilla* y los entremeses indubitados (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

CAPÍTULO 6

Aunque ningún par formado por *La cárcel de Sevilla* y otro entremés indubitado alcanza el 40%, seis de las diez obras con mayores porcentajes pertenecen a Cervantes (concretamente, sus seis entremeses en prosa). Mostramos a continuación un análisis pormenorizado de los tres pares más significativos (*El padre engañado*, *El triunfo de los coches* y *El vizcaíno fingido*, con un 37, 36 y 34% respectivamente), con el objetivo de conocer los porcentajes exactos de vocabulario compartido una sola vez (*once only in both*) y más de una vez (*shared > 1*) y de vocabulario único o exclusivo (*only in*).

Anon_EP_padre vs Dub_EP_carcel				
	Texto	Shared %	Content	Function
Total	Anon_EP_padre		1196	1477
Total	Dub_EP_carcel		1549	1693
Shared > 1	Anon_EP_padre	45%	542	1416
Shared > 1	Dub_EP_carcel	27%	423	1565
Only in both	Anon_EP_padre	4%	45	9
Only in both	Dub_EP_carcel	3%	45	9
Only in	Anon_EP_padre	51%	609	52
Only in	Dub_EP_carcel	70%	1081	119

Tabla 23. Estadísticas de *El padre engañado* vs *La cárcel de Sevilla*
(CopyCatch Gold. Statistics)

Barr_EP_triunfo vs Dub_EP_carcel				
	Texto	Shared %	Content	Function
Total	Barr_EP_coches		2019	2676
Total	Dub_EP_carcel		1549	1693
Shared > 1	Barr_EP_coches	38%	759	2538
Shared > 1	Dub_EP_carcel	30%	471	1608
Only in both	Barr_EP_coches	3%	62	16
Only in both	Dub_EP_carcel	4%	62	16
Only in	Barr_EP_coches	59%	1198	122
Only in	Dub_EP_carcel	66%	1016	69

Tabla 24. Estadísticas de *El triunfo de los coches* vs *La cárcel de Sevilla*
(CopyCatch Gold. Statistics)

Cerv_EP_vizcaino vs Dub_EP_carcel				
	Texto	Shared %	Content	Function
Total	Cerv_EP_vizcaino		1920	2497
Total	Dub_EP_carcel		1549	1693
Shared > 1	Cerv_EP_vizcaino	35%	667	2 371
Shared > 1	Dub_EP_carcel	30%	470	1606
Only in both	Cerv_EP_vizcaino	3%	61	11
Only in both	Dub_EP_carcel	4%	61	11
Only in	Cerv_EP_vizcaino	62%	1192	115
Only in	Dub_EP_carcel	66%	1018	76

Tabla 25. Estadísticas de *El vizcaíno fingido* vs *La cárcel de Sevilla*
(CopyCatch Gold. Statistics)

A pesar de que el porcentaje de similitud textual más alto se da en el par Anon_EP_padre vs Dub_EP_carcel (37%), un análisis desglosado de estos tres pares permite advertir que *La cárcel de Sevilla* presenta menor vocabulario único con respecto a *El vizcaíno fingido* y *El triunfo de los coches* (66% en los dos casos) que con respecto a *El padre engañado* (70%). Ocurre de igual manera con el vocabulario compartido una sola vez —que es más significativo, por su exclusividad, que el compartido más de una vez—, pues es del 3% con respecto al entremés anónimo y del 4% con respecto a los entremeses de Barrionuevo y Cervantes.

Antes de pasar al siguiente apartado, vamos a analizar el porcentaje de similitud textual que existe entre *La cárcel de Sevilla* y las tres partes de la *Relación* (Chav_R_carcel1, Chav_R_carcel2 y Anon_R_carcel3), ya que, como apuntaban Hernández Alonso y Sanz Alonso (y como hemos podido comprobar en el apartado 6.3.1.), el autor de este entremés dubitado tuvo que tener delante las dos primeras partes cuando redactó su pieza, ya que en ocasiones “es un calco fidelísimo” de aquéllas [1999: 399]. También analizaremos, por su utilidad interpretativa, la similitud textual que existe entre las tres partes que conforman la *Relación*:

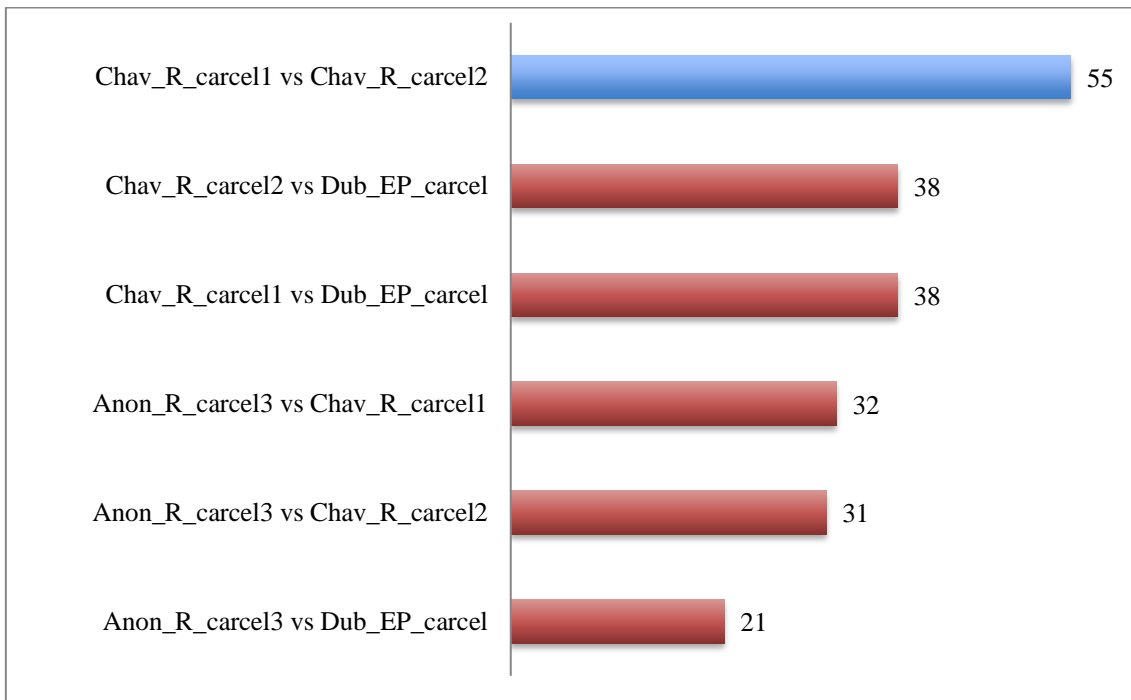


Gráfico 45. Nivel de similitud textual entre *La cárcel de Sevilla* y las tres partes de la *Relación* (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

A pesar de que todas las obras abordan una misma temática, y a pesar de que la fuente documental del *Entremés de la cárcel de Sevilla* son las dos primeras partes de la *Relación* (las que pertenecen a Chaves), el único par textual que supera el 40% es el conformado por las dos obras del procurador de la Audiencia de Sevilla (Chav_R_carcel1 vs Chav_R_carcel2), a las que las vincula un mismo idiolecto (55%). Aunque no es el objeto de este estudio, el bajo porcentaje existente entre la tercera parte de la relación y el entremés, ambos atribuidos a Cervantes, no invitan a pensar que sean de un mismo autor (21%).

6.3.8. Clasificación secuencial de aprendizaje automático (Rolling classify)

En este apartado vamos a utilizar Rolling classify, un método estilométrico que combina la clasificación de aprendizaje automático supervisado con el análisis secuencial, bajo la idea de que el estudio de los elementos que componen un texto es tan importante como el texto en sí mismo.

La función Rolling Classify (Stylo) divide el texto de prueba en bloques o porciones consecutivas de igual tamaño (*slices*) y realiza una clasificación supervisada de éstas en relación con los textos y/o autores contenidos en un corpus de entrenamiento previamente establecido. Como explica Eder, a diferencia de otros procedimientos, cuyo objetivo es evaluar las similitudes o diferencias estilísticas entre una serie de textos, Rolling classify “supported with compact visualization, tries to look inside a text represented as a set of linearly sliced chunks, in order to test their stylistic consistency” [Eder, 2016: 107]. Para poder interpretar correctamente el resultado gráfico, debemos saber que la franja primaria es la que figura en la parte inferior y que su nivel de certeza viene indicado por el grosor de la misma.

En el campo de la Estilometría, Rolling Classify ha sido utilizado principalmente para detectar autorías mixtas, aunque también puede usarse con otras finalidades. Nosotros lo hemos aplicado recientemente al estudio de los tiempos de composición de *Los santos inocentes* de Miguel Delibes [Celma Valero y Ruiz Urbón, 2021] y, en este caso, vamos a hacerlo para confirmar si realmente podemos hablar de una autoría única o si, por el contrario, la influencia de la obra de Chaves ha dejado una huella relevante en el *Entremés de la cárcel de Sevilla*.

Para ello, creamos un corpus de referencia (*reference set*) con los entremeses indubitados (excepto los anónimos) y las dos partes de la *Relación de la cárcel de Sevilla*, agrupados por autor, y un corpus de prueba (*test set*) con el entremés dubitado. Seleccionamos las 1000 palabras más frecuentes y pedimos a Rolling classify que segmente el texto objeto de estudio en muestras de 250 palabras (*slice size*), permitiendo una superposición de 200 (*slice overlap*). El algoritmo que utilizaremos en esta ocasión será Support Vector Machines (SVM)²⁰³.

²⁰³ Stylo permite la aplicación de Rolling classify con Delta (Classic Delta), Nearest Shrunken Centroids (NSC) y Support Vector Machines (SVM). Eder, en “Rolling stylometry” [2016], utiliza cada uno ellos al estudio de un caso concreto: Delta para una novela escrita en colaboración por Joseph Conrad y Ford Madox Ford en 1901; NSC para una traducción de la Biblia al polaco de mediados del siglo XV y SVM para un poema alegórico francés del siglo XIII.

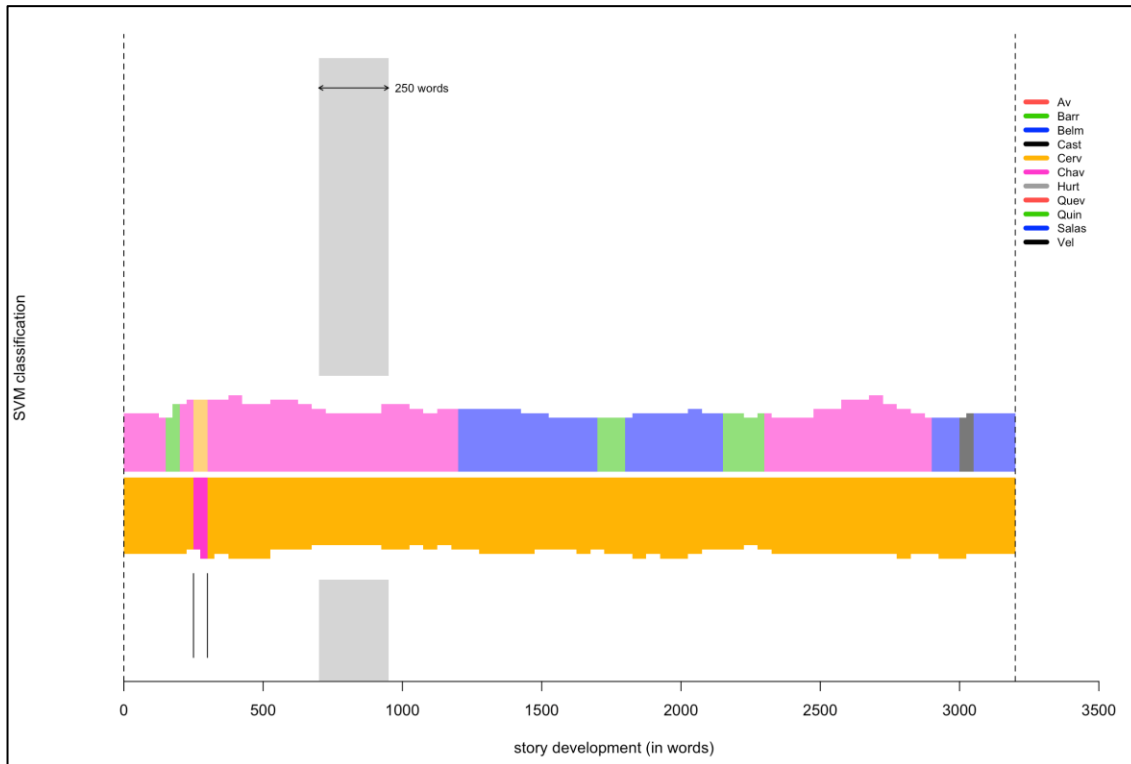


Gráfico 46. Clasificación secuencial de aprendizaje automático de *La cárcel de Sevilla* (RStudio. Stylo. Rolling classify. SVM. 1000 MFW. Slice size = 250. Slice.overlap = 200)

Como muestra el Gráfico 46, Rolling classify determina que *La cárcel de Sevilla* ha sido escrita mayoritariamente por una sola mano, que identifica de forma rotunda con Cervantes (color naranja), aunque señala una pequeñísima porción, en torno a las 230-280 primeras palabras del texto, que no parece identificarse con él, sino con Chaves (color rosa)²⁰⁴. Esta zona textual se corresponde con el momento en el que se inserta el romance de germanía “Alta mar esquiva”. Asimismo, el análisis secuencial determina los momentos en los que Chaves deja su huella en un segundo plano y que se corresponden con algunos de los pasajes señalados en nuestra Tabla 15 (por ejemplo, la comunicación de la pena de muerte y de la apelación se da entre las 950 y las 1200 palabras y la intervención de Escarramán a partir de las 2600 palabras).

²⁰⁴ Es importante advertir que el hecho de que Rolling classify identifique esa porción de texto con Chaves no indica necesariamente que este sea el autor del romance, sino que es el autor que más se aproxima en su comportamiento a él.

6.3.9. Verificación de autoría con General Imposters (GI)

Otra de las funcionalidades del paquete Stylo es el método General Imposters (GI), introducido por Koppel y Winter [2014] y aplicado poco después al estudio de varios textos dubitados de Julio César [Kestemont, Stover, Koppel, Karsdorp y Daelemans, 2016a]. El objetivo de la función GI:

is not to assess whether two documents are simply similar in writing style, given a static feature vocabulary, but rather, it aims to assess whether two documents are significantly more similar to one another than other documents, across a variety of stochastically impaired feature spaces (Eder, 2012; Stamatatos, 2006), and compared to random selections of so-called distractor authors (Juola, 2015), also called ‘imposters’ [Kestemont, Stover, Koppel, Karsdorp y Daelemans, 2016a: 88].

Este análisis de verificación de autoría —que trabaja con una métrica de “segundo orden”— compara, en varias iteraciones, un texto dubitado D con un conjunto de textos de varios autores distintos, a fin de determinar si alguno de esos autores pudo escribir D. A cada caso se le asigna una puntuación entre 0 y 1; en términos teóricos, cualquier puntuación superior a 0,5 sugeriría la verificación de autoría, pero lo cierto es que cualquier puntuación entre 0,39 y 0,63 debe ser considerada dudosa, pues indica que el clasificador tuvo problemas para tomar decisiones claras.

Este análisis es especialmente interesante en nuestro caso, ya que permitirá demostrar si las similitudes que estamos detectando hasta el momento se deben a que realmente Cervantes es el autor del texto dubitado o si, por el contrario, sólo es el autor que más se aproxima de nuestro corpus (*nearest neighbor problem*).

Para aplicar el método GI al estudio de autoría de *La cárcel de Sevilla*, configuramos un corpus de análisis en el que incluimos el texto dubitado, los entremeses de Ávila, Barrionuevo, Belmonte, Cervantes, Hurtado de Mendoza, Quevedo, Quiñones de Benavente, Salas Barbadillo y Vélez de Guevara y las dos partes de la *Relación* de Chaves. Siguiendo a Kestemont, Stover, Koppel, Karsdorp y Daelemans, vamos a elegir el algoritmo Minmax (Ruzicka Distance), ya que:

the results demonstrate the Ruzicka minmax metric returns the most stable results across the experiments and clearly has a regularizing effect across different hyperparametrizations [2016b: 248].

En la siguiente tabla mostramos los resultados para 500, 800 y 1000 MFW:

	Av	Barr	Belm	Cast	Cerv	Chav	Hurt	Quev	Quin	Salas	Vel
500 MFW	0.20	0.00	0.00	0.00	0.98	0.01	0.00	0.37	0.2	0.04	0.01
800 MFW	0.23	0.01	0.00	0.00	1.00	0.00	0.01	0.32	0.18	0.07	0.00
1000 MFW	0.18	0.01	0.00	0.00	1.00	0.02	0.00	0.35	0.18	0.06	0.01

Tabla 26. Aplicación de GI a *La cárcel de Sevilla*
(RStudio. Stylo. General Imposters. Ruzicka Delta distance. 500, 800 y 1000 MFW)

La función identifica a Cervantes como autor de *La cárcel de Sevilla* en todas las frecuencias analizadas (500, 800 y 1000 MFW), con un nivel de similitud entre 0,98 y 1; el resto de autores contemplados, por el contrario, no superan el 0,39.

6.3.10. Longitud de secuencias de palabras idénticas (*verbatim*)

Aunque hay muchos casos de *verbatim* de 2, 3, 4 y hasta 5 ítems entre *La cárcel de Sevilla* y las obras de Cervantes (Cerv_total), del tipo “cantan adentro”, “trato airado”²⁰⁵, “mal haya”, “me parece que”, “plega a Dios”, “que son amigos”, “que se acaba”, “la principal causa”, “os vienen a”, “este ladrón de”, “no es este tiempo”, “tengo dada la palabra”²⁰⁶, “lo que dijo el”, “que yo también he”, “no hay qué decir”, “que no ha de haber”, etc., el único *verbatim* realmente significativo es el que habíamos detectado en el capítulo anterior al hacer las comprobaciones previas para formular nuestras hipótesis de trabajo: “eso haré yo de muy buena gana” (7 ítems) (Tabla 12).

²⁰⁵ CORDE (1547-1617) sólo registra su uso en *La cárcel de Sevilla* (Dub_EP_carcel), Cervantes (Cerv_C_rufian) y en los romances de germanía de Juan Hidalgo.

²⁰⁶ CORDE (1547-1617) sólo registra su uso en *La cárcel de Sevilla* (Dub_EP_carcel) y Cervantes (Cerv_N_QUIJOTE1).

ESTUDIO DE AUTORÍA DEL *ENTREMÉS DE LA CÁRCEL DE SEVILLA*

AUTOR	OBRA	FRAGMENTO
Autor dubitado	<i>La cárcel de Sevilla</i>	ESCARRAMÁN. Eso haré yo de muy buena gana.
		- Eso haré yo de muy buena gana —dijo Vivaldo.
	<i>Quijote (I)</i>	- Eso haré yo de muy buena gana , señor mío —respondió Sancho—, y volvamos a mi aldea en compañía destes señores que su bien desean, y allí daremos orden de hacer otra salida que nos sea de más provecho y fama.
		- Eso haré yo de muy buena gana , señor mío —respondió Sancho—, con condición que vuestra merced no se ha de enojar de lo que dijere, pues quiere que lo diga en cueros, sin vestirlo de otras ropas de aquellas con que llegaron a mi noticia.
	<i>Quijote (II)</i>	- Eso haré yo de muy buena gana , señor mío —respondió la moza, que mostraba ser de edad de catorce años, poco más a menos.
Miguel de Cervantes		- Eso haré yo de muy buena gana —respondió don Álvaro—, puesto que cause admiración ver dos don Quijotes y dos Sanchos a un mismo tiempo tan conformes en los nombres como diferentes en las acciones; y vuelvo a decir y me afirmo que no he visto lo que he visto, ni ha pasado por mí lo que ha pasado.
	<i>Rinconete y Cortadillo</i>	- Eso haré yo de muy buena gana —respondió Monipodio.
	<i>El vizcaíno fingido</i>	CRISTINA. Eso haré yo de muy buena gana ; y aun, si el señor vecino quiere mi casa y cuanto hay en ella, aquí la hallará sola y desembarazada; que bien sé en qué caen estos negocios.
	<i>El juez de los divorcios</i>	JUEZ. Eso haré yo de muy buena gana ; y pluguiese a Dios que todos los presentes se apaciguasen como ellos.
Lope de Vega	<i>Pastores de Belén</i>	“ Eso haré yo de muy buena gana , dijo el pastor, aunque os prometo que las lenguas de los ángeles quedarán cortas, cuanto más la de tan rudo cronista; echad las ovejas por aquellos tomillos para que nos den más lugar y silencio, entretenidas en ellos, y estadme atentas”.
Jorge de Montemayor	<i>Los siete libros de la Diana</i>	“ Eso haré yo de muy buena gana —dijo Arsileo— porque ningún servicio se os puede hacer que no quepa en vuestro merecimiento”.

Tabla 27. Presencia del verbatim “eso haré yo de muy buena gana” en CORDE entre 1547 y 1617

Como muestra la tabla precedente, CORDE registra el uso de este *verbatim* en 11 ocasiones entre 1547 y 1617: una en el entremés dubitado de *La cárcel de Sevilla*, ocho en Cervantes, una en Lope de Vega y una en Jorge de Montemayor. Por tanto, aunque el *verbatim* no es exclusivo de Cervantes, sí es muy característico de su escritura.

6.3.11. Coocurrencias y redes léxicas

Como explica Fradejas Rueda en *Cuentapalabras. Estilometría y análisis de texto con R para filólogos* [2020], resulta interesante analizar las palabras de un texto de manera aislada, pero también en relación con sus compañeras. Así, en Lingüística de corpus, se habla de coocurrencia para referirnos a las palabras que coaparecen a una cierta distancia, sin un orden exacto y estable, como, por ejemplo, los vocablos *platos*, *sartén* y *asador*, presentes dentro de una misma frase tanto en *La cárcel de Sevilla* como en *El rufián viudo*:

Dub_EP_carcel	Cerv_EP_rufian
PAISANO. Seora Torbellina, voacé será testigo o testiga, lo que mejor le pareciere, cómo a esta mujer la hago heredera de todos mis bienes, muebles y raíces, de mi calabozo. Item, de cuatro o cinco platos y escudillas, taladro, barreno, un candelero de barro, una sartén y un asador . Item, una manta y un jergón servicio y pulidor.	TRAMPAGOS. Vuelve por el mortero y por el banco, y el broquel no se olvide, Vademécum. VADEMÉCUM. Y aun traíré el asador , sartén y platos .

Tabla 28. Ejemplo de coocurrencia léxica

El análisis automatizado de textos con lenguaje R facilita la localización de coocurrencias. Hemos analizado las palabras que se relacionan con algunos de los sustantivos que comparten con mayor frecuencia *La cárcel de Sevilla* y los entremeses de Cervantes —“Dios”, “hombre”, “muerte”, “mujer”, “mundo” y “vida”, con un mínimo de 5 repeticiones en cada uno de los corpus (filter: n > 5)—, a fin de advertir posibles concomitancias (Gráficos 47 y 48):

ESTUDIO DE AUTORÍA DEL ENTREMÉS DE LA CÁRCEL DE SEVILLA

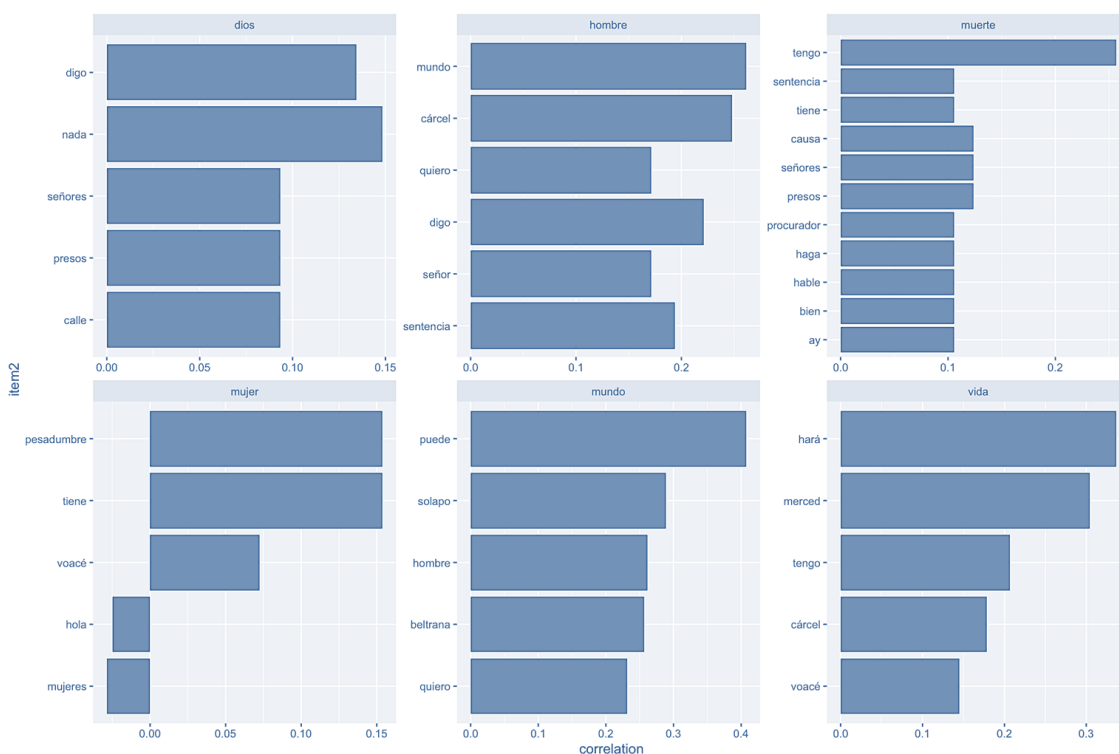


Gráfico 47. Cocurrencias de ciertos vocablos en *La cárcel de Sevilla* (RStudio)

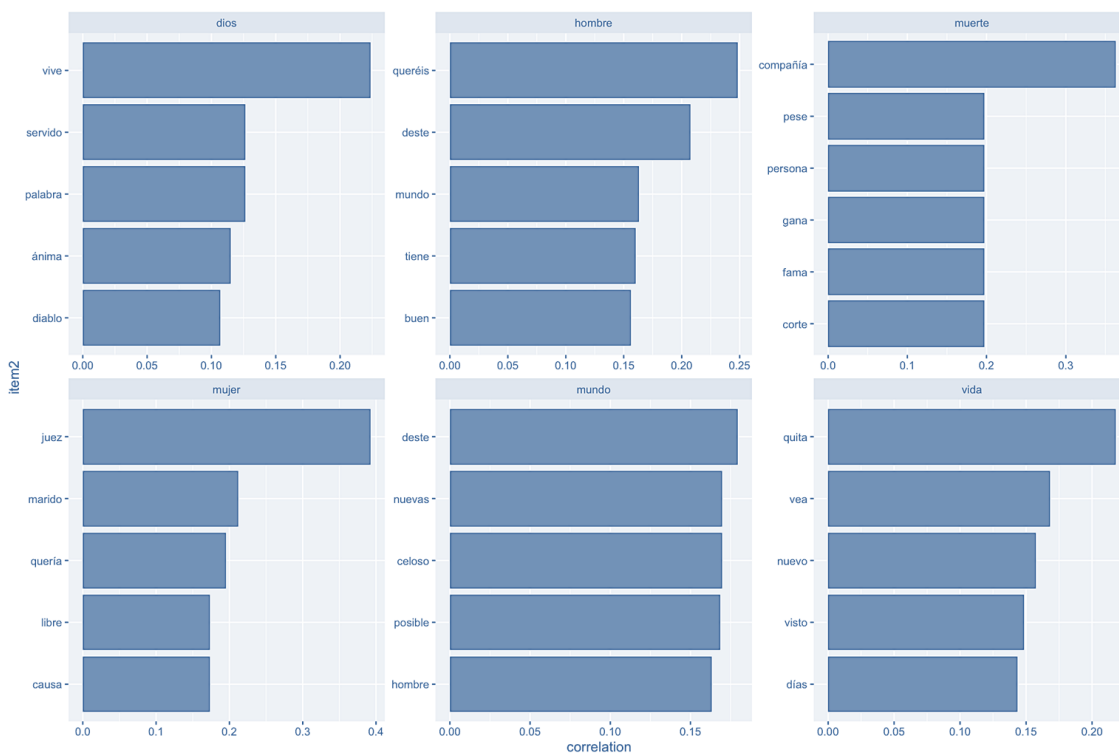


Gráfico 48. Cocurrencias de ciertos vocablos en los entremeses de Cervantes (RStudio)

Tal y como se desprende de la comparación de ambos gráficos, tanto en el entremés dubitado como en los entremeses de Cervantes hay cierta predilección a relacionar las palabras *hombre* y *mundo*, que coaparecen en varias ocasiones dentro de un mismo parlamento. Un análisis más detallado de esta coocurrencia nos muestra los siguientes datos:

	Nº de veces (n)		Coeficiente phi (correlation)	
	Dub	Cerv	Dub	Cerv
<i>hombre + mundo</i>	3	4	0.2616	0.1633
<i>mundo + hombre</i>				

Tabla 29. Análisis de la correlación *hombre + mundo* en *La cárcel de Sevilla* y los entremeses indubitados de Cervantes (Elaboración propia, a partir de los datos obtenidos en RStudio)

La coocurrencia *hombre + mundo* (o *mundo + hombre*) se da 3 veces en *La cárcel de Sevilla* y 4 en el total de entremeses indubitados cervantinos, lo que supone un coeficiente de correlación (o coeficiente phi) del 0,2616 para el primer caso y de 0,1633 para el segundo²⁰⁷. Siguiendo la tabla propuesta por Rumsey, la fuerza de la relación lineal entre estas dos palabras es ascendente, aunque muy débil, pues está relativamente próxima al 0 [Fradejas Rueda, 2020]. De hecho, esta coocurrencia no aparece en las redes de palabras (agrupaciones) que se producen con una correlación superior al 0,30 (correlación > 0.30):

²⁰⁷ El coeficiente de correlación o coeficiente phi permite medir la fuerza y la dirección de la relación lineal entre dos variables, en este caso, los vocablos *hombre* y *mundo* (es decir, cuántas veces aparecen juntos en relación con las veces en que lo hacen separados). Ambas variables pueden intercambiarse sin que el coeficiente varíe (la correlación *hombre + mundo* será igual que la de *mundo + hombre*). El valor del coeficiente de correlación está comprendido entre el -1 y el +1 (El 0 indicaría que no existe relación lineal alguna; cuanto más alejado de 0 sea el valor, mayor será la dependencia entre los dos vocablos, y a la inversa).

6.3.12. Estudio de las palabras de función: orden de frecuencia, frecuencia relativa y PCA

Las palabras de función son otro de los elementos clave para discriminar autorías, ya que escapan del control consciente del escritor y, además, su uso no viene condicionado por el tema del texto. Burrows, antes de desarrollar su método 'Delta' [2002], descubrió que algunos fragmentos interpolados en las novelas de Henry Fielding habían sido escritos en realidad por su hermana Sarah, basándose en el estudio de las palabras de función [1988]. Más tarde, su discípulo Craig [2009] resolvió siete casos de autoría dudosa relacionados con William Shakespeare utilizando, entre otros, un método basado en el PCA de las palabras de función (conocido hoy, en su honor, como método Craig)²⁰⁸.

En este apartado vamos a realizar una serie de estudios relacionados con este tipo de palabras sin significado léxico (preposiciones, conjunciones, determinantes, pronombres, verbos auxiliares etc.). Antes de realizar los análisis, es necesario tratar los textos para deshacer aquellas formas contractas (*al, del, deste, dello*, etc.) o enclíticas (como *cayóseme* o *doyle*) que puedan distorsionar los resultados.

Extraemos, en primer lugar, las diez palabras de función más frecuentes en *La cárcel de Sevilla* —que, en orden de mayor a menor frecuencia, son *que, de, y, la, el, no, en, a, voacé* y *me*— y las comparamos con las diez que aparecen en el conjunto de entremeses indubitados, agrupados por autores, a fin de comprobar si los comportamientos se asemejan o no (Tabla 30). Para facilitar la visualización, marcamos en verde las palabras que coinciden con el texto dubitado en la misma posición, en azul las que coinciden, pero en distinta posición, y en rojo las que no coinciden²⁰⁹.

²⁰⁸ Aunque citamos a Burrows y Craig por ser los responsables de estos análisis estilométricos basados en las palabras de función, sus trabajos fueron escritos en colaboración con reputados especialistas en historia literaria [Burrows y Hassal, 1988, y Craig y Kinney, 2009].

²⁰⁹ Los verbos copulativos están a medio camino entre las formas verbales con significado pleno, que indican una acción (*cantar*), un proceso (*aprender*) o un estado (*saber*), y las que solo tienen significado gramatical, como los verbos auxiliares. Hemos optado por incluir el verbo *ser* en este listado de palabras de función.

LISTADO DE LAS DIEZ PRIMERAS PALABRAS DE FUNCIÓN										
	1 ^a	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a	7 ^a	8 ^a	9 ^a	10 ^a
Dub	que	de	y	el	la	a	no	en	voacé	me
Av	que	de	y	el	a	la	en	no	por	es
Barr	que	y	de	a	no	v. m.	el	la	en	se
Belm	de	que	y	no	el	en	a	es	la	lo
Cast	de	el	que	y	en	a	no	un	con	la
Cerv	que	de	y	el	la	a	no	en	me	por
Hurt	de	que	el	y	en	a	no	la	es	que
Quev	de	y	que	el	la	a	no	en	lo	es
Quiñ	que	de	y	la	el	a	en	no	con	es
Salas	que	de	y	el	a	en	la	no	con	es
Vel	que	de	y	el	a	la	es	no	es	si

Tabla 30. Listado de las 10 primeras palabras de función
(Elaboración propia, a partir de los datos obtenidos en RStudio)

La tabla permite advertir que hay una alta coincidencia en la utilización de unas mismas palabras de función entre todos los entremesistas, pero, también, que no todos las usan con el mismo orden de frecuencia. De hecho, la cercanía más significativa se da entre el autor del texto dubitado y Cervantes, que comparten 9 de las 10 palabras de función más frecuentes, ocho de ellas en la misma posición²¹⁰. Para advertir la singularidad de este resultado, basta observar que ninguna otra pareja de entremesistas mantiene entre sí esos parecidos.

Vamos a calcular ahora la frecuencia relativa de las diez palabras de función más utilizadas en *La cárcel de Sevilla* y la frecuencia que tienen éstas en el resto de autores de nuestro corpus (Tabla 31). Para facilitar la visualización, acompañamos la tabla de un gráfico de barras para cada una de las palabras (Gráficos 51-60):

²¹⁰ Comentaremos, más adelante e l caso de *voacé*, que aparece en el texto dubitado en noveno lugar pero que en ninguno de los otros autores ocupa un lugar dentro del *top-ten*.

CAPÍTULO 6

FRECUENCIA RELATIVA DE LAS PALABRAS DE FUNCIÓN											
	Dub	Av	Barr	Belm	Cast	Cerv	Hurt	Quev	Quiñ	Salas	Vel
que	0.0549	0.0565	0.0549	0.0436	0.0360	0.0506	0.0411	0.0422	0.0418	0.0456	0.0452
de	0.0461	0.0469	0.0416	0.0488	0.0518	0.0502	0.0513	0.0505	0.0341	0.0445	0.0393
y	0.0451	0.0348	0.0469	0.0294	0.0338	0.0452	0.0330	0.0449	0.0341	0.0327	0.0344
el	0.0293	0.0322	0.0218	0.0220	0.0360	0.0262	0.0333	0.0320	0.0281	0.0316	0.0236
la	0.0281	0.0176	0.0180	0.0157	0.0125	0.0241	0.0160	0.0260	0.0286	0.0209	0.0209
a	0.0229	0.0262	0.0258	0.0163	0.0198	0.0239	0.0216	0.0253	0.0236	0.0243	0.0214
no	0.0198	0.0137	0.0249	0.0283	0.0151	0.0198	0.0214	0.0175	0.0151	0.0158	0.0184
en	0.0195	0.0154	0.0163	0.0199	0.0209	0.0193	0.0217	0.0166	0.0156	0.0213	0.0160
voacé	0.0134	0	0	0	0	0.0001	0	0	0	0	0
me	0.0122	0.0096	0.0076	0.0083	0.0073	0.0119	0.0057	0.0081	0.0125	0.0102	0.0105

Tabla 31. Frecuencia relativa de las 10 palabras de función más frecuentes en el texto dubitado (Elaboración propia, a partir de los datos obtenidos en RStudio)

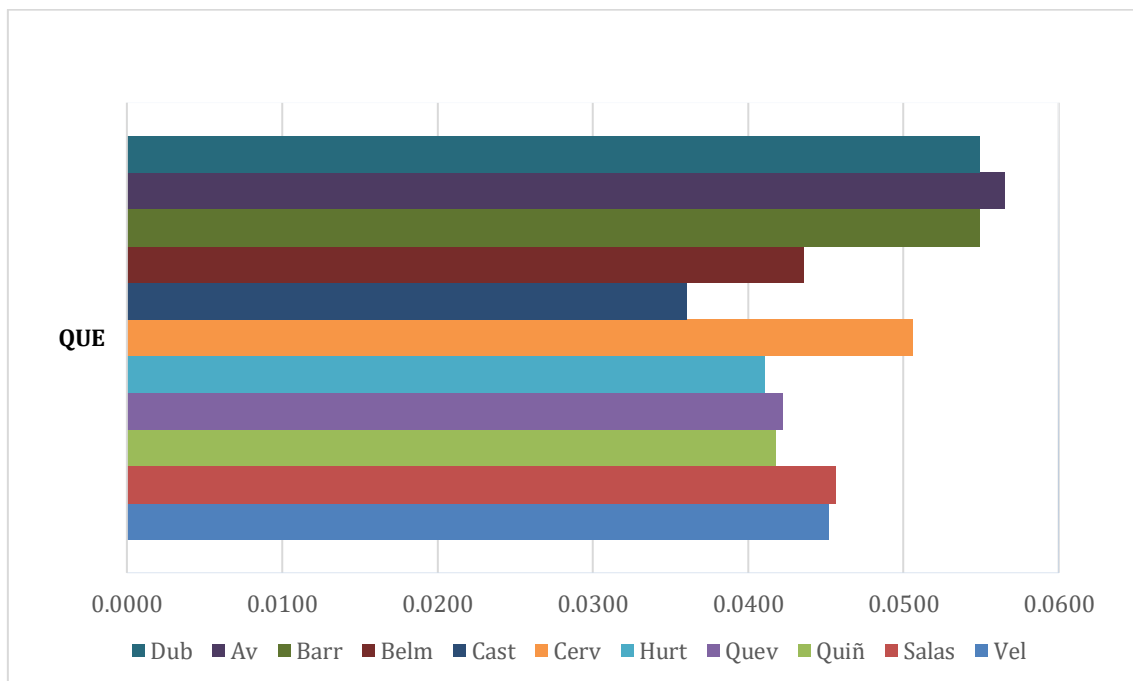


Gráfico 51. Análisis comparado de la frecuencia relativa de “que” (Elaboración propia, a partir de los datos obtenidos en RStudio)

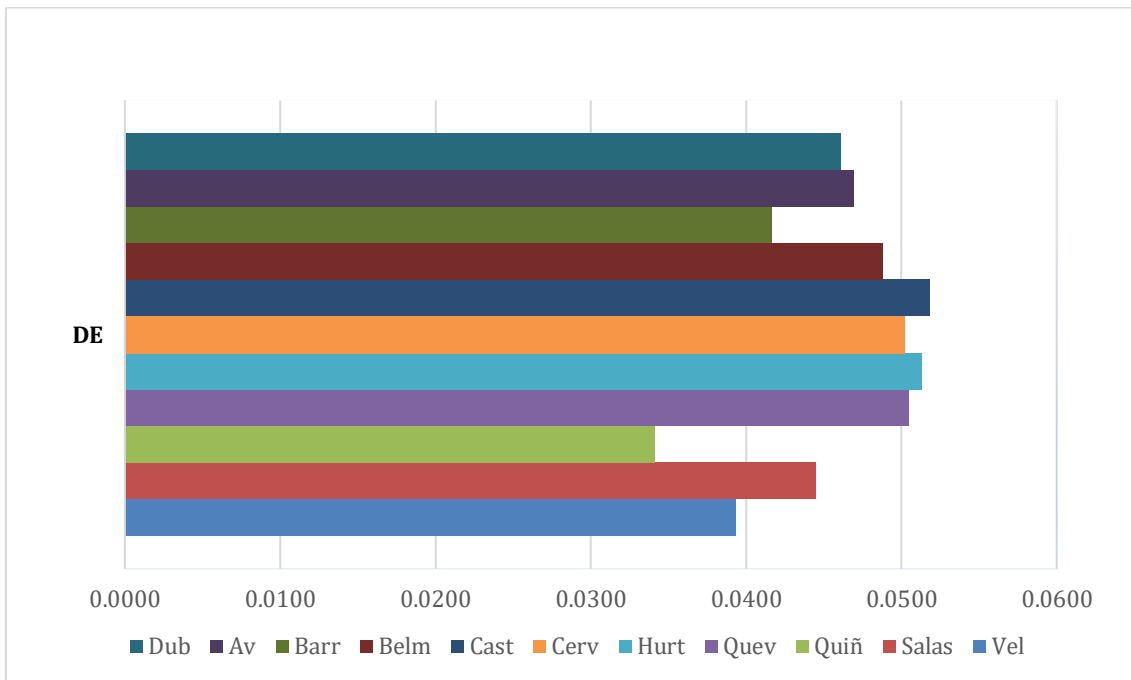


Gráfico 52. Análisis comparado de la frecuencia relativa de “de”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

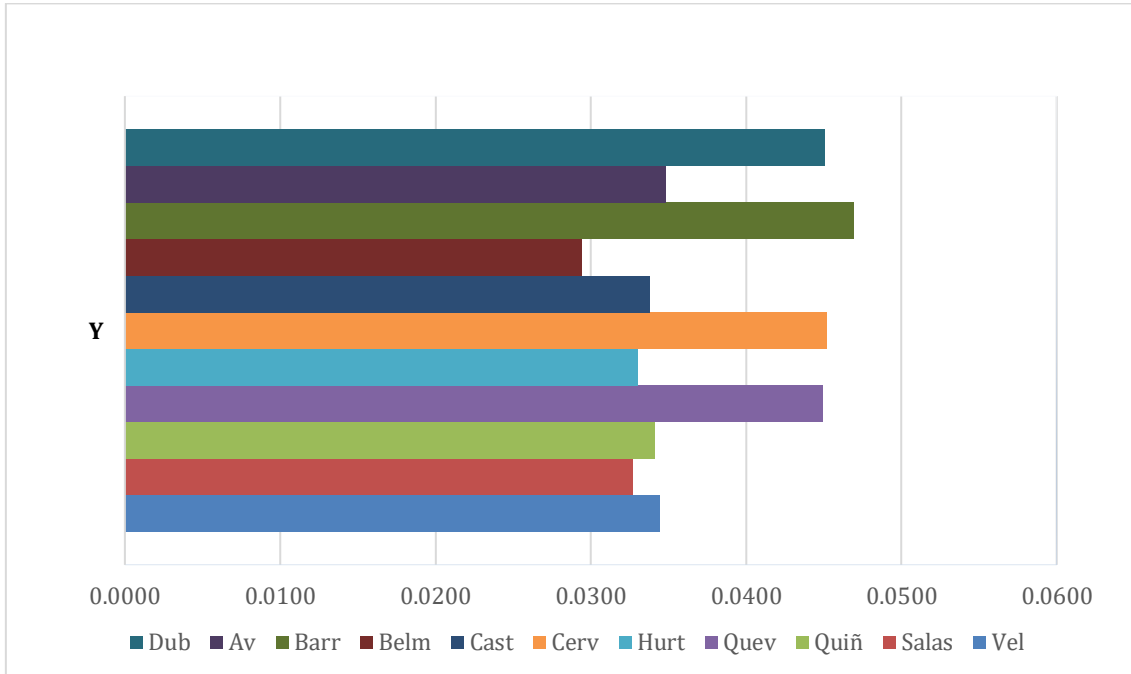


Gráfico 53. Análisis comparado de la frecuencia relativa de “y”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

CAPÍTULO 6

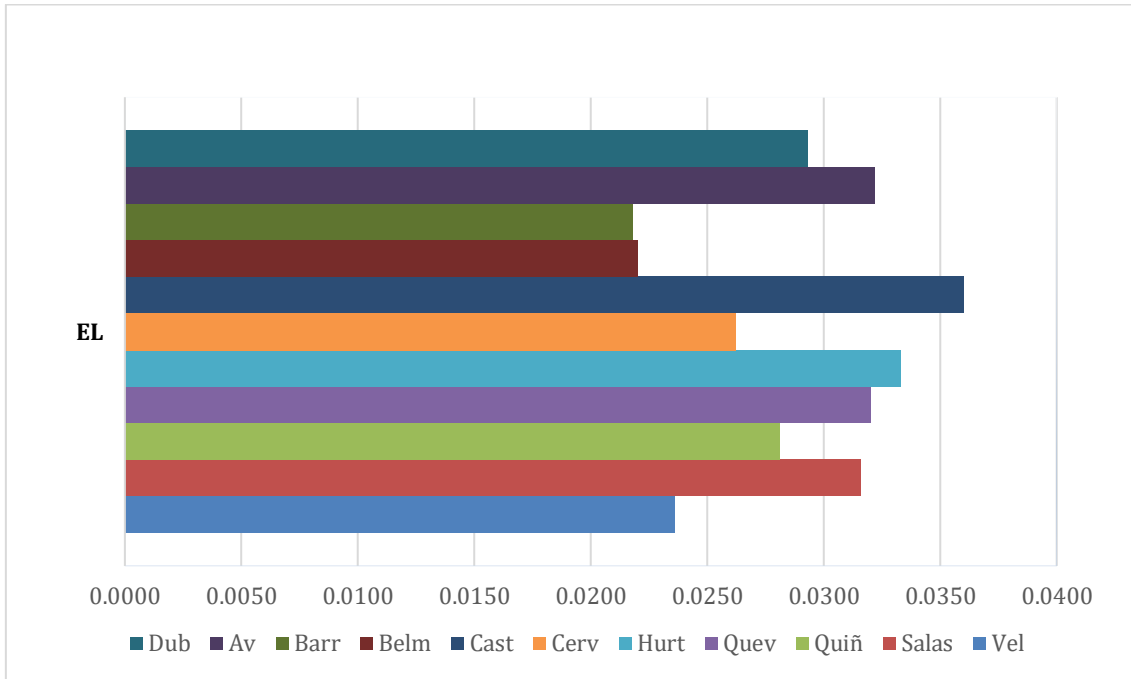


Gráfico 54. Análisis comparado de la frecuencia relativa de "el"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

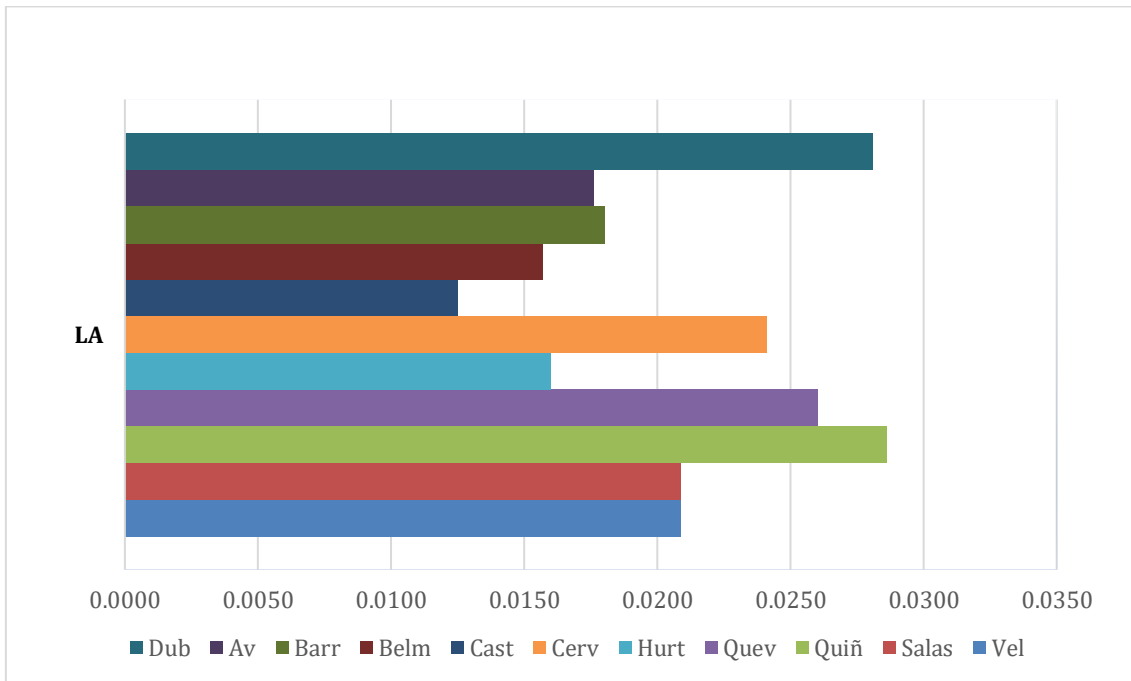


Gráfico 55. Análisis comparado de la frecuencia relativa de "la"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

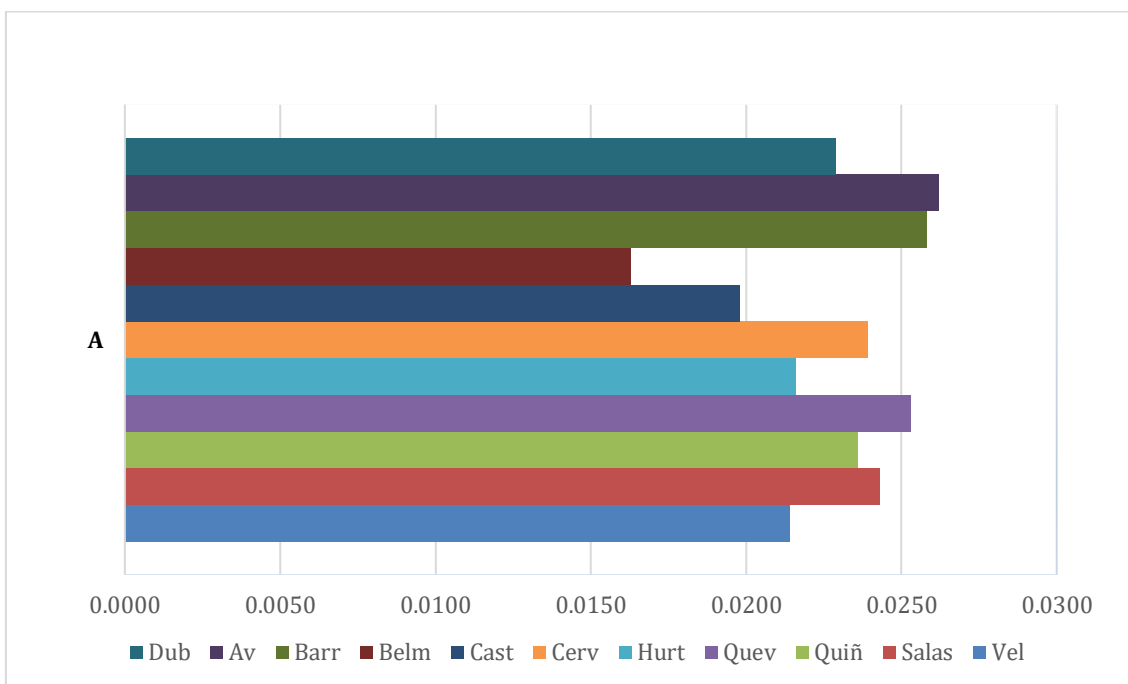


Gráfico 56. Análisis comparado de la frecuencia relativa de "a"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

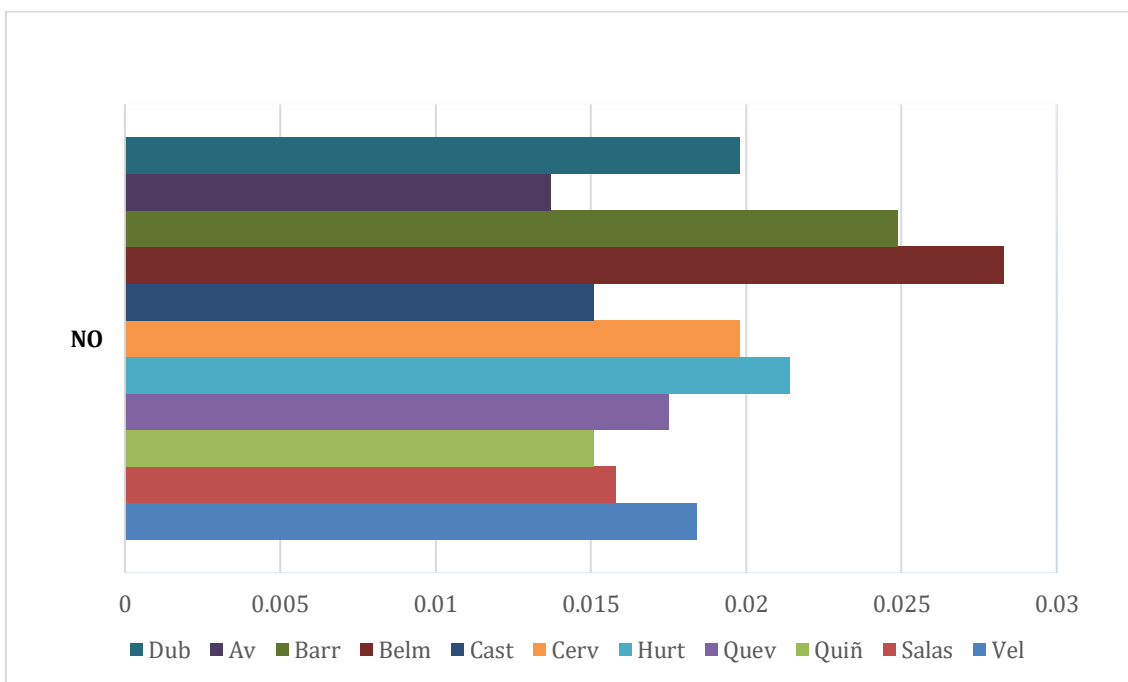


Gráfico 57. Análisis comparado de la frecuencia relativa de "no"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

CAPÍTULO 6

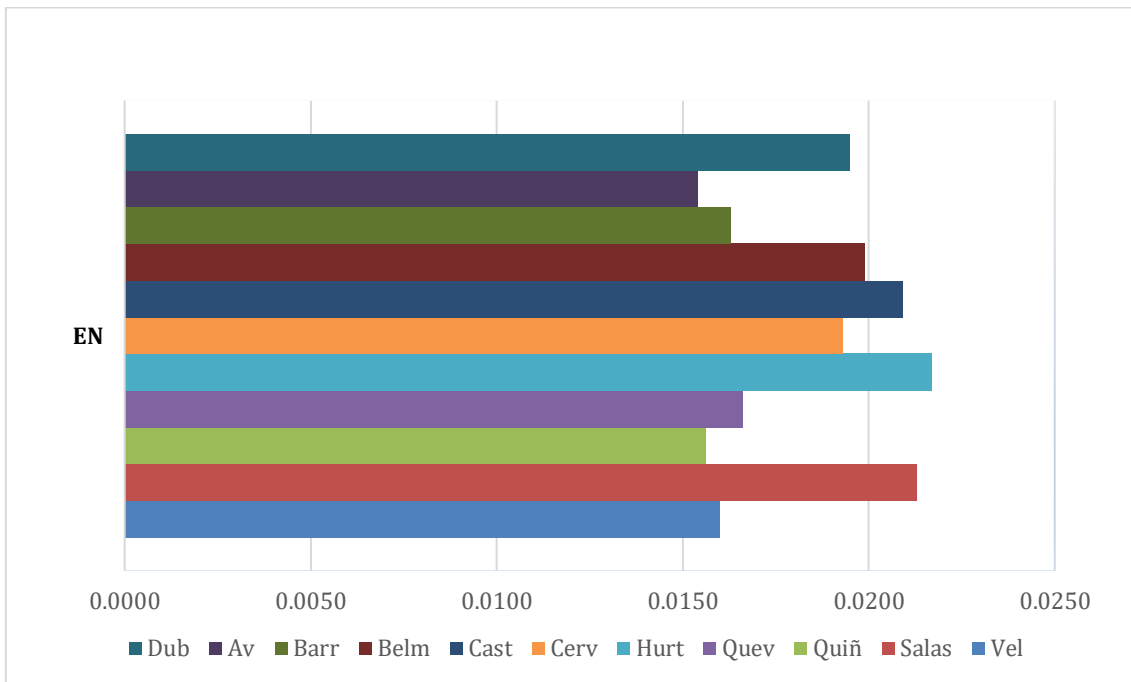


Gráfico 58. Análisis comparado de la frecuencia relativa de “en”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

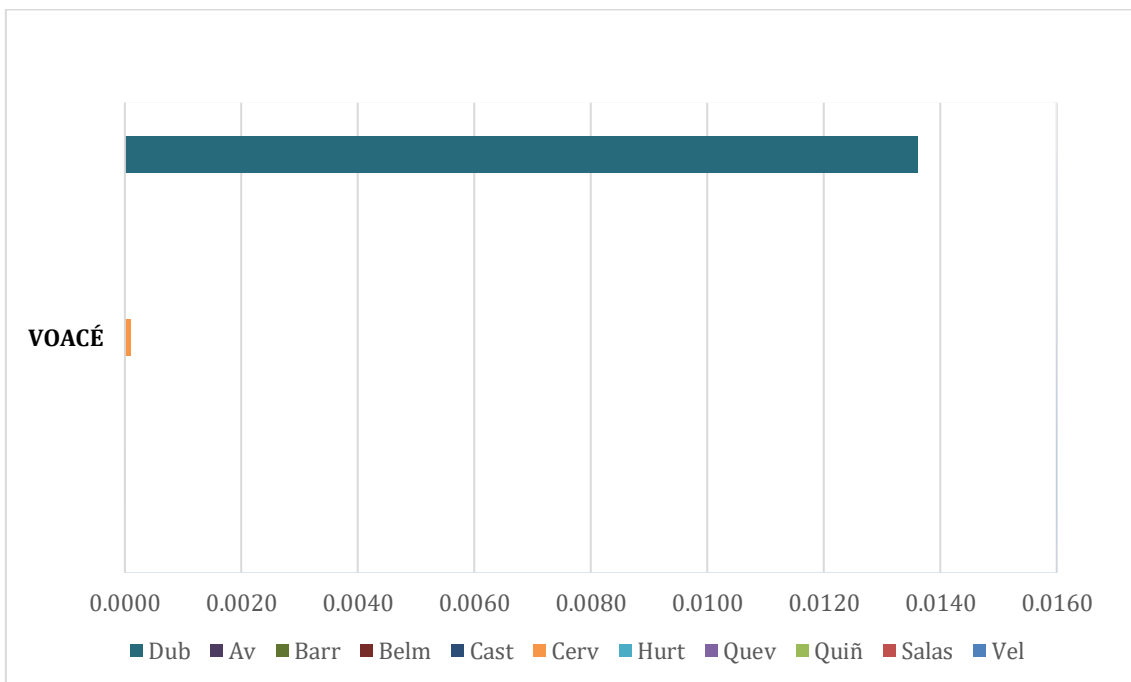


Gráfico 59. Análisis comparado de la frecuencia relativa de “voacé”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

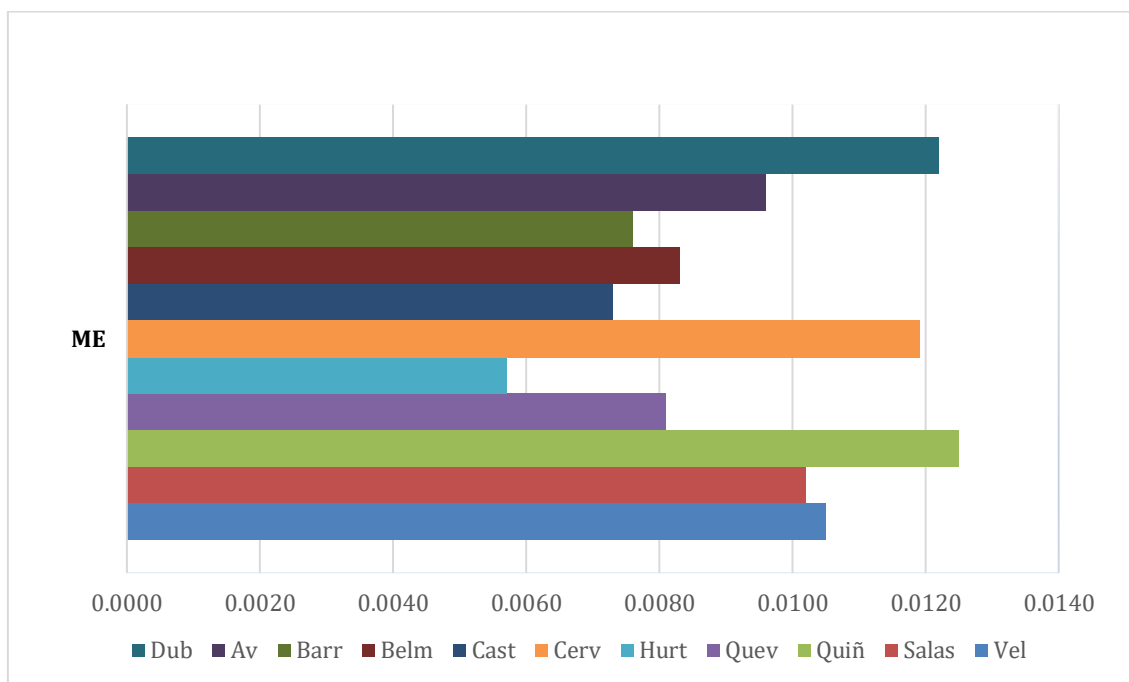


Gráfico 60. Análisis comparado de la frecuencia relativa de "me"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

Los diagramas de barras permiten advertir un comportamiento prácticamente idéntico entre el autor del texto dubitado y Cervantes en la frecuencia relativa de la conjunción copulativa "y", del adverbio "no", de las preposiciones "en" y "a" y del pronombre personal "me", así como frecuencias ciertamente cercanas en el resto de casos. Mención aparte merece el caso de "voacé", con una frecuencia de uso en *La cárcel de Sevilla* de 0,0134, en Cervantes de 0,0001 y de 0 en el resto de autores analizados. *Voacé* es la forma vulgar de *vusted*, usada en *La cárcel de Sevilla* para caracterizar la forma de hablar de bravos y rufianes (en contraposición, el alcaide usa "vuesa merced" para referirse al escribano), como también hace Cervantes en *El rufián viudo*, *Rinconete y Cortadillo* o la primera parte del *Quijote*, para caracterizar el habla de Ginés de Pasamonte²¹¹.

²¹¹ Quevedo, en sus *Capitulaciones de la vida de la Corte*, explica así el habla de los rufianes: "dicen voacé, so compadre, so camarada, y llaman media janega a la media azumbre".

Mostramos ahora el gráfico de tendencia de las diez palabras de manera conjunta para que puedan apreciarse mejor las conclusiones antes descritas:

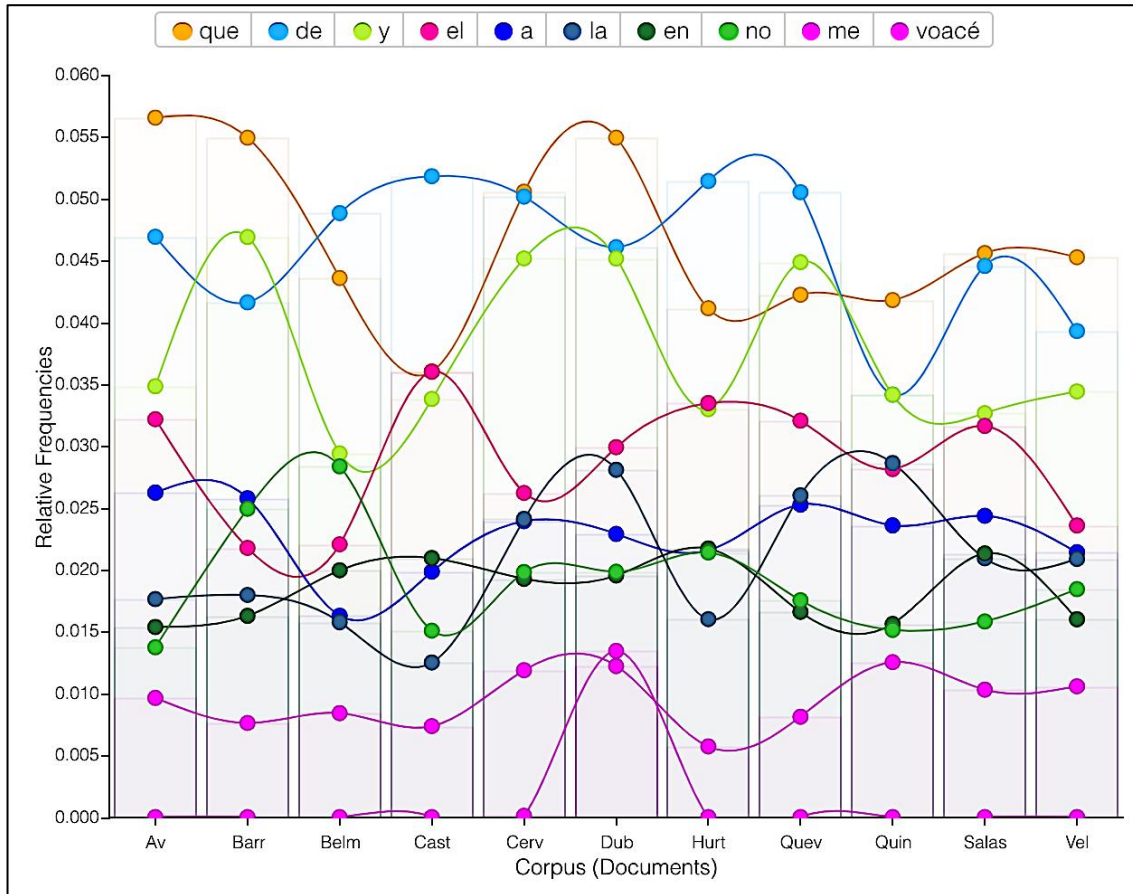


Gráfico 61. Análisis comparado de la frecuencia relativa de las diez primeras palabras de función en *La cárcel de Sevilla* (Voyant Tools. Tendencias)

Vamos a realizar ahora un PCA de la tabla de frecuencias relativas, con el objetivo de condensar la información en unos pocos componentes principales. En la primera visualización se muestra el gráfico de individuos (Gráfico 62); en la segunda, el gráfico de variables (Gráfico 63); y en la tercera, el biplot de individuos y variables (Gráfico 64)²¹². Para poder interpretarlos, debemos tener en cuenta las siguientes claves:

²¹² En estadística, un biplot es un tipo de gráfico exploratorio que sirve para representar tres o más variables en un espacio de dimensión reducida, normalmente de dimensión dos.

(1) Dub, (2) Av, (3) Barr, (4) Belm, (5) Cast, (6) Cerv, (7) Hurt, (8) Quev, (9) Quiñ, (10) Salas y (11) Vel.

El gráfico de individuos (autores) muestra la cercanía entre Ávila y Hurtado de Mendoza (2 y 8), entre el autor del texto dubitado y Cervantes (1 y 6) y entre Cervantes y Vélez de Guevara (1 y 11). Los autores más alejados del centro, es decir, los que representan una variabilidad más extrema, son Belmonte (4), Castillo de Solórzano (5) y Hurtado de Mendoza (7).

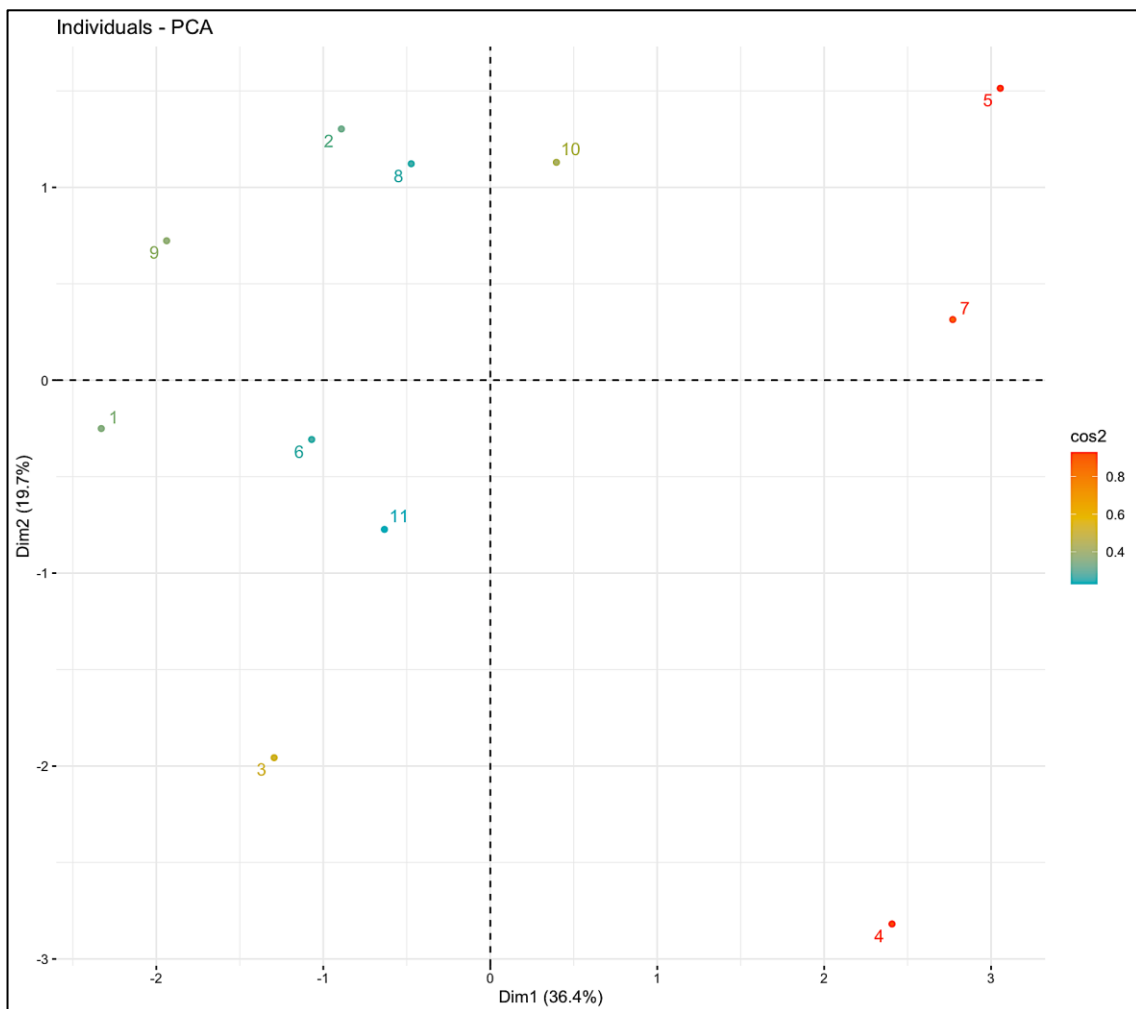


Gráfico 62. Visualización del PCA de las palabras de función. Gráfico de individuos.
(RStudio. PCA. Graph of individuals)

CAPÍTULO 6

El gráfico de variables ofrece la visualización de las diez palabras de función en vectores. Para interpretarlo, debemos fijarnos en la longitud y el ángulo de los vectores, que representan las variables, esto es, las palabras de función: cuanto más paralelo es un vector al eje de una componente, más ha contribuido a la creación de la misma; y cuanto más largo es, mayor es la variabilidad contenida en la representación de los dos componentes del biplot, es decir, mejor representada está su información en el gráfico.

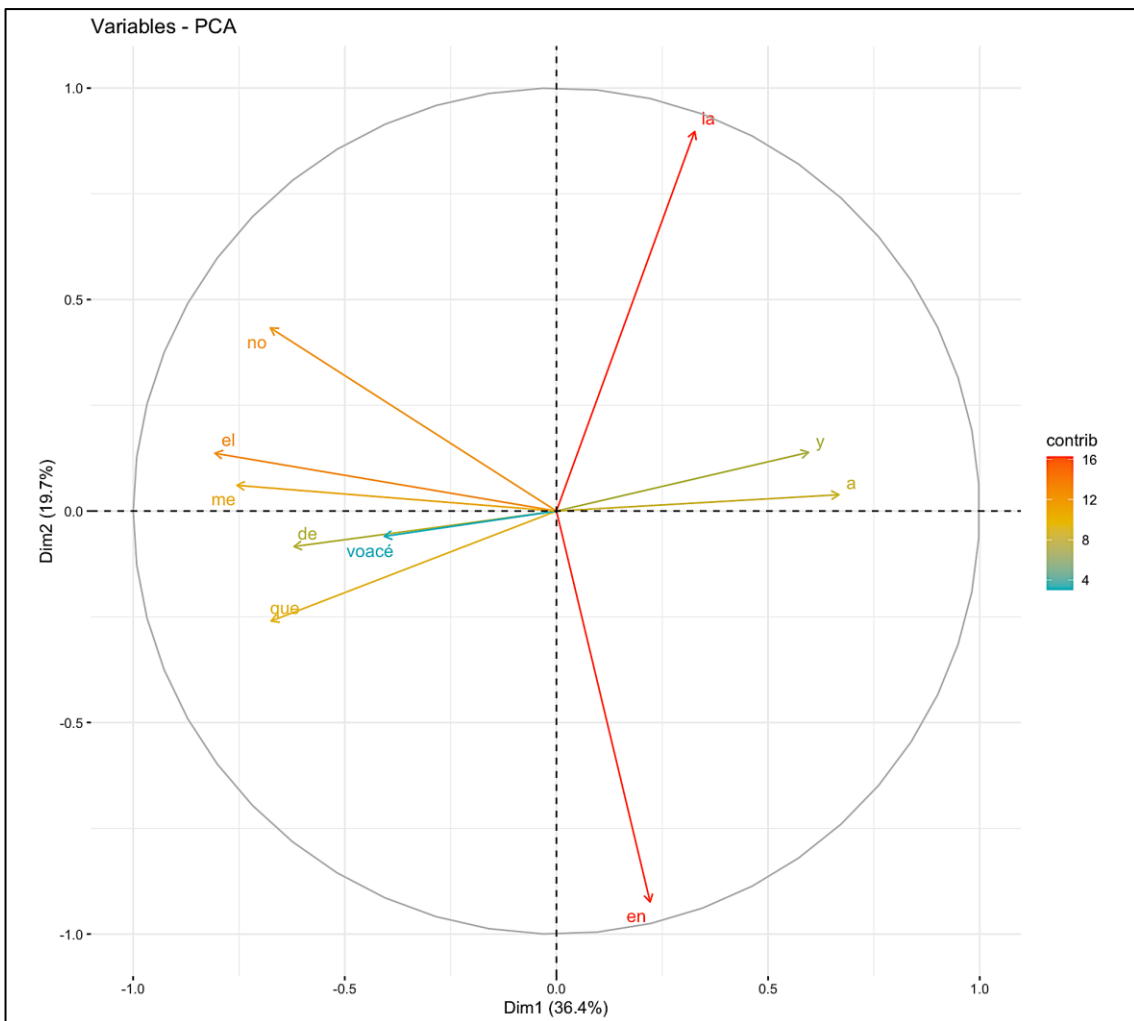


Gráfico 63. Visualización del PCA de las palabras de función. Gráfico de variables.
(RStudio. PCA. Graph of variables)

6.3.13. Ley de Zipf

La Ley de Zipf, formulada en 1932 por el lingüista norteamericano George Kingsley Zipf, hace referencia a una verdad empírica que se cumple en prácticamente todas las lenguas y que, en esencia, nos permite afirmar que la frecuencia con la que aparece una palabra es inversamente proporcional al rango que por frecuencia le corresponde en el listado de la totalidad de palabras de un texto²¹³.

En este apartado vamos a estudiar el comportamiento de la Ley de Zipf en *La cárcel de Sevilla* para, a continuación, comprobar si es similar al que presentan los entremeses en prosa de Cervantes (Gráfico 65):

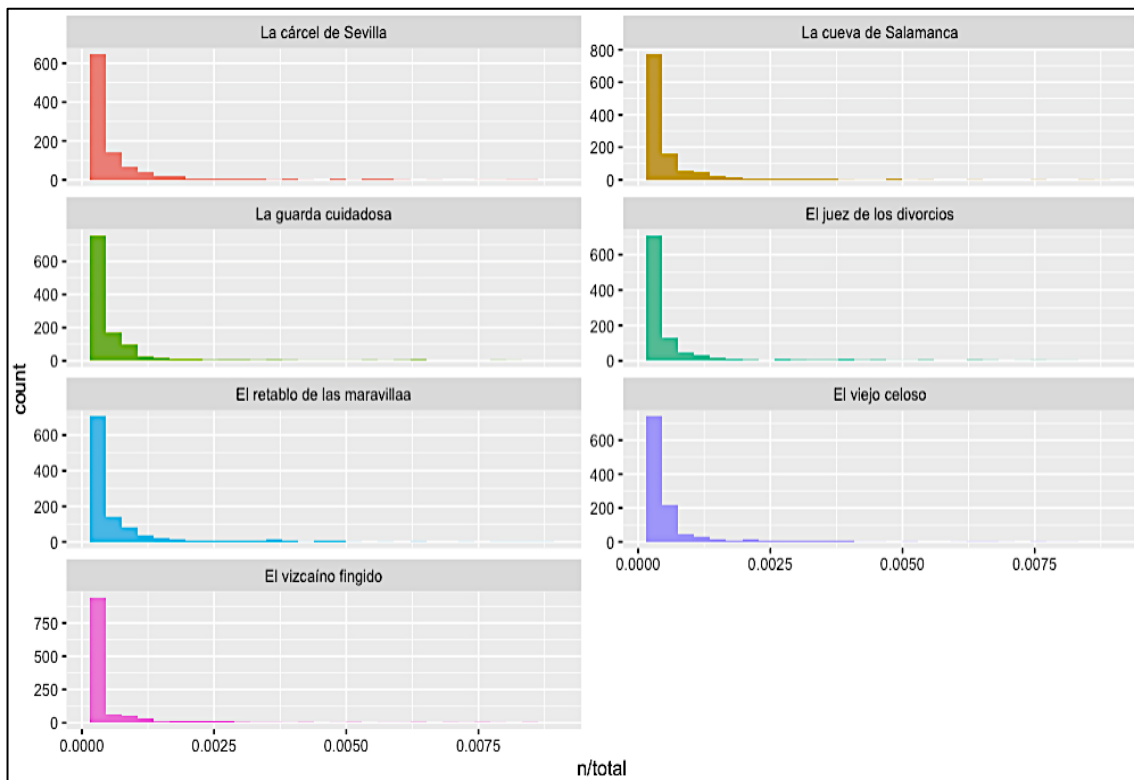


Gráfico 65. Visualización de la Ley Zipf en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes (RStudio. Histogramas)

²¹³ Basándose en esta regla, George Udny Yule intentó establecer una medida objetiva de la frecuencia de aparición de una palabra, conocida como la “característica K de Yule”, pero que ha resultado ineficaz para discriminar autorías.

La serie de histogramas recogida en el Gráfico 65 permite advertir un comportamiento análogo en la frecuencia con la que aparecen las palabras y su rango en el número de uso. Si trazamos una escala logarítmica, en donde el eje X sea el rango y el eje Y la frecuencia, observamos aún mejor esta similitud, sin que el *Entremés de la cárcel de Sevilla*, representado en color rojo, provoque distorsión alguna en el conjunto de entremeses en prosa de Cervantes (Gráfico 66).

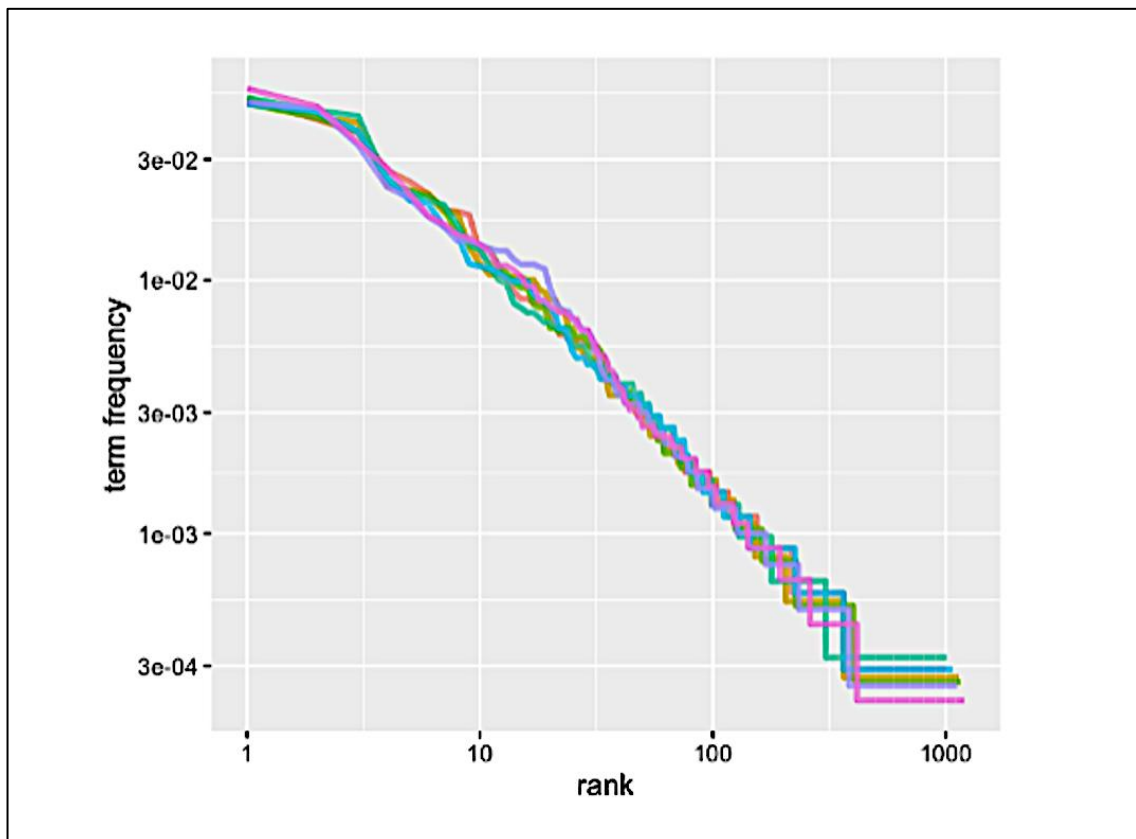


Gráfico 66. Visualización de la Ley Zipf en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes (RStudio. Escala logarítmica)

La relación entre rango y frecuencia tiene una pendiente negativa, pero discontinua, algo que se puede apreciar con mayor facilidad si lo ponemos en relación con la media del rango (Gráfico 67):

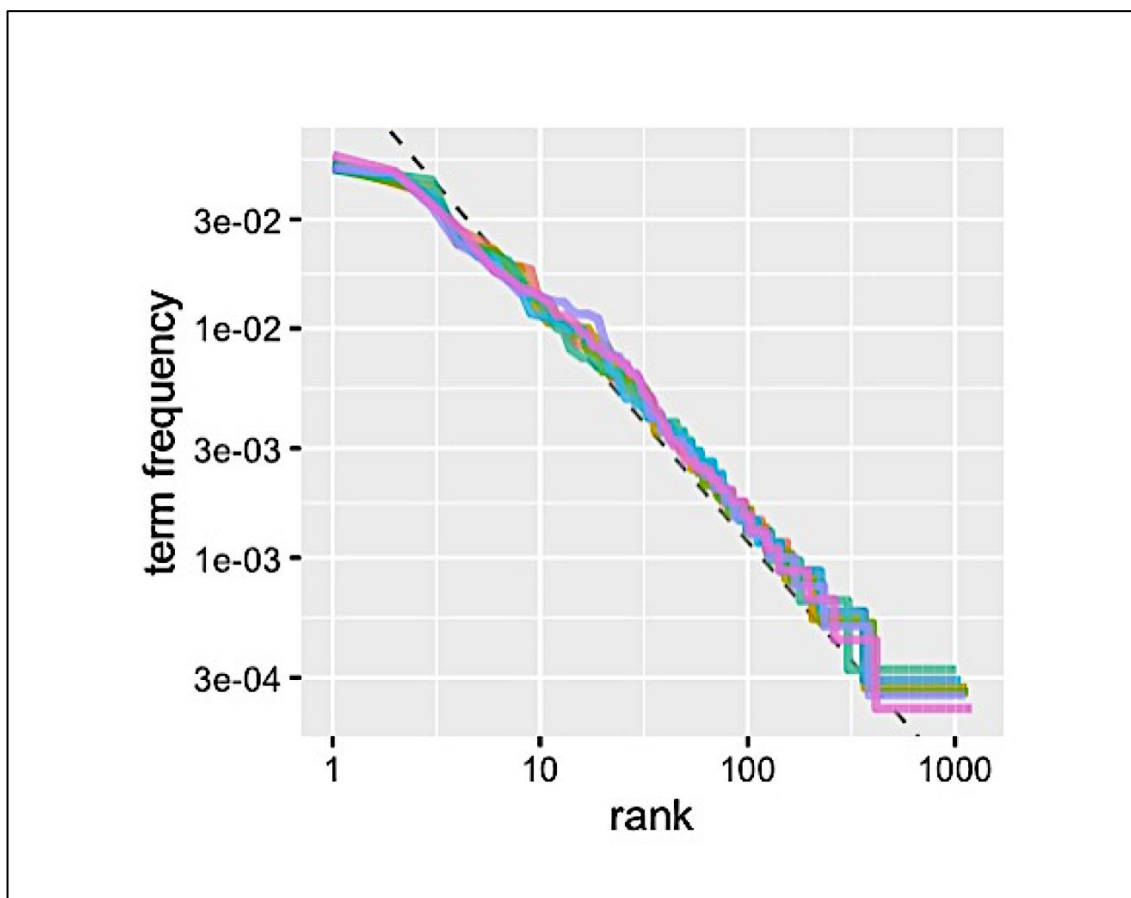


Gráfico 67. Visualización de la Ley Zipf en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes (Elaboración propia, a partir de los datos obtenidos en RStudio. Escala logarítmica con modelo)

En los seis entremeses en prosa de Cervantes, la frecuencia de las primeras palabras del ranking es más baja de lo esperado por la ley de Zipf; sin embargo, a partir del rango 25, aproximadamente, la frecuencia es superior a la media. Esta afirmación se cumple también en *La cárcel de Sevilla*, lo que supone un argumento más para sumar este texto a la producción del autor del *Quijote*.

6.3.14. Función tf-idf

La función tf-idf (del inglés, *Term frequency – Inverse document frequency*) está en el extremo opuesto a la Ley de Zipf, ya que su función es encontrar las palabras más importantes para el contenido de cada obra, disminuyendo el peso de las palabras de uso

común en la totalidad de los documentos del corpus y aumentando, en consecuencia, el peso de las palabras que menos representadas. En otras palabras, se trata de encontrar las palabras que son frecuentes en un texto, pero no *demasiado* frecuentes en el corpus total objeto de análisis.

La utilidad de la función `tf_idf` se hace visible en el siguiente gráfico:

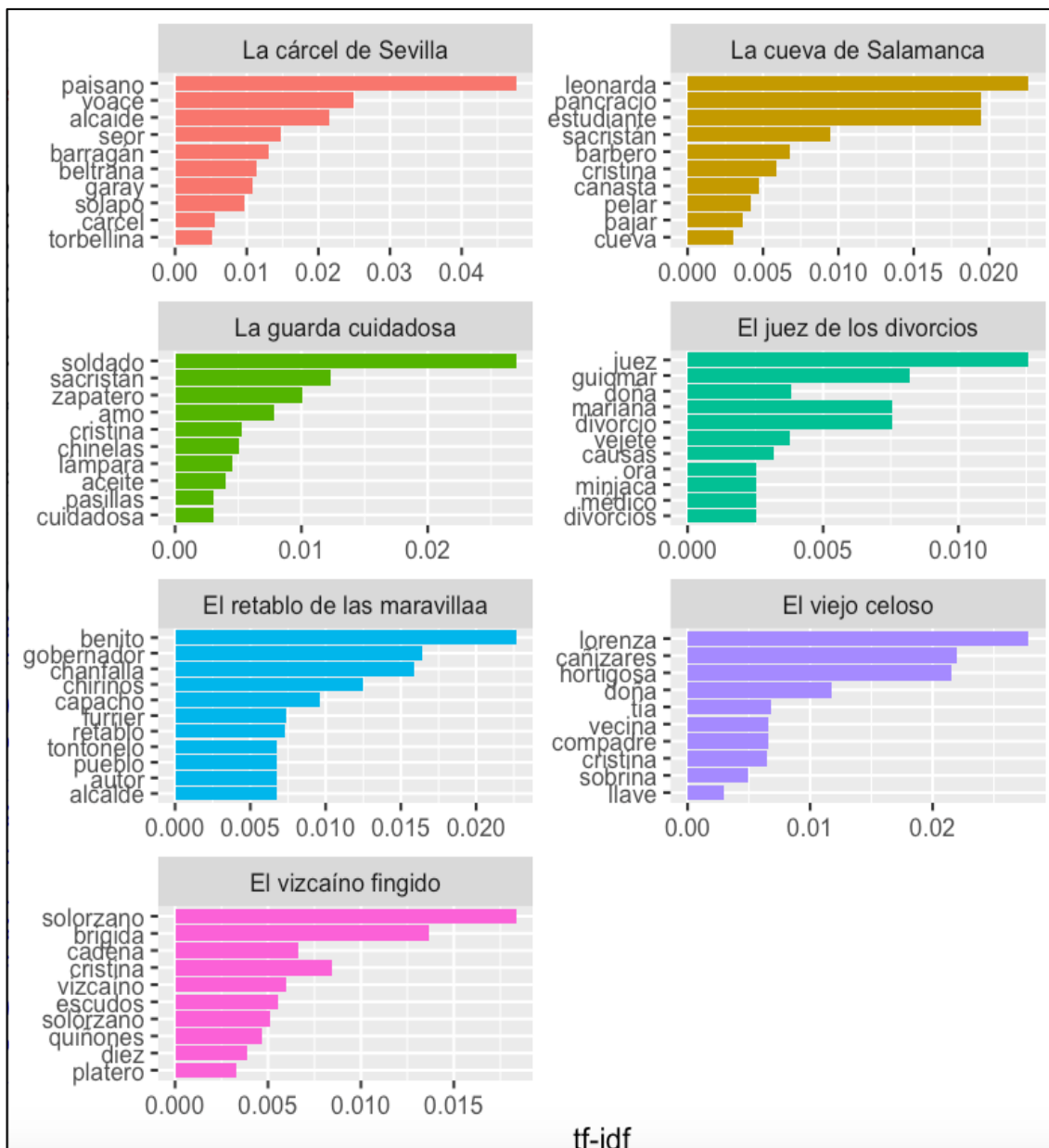


Gráfico 68. Visualización de la función `tf-idf` en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes (RStudio. Histograma)

La función *tf_idf* ha identificado aquellas palabras distintivas de cada documento, es decir, aquellas que tienen una frecuencia elevada en un documento pero que no aparecen en los restantes²¹⁴. Tal y como se desprende del gráfico 68, todos los entremeses analizados, incluido el de *La cárcel de Sevilla*, emplean un lenguaje similar, diferente esencialmente en los nombres de los personajes (Paisano, alcaide, Benito, Cristina, Quiñones, sacristán, Lorenza, Cañizares, estudiante...) o en los lugares y escenarios de cada obra (*cárcel, pueblo, pajar...*). En las voces representativas de *La cárcel de Sevilla* llaman la atención palabras como “voacé” y “seor” —que, como ya hemos advertido antes, son las formas sincopadas de “vuestra merced” y “señor”, características del habla de los hampones—, pero que no dejan de ser estar tan condicionadas por la temática rufianesca como lo están “divorcio” y “juez” en *El juez de los divorcios* o “cadena” y “escudo” en *El vizcaíno fingido*²¹⁵.

6.3.15. Frecuencia de clases de palabras y bigramas y trigramas de secuencias morfosintácticas mediante etiquetadores gramaticales automáticos (POS tagging)

Otra de las variables características del idiolecto de un determinado individuo es la frecuencia de las distintas clases de palabras (sustantivos, adjetivos, verbos, etc.), que puede calcularse fácilmente con la ayuda de los etiquetadores gramaticales automáticos, también llamados desambiguadores léxicos (en inglés, Part-Of-Speech tagging o POS-tagging). Los etiquetadores asignan a cada palabra una categoría gramatical, en función

²¹⁴ Este análisis debe ponerse en relación con el del vocabulario único y exclusivo que realizamos con CopyCatch en el apartado 6.3.6., en donde, como vimos, *La cárcel de Sevilla* tiene menor vocabulario único con respecto a *El vizcaíno fingido* o *El triunfo de los coches* (66% en los dos casos) que con respecto a *El padre engañado* (70%), a pesar de que con esta última mantenía mayor porcentaje de similitud textual.

²¹⁵ En *La guarda cuidadosa* aparecen “lámpara” y “aceite” por el famoso juego de palabras que protagonizan el mozo y el soldado, y que se encuentra ya en Lope de Rueda.

de su forma y su contexto²¹⁶ y, a partir de esa información, pueden elaborarse listados de frecuencias de clases de palabras o de secuencias morfosintácticas (bigramas, trigramas...), como veremos en el próximo apartado.

Existen varios etiquetadores automáticos (véase el apartado 2.3.4.), pero nosotros nos hemos decantado por FreeLing (versión 4.2_2), a pesar de que su utilización con R es compleja, ya que dispone de diccionario para el español. Utilizamos las etiquetas propuestas por el grupo EAGLES para la anotación morfosintáctica de lexicones y corpus de todas las lenguas europeas, agrupadas por categorías:

A = Adjetivos	C = Conjunciones
R = Adverbios	M = Numerales
T = Artículos	I = Interjecciones
D = Determinantes	Y = Abreviaturas
N = Nombres	S = Preposiciones
V = Verbos	F = Signos de puntuación
P = Pronombres	

En función de la categoría de la palabra, el etiquetado ofrece distinta información²¹⁷. Por ejemplo, en el caso de un nombre (N), si es común (C) o propio (P), si es masculino (M), femenino (F) o común (C), si es singular (S), plural (P) o invariable (N); en el caso de un determinante (D), si es demostrativo (D), posesivo (P), interrogativo (T), exclamativo (E) o indefinido (I), si es de la primera (1), segunda (2) o tercera persona (3), si es masculino (M), femenino (F) o común (C), si es singular (S), plural (P) o invariable (N) y, en el caso de los posesivos, si el poseedor es la 1ª persona

²¹⁶ Como explica Fradejas Rueda [2020] “se han desarrollado diversos algoritmos de aprendizaje automático (*Learning machine*) para realizar estas tareas. Unos se basan en corpus que han sido anotados manualmente y otros en reglas del tipo: si la palabra X es desconocida o ambigua, si le precede un determinante y le sigue un verbo, entonces es un sustantivo”.

²¹⁷ Para acceder al listado completo de etiquetas del grupo EAGLES, véase <https://www.cs.upc.edu/~nlp/tools/parole-sp.html>.

del singular (1), la 2ª persona del singular (2), la 3ª persona (0), la 1ª persona del plural (4) o la 2ª persona del plural (5); etc. Mostramos a continuación un ejemplo de análisis morfológico de una oración de *La cárcel de Sevilla*.

Lee	el	escribano	la	sentencia	en	voz	alta	.
VMIP3S0	DA0MS0	NCMS000	DA0FS0	NCFS000	SP	NCFS000	AQ0FS00	Fp

Tabla 32. Ejemplo de etiquetado gramatical automático (POS tagging)
(FreeLing. Etiquetado Eagles)

Aunque FreeLing funciona bastante bien para el español moderno, siempre es necesario supervisar el etiquetado, ya que es posible encontrar errores de catalogación en aquellas formas que pueden pertenecer a más de una categoría (pensemos, por ejemplo, en la palabra “bajo”, que puede ser adjetivo, sustantivo o preposición). En el caso del castellano del Siglo de Oro, el proceso de revisión es aún más necesario, pues encontramos errores frecuentes en aquellas palabras en desuso o en el caso de formas contractas y verbos con pronombres enclíticos. En nuestros textos, por ejemplo, FreeLing catalogaba “ahora” como sustantivo (NCFS000) en lugar de adverbio (RG), “detuviéredes” como sustantivo (NCFS000) en lugar de verbo (VMSF2P0), “voacé” como verbo (VMIS1S0) en lugar de pronombre (PP2CS0P), “Víctor” como nombre propio (NP00SP0) en lugar de interjección (I), o “destas” como sustantivo (NCMN000), sin separarlo en preposición (SP) y demostrativo (DD0FP0), como sí hace en las formas contractas actuales *del* o *al*.

Vamos a realizar la comparación de las clases de palabras y de los bigramas y trigramas de etiquetas morfosintácticas entre *La cárcel de Sevilla* y los seis entremeses en prosa de Cervantes. No utilizamos en esta ocasión los dos entremeses en verso (*El rufián viudo* y *La elección de los alcaldes de Daganzo*), ya que el comportamiento de las secuencias morfosintácticas puede estar condicionado por las exigencias marcadas por el ritmo o la rima.

Comenzamos por las clases de palabras²¹⁸:

CLASES DE PALABRAS	DUB_EP_CARCEL	CERV_EP
A. ADJETIVOS (ADJ)	84	868
CC. CONJUNCIONES COORDINADAS (CCONJ)	168	1289
CS. CONJUNCIONES SUBORDINADAS (SCONJ)	130	1007
D. DETERMINANTES (DET)	429	2722
F. SIGNOS DE PUNTUACIÓN (PUNT)	749	4606
I. INTERJECCIONES (INTJ)	35	131
M. NUMERALES (NUM)	13	155
NC. SUSTANTIVOS (SUST)	601	3761
NP. NOMBRES PROPIOS (NPROP)	102	533
P. PRONOMBRES (PRON)	504	3097
R. ADVERBIOS (ADV)	176	1309
S. PREPOSICIONES (PREP)	372	2715
VA. AUXILIARES (AUX)	98	676
VM. VERBOS (VERB)	609	3677

Tabla 33. Frecuencias absolutas por clases de palabras en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes (RStudio. FreeLing)

Como FreeLing ofrece frecuencias absolutas (Tabla 34), vamos a visualizar los resultados en diagramas de barras para poder advertir posibles similitudes o diferencias en el orden y frecuencia de uso de cada categoría (Gráfico 69).

Dejando a un lado la puntuación, que en el caso de nuestro corpus no es responsabilidad del autor sino del editor moderno, tanto *La cárcel de Sevilla* como los entremeses en prosa de Cervantes presentan un orden casi equivalente en la frecuencia

²¹⁸ Señalamos la nomenclatura de Eagles y también, entre paréntesis, la que utilizaremos nosotros para una mejor identificación: A (ADJ), CC (CCONJ), CS (SCONJ), D (DET), F (PUNT), I (INTJ), M (NUM), NC (SUST), NP (NPROP), P (PRON), R (ADV), S (PREP), VA (AUX), VM (VERB).

de utilización de las distintas clases de palabras. Las dos categorías de mayor frecuencia son los nombres comunes y los verbos (con un uso casi parejo), seguidas a una distancia equiparable de pronombres y determinantes. La utilización del resto de categorías es bastante similar, a excepción de las preposiciones y los adjetivos, ligeramente más frecuentes en los entremeses de Cervantes.

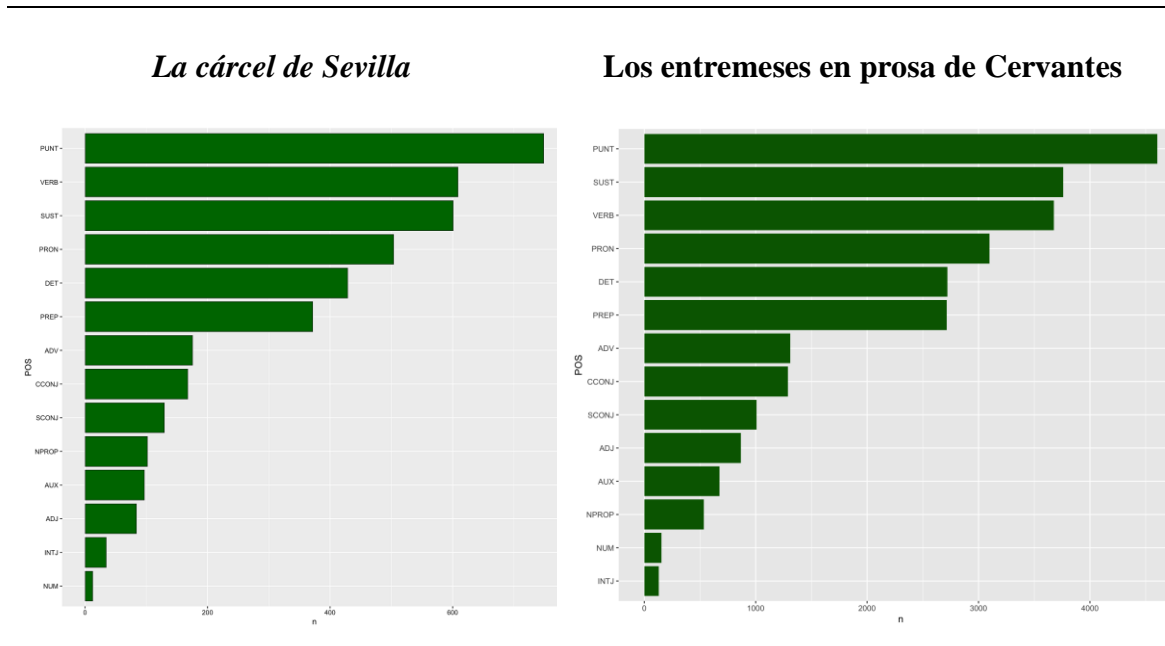


Gráfico 69. Comparativa de clases de palabras. Frecuencias absolutas.
(RStudio. FreeLing. Ggplot)

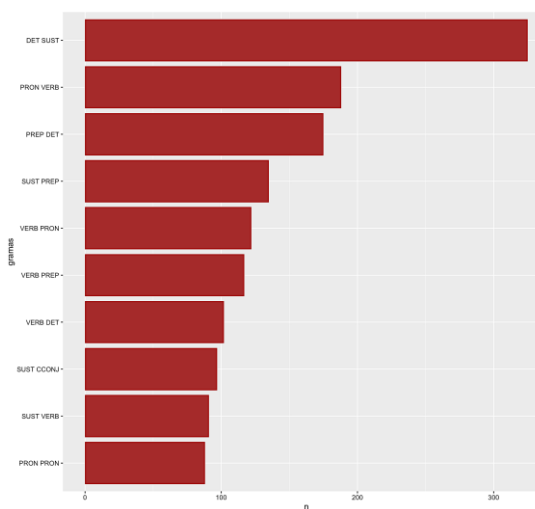
Vamos a utilizar el mismo procedimiento para comparar los bigramas y trigramas de etiquetas morfosintácticas más frecuentes en cada uno de los casos, obviando, por lo ya comentado, los signos de puntuación. Ofrecemos en primer lugar el ranking de los diez primeros casos, ordenado por frecuencia, en el que, para una mejor visualización, marcamos en color verde las coincidencias que se dan en la misma posición, en azul las que se dan en distinta posición y en rojo las no coincidentes (Tabla 34). Al igual que en el caso anterior, mostramos los resultados en diagramas de barras para apreciar mejor las tendencias de comportamiento (Gráficos 70 y 71).

ESTUDIO DE AUTORÍA DEL *ENTREMÉS DE LA CÁRCEL DE SEVILLA*

	BIGRAMAS		TRIGRAMAS	
	DUB_EP_CARCEL	CERV_EP	DUB_EP_CARCEL	CERV_EP
1	DET SUST (325)	DET SUST (2045)	PREP DET SUST (144)	PREP DET SUST (910)
2	PRON VERB (190)	PRON VERB (1325)	VERB DET SUST (78)	DET SUST PREP (465)
3	PREP DET (175)	PREP DET (1158)	DET SUST PREP (66)	VERB DET SUST (461)
4	SUST PREP (135)	SUST PREP (871)	SUST PREP DET (59)	SUST PREP DET (423)
5	VERB PRON (124)	VERB PREP (858)	PRON PRON VERB (57)	PRON PRON VERB (374)
6	VERB PREP (117)	VERB DET (589)	DET SUST PRON (57)	PRON VERB PREP (363)
7	VERB DET (102)	SUST CCONJ (581)	DET SUST CCONJ (56)	VERB PREP DET (327)
8	SUST CCONJ (97)	VERB PRON (582)	VERB PREP DET (55)	DET SUST CCONJ (321)
9	SUST VERB (92)	PRON PRON (553)	DET SUST VERB (54)	DET SUST PRON (311)
10	PRON PRON (88)	SUST PRON (542)	SUST PRON VERB (43)	DET SUST VERB (255)

Tabla 34. Ranking de bigramas y trigramas en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes (RStudio. FreeLing)

La cárcel de Sevilla



Los entremeses en prosa de Cervantes

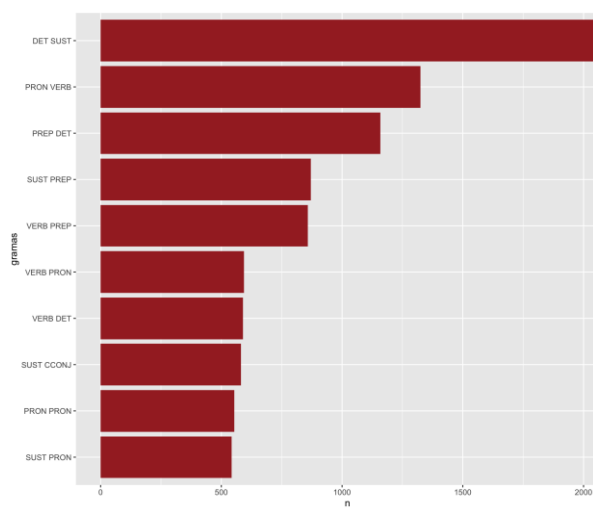


Gráfico 70. Comparativa de los diez bigramas más frecuentes en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes. Frecuencias absolutas. (RStudio. FreeLing. Ggplot)

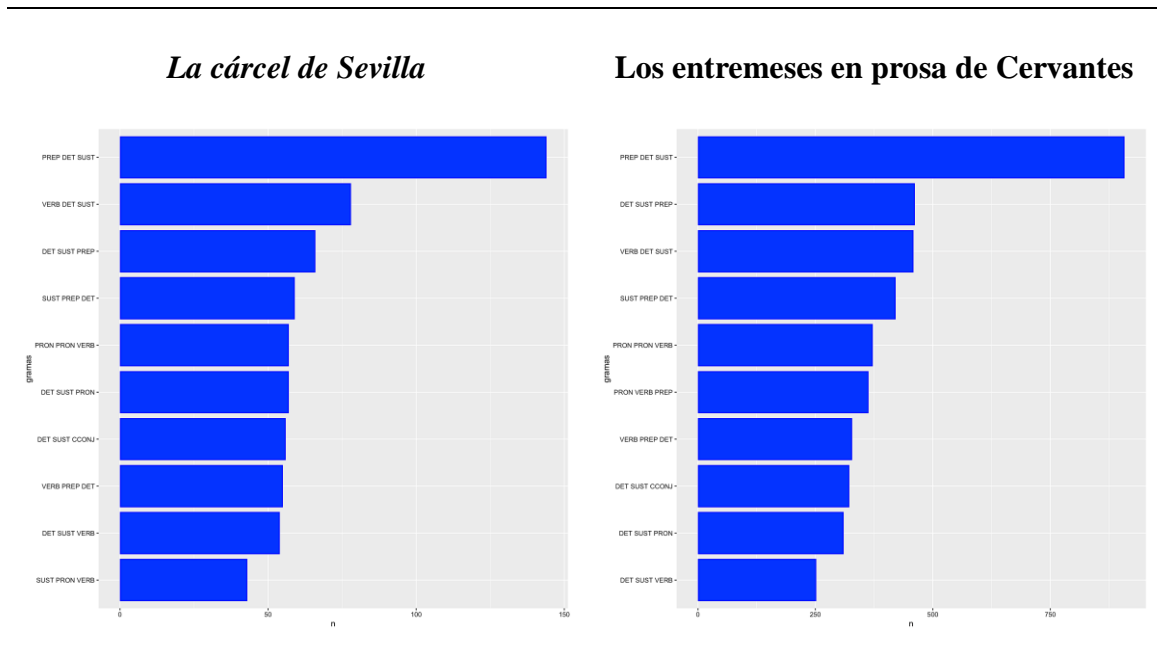


Gráfico 71. Comparativa de los diez trigramas más frecuentes en *La cárcel de Sevilla* y los entremeses en prosa de Cervantes. Frecuencias absolutas. (RStudio. FreeLing. Ggplot)

En el caso de los bigramas, el uso en Cervantes y en el autor del texto dubitado es prácticamente idéntico (la única diferencia llamativa se da en la secuencia VERB PRON, que es significativamente superior en *La cárcel de Sevilla*). En el caso de los trigramas, ocurre lo mismo (sólo es reseñable el mayor uso de la secuencia PRON VERB PREP en Cervantes, algo que va en consonancia con esa mayor presencia de preposiciones que mostraba el análisis de categorías de palabras)²¹⁹. El comportamiento morfológico de *La cárcel de Sevilla*, por tanto, se asemeja enormemente a Cervantes, reforzando de nuevo nuestra hipótesis de trabajo (que sean en realidad de la misma persona).

²¹⁹ La secuencia PRON VERB PREP ocupa la sexta posición en el ranking cervantino, pero no aparece en el *top-ten* del texto dubitado (ya que ocupa el puesto 12). Este hecho, sin embargo, no es representativo de un uso discordante, pues lo cierto es que las diferencias entre el décimo y el decimosegundo lugar son mínimas.

6.4. CONCLUSIONES SOBRE LA AUTORÍA CERVANTINA DE ESTE ENTREMÉS

El estudio realizado en el apartado precedente nos lleva a establecer las siguientes consideraciones en relación con la posible autoría cervantina del *Entremés de la cárcel de Sevilla*:

1. Los análisis del tema, el argumento, los personajes, la segmentación y la estructura externa e interna de *La cárcel de Sevilla* muestran un comportamiento propio del teatro menor de Cervantes, aunque también coincidente con el de otros autores del momento, ya que la dramaturgia del Siglo de Oro seguía, por lo general, unas características bien definidas y marcadas.
2. El análisis cualitativo de las once marcas idiosincrásicas detectadas por los expertos en la lengua de Cervantes [Agostini, 1964b, Rosenblat, 1978, Gutiérrez Cuadrado, 1988, Ariza Viguera, 1998, Maestro 1998 y Baras Escolá, ed., 2012] demuestra que el autor del texto dubitado tiene unos *usus scribendi* próximos a los del autor del *Quijote*. En *La cárcel de Sevilla* encontramos adverbios en *-mente* (*públicamente, naturalmente*), diminutivos en *-ico* con valor despectivo (*juececicos*), leísmos de persona y cosa, creación de nuevas palabras por mecanismos de composición (*perniquebrados, boquirrubio*), gusto por las figuras de dicción, las repeticiones y los sinónimos coordinados, utilización pragmática de enunciados en forma interrogativa y caracterización de los personajes a través del habla; no hay, sin embargo, casos de superlativo en *-ísimo*, latinajos o creación de nuevas palabras mediante derivación o parasíntesis, que suelen darse en Cervantes, (aunque no en todos sus entremeses).
3. Los distintos análisis de la distribución de la frecuencia de los 5-gramas de caracteres y 1-grama de palabra más empleados en el corpus de entremeses en prosa, realizados con el algoritmo Eder's Delta (*Analysis Cluster* y *Bootstrap Consensus Tree*), ponen de manifiesto la cercanía entre *La cárcel de Sevilla* y

Cervantes (especialmente con *El retablo de las maravillas*). *La cárcel de Sevilla* está más cerca de Cervantes que de ningún otro autor, incluido Barrionuevo, a pesar de que *El triunfo de los coches* también se ubica en la misma rama. Cabe notar, no obstante, que mientras el entremés dubitado mantiene asociaciones de primer nivel con los textos cervantinos, no ocurre así con el de Barrionuevo.

4. El análisis de componentes principales de la distancia Delta (Eder's Delta) aplicado a los ocho entremeses cervantinos y *La cárcel de Sevilla*, sitúa el texto dubitado cerca del centro de la gráfica de puntuaciones, lo que implica que sus componentes se identifican con los del grupo mayoritario. El entremés dubitado muestra mayor sintonía con el conjunto de entremeses cervantinos que *El rufián viudo*, que aparece en el extremo derecho del cuarto cuadrante.
5. La realización de clasificaciones supervisadas por aprendizaje automático (Classify) asocia correctamente todos los textos indubitados a partir de las 600 palabras más frecuentes. Al aplicar este análisis a *La cárcel de Sevilla*, vemos que este entremés siempre se asocia con uno de Cervantes, desde las 100 a las 1000 MFW, y nunca con los entremeses de Ávila, Quiñones de Benavente, Hurtado de Mendoza o Quevedo, ni con las dos partes de la *Relación* de Cristóbal de Chaves, con los que guarda una evidente relación temática.
6. Seis de las diez obras con mayores porcentajes de similitud textual con *La cárcel de Sevilla* son de Cervantes (concretamente, sus seis entremeses en prosa). El mayor porcentaje lo tiene con *El padre engañado*, con un 37%, seguido de cerca por *El triunfo de los coches* (36%) y *El vizcaíno fingido* (34%); un análisis minucioso de estos tres pares permite advertir que *La cárcel de Sevilla* tiene menos vocabulario exclusivo y más vocabulario compartido una sola vez con respecto a *El vizcaíno fingido* y *El triunfo de los coches* que con respecto a *El padre engañado*, lo que intensifica su cercanía léxica con estas dos obras. Cabe señalar, por otro lado, que el porcentaje de

similitud textual entre el entremés dubitado y la tercera parte de la *Relación de la cárcel de Sevilla* es tan bajo (21%) que parece poco probable que hayan sido escritos por la misma persona.

7. La clasificación secuencial de aprendizaje automático (Rolling classify), diseñada para detectar autorías múltiples, determina que *La cárcel de Sevilla* ha sido escrita en su mayoría por una única persona, a la que identifica con Cervantes. El examen es tan preciso que señala como ajeno el romance insertado (“Alta mar esquiva”) y detecta, en un segundo plano, la presencia de la *Relación* de Chaves, que, como hemos demostrado la Tabla 15, es la fuente temático-argumental de este entremés.
8. El análisis de verificación de autoría, realizado con el método General Imposters (GI), demuestra que Cervantes no sólo es el autor que más se aproxima en su comportamiento léxico al del texto dubitado (*nearest neighbor*), sino que lo hace con tal fuerza (entre 0,98 y 1 sobre 1) que es inevitable pensar que sea su autor.
9. El texto dubitado comparte varios *verbatim* con Cervantes, uno de ellos de 7 ítems (“eso haré yo de muy buena gana”). Según indica CORDE, esta secuencia de 7 palabras es muy característica del autor del *Quijote* (la usa hasta en 8 ocasiones), aunque también aparece una vez en Lope de Vega y en otra en Jorge de Montemayor.
10. El entremés dubitado comparte con Cervantes la predilección por relacionar las palabras *hombre* y *mundo*, que coaparecen en varias ocasiones dentro de un mismo parlamento.
11. El estudio del orden de frecuencia, la frecuencia relativa y el análisis de componentes principales (PCA) de las palabras de función también acercan el texto dubitado a los entremeses de Cervantes. Ambos corpus comparten 9 de las 10 palabras de función más frecuentes, ocho de ellas en la misma posición, algo que no ocurre en el resto de caso; además, presentan frecuencias relativas similares en el uso de ciertas conjunciones, preposiciones y adverbios; y están bastante próximos en el PCA.

12. El histograma de la ley Zipf refleja un comportamiento análogo entre *La cárcel de Sevilla* y cada uno de los entremeses de Cervantes en la frecuencia con la que aparecen las palabras y su rango en el número de uso.
13. El análisis de la función tf_idf demuestra que *La cárcel de Sevilla* emplea un lenguaje similar al de los entremeses cervantinos, diferenciándose esencialmente en los nombres de los personajes y en los lugares y escenarios de cada obra.
14. La comparación de las clases de palabras y de los bigramas y trigramas de etiquetas morfosintácticas, realizados con FreeLing, muestran un comportamiento armónico entre *La cárcel de Sevilla* y los seis entremeses en prosa de Cervantes.

A la luz de todo lo analizado, y siguiendo la escala de opinión verbal propuesta en el apartado 5.6., podemos determinar que Miguel de Cervantes Saavedra es el autor del entremés de *La cárcel de Sevilla*, con un grado de probabilidad (4) bastante alto.

CAPÍTULO 7

ESTUDIO DE AUTORÍA DEL *ENTREMÉS* *DE LOS ROMANCES (VERSO)*

LEONARDA: Después que das en leer,
Inés, en el Romancero,
lo que a aquel pobre escudero
te podría suceder.

INÉS: Don Quijote de la Mancha
(perdone Dios a Cervantes)
fue de los extravagantes
que la coronica ensancha.

(Lope de Vega, *Amar sin saber a quién*)

7.1. ANTECEDENTES Y ESTADO DE LA CUESTIÓN²²⁰

Otra de las piezas auriseculares que desde finales del siglo XIX ha visto planear sobre sí la sombra de la paternidad cervantina es el *Entremés famoso de los romances*, un intermedio en verso muy conocido —y también muy mal valorado literariamente—

²²⁰ Este capítulo supone una versión actualizada y mejorada de un trabajo publicado en 2010, dentro del volumen *Hos ego versículos feci... Estudios de atribución y plagio* [Ruiz Urbón, 2010]. Aunque ciertos análisis se mantienen, como los relativos a la métrica, las marcas idiosincrásicas, la proporción de palabras lexicales / palabras funcionales o la ratio *type/token* de palabras de contenido, hemos añadido otros (relacionados esencialmente con la Estilometría) que vienen a complementar y reforzar las conclusiones allí alcanzadas. Actualizamos, asimismo, la bibliografía sobre este asunto, haciendo especial hincapié en el último estudio atributivo de Rodríguez López-Vázquez [2019].

por la polémica que se mantiene desde antiguo en torno a si es la fuente inspiradora del *Quijote* o si, por el contrario, es una simple derivación de éste. El entremés, escrito entre la última década del siglo XVI y la primera del XVII e impreso en la *Parte tercera de las comedias de Lope de Vega y otros* (Valencia, 1611), recoge la historia de Bartolo, un labrador que tras enloquecer “de leer el Romancero” decide convertirse en soldado para ir a “reñir con los ingleses”, llevando como escudero a su convecino Bandurrio. El anónimo entremesista potencia la hilarante comicidad de la pieza mediante la inserción de un sinfín de romances populares fácilmente reconocibles por los espectadores de la época y que, a modo de calco cuasi literal, son puestos en boca de sus personajes.

Como recordábamos en el capítulo 4, el primero en vincular este entremés con la gran obra de Cervantes es Adolfo de Castro, quien en 1874 arguye que la pieza ha de ser anterior al *Quijote*, porque en ella se dice que Bartolo se marcha a Inglaterra “a matar el Draque / y a prender a la reina”, acontecimientos anteriores a 1602, año en que muere la reina Isabel I de Inglaterra²²¹. Además, tiene noticia de que se representó junto con *La noche toledana* de Lope de Vega, obra escrita en torno a 1604 y llevada a los teatros por esas mismas fechas, lo que le lleva a confirmar que “el *Entremés de romances* se escribió antes que el *Quijote*, y que antes de publicarse el *Quijote* se dio al teatro” [1874: 133].

Para Castro, las similitudes entre esta pieza y los capítulos 2, 3, 4 y 5 de la primera parte del *Quijote* no pueden ser fruto de la casualidad. Ofrece, además, lo que considera una prueba concluyente de su relación genética: ambos textos cometen el mismo error al citar dos versos del *Romance del Marqués de Mantua*, lo que le lleva a pensar que es obra de Cervantes, pues “es imposible que dos autores se equivoquen de idéntica manera, escribiendo así al propio tiempo” [1874: 139]²²². Concluye su reflexión con la siguiente afirmación:

²²¹ Adolfo de Castro se equivoca en la fecha, pues la soberana inglesa muere en 1603, no en 1602. En todo caso, estas palabras, puestas en boca de Dorotea, pertenecen en realidad al romance “Hermano Perico”, publicado en la Flor segunda (Valencia, 1593).

²²² Corrigiendo a Agustín Durán, que expuso en su *Romancero* que Cervantes hubo de seguir una lección más moderna del romance del Marqués de Mantua (puesto que dice “¿Dónde estás, señora mía, /

El *Entremés de los Romances* es verdaderamente el bosquejo del carácter de don Quijote y de la primera salida del ingenioso hidalgo. Cervantes hizo lo que los grandes pintores: trazó un borrón o un ligero dibujo de un gran cuadro, primitivo pensamiento que luego desarrolló en un libro admirable [1984: 140].

Las teorías de Castro sobre la génesis del *Quijote* y la autoría del *Entremés de los romances* son defendidas ese mismo año por Marcelino Menéndez Pelayo que, siendo todavía estudiante, escribe lo siguiente²²³:

Es verdad que faltan en la figura de Bartolo muchos rasgos que a la de don Quijote distinguen, razón que nos induce a mirar la primera como bosquejo de la segunda. La primera encarnación del pensamiento cervantino fue el rústico e ignorante Bartolo. Purificada y embellecida transformóse en don Quijote, tipo acabado de virtud e hidalguía, modelo de discreción en cuanto no se rozaba con su manía. El insignificante escudero Bandurrio vióse convertido en Sancho Panza, personificación del buen sentido y de la recta razón aplicada a las más prosaicas necesidades de la vida. En el *Entremés de los romances*, vemos a las claras el desarrollo del pensamiento fundamental del *Quijote* en la mente de Cervantes. ¿Podremos dudar de que es obra suya la pieccecita que analizamos? Pues qué ¿él, que a todos excedía en la invención, había de tomar la idea capital de su obra de un entremés conocidísimo y aplaudido ya en las tablas? ¿No están declarando la paternidad de la obra su estilo, su lenguaje y hasta los yerros en que el autor incurre al citar algunos romances, con la misma lección que Cervantes en el *Quijote*, en lugar de la propia y genuina con que aparecen en los pliegos sueltos y en el *Romancero*? Inútil sería detenernos en esta cuestión, cuando tan eruditamente la resuelve el señor Castro en su doctísima *Introducción* a la obrita que acabamos de examinar [1941: 285].

El primero en oponerse a Castro y Menéndez Pelayo es Armando Cotarelo Valledor, quien en su monografía sobre *El teatro de Cervantes* defiende que este entremés “fue obra de algún ingenio, ciertamente mediocre, que imitó muy de cerca la primera salida de don Quijote” [1915: 723]. Esta misma idea es la que mantiene Francisco Rodríguez Marín al editar la gran novela cervantina en 1916, aunque años más tarde, al hacer la “nueva edición crítica” de la obra, se muestra más prudente y asegura no disponer de ninguna prueba inescrutable que avale tal teoría [1928: 159-163].

que no te *duele* mi mal” en lugar de “¿Dónde estás, señora mía, / que no te *pesa* mi mal”), Castro asegura que lo que ocurrió en realidad fue que Cervantes confundió los dos versos citados con los del romance de Tirsi, que rezan así: “¿Dónde estás, señora mía, / que no te *pesa* mi mal? / O no lo sabes, señora / o eres falsa o desleal”. Hoy por hoy sabemos que el error al que alude Castro no es tal, y que, en realidad, ambos textos sí recogen una versión más moderna del conocido romance.

²²³ Citamos por la reimpresión de sus obras completas [1964].

Ramón Menéndez Pidal, en el discurso que pronuncia el 1 de diciembre de 1920 con motivo de la inauguración del curso académico en el Ateneo Científico, Literario y Artístico de Madrid, retoma la idea de que el anónimo entremés es la fuente literaria indiscutible del *Quijote* de 1605, describiendo el proceso de gestación del siguiente modo²²⁴:

Cervantes descubrió una gracia fecunda en el entremés que se burla del trastorno mental causado por la indiscreta lectura del *Romancero*. Esta sátira literaria le pareció tema excelente; pero la apartó del *Romancero*, género poético admirable, para llevarla a un género literario de muchos execrado, el de las novelas caballerescas, no menos en moda que el *Romancero* [1964: 24].

Sobre la fecha de composición de la pieza, sostiene que hubo de componerse en torno a 1591 (y nunca después de 1602), porque treinta de los treinta y tres romances que contiene sólo habían aparecido de manera conjunta en la *Flor de varios nuevos romances*, publicada en Valencia en 1591 y reimpresa en 1593²²⁵, y porque, además, en uno de ellos se hace alusión a la victoria de los ingleses en las Azores, que tuvo lugar ese mismo año de 1591; en su opinión, “sería salirse de lo ordinario y corriente creer que el autor hubiese colocado sus alusiones y su ambiente en un pasado histórico” ya que “hay que suponer que el teatro cómico se mueve dentro de la época actual y de la vida familiar de todos” [1964: 56].

Aunque parece coincidir en parte con las tesis de Castro y Menéndez Pelayo, Menéndez Pidal no acepta de modo alguno la identificación de Cervantes con el anónimo entremesista:

Ocurriósele a Adolfo de Castro exhumar esta pobre composición teatral, afirmando que Cervantes mismo era autor de ella, y se atrajo el más justo y general descrédito [1964: 20].

²²⁴ Citamos por la versión incluida en *De Cervantes y Lope de Vega* [1964], que ya contiene la “nota adicional” a la que hacemos alusión más adelante.

²²⁵ Según las investigaciones posteriores de Rodríguez-Moñino, la edición de la *Flor* es de 1593, pero la mala conservación del original motivó que Menéndez Pidal leyese un “1” donde en realidad hay un “3”, pensando, en consecuencia, que la edición original era de 1591 y que ésta fue reimpresa posteriormente en 1593 [1957].

Los postulados de Menéndez Pidal tienen respuesta inmediata en un trabajo de Emilio Cotarelo y Mori publicado ese mismo año, en el que vuelve a defender que fue el ignoto escritor el que redactó su intermedio lúdico a partir de la obra magna cervantina, al tiempo que aconseja “volver los ojos, no a los documentos literarios, sino a los archivos de Esquivias, a los documentos privados” [1920: 61].

Esta crítica será rebatida de nuevo por un resentido Menéndez Pidal que, de manera inmediata y contundente, publica en el nº 7 de la *Revista de Filología española* una nota bibliográfica sobre la obra de Cotarelo y Mori [1920: 389-392] —que a partir de 1924 se incluirá como “nota adicional” en su trabajo original—, en donde asegura que “no necesitamos documentos notariales para saber que el artista toma de personajes de carne y hueso el barro con que forma los tipos a los cuales infunde el soplo de la vida ideal”. Reafirma la hipótesis de que es Cervantes quien se inspirara en el entremés para escribir la primera salida de don Quijote, no a la inversa, y aclara que de sus palabras no se infería, de modo alguno, que pudiera ser su autor, aseverando que “lo único que da completa claridad al problema es suponer que el *Entremés* no es de Cervantes, pero que le inspiró” [1920: 390].

George T. Northup, al hacer en 1922 el *review* del trabajo de Menéndez Pidal, dice que la atribución del *Entremés de los romances* a Cervantes, a pesar de estar hecha por un académico de descrédito como es Adolfo de Castro, es “worthy of consideration”, pues de sobra es sabido que “when Cervantes had a good idea he often used it two or three times in various works”:

He was a writer of farces, and if we find in a farce ideas which he has used elsewhere, there is a possibility if not a presumption that he wrote the farce question [1922b: 436].

Aunque reconoce que basar la autoría de una obra tan breve en el estilo individual de un autor es peligroso —y más si tenemos en cuenta que la mitad de la pieza se ha construido sobre la cita de fragmentos de diversos romances y que las obras de Cervantes presentan muy variados estilos—, Northup acaba posicionándose del lado de la paternidad cervantina:

It is easier to accept the view that Cervantes wrote the entremés than to relieve that he followed so slavishly the invention of another [1922b: 436].

Ninguna de estas hipótesis ha sido aceptada unánimemente por la crítica moderna. Desde la acalorada discusión filológica entre Menéndez Pidal y Cotarelo y Mori, casi todos los estudios se han centrado en la datación de la pieza y su direccionalidad con respecto al *Quijote*, olvidándose, casi por completo, del misterio de su autoría (que sólo ha vuelto a interesar en los últimos tiempos).

Entre los que siguen los postulados de los Cotarelo (Valledor y Mori) y defienden la anterioridad del *Quijote* de 1605, encontramos los trabajos de Schevill, Asensio y Murillo.

Rudolph Schevill, en las notas que escribe para la edición de *Don Quijote de la Mancha* (cuando ya ha fallecido Bonilla y San Martín), presupone la anterioridad de la obra cervantina. Al referirse a la tesis de Menéndez Pidal, escribe:

Siento no poder estar conforme con esta conclusión de mi amigo y maestro. Es verdad que muchos entremeses fueron publicados años después de ser representados; sin embargo, creo que dicho *Entremés*, que fué publicado por primera vez en 1611, o sea unos diez años después de “engendrarse” el *Quijote*, parece ser más bien un remedo de los primeros capítulos de la obra cervantina. Precisamente el gran número de detalles en que se asemejan el *Entremés* y el *Quijote* aboga contra el argumento de la anterioridad de aquél. Cervantes no imita sus fuentes con tanta minuciosidad, y hemos de creerle en este caso que sus invenciones no eran “imitadas ni hurtadas [1928: tomo I, 417].

Eugenio Asensio cree que la presunción de actualidad de la que habla Menéndez Pidal no es del todo acertada, ya que el entremesista debía reflejar en sus versos romances manidos para asegurarse de que el público de los corrales los conociesen, pues sólo así se podía garantizar la comicidad de la pieza. Además, el empleo del verso y del romance en diálogo le hacen “retrasar la fecha del *Entremés de los Romances* a los primeros años del siglo XVII” [1971: 75].

Luis Andrés Murillo, por su parte, supone que “el entremés se habría compuesto [...] después de 1605, es decir, después de salir impreso el *Quijote*”, pues, teniendo en cuenta que la comedia lopesca no introduce el uso del romance en diálogo antes del 1600, parece “improbable que se pudiese haber escrito o representado [...] antes de 1604” [1986: 355-356], y más aún si tenemos en cuenta la escasa valía de la pieza. Defiende que *Los romances* es herencia temática de la primera salida de don Quijote y herencia

formal de la técnica innovadora del entremés de *Melisendra* (primer entremés publicado en verso); y sentencia:

el entremesista se pudo haber servido de la *Flor* de 1593 para una parodia inspirada por la boga del *Romancero general* ya impreso, habiendo visto o teniendo noticias del libro (o manuscrito) de Cervantes en que un hidalgo de aldea, enloquecido por las lecturas de libros de caballerías, se nombrara a la usanza de caballerías en poesía y prosa y las imitaba. Desde luego, quedaría enigmático que no hiciese ninguna alusión al *Quijote*” [356].

Los que secundan a Castro y los Menéndez (Pelayo y Pidal) defienden, por el contrario, la prioridad del *Entremés de los romances*. Es la línea mayoritaria, representada, entre otros, por Alonso, Millé y Giménez, López Navío, Varo, Baras Escolá, Pérez Lasheras y Rey Hazas.

Dámaso Alonso, en la “nota preliminar” que encabeza su edición de los entremeses atribuidos a Cervantes, habla de la patente relación entre *Los romances* y “algunos capítulos de nuestra mayor novela”. Según su opinión, “se puede asegurar, casi con certeza absoluta que Cervantes imitó ese pobre entremés” [1986: 12-17].

Más interesante es el estudio de Juan Millé y Giménez, que piensa que todo el entremés se configura como una parodia contra Lope de Vega. Aunque asume que pudo ser escrito entre 1585 y 1611, propone como fecha de composición el año 1588, lo que le lleva a plantearse la cuestión de la autoría:

¿Quién es el desconocido entremesista? ¿El mismo Góngora o el travieso Salinas (también enemigo de Lope), o nuestro grande Cervantes, o algún otro de los ingenios que picardean por entonces alrededor del palco escénico madrileño? Confesemos sencillamente que es una cuestión que, con los datos conocidos no nos es posible afrontar [1930: 120].

En cuanto a la autoría, descarta la hipótesis de que Lope de Vega sea el autor del entremés, ya que resultaría ilógico que él mismo escribiese una burla tan mortificante hacia su figura; la idea de que sea de Juan de Salinas, sustentada únicamente en el hecho de que éste era enemigo declarado de Lope, no autoriza “a formular otra cosa que una vaga sospecha”; más verosímil le parece que sea de Góngora, aunque echa en falta “la acritud, la mordacidad corrosiva de la sátira gongorina” [1930: 121].

José López Navío refrenda la teoría de la anterioridad del *Entremés*. En su opinión, esta pieza debe ser entendida como una parodia satírica contra Lope de Vega, escrita por alguno de sus enemigos [1959-1960: 151-239].

Carlos Varo también se posiciona en “pro de la precedencia del Entremés”, argumentando que es absurdo pensar que Cervantes, pretendiendo parodiar los libros de caballerías, configure el comienzo de su novela como una parodia del Romancero, “si no es que hubo antes una composición que glosara los romances y que se le quedó demasíadamente marcada” [1968:132]:

Cervantes descubrió una gracia fecunda en el *Entremés*, que se burla del trastorno mental causado por la indiscreta lectura del Romancero. Esta sátira literaria le pareció tema excelente, pero la apartó del Romancero, género poético admirable, para llevarla a un género literario de muchos execrado, el de las novelas caballerescas, no menos en moda que el Romancero [1968: 133].

Por su parte, Antonio Pérez Lasheras hace la siguiente afirmación categórica, no sin pedir perdón por su “*quijotesco* atrevimiento”:

Nuestra conclusión es clara: el *Entremés de los romances* tuvo que escribirse poco después de febrero de 1592 [...] Al mismo tiempo, creemos que, al fijarse la fecha del *Entremés*, puede asentarse la teoría de Stagg sobre el comienzo de la escritura del *Quijote* en septiembre de 1593, entre otras cosas porque la literatura paródica necesita de una intertextualidad que juegue con la memoria de un público al que se exige un acto de descodificación complejo [1988: 76].

Ya en el siglo XXI, Antonio Rey Hazas sostiene que el *Quijote* se parece a *Los Romances* “como una gota de agua a otra, lo que sugiere la imitación cervantina” [2007: 1]. Retrasa la datación del entremés a los años comprendidos entre 1593 y 1597 (antes en todo caso de la redacción del *Quijote* que, según su criterio, empezó a gestarse hacia 1598) y propone como firme candidato a autor a Gabriel Lobo Lasso de la Vega, un escritor muy cercano a Cervantes, autor de alguna tragedia de corte neosenequista y de un buen número de romances, en alguno de los cuales manifiesta abiertamente su repulsa hacia Lope por su obsesión por los romances de tema morisco:

En tanto que aparece o no un documento revelador que lo demuestre con objetividad, no es descabellado pensar que Gabriel Lobo Lasso de la Vega pudo ser el autor del *Entremés de los romances*, quizá en conveniencia expresa con el propio Miguel de Cervantes [2007: 42].

José Luis Pérez López, contradiciendo a Millé Giménez y López Navío, cree que el texto es un sentido homenaje a Lope de Vega. Propone como candidato a autor a Liñán de Rianza, gran amigo del Fénix, que tal vez lo escribió en connivencia con éste [2009]. Y lo mismo opina Javier Blasco tras llevar a cabo una serie de análisis relacionados con la riqueza léxica, el orden y frecuencia de las palabras de función, los trigramas de palabra y el uso de léxico exclusivo [2013].

Aunque Baras Escolá considera que “ya nadie es partidario de la autoría de Cervantes” [2012: 200], lo cierto es que recientemente Alfredo Rodríguez López-Vázquez ha vuelto a poner el foco sobre él. En su opinión, el *Entremés de los romances* fue compuesto “en el trienio 1605-8” [2016: 93], es decir, después del *Quijote* de 1605, y nos invita a pensar que es de Cervantes

en tanto no se proponga a un autor que en el período 1590-1600 use secuencias de quintillas del tipo *aabba* y use también la forma rústica “Dolo (a huego)”, manteniendo además un uso similar al de Cervantes para esos otros trece [sic] índices autoriales [2019: 238].

Una vez más, los tres argumentos atributivos utilizados por Rodríguez López-Vázquez nos parecen insuficientes para ahijar una obra a un determinado autor.

- Argumenta, en primer lugar, que la utilización de una secuencia de seis quintillas seguidas de tipo 5 (*aabba*) es una rareza, pues, tomando como referencia el estudio de Morley y Bruerton, este tipo de quintillas en Lope de Vega siempre aparecen en combinación con una quintilla de tipo 1 (*ababa*) o de tipo 2 (*abbab*); este uso, sin embargo, sí se da en Cervantes, en concreto en su comedia *La casa de los celos*²²⁶. Es cierto que ocurre así, aunque creemos

²²⁶ El fragmento de *La casa de los celos* al que se refiere Rodríguez López-Vázquez es éste:

[ESPÍRITU]:	A tu primo entre esa yerba	a
	pondrás, que a mi se reserva	a
	y a mi fuente su salud,	b
	que hasta agora su virtud	b
	en ella el cielo conserva.	a
MALGESÍ:	Volveos por do vinistes	c
	figuras feas y tristes,	c

que su estudio métrico debería abarcar el análisis global del metro, la estrofa y la rima del entremés, y no sólo el de unos cuantos versos seleccionados aleatoriamente.

- En segundo lugar, destaca la utilización de la forma rústica *dolo* como pregunta y no como exclamación, que, según explica, era su uso habitual en el teatro popular breve del siglo XVI, a partir de la expresión “Doylo al hugo > Doylo al fuego (del infierno)”. Sin embargo, tal y como anota Rey Hazas en su edición de este entremés, hay otra alternativa: efectivamente podría “tratarse de *¡dolo!*, en el sentido de «¡malo, mala cosa!»”, pero también de “¿*dólo?* [...], es decir, «¿por dónde [viene]?», que es lo más probable” [2006: 88]. Sea como fuere, Rodríguez López-Vázquez reconoce que “ni Cervantes, ni Góngora, ni Gabriel Lobo Lasso de la Vega usan esta forma arcaizante” [2019: 235].
- En tercer lugar, realiza un análisis de diez unidades léxicas (entre las que se incluye *dolo*). Al igual que ocurría en su estudio sobre *La cárcel de Sevilla*, los resultados deben interpretarse dentro de la realidad que nos ofrece CORDE como repositorio limitado de textos. Rodríguez López-Vázquez decide restringir su búsqueda a un intervalo de tan sólo trece años, entre 1600 y 1613, intervalo en el que, por ejemplo, no se recoge ninguna obra dramática de Lasso de la Vega. Pero es que, además, y tal vez esto es lo más relevante, sólo da cuenta de los resultados para Cervantes, Quevedo, Góngora, Vélez de Guevara y Lasso de la Vega, pero no especifica los resultados obtenidos al margen de estos cinco autores. En definitiva, aunque Rodríguez López-Vázquez pone en valor que nueve de los diez índices están en Cervantes [2019: 235], su selección nos parece arbitraria, pues —como analizaremos más tarde—, podemos citar otros muchos índices que aparecen en el texto dubitado y en otros autores de la época, pero no en Cervantes: “que no lo digo por tanto”,

que mi primo quedará	d
adonde esperar podrá	d
el remedio que no distes.	c

“partirse a la guerra”, “está fuera de sí”, “para alegrar la boda”, etc. [2010: 248-251].

El método de Rodríguez-López Vázquez sí parece convencer a Adrián J. Saéz, que en su trabajada edición de los *Entremeses* de Cervantes para Cátedra considera que esta pieza tiene su ADN, arguyendo “parecidos estilísticos, intertextuales, léxicos, métricos, temáticos y demás” [2020: 30].

Este breve repaso por la historia crítica del *Entremés de los romances* permite advertir que, tras casi ciento cincuenta años de diatriba, la polémica en torno a su fecha de composición y su anterioridad o posterioridad respecto del *Quijote* sigue sin resolverse. Y lo mismo sucede en relación con su autoría: “the question remains open” [Stagg, 2002: 132].

7.2. ¿PUDO CERVANTES ESCRIBIR ESTE ENTREMÉS?

Lo único que podemos afirmar con cierta seguridad del autor del *Entremés de los romances* es que era un dramaturgo en activo entre finales del XVI y principios del XVII, posiblemente admirador de Góngora y enemigo de Lope de Vega, que estaba al corriente de las nuevas tendencias que había adoptado el entremés por esos años y que conocía prolijamente el Romancero nuevo.

Aunque no sabemos la fecha exacta de composición de la pieza, ni la direccionalidad de la imitación con la gran obra de Cervantes, podemos manejar una horquilla temporal que abarca desde la década de los 90 del siglo XVI hasta el año 1611, en que sale impreso en la *Parte tercera* de las comedias de Lope²²⁷. El autor hubo de ser un dramaturgo experimentado o, al menos, un gran conocedor de las novedades dramáticas del teatro breve, como prueba el hecho de que la pieza introduzca varias de

²²⁷ Es importante aclarar que, aunque la edición conservada de la *Parte tercera de las Comedias de Lope de Vega y otros* es de 1613, en casa de Miguel Serrano Vargas, parece probada la existencia de una edición de 1611.

las novedosas fórmulas entremesiles que se habían puesto en boga por esos años, como la forma versificada²²⁸, el romance dialogado y la duplicidad de la trama.

No cabe duda de que el autor conocía el Romancero nuevo, es decir, el compuesto por autores cultos, o, más exactamente, que tuvo acceso directo a él a través de algún cuaderno de romances, una flor o alguna de las partes del Romancero general, ya que tan sólo cuatro de las treinta y tres composiciones que contiene la pieza dubitada pertenecen al Romancero viejo popular, de transmisión oral. Aunque cabe la posibilidad de que el anónimo autor las conociera de memoria, los dieciocho versos de diferentes romances que Bartolo ensarta magistralmente cuando es llevado maltrecho a casa parecen haber sido compuestos teniendo delante alguna de esas colecciones de romances, muy probablemente, como apunta Rey Hazas, “una edición, hoy perdida o no localizada, de las *Flores*, basada en un texto próximo al de las *Flores primera, segunda y tercera* (Lisboa, 1592; Valencia; 1593)”, pero que, a diferencia de éste, incluía el romance “Si tienes el corazón” y la versión “Dígame tú la serrana”, en lugar de “Dígame tú, el aldeana” [2007: 6].

Todo parece indicar, además, que el anónimo entremesista era simpatizante de Góngora, ya que la pieza incluye varios romances suyos: dos de autoría segura (“Ensílleme el potro rucio”, que parodia uno de Lope de Vega, y “La más bella niña”) y algún otro de atribución dudosa (“Hermano Perico” y “Cabizbajo y pensativo”). Y que era enemigo de Lope de Vega, a quien ridiculiza, como apuntó Millé, tanto en la figura de Bartolo y Simocho como en la inclusión de un sinfín de romances de tema morisco, muy del gusto del *Fénix*.

Este vago perfil parece coincidir con el que hemos conformado de Cervantes, un dramaturgo que durante algunos años vio representar sus comedias “con general y gustoso aplauso de los oyentes”, que escribió al menos dos entremeses en verso, que fue admirador confeso de Góngora (“aquel que tiene de escribir la llave”) y enemigo

²²⁸ El primer entremés publicado en verso, como ya hemos recordado a lo largo de este trabajo, es el de *Melisendra*, que aparece en la *Primera parte de las comedias de Lope* (Valencia, 1605). A partir del año 1620, la forma versificada desbancará a la primitiva prosificada, sobre todo de la mano de Quiñones de Benavente.

declarado de Lope y “que conocía bien los romances, algunos de oídas, y los citaba de memoria” [Eisenberg, 1991: 57]²²⁹. Sin duda, las pequeñas pinceladas histórico-biográficas que ofrece la obra dubitada pueden corresponderse con la imagen del autor del *Quijote*, pero, también, con las de otros muchos autores del momento.

7.3. ANÁLISIS DE AUTORÍA BASADO EN LA COMPARACIÓN FORENSE DE TEXTOS

Tal y como hicimos en el capítulo anterior, vamos a llevar a cabo un estudio comparado que permita establecer el grado de similitud o disimilitud existente entre *Los romances* y las obras de Cervantes (esencialmente las de su teatro breve), para poder establecer, con mayor o menor grado de probabilidad, si es o no su autor²³⁰. En este caso, y dado que la comicidad de la pieza se sustenta en la inserción de varios fragmentos de una treintena de romances, nos vemos obligados a expurgarlos del texto, para que no interfieran en el análisis de ciertos rasgos²³¹.

²²⁹ Cervantes, en varias ocasiones, da muestras evidentes de su gusto por los romances. Menciona el Romancero general en *La gitanilla*, cita varios romances viejos en el *Quijote* —uno de los cuales es el de la versión moderna del antiguo del *Marqués de Mantua*, que también recoge el *Entremés de los romances*—, y asegura ante Apolo en el *Viaje del Parnaso* haber compuesto “romances infinitos”, siendo el mejor de todos el de los *Celos*. Como apunta Eisenberg, “la familiaridad de Cervantes con el cuerpo de poemas conocidos como el Romancero viejo es tan excepcional” que sólo es comparable al de “otros autores de comedias, como Juan de la Cueva o Lope de Vega” [1991: 57]. Parece destacar además “entre los iniciadores de la nueva manera de entender el romance, un género ya viejo en aquella época” y perfilarse como uno de los grandes romanceristas del momento, a la altura de otros como Liñán de Rianza, Lope de Vega, Góngora, Arias Girón, Luis de Vargas o Salinas [González, 1993: 610], no pudiendo olvidar que su *Romance de los celos* se cuenta entre los mejores de la producción [613].

²³⁰ Para todo lo referente a la delimitación del corpus de análisis y el tratamiento de los textos, véase el apartado 5.3.1.

²³¹ Estos son los 32 romances que se utilizan en el *Entremés de los romances*, por orden de aparición: (1) “Ensíllenme el asno rucio (vv. 31–45, 63–70, 71–74, 76–78); (2) “La más bella niña” (vv. 93–94, 97–98, 101–02, 105–06, 107–110, 126–127); (3) “Hermano Perico” (vv. 132–35, 136–39, 140–193); (4) “Cabizbajo y pensativo” (vv. 204–205, 206–221, 226–229); (5) “Mira, Tarfe que a Daraja” (vv. 238–251, 258–261, 401); (6) “Mal hubiese el caballero” (vv. 287–288); (7) “De Mantua salió el marqués” (vv. 292–293, 306–313, 314–317, 318–321, 324–325, 324–339, 344–345, 349–355, 358–361); (8) “¿Dónde estás, señora mía?” (vv. 294–301); (9) “Dime, Bencerraje amigo” (vv. 384–389); (10) “Galiana está en

Tras eliminar los versos que son reproducción literal o cuasi literal de conocidos romances, el texto se reduce casi a la mitad, lo que complica enormemente el estudio de autoría²³². Cabe destacar, no obstante, que el texto resultante —al que asignamos la etiqueta *Dub_EV_romances_R*— supera la frontera de las 1000 palabras en la que, según Hirst y Feiguina, está el límite para poder ofrecer análisis cuantitativos de cierta fiabilidad [2007].

OBRA	EXTENSIÓN	OBRA	EXTENSIÓN
Dub_EV_romances	2243 palabras	Dub_EV_romances_R	1164 palabras

Tabla 35. Diferencias entre la extensión del *Entremés de los romances* (*Dub_EV_romances*) y la extensión real, resultante de eliminar los poemas insertados (*Dub_EV_romances_R*)

Teniendo en cuenta las particularidades del vehículo de expresión de este entremés (verso) y la extensión real con la que podemos trabajar para identificar a su autor (1164 palabras), nuestro análisis va a contemplar las siguientes variables:

- Tema, argumento, personajes, segmentación y estructura
- Análisis cualitativo de marcas idiosincrásicas
- Análisis métrico (metros, estrofas y rimas)

Toledo” (v. 393); (11) “Si tienes el corazón” (vv. 396–397); (12) “Por una nueva ocasión” (v. 400); (13) “Rendido está Reduán” (v. 402); (14) “De las montañas de Jaca” (v. 403); (15) “Elicio, un pobre pastor” (v. 404); (16) “En una pobre cabaña” (v. 405); (17) “Con semblante desdeñoso” (v. 406); (18) “De pechos sobre una vara” (v. 407); (19) “Bravonel de Zaragoza” (v. 408); (20). “Discurriendo en la batalla” (v. 409); (21) “Por muchas partes herido” (v. 410); (22) “Rotas las sangrientas armas” (v. 411); (23) “Sale la estrella de Venus” (v. 412); (24) “Rompiendo la mar de España” (v. 413); (25) “Después que con alboroto” (v. 414); (26) “Entró la malmaridada” (v. 415); (27) “En un caballo ruano” (v. 416); (28) “¡Afuera, afuera! ¡Aparta, aparta!” (v. 417); (29) “Todos dicen que soy muerto” (v. 424); (30) “Dígame tú la serrana” (v. 425); (31) “Azarque, indignado y fiero” (vv. 426–427); (32) “Azarque vive en Ocaña” (v. 429); (33) “Ardiéndose estaba Troya” (v. 470).

²³² Eliminamos los versos que son reproducción literal y también aquellos en los que la copia es más que evidente, aunque se alteren algunas palabras, como ocurre, por ejemplo, con la afirmación de Bartolo “veintidós palos me han dado, / que el menor era mortal”, que en boca de Valdovinos era “veinte y dos heridas tengo / que cada una es mortal”.

- Distribución de la frecuencia de los 5-gramas de caracteres más empleados en los textos (*character 5-grams*)
- Distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*)
- Análisis de componentes principales (PCA) aplicado a la *Delta distance*
- Clasificación supervisada por aprendizaje automático (Classify)
- Proporción palabras de contenido / palabras de función
- Medidas de riqueza y densidad léxica
- Porcentaje de similitud textual (*similarity*) y porcentajes de vocabulario compartido una sola vez (*Once only in both*), de vocabulario compartido más de una vez (*Shared > 1*) y de vocabulario único y exclusivo (*Only in*)
- Verificación de autoría con General Imposters (GI)
- Longitud de secuencias de palabras idénticas (*verbatim*)
- Estudio de las palabras de función: orden de frecuencia, frecuencia relativa y PCA
- Frecuencia de clases de palabras
- Uso y frecuencia relativa de las conjunciones coordinadas y subordinadas adverbiales
- Búsqueda de palabras y expresiones del texto dubitado en corpus de referencia de la época

Aunque el estudio de autoría de este entremés contempla, en esencia, las mismas variables que el de *La cárcel de Sevilla*, las diferencias de extensión (3240 palabras / 1164 palabras) y composición (prosa / verso) entre ambos textos condiciona la sustitución de alguna de ellas. En este caso, como se ve en el listado, no analizamos los bigramas y trigramas de clases de palabras, ya que el orden de estos elementos podría venir condicionado por la forma versificada, y añadimos, además del estudio métrico,

otras variables menos susceptibles a la extensión, como aquéllas relacionadas con la proporción de palabras de contenido y palabras de función, las medidas de riqueza y densidad léxica o el uso y la frecuencia relativa de ciertas conjunciones. Tampoco utilizamos en este caso las variables relacionadas con la puntuación, a pesar de su fiabilidad [Blasco y Ruiz Urbón, 2009: 42; De la Rosa y Suárez, 2016: 410], pues, desgraciadamente, no conservamos los manuscritos autógrafos.

Una vez más, nos vemos en la obligación de recordar que será la acumulación de varios rasgos o factores (y no sólo de uno o dos aislados) los que permitirán apoyar o rechazar la autoría cervantina.

7.3.1. Tema, argumento, personajes, segmentación y estructura

7.3.1.1. Tema central y argumento

El *Entremés de los romances* recoge la historia de Bartolo, un pobre labrador tan obsesionado con el Romancero que llega a creerse uno de esos héroes que aparecen en los textos. Arrastrado por su locura, que le lleva a recitar versos romanceriles sin parar, decide abandonar su casa, su hacienda y a su mujer Teresa para hacerse soldado y marcharse a la guerra en compañía de su escudero Bandurrio. Tras pasar varias aventuras y sufrir algún que otro varapalo, la familia de Bartolo consigue llevarlo de vuelta a casa para que descanse en su cama y se recupere del molimiento.

El motivo de esta piecilla teatral recuerda, como hemos visto, al del *Quijote*. Si Bartolo enloquece “de leer el romancero” hasta creerse el caballero de algún romance morisco y “quiere embarcase en la guerra” para luchar contra los ingleses, a Alonso Quijano se le seca el cerebro “del poco dormir y del mucho leer” libros de caballerías hasta decidir “hacerse caballero andante, y irse por todo el mundo con sus armas y caballo a buscar las aventuras y a ejercitarse en todo aquello que él había leído que los caballeros andantes se ejercitaban” (*Quijote*, I, 1).

Tenemos noticias documentadas de la época que recogen ciertos casos reales de personas a quienes la lectura de libros condujeron a la locura: Alonso de Fuentes relata

en la *Summa de philosophia natural* (1547) que conocía el caso de un hombre que se sabía de memoria el *Palmerín de Oliva*; López Pinciano recuerda en su *Philosophia antigua poética* (1596) que un amigo suyo quedó largo tiempo inconsciente tras leer la muerte de Amadís; y Gaspar Galcerán de Pinós, conde de Quimera, explica en el 1600 que un estudiante de Salamanca:

en lugar de leer sus liciones, leía en un libro de caballerías, y como hallase en él que uno de aquellos famosos caballeros estaba en aprietos por unos villanos, levantóse de donde estaba, y empuñando un montante, comenzó a jugarlo por el aposento y esgremir en el aire, y como lo sintiesen sus compañeros, acudiendo a saber lo que era, y él respondió: “Déjenme vuestras mercedes, que leía esto y esto, y defiendo a este caballero [Américo Castro, 1960: 295].

Pero si las similitudes temáticas entre el anónimo y la obra magna cervantina parecen formar parte del imaginario común de la época, las concomitancias argumentales son tan notorias que nos obligan a pensar en una relación inter-textual entre ambos escritos, mucho mayor que la que pudiera desprenderse de las anécdotas reales contadas por autores coetáneos. Como ya han señalado los críticos en numerosas ocasiones, los paralelismos entre la trama de uno y otro texto son altamente significativos: tanto Quijote como Bartolo enloquecen a causa del consumo excesivo de cierto género literario; se visten con armaduras para emular a los héroes de sus lecturas²³³; abandonan sus hogares y se ponen al servicio de la sociedad; en su lucha por defender a los débiles y desprotegidos, se enfrentan a todo aquel que consideran que está cometiendo una injusticia; en una de sus aventuras, caen de su caballo y son apaleados con su propia lanza, culpando luego al pobre equino de su derrota; se creen el enamorado Valdovinos y pronuncian los mismos versos que éste entona en el famoso romance del Marqués de Mantua²³⁴; son encontrados por uno o varios lugareños que

²³³ Recordemos, no obstante, que la imagen de Arlecchino montado sobre un asno y portando una vieja armadura, con una espada rota, una olla de hierro en la cabeza y una lanza en lo alto de la mano, se conserva entre los *scenari* de la compañía *I Gelosi* (Imagen 1).

²³⁴ La citación del mismo romance no es tan significativa —pues era “historia sabida de los niños, no ignorada de los mozos, celebrada y aun creída de los viejos” (*Quijote*, I, 5)— como el hecho de que ambos no citen el viejo romance sino una nueva versión del mismo. En todo caso, Azcune y Fernández Nieto advierten que Cervantes “no sigue el texto del *Entremés*, sino una versión distinta: la de la *Flor II...*, prueba evidente de que aún no conocía el texto del *Entremés*. Es claro que si le estuviese imitando, no

deciden llevarlos de nuevo a su pueblo; reviven por el camino el pasaje de un romance morisco entre Bencerraje y algún alcalde; son llevados a casa y acostados en su cama, en donde duermen su locura; se despiertan sobresaltados y se levantan de la cama dando voces y gritos, interrumpiendo la acción que en ese momento estaban desempeñando sus familiares. Argumentalmente, lo que más distancia a ambos textos es la figura del escudero, ya que, si Bartolo cuenta con Bandurrio, Quijote no va acompañado aún del fiel Sancho Panza.

El hecho de que Cervantes parodie ciertos romances en los capítulos concernientes a la primera salida de don Quijote cuando, en realidad, su obra acomete parodiar los libros de caballería, y que tal citación aparezca en mucha menor medida en los capítulos posteriores, en donde llega a desvanecerse casi por completo, parece acercar aún más ambos escritos, induciéndonos a presuponer, como apuntaba Varo, la prioridad del entremés y a conjeturarlo como bosquejo de la pieza cervantina. Sin embargo, no nos parece lógico que, si Cervantes se basó en este entremés para componer los primeros capítulos del Quijote de 1605, eliminara la figura del escudero, que es, sin duda, una de las claves del éxito de su novela. Este matiz, por tanto, nos lleva a pensar en que es el anónimo entremesista el que se inspira en el texto de Cervantes, y no a la inversa.

Las analogías temático-argumentales entre el entremés y los primeros capítulos del *Quijote* sirven para demostrar la influencia de uno en el otro, pero no su autoría común. De hecho, advertimos que existen aspectos relacionados con los mecanismos de operación de la parodia y la elección y utilización de los materiales insertados que parecen alejar a Cervantes de la redacción del anónimo.

La elaboración de un mismo tema aparece de modo tan diverso en los dos textos que parece poco probable que hayan sido compuestos por la misma mano, siendo notorio que la parodia del anónimo es más simple y menos efectiva que la cervantina. La funcionalidad de las piezas insertadas en el *Entremés de los romances* es desatinada,

tendría necesidad de acudir a otra fuente para copiar los únicos cuatro versos del romance que se incluyen en la novela” [2004: 108].

incurriendo su autor en un gravísimo error de base: si el personaje enloquecido por la lectura del Romancero es Bartolo, ¿por qué todos los personajes acaban recitando conocidos versos entremesiles incluso cuando el labrador no está en escena? La finalidad paródica de la pieza dramática pierde eficacia en el momento en el que participan del recitado todos los demás, pues lo que empieza presentándose como signo manifiesto de la locura de un personaje concreto acaba convirtiéndose en rasgo común dialógico, anulando en parte la comicidad y transformando el entremés en un incrustado de romances —ora en boca de Bartolo, ora de Mari Crespa, Teresa, Pero Tanto, Perico, Dorotea, Simocho o Antón— con una intención poco clara.

No ocurre lo mismo en Cervantes, en donde se respeta y mantiene la coherencia paródica en todo momento. Si el enloquecido es don Quijote, él y sólo él debe hablar como caballero andante y, cuando el resto de los personajes traspasan la frontera del mundo real para introducirse en su mundo ideado, hay una serie de fines prácticos que lo justifican; así, por ejemplo, durante esta primera salida, se nos dice que el ventero le sigue el juego y determina “hablarle comediantemente” (I, 2) y que “por tener que reír aquella noche, determinó de seguirle el humor” (I, 3) y armarle caballero andante.

Si el entremesista fundamenta su parodia en la inserción *a trochemoche* de versos romanceriles, Cervantes lo hace en el modo de elaboración del discurso de don Quijote, construido a partir del lenguaje arcaico y rimbombante que utilizaban los caballeros andantes y de la alusión a pasajes y personajes conocidos de tales libros (y, en muy escasa proporción, se sirve de calcos literales, que, además, son de mucha menor extensión que los insertados en el *Entremés de los romances*). Los recursos lingüísticos de estilización paródica cervantina superan el mero *copy-paste* del entremés y se conforman a través del remedo de la lengua antigua castellana, que contribuye a la formación del discurso polifónico del que hablaba Bajtin. En este sentido, la novela cervantina mantiene una coherencia pragmática de la que adolece el entremés.

Lo que resulta más curioso, y sobre lo que nadie ha llamado su atención, es que el único romance que sabemos de Cervantes, el *Romances de los celos*, parece identificarse con uno recogido por primera vez en la *Flor* publicada en Valencia por Felipe Mey en 1593 [González, 1993: 613], casualmente la misma *Flor* que parece

haber tenido delante el anónimo entremesista al escribir el *Entremés de los romances*²³⁵. Si Cervantes fuese el autor de un entremés que construye su particular parodia al Romancero a través de la inserción de varios romances que parece calcar teniendo delante un ejemplar de la *Flor* de 1593 en la que, además, se recoge su romance máspreciado, ¿no resulta chocante que, dado el gusto del alcalaíno a recursos como la intertextualidad y la metaliteratura entre sus propios escritos, no introdujese algunos versos de *El romance de los celos* en lugar de haberse centrado en los de Góngora, Lope y otros autores?

Por tanto, y a pesar de la coincidencia temático-argumental entre ambos textos, la elaboración y el ensamblaje de los materiales y los mecanismos de operación de la parodia, no permiten identificar al anónimo entremesista con Cervantes²³⁶.

7.3.1.2. Personajes

En este entremés intervienen cuatro mujeres (Mari Crespa, Teresa, Dorotea y Marica), seis hombres (Bartolo, Bandurrio, Pero Tanto, Perico, Antón y Simocho) y un

²³⁵ Dice González: “El único romance atribuible indudablemente a Cervantes que se recoge en el *Romancero general* de 1600 es el *Romance de los celos*, lo cual no quiere decir que no haya otros suyos” [1993: 610]. Aunque es cierto que tal romance comparte título con el citado por Cervantes y que en él aparecen como protagonistas Lauso y Silena, los pastores de la *Galatea*, no podemos calificar su autoría como indubitada.

²³⁶ Aunque el objetivo de este trabajo no es establecer la direccionalidad de *Entremés de los romances* con respecto a los primeros capítulos del *Quijote*, sí nos gustaría dar nuestra opinión al respecto, pues este hecho afecta, de algún modo, al asunto de la autoría. Los cuatro argumentos que nos llevan a presuponer que el *Entremés* es posterior, tres de los cuales ya hemos ido desgranando en las páginas precedentes, son los siguientes:

- la presencia del escudero;
- la falta de coherencia en la caracterización del habla de los personajes cuerdos, que hablan en romance incluso cuando Bartolo no está en escena;
- el empleo de la forma versificada, que empieza a ponerse de moda tras la publicación de *Melisendra* (1605), y del romance dialogado, que no se da en las comedias hasta principios del XVII [Asensio: 1971];
- el empleo de la fórmula sintética *vuesastedes*, frente al uso analítico anterior *vuesa/s merced/es*, que es el que encontramos en el *Quijote*.

número indeterminado de músicos, por lo que el cómputo rondaría los 12 personajes. El número total es acorde al de los entremeses de Cervantes, que, recordemos, fluctúa entre siete y catorce (véase apartado 6.3.1.2.), pero no lo es la distribución, pues ninguno de los suyos incluye tantos personajes femeninos [Agostini, 1964: 275].

Los nombres de Teresa, Dorotea y Crespo (que no Crespa) son propios del teatro cervantino (Teresa aparece en *El retablo de las maravillas*, Dorotea en *La entretenida* y Crespo en *Pedro de Urdemalas*); nombres que, por otro lado, son habituales de la literatura del momento. De hecho, la elección de los nombres viene condicionada por los romances que se insertan en él, por lo que no pueden achacarse al ingenio del autor (en “Ensílleme el potro rucio” aparecen Teresa la del Villar y Antón Morente; en “Hermano Perico”, Bartolo, Perico y Dorotea; y en “Cabizbajo y pensativo”, Cimocho y Marica).

7.3.1.3. Segmentación y estructura externa e interna

Tomando como referencia la clasificación de los ocho entremeses de Cervantes de Asensio [1970: 19-41], *Los romances* se incluiría dentro de los entremeses de acción (junto a *El vizcaíno fingido*, *La cueva de Salamanca* y *El viejo celoso*), ya que muestra predilección por las historietas moldeadas en la imaginación colectiva.

Si la comparación entre el tema y el argumento de la pieza anónima y los capítulos de la primera salida del hidalgo manchego hacían pensar en una dependencia textual, lo mismo ocurre al analizar su segmentación. Según ha señalado Baras, “del análisis de la forma interna del *Quijote* se desprende la conclusión de que la obra está constituida, no como parodia estricta de los libros de caballerías, sino como una serie de secuencias entremesiles”, lo que le lleva a calificarlo de “entremés narrado en prosa” [1993: 331-334]. El análisis de las funciones de los dos textos revela su analogía, pues ambos están constituidos por cuatro unidades de farsa que forman una misma secuencia, tal y como mostramos en la siguiente tabla:

<p><i>Entremés</i> de los romances 1 secuencia → 4 unidades de farsa</p>	<ol style="list-style-type: none"> (1) Salida de Bartolo de su pueblo, convertido en caballero. (2) Aventura de la riña de los pastores Simocho y Marica. (3) Bartolo es encontrado por sus familiares y amigos y llevado de vuelta a casa. (4) Incidente erótico a cargo de los mozos Perico y Dorotea / Bartolo se levanta de la cama e interrumpe el baile nupcial con unos versos sobre la destrucción de Troya, en los que alude al fuego.
<p><i>Quijote</i> (I, 1-6) 1 secuencia → 4 unidades</p>	<ol style="list-style-type: none"> (1) Salida de don Quijote de su pueblo, convertido en caballero andante. (2) Aventura de los mercaderes de Toledo. (3) Don Quijote es hallado por un vecino y conducido a casa. (4) Escrutinio de la biblioteca / Los gritos de don Quijote, que confunde al cura con el arzobispo Turín, interrumpen la quema de libros.

Tabla 36. Segmentación del *Entremés de los romances* y de la primera salida de Don Quijote

El análisis de la estructuración externa de *Los romances*, que, recordemos, es la que tiene que ver con las partes que conforman el entremés (comienzos, movimientos escénicos, acotaciones, lugares de acción y desenlaces) [Baras Escolá, ed., 2012: 176], revela lo siguiente:

- El entremés comienza con la salida de Mari Crespa, Teresa, Perico y Pero Tanto, que conversan sobre la descabellada intención de Bartolo de hacerse soldado. Bartolo, el protagonista, no hace su aparición hasta la segunda escena (verso 31), una vez que los espectadores, gracias a la conversación que han mantenido sus amigos y familiares, ya tienen una idea clara de quién es y qué le pasa. No es la manera en la que suelen empezar los entremeses cervantinos de protagonista único, en los que la acción casi siempre arranca con éste ya en escena, al igual que ocurre con don Quijote. No es habitual, tampoco, que el entremés arranque con cuatro personajes en escena, algo que sólo se da en *La elección de los alcaldes de Daganzo*.

- La pieza es de un altísimo movimiento escénico, con hasta 22 salidas y entradas de personajes; muy por encima, por tanto, de Cervantes, en donde los entremeses con mayor movimiento escénico son *La guarda cuidadosa* y *El vizcaíno fingido*, con 18 [Agostini, 1964b: 505-509].
- La gran mayoría de las acotaciones sirven para marcar la salida y entrada de los personajes (“sale(n)”, “va(n)se”, “llévale”), sus acciones (“dale [...] de palos”, “cantan”, “baila”...) o su indumentaria (“vestidos de labradores”, “armado de papel, de risa, y en un caballo de caña”, “en camisa”), como también hace Cervantes en sus entremeses. No encontramos, sin embargo, ninguna acotación que aluda al estado anímico de los personajes.
- El entremés se desarrolla entre el pueblo de Bartolo y las tierras a las que sale en busca de aventuras. Los dos primeros cambios de escena coinciden con el vaciado del escenario, como hace siempre Cervantes en los cuatro entremeses en los que no se cumple la unidad de lugar (*El vizcaíno fingido*, *El retablo de las maravillas*, *La cueva de Salamanca* y *El viejo celoso*) [Agostini, 1964b: 513-514]; el último, sin embargo, es distinto, ya que el movimiento se marca con la salida de Teresa, que se dirige al pueblo de manera anticipada, y es su vuelta a escena la que implica la llegada del resto de personajes a ese lugar:

ANTÓN. Tened, Bandurrio, de ahí,
y empezad a caminar.

CRESPA. Adelántate tú, hija.

TERESA. Yo voy volando al lugar.

(*Vase*).

[...]

(*Sale Teresa*)

TERESA. Ellos sean bien llegados,
que ya está hecha la cama.

- El entremés finaliza con música, canto y baile, como es lo propio en Cervantes y en todas las obras de este género.

La estructura interna de *Los romances* parece corresponderse con la de Cervantes, caracterizada, como explicamos en el capítulo anterior, por la utilización de unos “protagonistas destemplados en situaciones límites que la máquina social del resto de personajes pretende convertir en un estado más o menos ‘templado’ o armonioso” [Pérez de León: 2005: 109]. No cabe duda de que Bartolo es un personaje destemplado y que la misión de sus familiares y amigos es la de reconducir su conducta; ahora bien, el autor comete la incoherencia de dotar a algunos personajes del rasgo destemplado de Bartolo (el habla arromanzada), incluso cuando él está fuera de escena. Así ocurre, por ejemplo, en el siguiente fragmento, en el que Perico, Mari Crespa, Teresa y Pero Tanto, sin justificación lógica, reproducen los primeros catorce versos de un romance de Góngora²³⁷:

PERICO.	Si es verdad aqueso. mi hermana será <i>la más bella niña</i> <i>de nuestro lugar.</i>
CRESPA.	¡Pobre de la triste, pues para su mal <i>hoy es viuda y sola</i> <i>y ayer por casar!</i>
TERESA.	¿Quién, señora madre, muerta no se cae, <i>viendo que sus ojos</i> <i>a la guerra van?</i>
TANTO.	La pobre Teresa, harta de llorar, <i>a su madre dice</i> <i>que escuche su mal.</i>
TERESA.	<i>Dulce madre mía,</i> <i>¿quién no ha de llorar,</i> <i>aunque tenga el pecho</i> <i>como un pedernal?</i>
CRESPA.	Calla, por tu vida, que remedio habrá.
	[...]

²³⁷ Los versos marcados en cursiva se corresponden con el comienzo del romance de Góngora “La más bella niña”, según la versión aparecida en el *Romancero general*.

TERESA. Cuando no le hallemos
dejadme llorar,
orillas del mar.

7.3.2. Análisis cualitativo de marcas idiosincrásicas

En este apartado vamos a analizar los mismos rasgos que identificamos en el capítulo anterior como marcas características del idiolecto entremesil cervantino, tomando como referencia los trabajos de Agostini [1964b], Rosenblat [1978], Gutiérrez Cuadrado [1988], Ariza Viguera, [1998], Maestro [1998] y Baras Escolá [ed., 2012]. Será la presencia o ausencia de varias de estas marcas en el *Entremés de los romances* lo que permita establecer si su comportamiento es afín o no a Cervantes.

7.3.2.1. Presencia del superlativo en *-ísimo*

Como explicamos en el apartado 6.3.2.1., el superlativo absoluto en *-ísimo*, heredado de la lengua italiana, es habitual en Cervantes, a pesar de que su uso todavía no era frecuente a finales del XVI y principios del XVII [Rosenblat, 1978: 190; Gutiérrez Cuadrado: 1998, 819-856].

Los entremeses cervantinos presentan una frecuencia notoria del adjetivo en grado superlativo, principalmente del absoluto en forma analítica *-ísimo*, que aparece de 1 a 3 veces en todos ellos (Tabla 37). Cervantes se sirve de modo sistemático de este sufijo para intensificar los adjetivos hasta su grado máximo (sintiendo una predilección especial por las formas *bonísimo* y *grandísimo*), incrementar el valor expresivo de adverbios de adjetivales en *-mente* (*bastantísimamente*, *discretísimamente*) o reforzar algún pronombre (*infinitísimos*), en contraposición con la nula ocurrencia que tiene *-ísimo* en el texto dubitado real (Dub_EV_romances_R). El anónimo tiene preferencia por el adjetivo en grado positivo o, a lo sumo, en grado comparativo de igualdad (ya que tampoco aparecen comparativos de inferioridad o superioridad), pero no por la forma analítica del superlativo.

SUPERLATIVO ABSOLUTO EN <i>-ÍSIMO</i> ²³⁸		
	Casos	N total
Cerv_EP_juez	sanísimo; *infinitísimos; *bastantísimamente	3
Cerv_EP_guarda	bonísima	1
Cerv_EP_vizcaino	bonísima (2); *discretísimamente	3
Cerv_EP_retablo	dignísimo; valentísimo; grandísimos	3
Cerv_EP_cueva	grandísimo	1
Cerv_EP_viejo	bonísimo	1
Cerv_EV_rufian	clarísima	1
Cerv_EV_eleccion	benditísimo; rarísima	2
Dub_EV_romances	∅	0
Dub_EV_romances_R	∅	0

Tabla 37. Estudio del superlativo absoluto *-ísimo*

7.3.2.2. Casos de *leísmo*

En el *Entremés de los romances* aparecen casos de *leísmo*, como también ocurre en Cervantes y, en general, en los autores del centro peninsular. Pero mientras en los textos de Cervantes encontramos *leísmos* tanto de persona como de cosa, en el dubitado “real” sólo de persona (“le persuadí”, “le arreaba”, etc.), nunca de cosa.

7.3.3.3. Utilización de adverbios en *-mente*.

Como vimos en el apartado 6.3.3.3., la formación de adverbios modales con el sufijo *-mente* es otro de los rasgos propios de la escritura cervantina. Este tipo de adverbios aparece en siete de sus ocho entremeses, pero no en el *Entremés de los romances*.

²³⁸ El asterisco señala los casos en los que el sufijo *-ísimo* no se ha aplicado a adjetivos.

ADVERBIOS EN <i>-MENTE</i>		
	Casos	N total
Cerv_EP_juez	encarecidamente, rectamente, bastantísimamente	3
Cerv_EP_guarda	altamente	1
Cerv_EP_vizcaino	verdaderamente, discretísimamente, ligeramente	3
Cerv_EP_retablo	principalmente, hondamente, especialmente, finalmente, verdaderamente,	5
Cerv_EP_cueva	∅	0
Cerv_EP_viejo	forzosamente, solamente	2
Cerv_EV_rufian	piadosamente	1
Cerv_EV_eleccion	bravamente, solamente	2
Dub_EV_romances	∅	0
Dub_EV_romances_R	∅	0

Tabla 38. Estudio del adverbio en *-mente*

7.3.2.4. *Uso del diminutivo en -ico*

El diminutivo en *-ico* era muy frecuente en época de los Reyes Católicos, pero su uso se ve muy relegado a partir del siglo XVI, en donde *-ito* pasa a ser la variante predominante. Como ha estudiado Ariza Viguera, [1998], este sufijo sigue teniendo en Cervantes una alta vitalidad, muy superior a la de otros autores de su época²³⁹.

Es habitual encontrar en Cervantes sufijos diminutivos en *-ico*, con valor afectivo o despectivo. Este rasgo aparece en sus ocho entremeses, aplicado tanto a adjetivos y nombres comunes (*bonica, morenico, polvico, borrico, demonicos*, etc.) como a nombres propios (un Perico, una Lorencica y dos Cristinicas). En el *Entremés de los romances* aparecen los personajes de Perico y Marica; no obstante, estos nombres son utilizados para parodiar a los personajes de dos conocidos romances (“Hermano Perico” y “Cabizbajo y pensativo”, respectivamente), por lo que, en ningún caso, son invención

²³⁹ De todos los autores analizados por Ariza Viguera, Cervantes es el único que usa más la variante *-ico* (22,70%) que *-ito* (17,40%). El porcentaje de Cervantes contrasta con el escaso 8,21% de Góngora, el 8,8% de Lope de Vega y el 13% de Quevedo [1998: 353].

del anónimo autor. Al margen de estos dos nombres, tomados del Romancero nuevo, no encontramos ningún otro diminutivo en *-ico*, ni en el texto completo ni en el real.

SUFIJO DIMINUTIVO EN <i>-ICO</i>		
	Casos	N total
Cerv_EP_juez	bonica, Perico, cedacico	3
Cerv_EP_guarda	Cristinica (7), añicos,	8
Cerv_EP_vizcaino	polvico, menudico, borrico	3
Cerv_EP_retablo	morenico,	1
Cerv_EP_cueva	Cristinica (5), demonicos	6
Cerv_EP_viejo	Cristinica (9), Lorencica (3), tortolica	13
Cerv_EV_rufian	levadicas, melindricos	2
Cerv_EV_eleccion	polvico (2)	2
Dub_EV_romances	Perico, Marica	2
Dub_EV_romances_R	Perico, Marica	2 > Ø

Tabla 39. Estudio del sufijo diminutivo *-ico*

7.3.2.5. Empleo de reduplicaciones, anáforas y repeticiones diseminadas

Ya vimos en el apartado 6.3.2.5. el gusto de Cervantes por las repeticiones léxicas. En sus entremeses es habitual encontrar reduplicaciones de nombres, verbos, interjecciones o modos adverbiales (“al cabo, al cabo”, “luego, luego”, “duelos, duelos”, duelos, duelos”, “échense todos, échense todos”, etc.), reiteraciones de los comienzos de las intervenciones de los distintos personajes (“Séanme testigos que...”, “abre aquí, alcaide, que nos comen...”), a modo de anáfora, y repeticiones diseminadas, en donde lo reiterado actúa a modo de estribillo, como ocurría con el “*de ex illis*” en un pasaje de *El retablo de las maravillas*. En el texto dubitado “real” aparece la reduplicación de palabras en varias ocasiones (“celuchos, celuchos”, “arre arre” y el enfático “¡Fuego, fuego!, ¡Fuego, fuego!” final); también hay dos casos del tipo de anáfora antes descrito, aunque mucho menos significativos, ya que la reduplicación sólo afecta a un elemento (“ay”, “adiós” o “teneldo”). No hay, sin embargo, ninguna repetición diseminada.

7.3.2.6. *Uso frecuente de figuras de dicción*

Otra característica de Cervantes, estrechamente relacionada con el punto anterior, es su predilección por las figuras de dicción y, en especial, por la duplicación de palabras con variación de género gramatical, la sucesión de variantes flexivas de un mismo verbo, la utilización de vocablos con igual significante pero diferente significado y la proximidad de palabras con un mismo lexema o con sonidos afines. El texto dubitado real es mucho más pobre en estos recursos, pues sólo encontramos un caso de duplicación con alternancia genérica (“moro ni mora”) y el juego de palabras “Tanto por tanto”, en donde la forma *tanto* hace referencia primero al nombre del personaje y luego al cuantificador.

7.3.2.7. *Creación de nuevas palabras*

Como ya explicamos en el capítulo anterior, es de sobra conocido que Cervantes crea palabras nuevas con finalidad humorística, ya sea por composición, derivación o parasíntesis. En el texto dubitado “real” no encontramos ninguna palabra creada a través de estos mecanismos, lo que contrasta con el ingenio creador de Cervantes.

7.3.2.8. *Inclusión de expresiones latinas*

En el texto dubitado no hay expresiones o palabras en latín, que sí aparecen en seis de los ocho entremeses de Cervantes (véase Tabla 21).

7.3.2.9. *Predilección por los sinónimos coordinados*

La acumulación coordinada de sinónimos, que Rosenblat resaltaba como propia de la lengua cervantina [1978], aparece en una ocasión en el dubitado “real”, cuando Bandurrio dice: “yo no hallo / senda alguna ni vereda”. No obstante, la inserción del indefinido detrás del primer elemento de la bimetración la distancia del uso generalizado que veíamos en las obras cervantinas, del tipo “hospitalera ni enfermera”,

“muela ni diente”, “vaivenes y empujones”, etc. El resto de estructuras que recogen la coordinación de dos elementos no son propiamente sinonímicas (“sola y moza”, “estoque ni puñal”, etc.).

6.3.2.10. Utilización pragmática de los diálogos de enunciados interrogativos

Otro de los rasgos característicos de los entremeses de Cervantes, como ha estudiado Maestro [1998], es el empleo recurrente de diálogos fundamentados en enunciados y discursos interrogativos. Este rasgo está relacionado con los vínculos interpersonales que se establecen entre los sujetos dramáticos, siendo generalmente el personaje débil el que pregunta y el dominante el que responde.

Las interrogaciones del texto dubitado “real” son generalmente preguntas retóricas que sólo sirven para enfatizar un enunciado (como el “¿qué he de hacer yo sola?” de Teresa o el “¿dejar a mi hija quieres?” de su madre, tras conocer que Bartolo se quiere marchar a la guerra), aunque en otros casos sí son preguntas que sirven para la obtención de información y muestran las relaciones de superioridad/inferioridad entre los personajes; así ocurre, por ejemplo, en la escena en la que Mari Crespa y Antón interrogan a Teresa sobre el incidente que está ocurriendo en la azotea entre sus hijos, Perico y Dorotea. Aunque no de manera tan significativa, tal vez por su menor extensión, este rasgo pragmático sí aparece en cierta medida en el dubitado “real”.

6.3.2.11. Adaptación del lenguaje al nivel sociocultural de los personajes

Es de sobra conocido que Cervantes caracteriza a sus personajes a través del lenguaje y que éstos, además, no se expresan de igual modo ante un interlocutor que ante otro [Agostini, 1964b]. Esta coherencia cervantina no se da en el entremés dubitado, en donde el habla arromanzada, que debía individualizar a Bartolo, acaba siendo adoptada por todos los personajes, a pesar de que éstos siguen manteniendo la cordura. Algo que ya había advertido Asensio en su *Itinerario*:

El autor vacila entre dos actitudes mal armonizadas: el poner en boca de Bartolo, como producto de su locura, trechos diversos del Romancero; y el escenificar, alterando

levemente lo narrativo, poemas enteros dialogados por los restantes personajes, a los que pinta normales y sensatos. Los mismos que acaban de maldecir el Romancero, se enzarzan en recitar un papel de la historia del Marqués de Mantua o una letrilla de Góngora. La locura literaria reclama método y continuidad [1971: 75].

Sí hay coherencia en las fórmulas de tratamiento con las que se relacionan los personajes, pues Antón sólo usa el tratamiento *vuesastedes* en la parte final, para dirigirse a los músicos, pero nunca cuando habla con sus hijos, su nuera o su consuegra. Es importante notar, no obstante, que ésta no es la variante utilizada por Cervantes ni en los entremeses ni en los capítulos de la primera salida de don Quijote, en donde todavía se opta por las formas analíticas previas (*vuestra/s merced/es*, *vuesa/s merced/es* y *sus mercedes*), lo que nos lleva a pensar que el *Entremés de los romances* es posterior en el tiempo²⁴⁰.

FÓRMULA DE TRATAMIENTO “VUESTRA MERCED”		
	Formas analíticas (A) y sintéticas (S)	N total
Cerv_EP_juez	vuesa merced (13), vuestas mercedes (2)	A = 16 S = Ø
Cerv_EP_guarda	vuesa merced (24), vuestas mercedes (2)	A = 26 S = Ø
Cerv_EP_vizcaino	vuesa merced (50), vuestas mercedes (3)	A = 53 S = Ø
Cerv_EP_retablo	vuesa merced (13), vuestas mercedes (4)	A = 17 S = Ø
Cerv_EP_cueva	vuesa merced (7)	A = 7 S = Ø
Cerv_EP_viejo	vuesa merced (10), vuestas mercedes (3)	A = 13 S = Ø
Cerv_EV_rufian	vuesa merced (1), voacé (3)	A = 1 S = 3
Cerv_EV_eleccion	vuestas mercedes (2), sus mercedes (2)	A = 4 S = Ø
Cerv_N_quijote1_1.6	vuestra merced (15), vuestras mercedes (3) su merced (2)	A = 20 S = Ø
Dub_EV_romances	vuesastedes (1)	A = 0 S = 1
Dub_EV_romances_R	vuesastedes (1)	A = 0 S = 1

Tabla 40. Tipo de variante de la fórmula de tratamiento *vuestra merced*

²⁴⁰ En época cervantina empezó a generalizarse el empleo de la fórmula de tratamiento *vuestra merced*, “que por desgaste, a través de una serie de variantes, se fue transformando en *usted*” [Rosenblat, 1978:180]. Cervantes sólo usa la forma analítica *voacé* para caracterizar el habla de germanía, pero nunca la forma *vuesasted*, que es la que aparece en *Los romances*.

De la observación de todos estos rasgos podemos concluir que, aunque alguno está presente en la parte “real” del *Entremés de los romances*, la ausencia de superlativos en *-ísimo*, de diminutivos en *-ico*, de adverbios en *-mente* y de latinajos, la falta de creatividad a la hora de componer palabras nuevas con fines humorísticos, la escasa presencia de sinónimos coordinados, la falta de coherencia en la caracterización de los personajes a través del lenguaje y, sobre todo, la elección de la variante *vuesasted/es* en lugar de *vuestra/s merced/es*, *vuesa/s merced/es* y *sus mercedes*, alejan la posible identificación del autor anónimo con Cervantes.

7.3.3. Análisis métrico (metros, estrofas y rimas)

La métrica, entendida ésta como el conjunto de principios que rigen el lenguaje versificado, se ha establecido como otro de los rasgos que permiten determinar la autoría de un texto. Para Plecháč, Bobenhausen y Hammeric, “versification is a reasonable stylometric marker, the strength of which is comparable to the other markers traditionally used in stylometry (such as the frequencies of the most frequent words and the frequencies of the most frequent character n-grams)” [2018: 29]. Según estos tres autores, la ventaja de las marcas relacionadas con el ritmo y la rima es triple: (i) pueden aplicarse a textos breves; (ii) están menos condicionadas por el tema que la frecuencia de palabras o caracteres; y (iii) son más subconscientes y, por tanto, difícilmente imitables²⁴¹.

Desde este enfoque, la comparación entre la métrica utilizada por Cervantes en sus obras dramáticas y la utilizada por el autor anónimo permitirá establecer si siguen o no un mismo patrón. Centraremos nuestro estudio contrastivo en el metro, la estrofa y la

²⁴¹ Ya a principios la década de 1920, Tomashevsky usó la versificación para demostrar que el final del poema inacabado *Sirena*, de Pushkin, que supuestamente Dmitry Zuev había encontrado en 1889, era en realidad una falsificación [1923]. Más recientemente la ha aplicado Tarlinskaja al estudio de Shakespeare, para saber qué partes del poema *Enrique VIII* fueron escritas por él y cuáles por John Fletcher [2014]. En el ámbito de la literatura española, García-Reidy ha usado la comparación de estrofas, en combinación con otras evidencias, para atribuir la comedia *Siempre ayuda a la verdad* a Lope de Vega [2019].

rima como variables a analizar, haciendo especial hincapié en los entremeses, de los que Schevill, en su “Introducción” a *El teatro de Cervantes*, dice lo siguiente:

Por lo que atañe a la versificación, ha de notarse que los dos entremeses escritos en verso, *El Rufián viudo* y *La Elección de los alcaldes de Daganzo*, pertenecen al estilo del siglo XVI, habiendo en el primero 339 versos sueltos, endecasílabos, y 315 en el segundo. Los romances que en ambas obras se encuentran, son cantados, reflejando en esto el gusto popular, que de seguro tenía rancio abolengo en la materia. Los versos que se leen en los demás entremeses, son glosas y bailes cantados, con su correspondiente estribillo (como en *El Viejo celoso*), y no carecen de viveza y brío [1922: 158-159].

Aunque a día de hoy los análisis métricos automáticos ofrecen una alta fiabilidad, de en torno al 95%²⁴², para nuestro estudio hemos optado por el análisis manual, ya que la métrica de todas las obras cervantinas ha sido excelentemente notada por Domínguez Caparrós [2002]²⁴³. Por tanto, en los análisis que mostramos sobre el metro, la estrofa y la rima, nosotros sólo aportamos lo referente al texto dubitado y el calculo de ratios y promedios.

7.3.3.1. Estudio del metro

Tal y como reflejamos en las siguientes tablas, el metro más empleado en las ocho comedias y los dos entremeses en verso de Cervantes es el octosílabo (62,13%), seguido por el endecasílabo (33,07%) y, ya muy de lejos, por el heptasílabo (2,83%), el hexasílabo (1,64%) y otras formas métricas de muy escasa significación (0,90%). Las ratios muestran un comportamiento muy diferenciado entre el tipo de metro empleado en las comedias, en las que predomina el octosílabo (75,08%), y el empleado en los entremeses, en los que lo hace el endecasílabo (85,02%).

²⁴² Navarro Colorado ha estudiado los patrones acentuales del endecasílabo español en el Siglo de Oro mediante análisis distante o macroanálisis [2016]. Según explica en su estudio, el sistema automático de escansión métrica tiene una fiabilidad del 95%.

²⁴³ Sólo hemos tenido en cuenta las piezas teatrales de autoría indubitadamente cervantina, omitiendo, por tanto, los datos relativos a *El trato de Argel* y la *Numancia*.

CAPÍTULO 7

El *Entremés de los romances* consta de 476 versos y se compone de 361 octosílabos, 112 hexasílabos, 2 heptasílabos y 1 pentasílabo²⁴⁴. Tiene lógica que el metro más utilizado sea el octosílabo (83,98%), si tenemos en cuenta que éste era el metro predominante en el Romancero.

	TIPO DE METRO					Total
	Hex.	Hept.	Oct.	End.	Otros	
Cerv_C_gallardo	---	16	2481	637	---	3134
Cerv_C_casa	36	208	1632	874	6	2756
Cerv_C_baños	4	153	2511	410	15	3093
Cerv_C_rufian	---	30	2156	660	---	2846
Cerv_C_sultana	16	80	2362	499	4	2961
Cerv_C_laberinto	---	102	1886	1039	51	3078
Cerv_C_entretenida	95	226	2332	434	---	3087
Cerv_C_pedro	77	30	2822	238	13	3180
Σ TOTAL C	228	845	18182	4791	89	24135
Cerv_EV_rufian	---	---	61	339	---	400
Cerv_EV_eleccion	12	---	20	313	22	367
Σ TOTAL V	12	---	81	652	22	767
DUB_EV_romances	112	2	361	0	1	476
DUB_EV_romances_R	37	0	194	0	0	231

Tabla 41. Distribución del metro (Nº total de versos)

²⁴⁴ Como explica Pérez Lasheras, “la irregularidad radica en la composición que «Cantan los músicos» (vv. 462-65): «Frescos ventecillos / favor os pido / que me anego en las olas / del mar del olvido». Letrilla heterométrica que no ha sido recordada por ninguno de los estudiosos del *Entremés*. Plantea algunos problemas, ya que la versión que encontramos en nuestra obra la recoge J. M. Alín (*El Cancionero español de tipo tradicional*, Madrid, Taurus, 1968, 689, p.684), que cita como fuente el *Entremés*. Con la variante «Frescos airecitos» la encontramos en el *Laberinto Amoroso de Juan de Chen*, 1618 (ed. facsímil de Blecua, Valencia, Castalia, 1953, pp. 26-28, de la ed. de Zaragoza, 1638) y en la *Primavera y flor de romances, segunda parte* (Zaragoza, 1629; ed. de Rodríguez-Moñino, Madrid, castalia, 1972, p.142). Los versos comentados constituyen el estribillo de un villancico, glosado por las mudanzas. De su carácter tradicional nos habla el hecho de que Góngora escribiera el romance «Frescos airecillos» (Millé, 29, 1590)” [1988: 61].

	TIPO DE METRO				
	Hex.	Hept.	Oct.	End.	Otros
Cerv_C_gallardo	0	0,51	79,16	20,33	0
Cerv_C_casa	1,30	7,55	59,22	31,71	0,22
Cerv_C_baños	0,13	4,95	81,18	13,26	0,48
Cerv_C_rufian	0	1,05	75,76	23,19	0
Cerv_C_sultana	0,54	2,70	79,77	16,85	0,14
Cerv_C_laberinto	0	3,31	61,27	33,76	1,66
Cerv_C_entretenida	3,08	7,32	75,54	14,06	0
Cerv_C_pedro	2,42	0,94	88,74	7,49	0,41
Σ TOTAL C	0,93	3,54	75,08	20,08	0,36
Cerv_EV_rufian	0	0	15,25	84,75	0
Cerv_EV_eleccion	3,27	0	5,45	85,29	5,99
Σ TOTAL V	1,64	0	10,35	85,02	3,00
\bar{X} PROMEDIO C + V	1,07	2,83	62,13	33,07	0,90
DUB_EV_romances	23,74	0,21	75,84	0	0,21
DUB_EV_romances_R	15,02	0	83,98	0	0

Tabla 42. Ratios de distribución del metro (%)

El análisis de las ratios de distribución del metro en el entremés dubitado y en los dos indubitados cervantinos revela un comportamiento altamente diferenciado. Todo el *Entremés de los romances* adopta metros de arte menor (penta, hexa, hepta y octosílabos), algo que lo distancia de *El rufián viudo* y *La elección de los alcaldes de Daganzo*, en donde el endecasílabo es el metro predilecto (85,02%); además, la ratio del octosílabo en el dubitado real (83,98%) está muy alejada de la de los dos entremeses indubitados en verso, en donde apenas supera el 10% (10,35%). Todas estas disonancias, no obstante, podrían estar condicionadas por el tipo de metro que exige un entremés

como el de *Los romances*, en donde, recordemos, se pretenden parodiar las composiciones del Romancero²⁴⁵.

7.3.3.2. Estudio de la estrofa y de los tipos de poema

Aunque los conceptos de poema y estrofa se deslindan muy bien teóricamente, en la práctica ambas nociones están entrelazadas de modo inextricable, más aún en el teatro²⁴⁶. En este apartado, por tanto, vamos a realizar un análisis comparado de formas, que incluyen poemas estróficos (soneto, copla real...), poemas no estróficos (romance, romancillo...) y estrofas propiamente dichas (quintilla, redondilla...).

El *Entremés de los romances* se caracteriza por el empleo abusivo de versos romanceriles —como era esperable—, aunque también tienen cabida otras formas²⁴⁷. Entre las formas más independientes, es decir, aquellas menos condicionadas por la inserción de versos del Romancero (↑), encontramos las coplas reales (quintillas) y las redondillas:

²⁴⁵ La identificación del romance con el verso octosílabo no se rastrea en ningún teórico hasta el siglo XIX. Los manuales y preceptivas del siglo de Oro aceptan el uso del verso hexasílabo (Díaz Rengifo así lo contempla en su *Arte poética española*) o endecasílabo (*Diccionario de autoridades*) para este tipo de composiciones; es más, Cervantes incluye en *La gitanilla* un romance hexasílabo (al que él mismo llama “romance”) y el recolector de los poemas del *Romancero general* inserta hasta 100 folios de poesías en versos endecasílabos. Por tanto, si bien teóricamente el deseo de parodiar el romance no parece implicar el mayor uso del octosílabo, ya que podría haberse satirizado este género literario recurriendo al verso endecasílabo (muy habitual en la métrica cervantina, aunque nunca en forma métrica de romance), si lo es la práctica, pues la realidad es que su presencia no parece documentarse hasta finales del siglo de Oro [Clarke, 1952].

²⁴⁶ “La estrofa es cada uno de los grupos de versos o subdivisiones de un poema que las contenga (p. ej. el *Laberinto de Fortuna* de Juan de Mena, escrito en estrofas de arte mayor), mientras el poema es el conjunto total” [Paraíso, 2000: 144].

²⁴⁷ En este esquema hemos señalado con ↓ los poemas o estrofas en las que predominan los versos tomados de piezas romanceriles (y que, por tanto, el empleo de esa forma métrica parece estar impuesta y no corresponder a la voluntad del autor), con ↑ los poemas o estrofas en los que predominan los versos originales (y que, por tanto, pueden ser tomadas como marcas idiosincrásicas de autoría) y con ↔ los poemas o estrofas en las que la combinación parece equilibrada.

<i>ENTREMÉS DE LOS ROMANCES</i>			
Tipo de estrofa	Nº de verso	Versos totales	Autor
Copla real: quintillas octosílabas consonantes (<i>aabba</i>) ²⁴⁸	1-30	30	↑
Romance octosílabo asonante en <i>é-e</i>	31-82	52	↓
Romancillo hexasílabo agudo en <i>-á</i>	83-131	49	↔
Romancillo hexasílabo agudo en <i>á-e</i>	132-193	62	↓
Redondillas octosílabas consonantes (<i>abba</i>)	194-205	12	↑
Romance octosílabo asonante en <i>ó-o</i>	206-237	32	↓
Romance octosílabo asonante en <i>á-e</i>	238-271	34	↓
Redondillas octosílabas consonantes (<i>abba</i>)	272-291	20	↑
Romance octosílabo agudo en <i>-á</i>	292-383	92	↓
Romance octosílabo asonante en <i>á-a</i>	384-465	82	↓ / ↑
Letrilla heterométrica	466-469	4	↓
Romance octosílabo asonante en <i>é-a</i>	470-476	7	↓

Tabla 43. Formas empleadas en el *Entremés de los romances*

Como mostramos en la siguiente tabla, los poemas y estrofas utilizados por Cervantes en sus composiciones teatrales son muy variados, e incluyen desde formas propias de la métrica castellana, como la copla real, el villancico, el romance, la redondilla, el perché, la seguidilla o el zéjel, hasta formas heredadas de la italiana, como el sexteto lira, la estancia o el soneto. Las formas más complejas —tales como los ovillejos, los estrambotes o los versos de cabo roto— aparecen en sus obras en prosa o en sus composiciones líricas, pero no en sus comedias y entremeses:

²⁴⁸ Las coplas reales se componen de dos quintillas simétricas (5+5) con rimas independientes, es decir, con cuatro rimas en diversas combinaciones [Paraíso, 2000: 277-278]. En este caso, son tres coplas reales de dos quintillas cada una, todas ellas de tipo 5 (*aabba*). El esquema métrico de estos primeros treinta versos —sobre el que, como vimos, ha llamado la atención Rodríguez López-Vázquez— sería éste: *aabba:ccddc:eeffe:gghhg:ijji:kkllk*.

CAPÍTULO 7

TEATRO COMPLETO DE CERVANTES				
Estrofas	C	EV	EP	Total
Canción medieval	0	0	20	20
Cantar + glosa	0	18	0	18
Copla de arte mayor	0	0	8	8
Copla mixta novena	90	0	0	90
Copla real	6658	0	0	6658
Cuarteto-lira suelto	578	0	0	578
Glosa (~ copla real)	0	0	11	11
Endecasílabo suelto	116	596	0	712
Endecasílabo suelto + 2 pareados	43	0	0	43
Endecasílabo suelto + pareado	1543	56	0	1599
Estancia canción italiana	182	0	0	182
Estríbillo + romance	0	0	52	52
Lira	15	0	0	15
Octava real	1611	0	0	1611
Octava-lira	8	0	0	8
Octavilla de pie quebrado	72	0	0	72
Pareado	2	0	0	2
<i>Perché</i>	30	0	0	30
Quintilla	35	0	0	35
Redondilla	9154	0	0	9154
Romance (hexasílabo) / romancillo	125	12	0	137
Romance (octosílabo)	2091	81	0	2172
Romance + estríbillo	26	0	0	26
Romance consonante	0	0	36	36
Seguidilla	4	4	8	16
Seguidilla + glosa	75	0	0	75
Serventesio	4	0	0	4
Sexteto-lira	636	0	0	636
Soneto	98	0	0	98
Soneto cabo roto	14	0	0	14
Soneto estrambote	34	0	0	34

ESTUDIO DE AUTORÍA DEL *ENTREMÉS DE LOS ROMANCES*

Tercetillo	3	0	0	3
Terceto encadenado	793	0	0	793
Villancico	57	0	26	83
Zéjel	25	0	0	25
¿Versos?	13	0	0	13
Σ TOTAL	24135	767	161	25063

Tabla 44. Formas empleadas en el teatro cervantino (Nº de versos)

Este cuadro esquemático de formas es bastante significativo, aunque no por ello debemos olvidar que el número total no es proporcional a la frecuencia de empleo, pues, por definición, la redondilla consta de cuatro versos, la quintilla de cinco, el sexteto-lira de seis, la octava real de ocho, el soneto de catorce y el romance ni siquiera tiene un número determinado de versos. A pesar de esta limitación, el recuento nos permite advertir que los poemas o estrofas con más peso en la obra dramática de Cervantes son la redondilla (9154 vv.), la copla real (6658 vv.), diferentes variedades de romance (2371 vv.) y el endecasílabo suelto (2354 vv.). Centrándonos en los dos entremeses en verso, la variedad se reduce a cuatro formas estróficas: el endecasílabo suelto (652 versos), el romance octosílabo y hexasílabo (93 versos), el cantar con glosa (18 versos) y la seguidilla (4 versos), siendo sólo la primera la que se utiliza en pasajes no cantados.

Para poder profundizar en la comparación, vamos a mostrar más detalladamente el esquema métrico de los entremeses en verso de Cervantes:

<i>ENTREMÉS DE EL RUFIÓN VIUDO</i>		
Estrofa	400 VERSOS	
	Nº de verso	Versos totales
Endecasílabo suelto	1-326, 331-343	339
Romance octosílabo asonante en -á	327-330, 344-400	61

Tabla 45. Formas empleadas en el *Entremés de el rufián viudo*

<i>ENTREMÉS DE LA ELECCIÓN DE LOS ALCALDES DE DAGANZO</i>		
Estrofa	367 VERSOS	
	Nº de verso	Versos totales
Endecasílabo suelto	1-257	257
Romance octosílabo asonante en <i>á-o</i>	258-269, 273-280	20
Seguidilla asonante / Estribillo (de 11 y 12 sílabas)	281-282, 299-300	4
Endecasílabo suelto con pareado final en <i>-ánta</i>	270-272, 283-286, 301-302, 307-308, 323-367	56
Romancillo hexasílabo asonante en <i>ó-e</i>	287-298	12
Cantar con glosa (de 6, 7 y 8 sílabas)	303-306, 309-322	18

Tabla 46. Formas empleadas en el *Entremés de la elección de los alcaldes de Daganzo*

El esquema de *El rufián viudo* se reduce al endecasílabo suelto (1-326 y 331-343) y el romance octosílabo asonante en *-a* (327-330 y 344-400) y el de *La elección de los alcaldes de Daganzo*, algo más variado, se compone de endecasílabo suelto (1-257), romance octosílabo asonante en *a-o* (258-269, 273-280), seguidilla asonante (281-282 y 299-300), endecasílabo suelto con pareado final en *-anta* (270-272, 283-286, 301-302, 307-308, 323-367), romancillo hexasílabo asonante en *o-e* (287-298) y cantar con glosa (303-306 y 309-322). Sin embargo, es importante señalar que, en realidad, salvo en los pasajes cantados (en donde introducen romances, glosas y la seguidilla a modo de estribillo), los dos entremeses en verso de Cervantes usan en exclusividad el endecasílabo suelto.

El análisis contrastivo del texto dubitado y los indubitados cervantinos nos lleva a establecer las siguientes consideraciones:

- La principal forma métrica del texto dubitado es el romance, que aparece unas veces en hexasílabos y otras en octosílabos. Los romances son habituales en Cervantes, como en el resto de los autores de la época, por lo que su empleo no resulta muy idiosincrásico (menos aún dentro en una obra que se construye como parodia de éstos). Cabe hacer notar, no obstante, que a pesar del gusto de Cervantes por el romance —que soporta el 50% del peso de los versos de ocho sílabas incluidos en el *Quijote* [Domínguez Caparrós, 2002: 60]—, éste sólo

está presente en tres ocasiones en sus entremeses en verso, en donde, además, es cantado. También nos resulta llamativo que el texto dubitado incluya dos largas tiradas de romances hexasílabos, primero en *-á* y luego en *á-e*, ya que Cervantes casi nunca usa este tipo de composición y, cuando lo hace, emplea una rima distinta; así, entre sus obras dramáticas, encontramos un romancillo en *La entretenida* (en *é-o*), dos en *Pedro de Urdemalas* (en *-ó* y *-é*) y uno en *La elección de los alcaldes de Daganzo* (en *ó-e*).

- La redondilla tiene una presencia notoria en todo el teatro de la época, también en el de Cervantes, por lo que su empleo *a priori* no parece muy significativo. Aún así, el uso tiende a ser diferente: Cervantes generalmente ofrece largas tiradas de redondillas, de entre 5 y 156 estrofas consecutivas (es decir, de 44 y 624 versos seguidos en redondillas), y, en rarísimas ocasiones, aparecen cantidades inferiores a 5, que es lo que ocurre en *Los romances*, con una tirada de 3 redondillas (12 versos)²⁴⁹. También es importante señalar que ninguno de los dos entremeses en verso de Cervantes tiene redondillas.
- La copla real, que abre el entremés dubitado, es la forma preferida de Cervantes, aunque no para el género entremesil. En los entremeses en verso, los personajes optan por el endecasílabo suelto, nunca por la copla real (ni por la quintilla suelta). Teniendo en cuenta que los dos entremeses en verso de Cervantes se decantan por el uso del endecasílabo suelto, choca que este entremés, de ser suyo, empiece con tres coplas reales (nunca usadas por él en este género); más aún si tenemos en cuenta que en ese pasaje inicial no ha aparecido todavía Bartolo, eje central de la parodia, por lo que la contraposición de arte mayor vs arte menor hubiese sido más efectiva desde el punto de vista dramático. Además, el esquema rítmico elegido por el autor anónimo del entremés para todas las quintillas (tipo 5) no es el habitual del teatro cervantino, si bien se da en *La casa de los celos* (véase nota 226).

²⁴⁹ Este uso, aunque no es el habitual en Cervantes, sí se da en alguna ocasión, En *El rufián dichoso* encontramos una tirada de 2 redondillas; en *La gran sultana*, una de 4 y otra de 3; en *Pedro de Urdemalas*, una de 3; y en *El gallardo español*, una de 2.

7.3.3.3. *Estudio de la rima*

La rima del texto dubitado alterna consonancia y asonancia, aunque con un claro predominio de la segunda sobre la primera. Los romances son asonantes y combinan las rimas oxítonas (en *-á*) con las paroxítonas (en *á-e*, *é-e*, *é-a* y *ó-o*); las redondillas y las quintillas, como exige la norma, llevan rima consonante, ya sea aguda (en *-í*, *-ó* y *-él*) o llana (en *-ánton*, *-éis*, *-ígo*, *-érra*, *-éro*, *-ánces*-, *-ésa*, *-ádo*, *-ár*, *-ímos*, *-érba*, *-ába*, *-álo*, *-éda*, *-érme*, *-álga*, *-ínos*, *-ésta* y *-éndo*). En Cervantes prácticamente todos los romances son asonantes, con predominio de la rima llana sobre la aguda. En los dos entremeses en verso, en donde los romances son cantados, hay uno en oxítona (en *-á*)²⁵⁰ y el resto en paroxítona (en *-á-o* y *ó-e*).

La falta de ingenio del entremesista anónimo queda de manifiesto si tenemos en cuenta que, a pesar del escaso número de versos de cosecha propia, recurra en más de una ocasión a la rima de palabras idénticas, como *muchacha* (vv. 435 y 443), *hablar* (vv. 299 y 303), *muestra* (vv. 289 y 290) o *tal* (vv. 374 y 379). La pobreza de la rima cervantina también es bastante notoria, pero está algo más trabajada que la del entremés anónimo. Cervantes suele optar por el mozdobre, es decir, por la rima de palabras emparentadas etimológicamente (como *culpa* / *disculpa*, *cubren* / *descubren*, *corta* / *acorta*, *amigo* / *enemigo*, *cordura* / *locura*, etc.), la rima homónima (*mora*) o la utilización de unos mismos tiempos verbales (*paga* / *satisfaga*) o pronombres enclíticos (*notólo* / *mirólo*), y aunque también utiliza la rima idéntica, no lo hace en un porcentaje tan llamativo como el anónimo. También encontramos en Cervantes casos de rimas ricas, es decir, aquellas en las que la reiteración de unos mismos sonidos se extiende más allá de la vocal tónica (*demonio* / *testimonio*, *afeites* / *aceites* o *beldad* / *frialdad*), que no aparecen en el dubitado real.

²⁵⁰ Aunque en el esquema métrico de *El rufián viudo* reflejamos dos romances cantados en *-á*, se trata en verdad de la misma composición. Los cuatro versos que aparecen en primer lugar (vv. 327-330) se vuelven a repetir más adelante (vv. 344-347).

7.3.4. Distribución de la frecuencia de los 5-gramas de caracteres más empleados en los textos (*character 5-grams*)

Como hemos visto al testar el corpus indubitado, la aplicación del algoritmo Eder's Delta a los 5-gramas de caracteres resulta bastante fiable para discriminar autorías (Gráficos 16 y 17), aunque no tanto como el análisis con 1-grama de palabra. Vamos a repetir ese análisis incorporando el dubitado real (Dub_EV_romances_R):

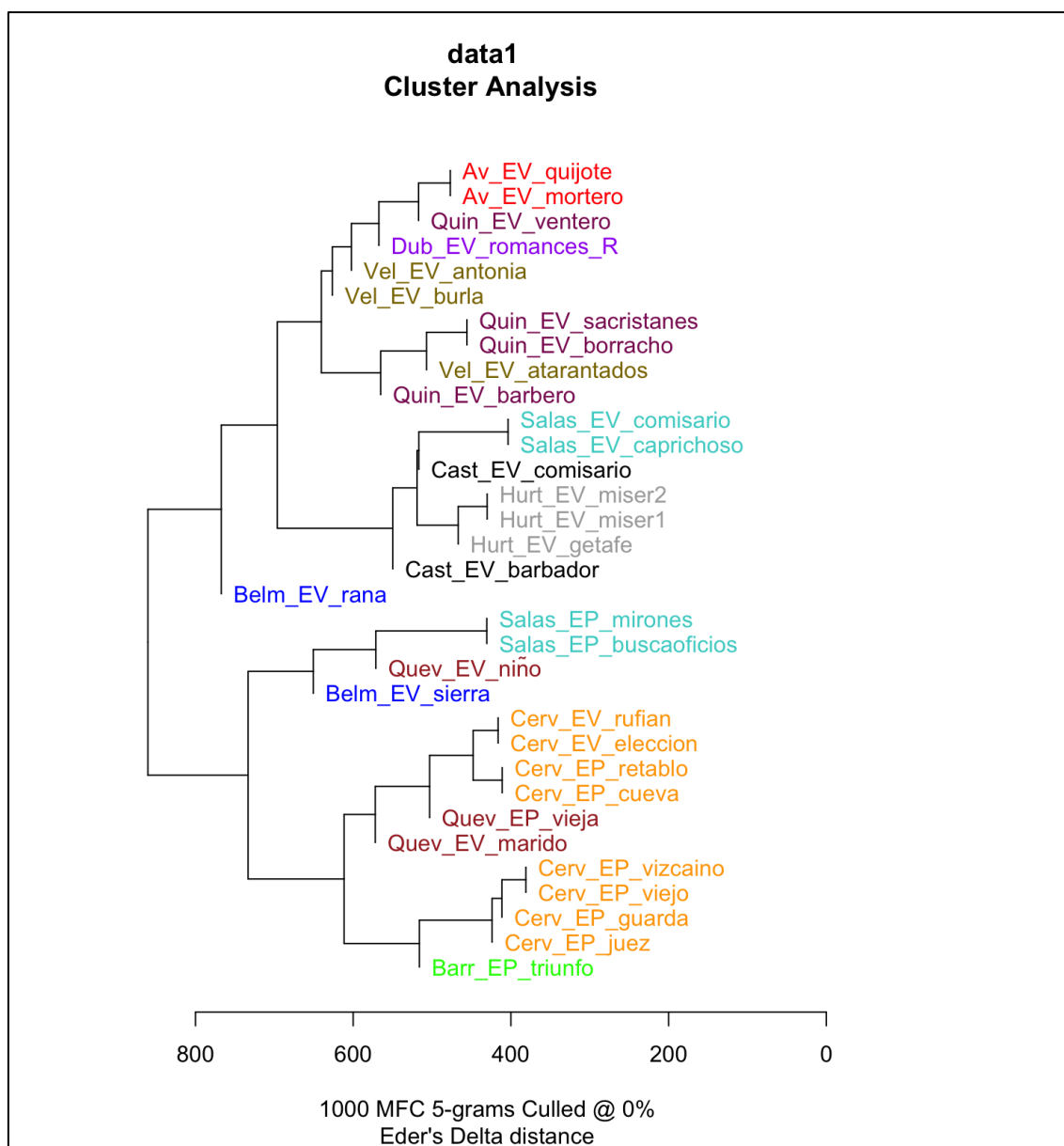


Gráfico 72. Clúster basado en la frecuencia de 5-gramas de caracteres para *Los romances* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFC 1000 5-grams. Culled 0%)

Tal y como muestra el dendrograma anterior, la parte real del *Entremés de los romances* presenta un comportamiento ortográfico muy diferente al de los ocho entremeses cervantinos. Es cierto que el análisis no funciona al 100%, pero los fallos de agrupación nunca se producen en las asociaciones más directas; algo que puede verse también en el siguiente árbol de consenso, que contiene un resumen de las iteraciones que permanecen estables entre las 800 y las 1000 MFW (*Bootstrap*):

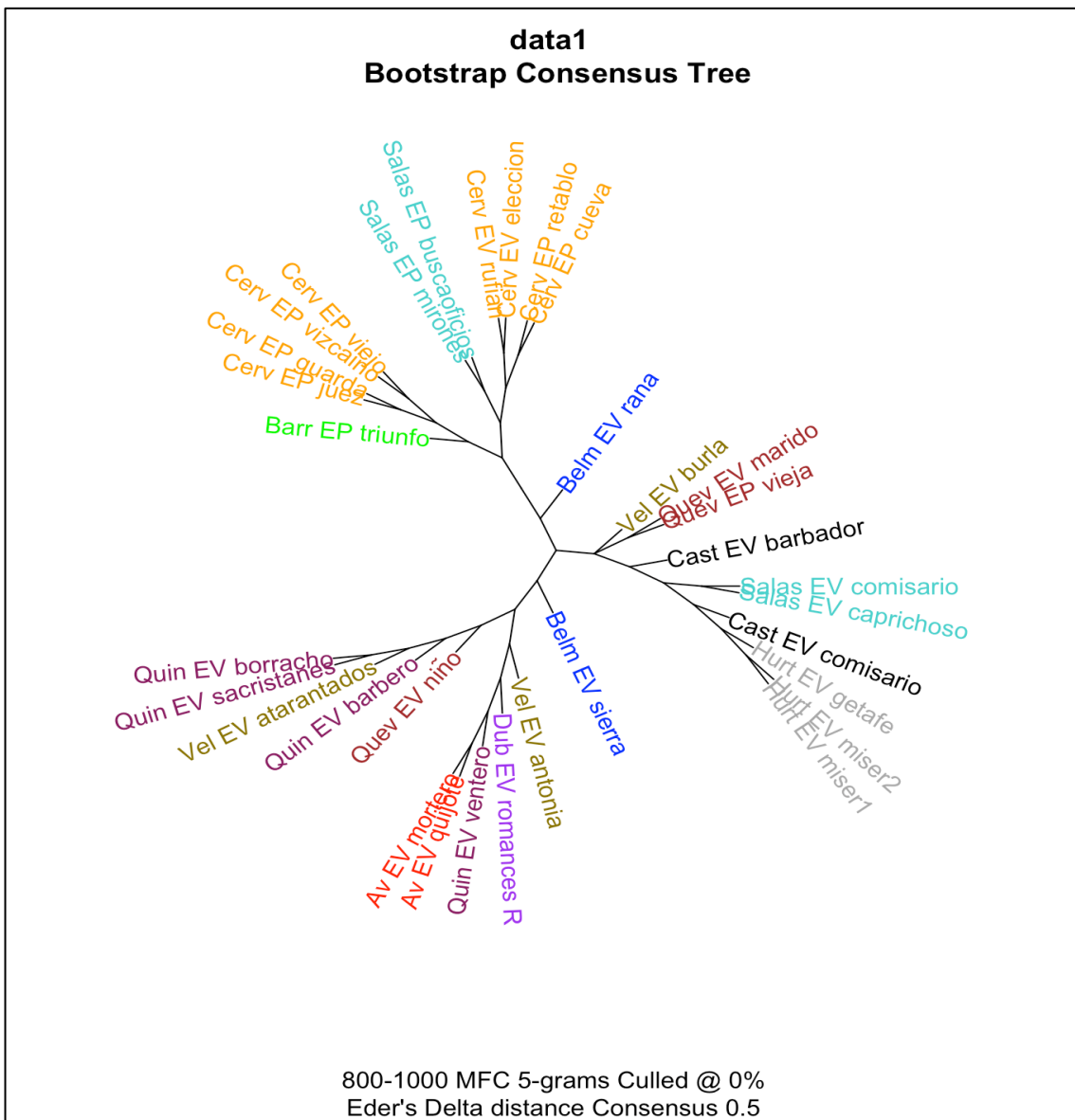


Gráfico 73. Árbol de consenso basado en la frecuencia de 5-gramas de caracteres para *Los romances* (RStudio. Stylo. Bootstrap Consensus Tree. 0.5 consensus strength. Eder's Delta. MFW 800-1000)

7.3.5. Distribución de la frecuencia de las palabras más empleadas en los textos (*Delta distance*)

En el capítulo 5, al realizar el estudio previo de *El entremés de los romances*, llevamos a cabo el análisis clúster de la distribución de la frecuencia de las 1000 palabras más frecuentes (MFW), aplicando el algoritmo Eder's delta. Como se aprecia en dicho análisis (Gráfico 31), el entremés dubitado se encuentra en una rama totalmente opuesta a la cervantina, asociándose directamente con los dos entremeses de Ávila. En aquella ocasión utilizamos el texto completo (DUB_EV_romances), por lo que vamos a repetir ahora el análisis con el texto real (DUB_EV_romances_R), primero con el corpus de indubitados de autor conocido y después sólo con los indubitados en verso.

Como puede verse en los dos dendrogramas resultantes (Gráficos 74 y 75), el comportamiento del texto atribuido a Cervantes es muy distinto al de sus entremeses de autoría cierta. En el Gráfico 74, *Los romances* se enmarca en una rama en la que están todos los entremeses de Quevedo, Quiñones de Benavente, Vélez de Guevara y Belmonte, pero sin asociarse directamente con ninguno de ellos; en la rama opuesta, se encuentran los entremeses de Cervantes, Hurtado de Mendoza, Ávila, Barrionuevo, Salas Barbadillo y Castillo de Solórzano. En el Gráfico 75, sin embargo, *Los romances* se aproxima a los dos entremeses de Ávila, aunque la asociación en este caso tampoco es directa. Esta variabilidad, unida al hecho de que la agrupación se produzca siempre muy a la izquierda, nos lleva a pensar que el verdadero autor del texto anónimo no está entre ninguno de los que conforman este corpus.

En todo caso, el *Entremés de los romances* (DUB_EV_romances_R) está siempre muy alejado de los textos cervantinos. Así queda patente en los dos dendrogramas analizados y también en el árbol de consenso que hemos obtenido tras analizar la distribución de las frecuencias de las 800-1000 MFW entre este texto y los ocho entremeses de Cervantes (Gráfico 75).

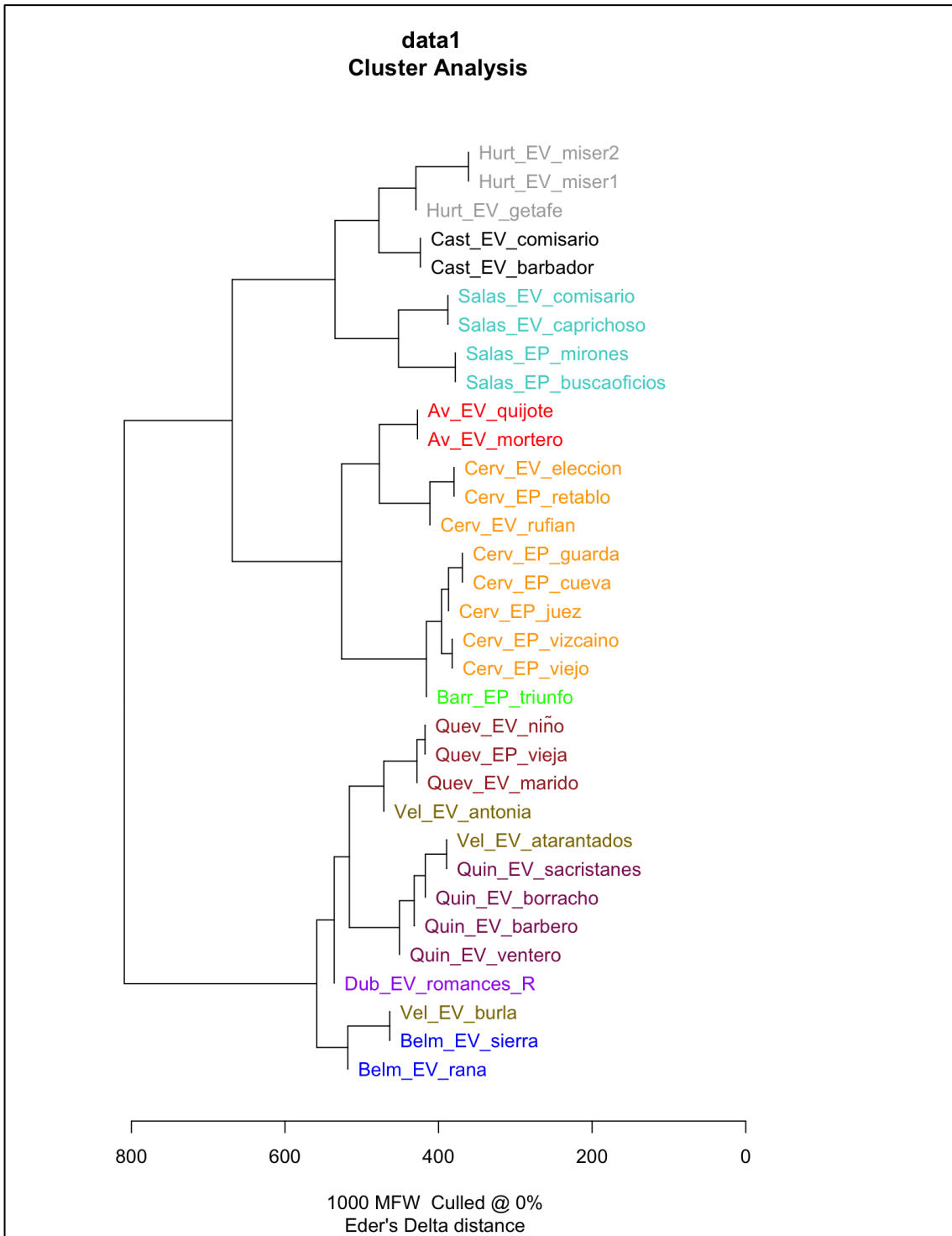


Gráfico 74. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses de autor conocido y *Los romances* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

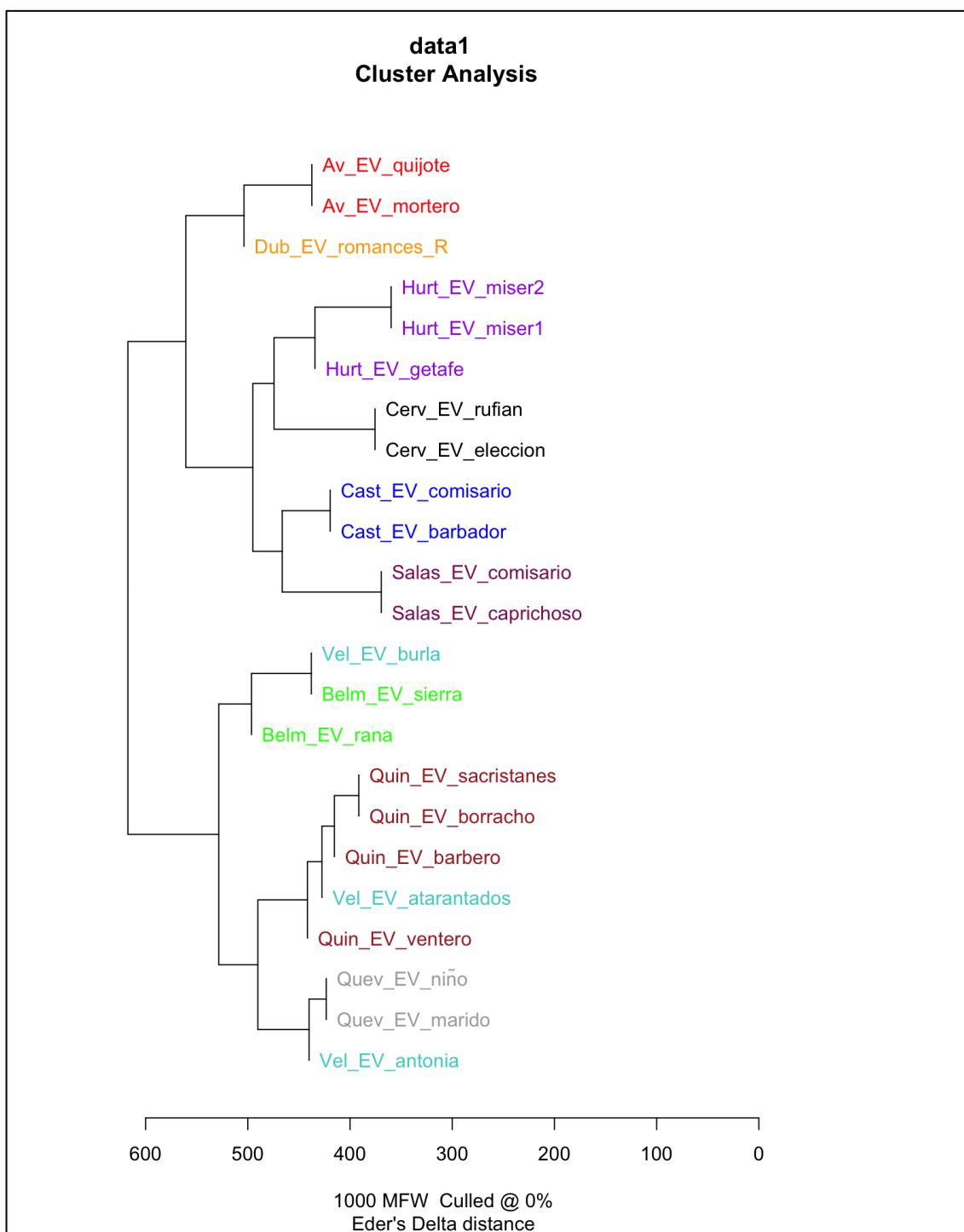


Gráfico 75. Clúster basado en la frecuencia de las palabras más empleadas en los entremeses en verso de autor conocido y *Los romances* (RStudio. Stylo. Cluster Analysis. Eder's Delta. MFW 1000. Culled 0%)

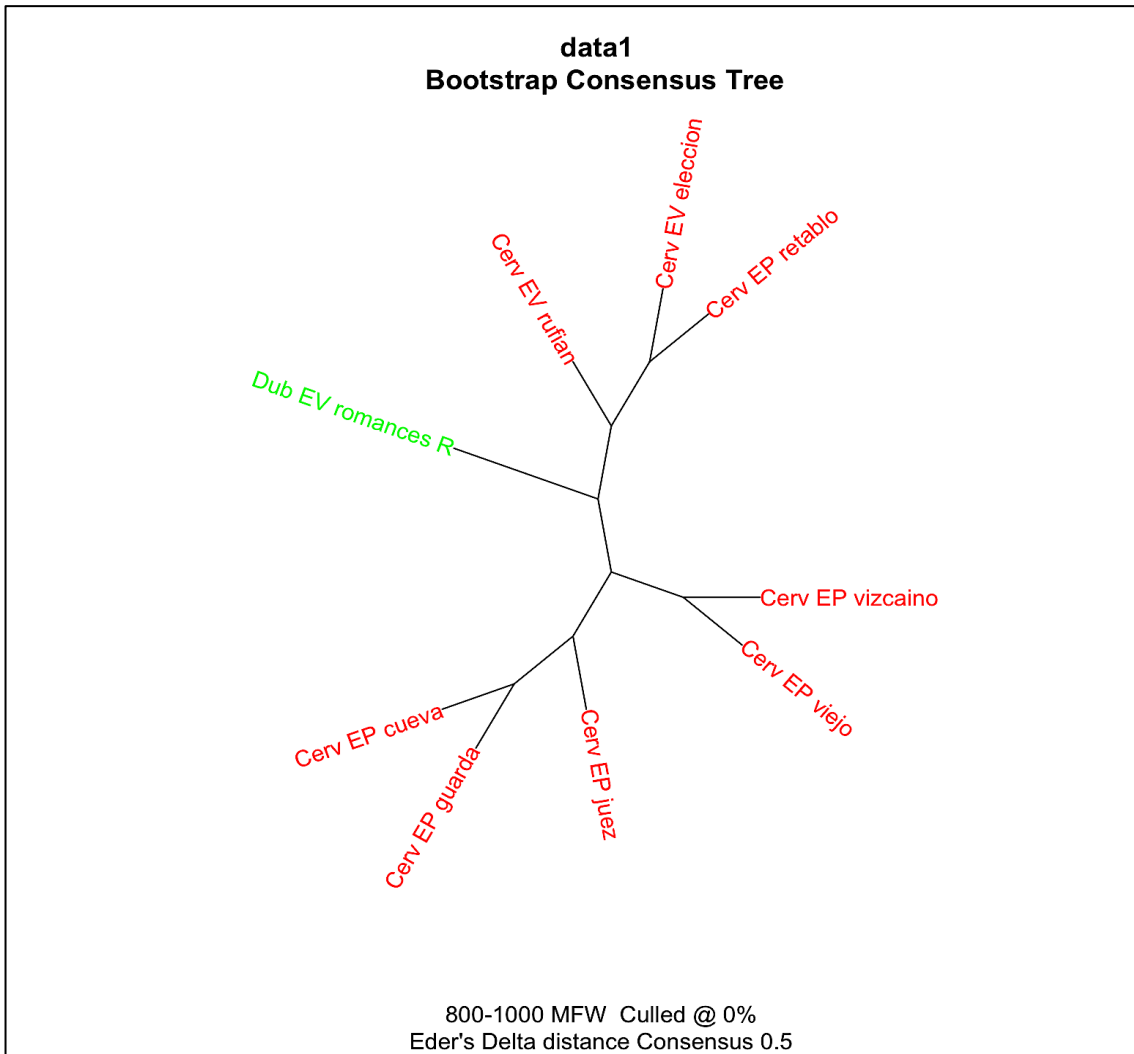


Gráfico 76. Árbol de consenso basado en la frecuencia de las palabras más empleadas (RStudio. Stylo. Bootstrap Consensus Tree. 0.5 consensus strength. Eder's Delta. MFW 800-1000)

7.3.6. Análisis de componentes principales (PCA) aplicado a la *Delta distance*

Como vimos en el quinto capítulo (Gráfico 20), el análisis de componentes principales (PCA) de las 1000 palabras más frecuentes aplicado al conjunto de entremeses dubitados e indubitados cervantinos ponía de manifiesto que el de *Los romances* era bastante particular; como era esperable, se encuadraba en el lado de los entremeses en verso (PC1), pero ciertamente alejado de *El rufián viudo* y *Los alcaldes de Daganzo*.

Si repetimos el análisis con parte real del dubitado, es decir, aquella resultante de eliminar los romances insertados (Dub_EV_romances_R), vemos que la disonancia se acentúa cuantiosamente (Gráfico 77); todo lo contrario de lo que sucedía en el PCA de *La cárcel de Sevilla* (Gráfico 43)²⁵¹.

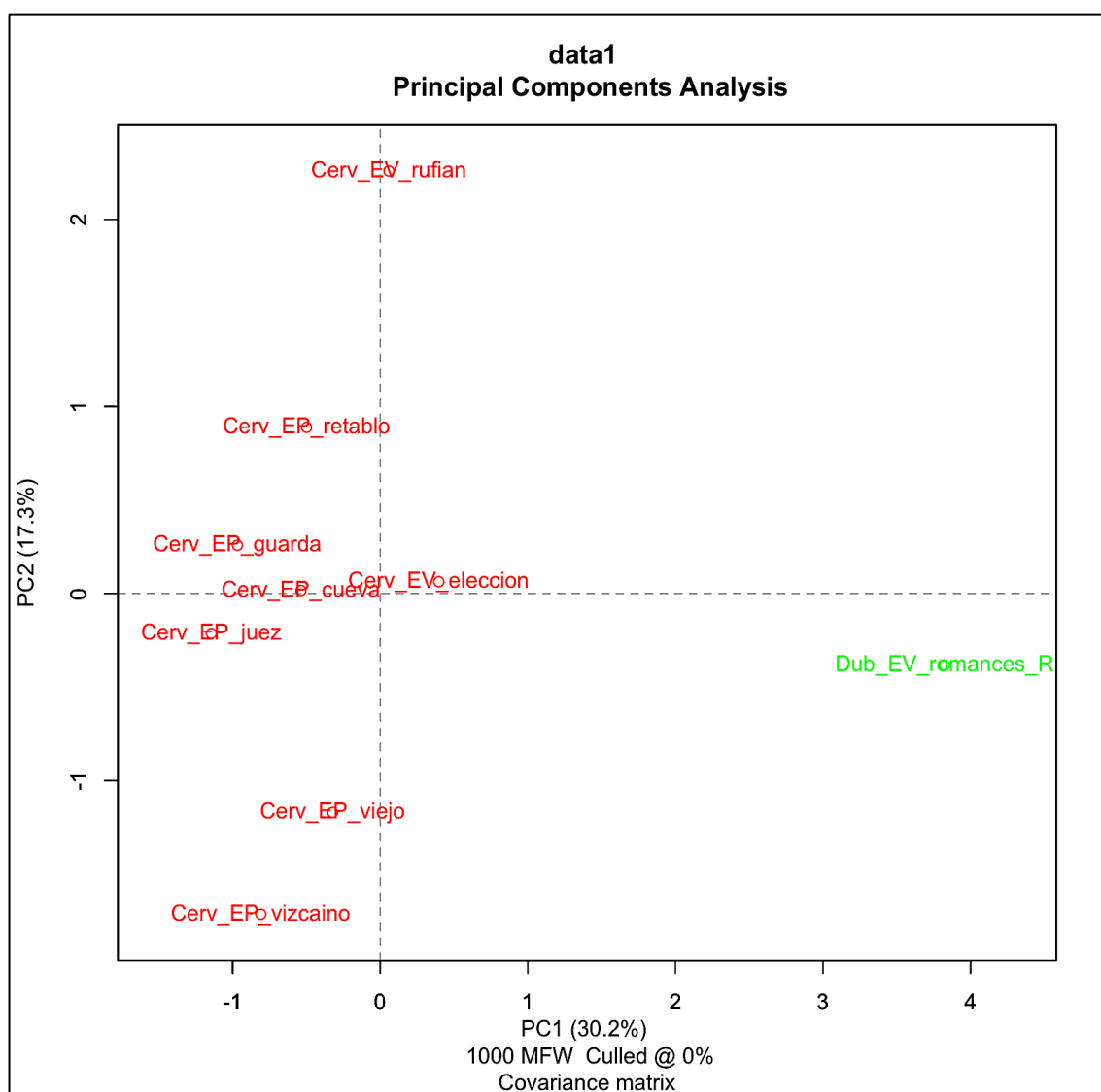


Gráfico 77. Gráfica de puntuaciones del PCA de las 1000 palabras más frecuentes en los entremeses dubitados e indubitados de Cervantes (RStudio. PCA. Eder's Delta. 1000 MFW)

²⁵¹ El PC1 sigue discriminando entre entremeses en verso y entremeses en prosa, con un 30,2% de la varianza total.

7.3.7. Clasificación supervisada por aprendizaje automático (Classify)

En este apartado, vamos a aplicar la clasificación supervisada por aprendizaje automático (Classify) al *Entremés de los romances*, seleccionando el algoritmo Eder's Delta y un intervalo de 100 a 1000 palabras más frecuentes (con un incremento de 100 en 100). El corpus de entrenamiento estará conformado por una serie de textos indubitados de Cervantes (Cerv_EV_eleccion y Cerv_EP_retablo), Ávila (Av_EV_mortero), Salas Barbadillo (Salas_EP_mirones), Quiñones de Benavente (Quin_EV_borracho), Hurtado de Mendoza (Hurt_EV_getafe) y Quevedo (Quev_EV_marido) y el de prueba por el texto dubitado (Dub_EV_romances) y otros tantos indubitados de los mismos autores (Cerv_EV_rufian y Cerv_EP_guarda, Av_EV_quijote, Salas_EV_caprichoso, Quin_EV_barbero, Hurt_EV_miser1 y Quev_EV_vieja). Éstos son los resultados:

RESULTADOS			
100 MFW	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Dub_EV_romances	-->	Cerv_EV_eleccion
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Cerv_EV_eleccion
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
200 MFW	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Quin_EV_borracho
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero

ESTUDIO DE AUTORÍA DEL *ENTREMÉS DE LOS ROMANCES*

300 MFW	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Quin_EV_borracho
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
400 MFW	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Quin_EV_borracho
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
500 MFW	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Dub_EV_romances	-->	Quin_EV_borracho
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
600 MFW	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Dub_EV_romances	-->	Quin_EV_borracho
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
700 MFW	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones

CAPÍTULO 7

800 MFW	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
900 MFW	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Av_EV_morterO
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
Salas_EV_caprichoso	-->	Salas_EP_mirones	
1000 MFW	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Dub_EV_romances	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Cerv_EP_guarda	-->	Av_EV_mortero
	Quev_EP_vieja	-->	Quev_EV_marido
	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones

Tabla 47. Clasificación supervisada por aprendizaje automático de Delta para Dub_EV_romances (RStudio. Classify. Eder's Delta. MFW 100 - 1000)

Al igual que ocurría con *La cárcel de Sevilla*, la función Classify identifica correctamente todos los textos indubitados del corpus de prueba a partir de las 700 palabras más frecuentes. El *Entremés de los romances* se asocia con Cervantes en las 100 MFW (Cerv_EV_eleccion), con Quiñones desde las 200 hasta las 600 MFW (Quin_EV_borracho) y con Ávila desde las 700 hasta las 1000 MFW (Av_EV_mortero). La asociación con Cervantes no es significativa, pues con ese número de palabras el éxito atributivo, calculado en base a las asociaciones entre indubitados (2 de 7), se sitúa en el 28,5%. Vamos a repetir el análisis con el dubitado real (Dub_EV_romancesR):

RESULTADOS			
100 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Dub_EV_romances_R	-->	Cerv_EV_eleccion
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Quev_EP_vieja	-->	Cerv_EV_eleccion
200 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Av_EV_mortero
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Quev_EP_vieja	-->	Quev_EV_marido
300 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Quev_EP_vieja	-->	Quev_EV_marido
400 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Quev_EP_vieja	-->	Quev_EV_marido

CAPÍTULO 7

500 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Cerv_EP_retablo
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Quev_EP_vieja	-->	Quev_EV_marido
600 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Quin_EV_borracho
	Quev_EP_vieja	-->	Quev_EV_marido
700 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Av_EV_mortero
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Quev_EP_vieja	-->	Quev_EV_marido
800 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Cerv_EV_eleccion
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Quev_EP_vieja	-->	Quev_EV_marido

900 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Cerv_EV_eleccion
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Quev_EP_vieja	-->	Quev_EV_marido
1000 MFW	Quin_EV_barbero	-->	Quin_EV_borracho
	Salas_EV_caprichoso	-->	Salas_EP_mirones
	Dub_EV_romances_R	-->	Av_EV_mortero
	Cerv_EP_guarda	-->	Cerv_EV_eleccion
	Hurt_EV_miser1	-->	Hurt_EV_getafe
	Av_EV_quijote	-->	Av_EV_mortero
	Cerv_EV_rufian	-->	Cerv_EV_eleccion
	Quev_EP_vieja	-->	Quev_EV_marido

Tabla 48. Clasificación supervisada por aprendizaje automático de Delta para Dub_EV_romances_R (RStudio. Classify. Eder's Delta. MFW 100 - 1000)

En este caso, el entremés dubitado se asocia con Av_EV_mortero desde las 200 a las 1000 MFW y con Cervantes, de nuevo, sólo en las 100 MFW (28,5% de fiabilidad). Classify, por tanto, tampoco identifica al autor de este texto anónimo con Cervantes.

7.3.8. Proporción palabras de contenido / palabras de función

Otra prueba tradicional para discriminar autorías, que de un tiempo a esta parte ha ido cayendo en desuso, consiste en calcular la proporción entre las palabras lexicales y las palabras funcionales de un texto. El siguiente gráfico permite ver la comparación de esta variable en el texto dubitado real y en diez indubitados en verso de cinco autores distintos (Cervantes, Ávila, Hurtado de Mendoza, Quiñones y Salas Barbadillo):

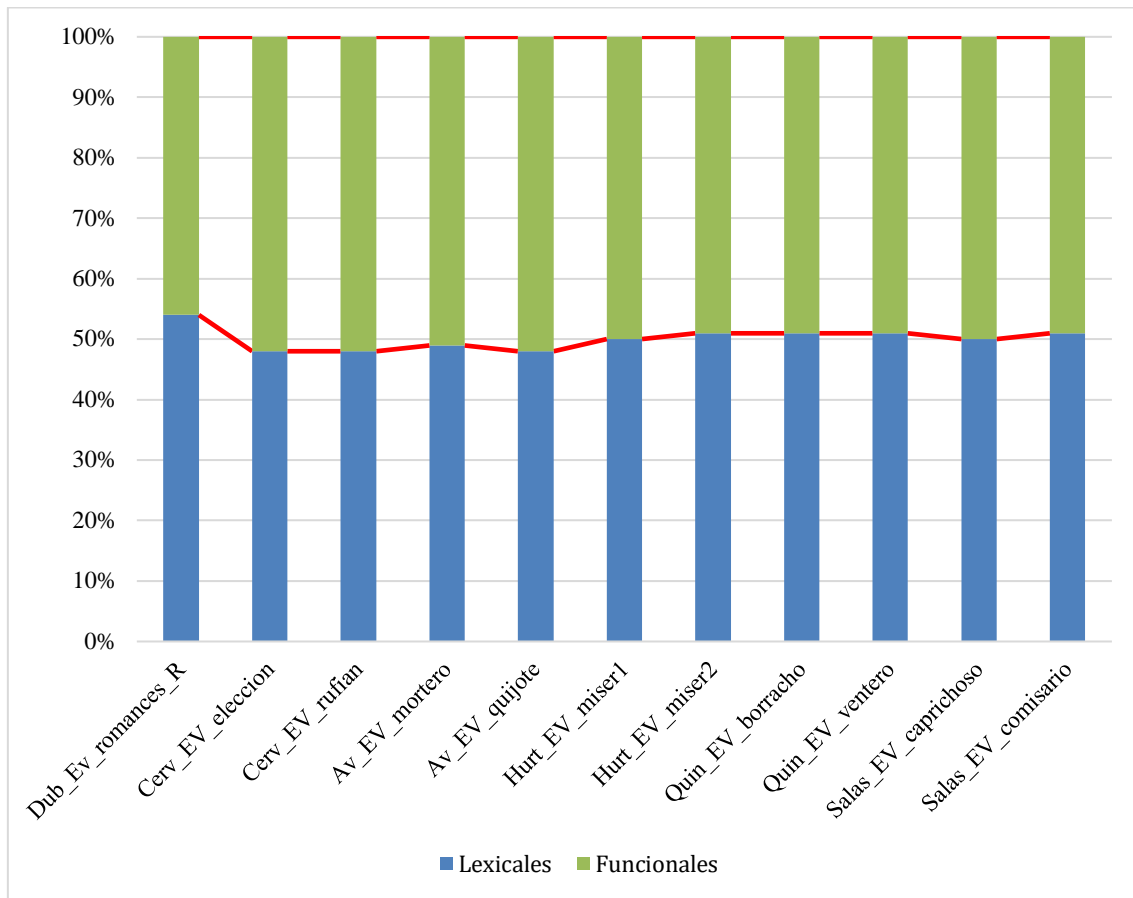


Gráfico 78. Proporción palabras de contenido / palabras de función
(Elaboración propia, a partir de los datos obtenidos en Vocalyse Toolkit)

El análisis de la proporción palabras léxicas / palabras funcionales vuelve a poner de manifiesto la disparidad entre el idiolecto contenido en el texto dubitado real y los dos entremeses en verso de Cervantes. Si en el texto dubitado real hay un ligero predominio de las palabras de función sobre las de contenido (54/46), en los dos entremeses en verso de Cervantes ocurre justo lo contrario (48/52); son dos posiciones opuestas, como deja patente el hecho de que los otros cuatro autores analizados presenten una proporción mucho más equilibrada, tendente al 50/50. Es importante advertir, además, que el comportamiento de los dos entremeses de un mismo autor es siempre igual o prácticamente igual.

7.3.9. Medidas de riqueza y densidad léxica²⁵²

7.3.9.1. La ratio type/token de palabras de contenido (TTR Content)

La densidad de un determinado texto refleja la relación existente entre el número de palabras distintas (*types*) y el número total de palabras que lo conforman:

$$\text{Densidad (TTR)} = \frac{N_{\text{type}}}{N_{\text{token}}} \quad \text{donde} \quad \begin{array}{l} N_{\text{type}} = \text{N}^{\circ} \text{ total de tipos (palabras diferentes)} \\ N_{\text{token}} = \text{N}^{\circ} \text{ total de formas (total de palabras)} \end{array}$$

El resultado se mueve inevitablemente entre los valores 0 y 1, de manera que, cuanto más cercano sea a 1, mayor será la densidad léxica de ese texto.

La relación entre *types* y *tokens* se ha usado desde la primera mitad del siglo XX para calcular el dominio del vocabulario de los hablantes de una lengua o la progresión de los alumnos en la adquisición de segundas lenguas, pero, desde hace unos años, también se ha estudiado su utilidad para identificar autorías. Grieve calculó que el poder discriminatorio de la *type/token ratio* (TTR) para la lengua inglesa era del 75% entre dos candidatos a autor, pero sólo del 8% entre cuarenta [2007]; para no desvirtuar los resultados, insistió en la necesidad de calcular la medida teniendo en consideración porciones iguales de cada texto, ya que, de no hacerlo así, el resultado podría verse condicionado por la extensión de las muestras (pues a partir de un determinado momento, las palabras funcionales tienden a repetirse, pero no necesariamente las lexicales). Hace unos años, testamos la operatividad de este parámetro en español con

²⁵² En nuestro trabajo del año 2010 sobre el *Entremés de los romances*, trabajamos con diferentes medidas de riqueza y densidad léxica, aplicadas a muestras de similar extensión [Ruiz Urbón, 2010: 228-237]. Ya entonces pusimos de manifiesto las limitaciones de algunas de estas medidas para la lengua española. En este apartado vamos a centrarnos únicamente en las dos medidas que, a pesar de estas limitaciones, resultan más interesantes para discriminar autorías: la ratio type/token de palabras de contenido (TTR Content) y la riqueza léxica (*Lexical richness*) en muestras de +/-580 palabras.

muestras de 500 palabras, pero los resultados fueron absolutamente decepcionantes, con un porcentaje de acierto del 0% entre 10 candidatos [Blasco y Ruiz Urbón, 2009]²⁵³.

No obstante, tal y como hemos podido comprobar en experiencias forenses posteriores, la TTR de las palabras con significado pleno (TTR Content), si bien no discrimina inequívocamente entre varios candidatos —básicamente porque muchos de ellos pueden tener densidades léxicas ciertamente próximas—, sí permite establecer un patrón de conducta más o menos estable entre las obras de un mismo autor, siempre que éstas pertenezcan al mismo género textual y se hayan escrito en fechas relativamente próximas, y comprobar, en consecuencia, si el texto dubitado se ajusta o no a dicho patrón.

Vamos a calcular la TTR Content de los dos entremeses en verso de Cervantes y de *Los romances* para comprobar la especificidad de las palabras lexicales y comprobar si hay cierta proximidad o no. Siguiendo la propuesta de Johnson [Bailey 1979: 212], hemos fragmentado nuestros textos en muestras de similar extensión, de unas 580 palabras aproximadamente, para evitar que la longitud influya en el resultado:

- Dub_EV_romances_R_a = 575 palabras
- Dub_EV_romances_R_b = 589 palabras
- Cerv_EV_rufian_a = 607 palabras
- Cerv_EV_rufian_b = 593 palabras
- Cerv_EV_rufian_c = 599 palabras
- Cerv_EV_rufian_d = 587 palabras
- Cerv_EV_eleccion_a = 597 palabras

²⁵³ Aunque en aquel entonces no analizamos las posibles causas, muy probablemente estén relacionadas con el carácter flexivo de la lengua española, ya que la TTR considera como palabras distintas aquellas formas de un mismo lema que son diferentes; por ejemplo, mientras en inglés un adjetivo tan ubicuo como *bad* es invariable, en español puede aparecer bajo las formas *malo*, *mala*, *malos* o *malas*, según lo exija la concordancia gramatical. Es un asunto sobre el que hay que seguir investigando, ya que probablemente sea más operativo calcular la TTR teniendo en cuenta las formas lematizadas.

- Cerv_EV_eleccion_b = 578 palabras
- Cerv_EV_eleccion_c = 580 palabras
- Cerv_EV_eleccion_d = 583 palabras
- Cerv_EV_eleccion_e = 410 palabras (que descartamos para este análisis)

Tal y como muestra el siguiente gráfico, basado en la TTR Content, los dos entremeses en verso de Cervantes presentan, dentro de su variación ([0,8047-0,9110]), una mayor densidad léxica que la que muestra el texto dubitado ([0,7296-0,7590]); dicho de otro modo, el autor del texto anónimo parece tener menor dominio léxico que Cervantes:

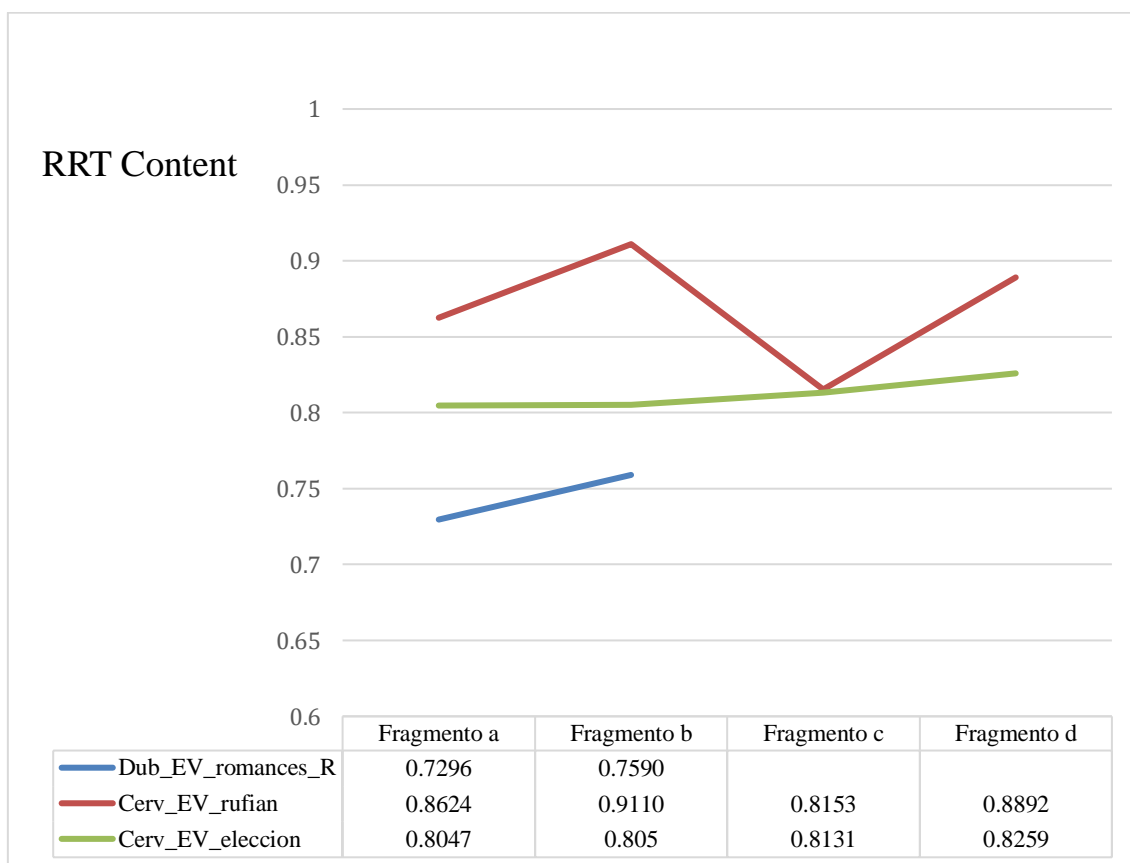


Gráfico 79. Ratio type/token de palabras de contenido (RRT Content) en muestras de +/-580 palabras (Elaboración propia, a partir de los datos obtenidos en Vocalyse Toolkit)

7.3.9.2. *La riqueza léxica (lexical richness)*

La riqueza de un texto (*richness*), establecida por Honoré en 1979, se calcula mediante una fórmula matemática que contempla el número total de palabras de un texto (*tokens*), el número total de palabras distintas (*types*) y el número total de palabras usadas una sola vez (*hapax legomena*):

$$\text{Riqueza} = \frac{100 \times \log N \text{ tokens}}{1 - (\text{HL} / N \text{ types})} \quad \text{donde} \quad \begin{array}{l} N \text{ tokens} = \text{N}^\circ \text{ total de palabras} \\ N \text{ types} = \text{N}^\circ \text{ de palabras distintas} \\ \text{HL} = \text{N}^\circ \text{ de } \textit{hapax legomena} \end{array}$$

La medida de Honoré, sin embargo, no es del todo precisa, ya que no distingue entre palabras lexicales y funcionales. Si tenemos en cuenta que la lista de palabras vacías de una lengua (*closed set*) es mucho más reducida que la de palabras llenas (*open set*), advertiremos que la distribución de los hapaxes de un texto será desproporcionada en función de su longitud, pues a partir de una determinada extensión las palabras funcionales tenderán a repetirse, pero no necesariamente las palabras lexicales; en consecuencia, “the method of calculation gives a higher score when the proportion of once-only use increases” [Woolls y Coulthard, 1998: 52]. Woolls y Coulthard, en un intento por contrarrestar el efecto de los diferentes índices de crecimiento en textos de diferente longitud, proponen sustituir en la fórmula de Honoré los *hapax legomena* totales por los *hapax legomena* de contenido. A este nuevo índice lo denominan *riqueza léxica (lexical richness)*:

Investigations into operation of the closed set revealed that in texts longer than 500 words the number of closed set items used only once tended to stay within a narrow band of 35 to 50 items, irrespective of the length of the text, while the number of lexical items continued to grow roughly in proportion to the length of the text [1998: 52].

Es preferible, siempre que la longitud sea similar, usar textos completos, ya que, aunque la distribución de hapaxes de un texto es regular, no es claramente uniforme, y, por tanto, un muestreo incompleto puede provocar una variabilidad de los hapaxes. No obstante, en nuestro caso, en donde la extensión del dubitado real es muy inferior a la de *El rufián viudo* y *La elección de los alcaldes de Daganzo*, usaremos muestras de aproximadamente 580 palabras.

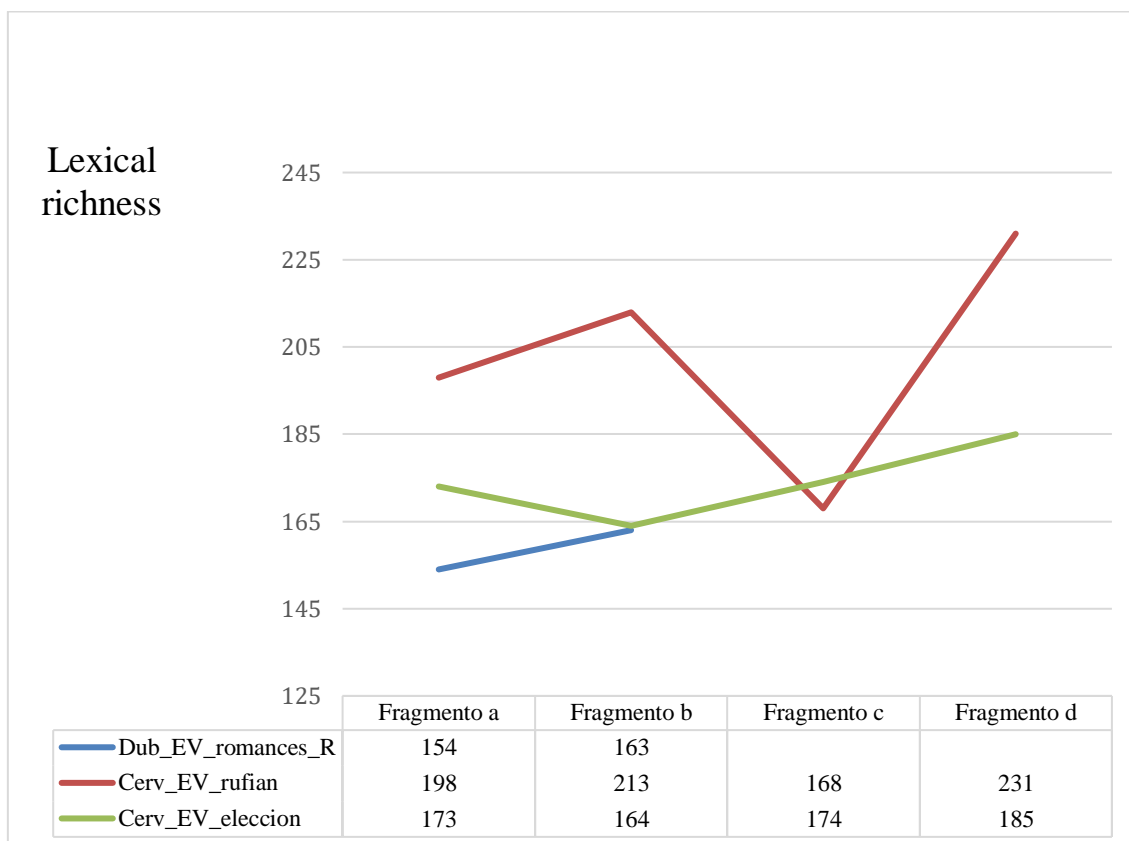


Gráfico 80. Riqueza léxica (*Lexical richness*) en muestras de +/-580 palabras (Elaboración propia, a partir de los datos obtenidos en Vocalyse Toolkit)

El Gráfico 80 pone de manifiesto que la riqueza léxica del autor del texto anónimo es, al menos en los casos analizados, inferior a la de Cervantes (con [154-163] frente a [164-231]).

7.3.10. Porcentaje de similitud textual (*similarity*), de vocabulario compartido una sola vez (*Once only in both*), de vocabulario compartido más de una vez (*Shared > 1*) y de vocabulario único y exclusivo (*Only in*)

Como analizamos en el Gráfico 32, la similitud textual entre *Los romances* y cada uno de los entremeses de Cervantes no era excesivamente significativo, con un intervalo del [22 , 34] %, frente al [26 , 44] % que mantienen los ocho indubitados entre sí. Para profundizar en este resultado, vamos a comprobar qué similitud textual tiene la parte

real del texto dubitado (Dub_EV_romances_R) con cada uno de los entremeses del corpus indubitado, tanto los de autor conocido como los anónimos (Gráfico 81):

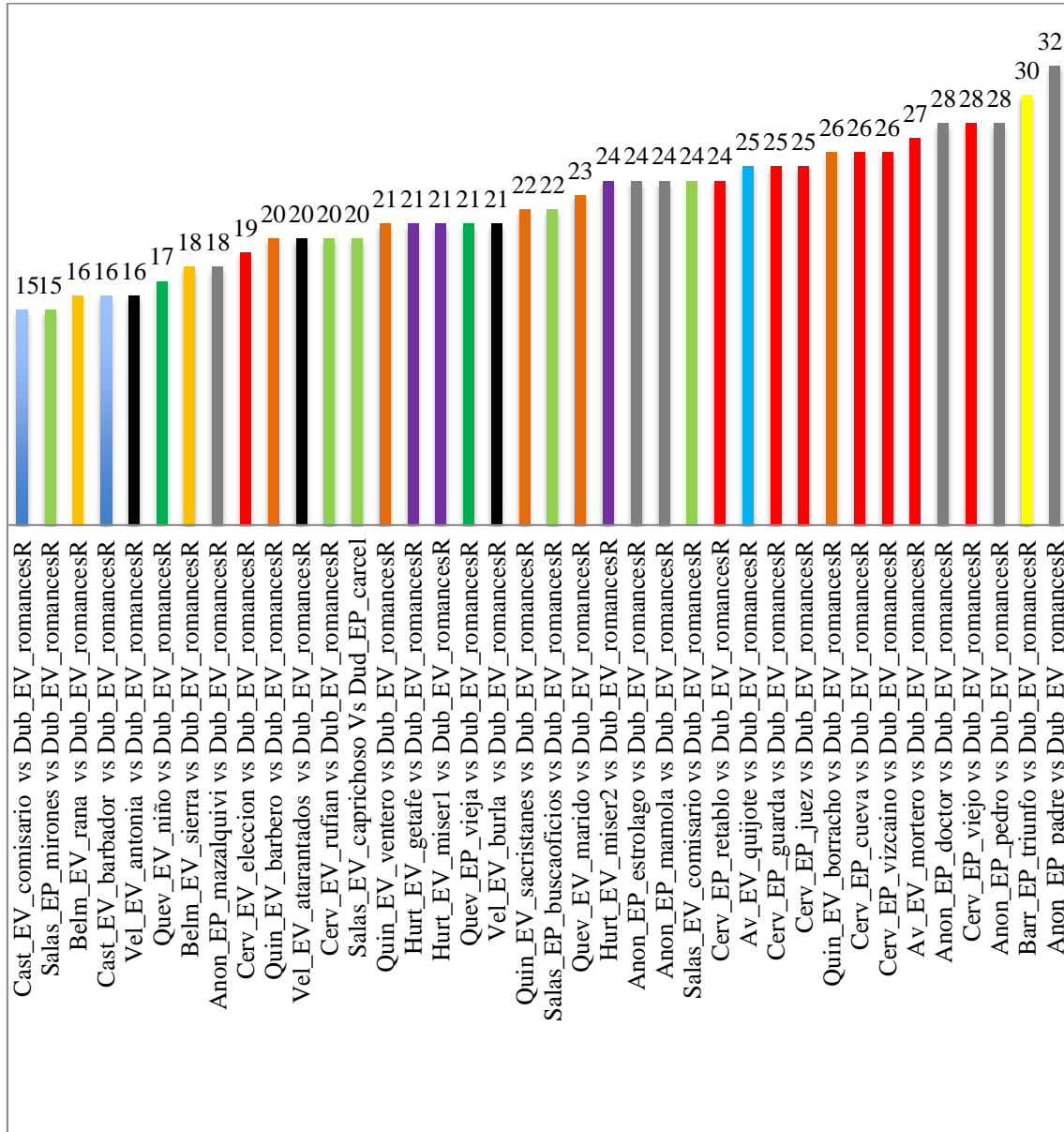


Gráfico 81. Nivel de similitud textual entre *Los romances* y los entremeses indubitados (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

El diagrama de barras refleja que el *Entremés de los romances* no alcanza similitudes léxicas significativas con ninguno de los entremeses de Cervantes, ya que todos ellos son inferiores al 32%.

Más revelador resulta que, teniendo en cuenta la vinculación genética que existe entre el *Entremés de los romances* y los seis primeros capítulos de la primera parte del *Quijote* (ya sea en una dirección u otra), ambos textos sólo tengan entre sí un porcentaje de similitud léxica del 26%; y que éste disminuya hasta el 18% si tomamos sólo la parte real²⁵⁴. Un porcentaje ciertamente bajo si, además de abordar un mismo tema, fueran realmente del mismo autor:

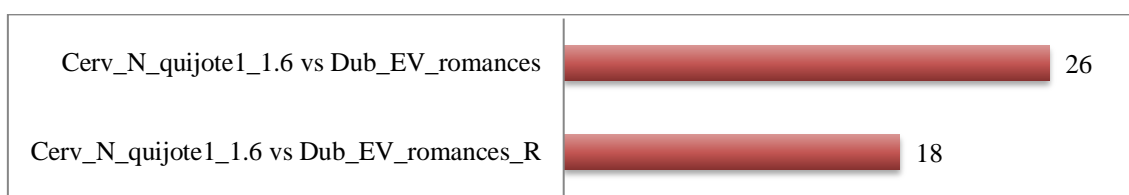


Gráfico 82. Nivel de similitud textual entre *Los romances* y los capítulos 1-6 del *Quijote* de 1605 (Elaboración propia, a partir de los datos obtenidos en CopyCatch Gold)

El análisis pormenorizado de este par no deja lugar a dudas:

Cerv_N_quijote1_1.6 vs Dub_EV_romances_R				
	Texto	Shared %	Content	Function
Total	Cerv_N_quijote1_1.6		5368	6437
Total	Dub_EV_romances_R		638	524
Shared > 1	Cerv_N_quijote1_1.6	16%	859	5622
Shared > 1	Dub_EV_romances_R	37%	236	499
Only in both	Cerv_N_quijote1_1.6	1%	37	3
Only in both	Dub_EV_romances_R	6%	37	3
Only in	Cerv_N_quijote1_1.6	83%	4472	812
Only in	Dub_EV_romances_R	57%	365	22

Tabla 49. Estadísticas de Cerv_N_quijote1_1.6 vs Dub_EV_romances_R (CopyCatch Gold. Statistics)

²⁵⁴ El porcentaje de similitud del 26% se debe al hecho de que tanto Bartolo como Quijote reciten unos mismos versos del *Romance del Marqués de Mantua*: “¿Dónde estás, señora mía, / que no te duele mi mal?” y “¡Oh noble marqués de Mantua, / mi tío y señor carnal!” (éste último verso, con una ligera alteración en el orden de los elementos).

Ambos pares textuales tienen bajos porcentajes de vocabulario léxico compartido una sola vez (*once only in both*) y altos porcentajes de vocabulario léxico único (*only in*), lo que viene a reforzar la tesis de la autoría independiente²⁵⁵. Si entramos a analizar el léxico único del *Entremés de los romances*, que supone el 57% del total, encontramos palabras que no sólo no están en esos seis primeros capítulos del Quijote de 1605, sino que nunca aparecen en las obras de Cervantes. Tal es el caso de:

- el sustantivo “nalgas” (Cervantes usa *posaderas*, pero nunca *nalgas*).
- la interjección “¡arre!” (Cervantes usa el verbo *arrear*, pero nunca la interjección “¡arre!” para estimular a las bestias).
- el adjetivo “proterva” (Cervantes elige sinónimos como *perverso*, *malvado*, *maligno*, etc., pero nunca *proterva*).
- el verbo “descabargar” (Cervantes se decanta por expresiones del tipo “apearse”, “saltar” o “arrojarse abajo”, pero nunca *descabargar*).

El análisis realizado por *CopyCatch* no revela similitudes textuales significativas entre la parte real del *Entremés de los romances* y la primera salida de don Quijote, a pesar de que ambos textos compartan un mismo tema, un argumento similar y una estructura subyacente análoga, lo que nos lleva a dudar de la autoría cervantina. Resulta paradójico pensar que si Cervantes retomó un texto suyo para crear otro (auto-copia de ideas), no apareciese en este segundo, ya fuese por elaboración consciente o subconsciente, unas mismas palabras (auto-plagio textual). En conclusión, la confrontación entre pares parece desvincular el texto dubitado “real” de la producción cervantina.

²⁵⁵ Es lógico que los porcentajes de *Dub_EV_romances_R* sean más altos que los *Cerv_N_quijote1_1.6*, ya que el texto dubitado real es mucho más breve que el de la primera salida del Quijote.

7.3.11. Verificación de autoría con General Imposters (GI)

Vamos a realizar ahora un análisis de verificación de autoría con el método General Imposters (GI) [Koppel y Winter, 2014], que, como explicamos en el apartado 6.3.9., es especialmente interesante para descubrir si el verdadero autor de un texto está dentro de los candidatos seleccionados, evitando así el “problema del vecino más cercano” (*nearest neighbor problem*). Tal y como vemos en la siguiente tabla, la comparación, en varias iteraciones, entre el texto dubitado real y los distintos autores del corpus de entremeses (Ávila, Barrionuevo, Belmonte, Cervantes, Hurtado de Mendoza, Quevedo, Quiñones de Benavente, Salas Barbadillo y Vélez de Guevara) no revela vinculaciones significativas:

	Av	Barr	Belm	Cast	Cerv	Hurt	Quev	Quin	Salas	Vel
500 MFW	0.43	0.00	0.19	0.19	0.00	0.17	0.11	0.47	0.00	0.22
800 MFW	0.29	0.00	0.39	0.44	0.00	0.19	0.22	0.19	0.00	0.34
1000 MFW	0.21	0.01	0.74	0.00	0.00	0.03	0.10	0.04	0.00	0.32

Tabla 50. Aplicación de GI a *Los romances*
(RStudio. Stylo. General Imposters. Ruzicka Delta distance. 500, 800 y 1000 MFW)

La función GI no identifica a Cervantes como autor de *Los romances* en ninguna de las frecuencias analizadas (500, 800 y 1000 MFW), estableciendo dicha posibilidad en 0 sobre 1. Tampoco parece hacerlo con ningún otro autor del corpus de análisis, ya que sólo en un caso se supera el umbral del 0,63 (Belmonte).

7.3.12. Longitud de secuencias de palabras idénticas (*verbatim*)

Como hemos explicado en varias ocasiones a lo largo del trabajo, Coulhard demostró hace años la baja probabilidad de que dos personas distintas produjesen secuencias de siete palabras consecutivas iguales [2006] y, tiempo después, Marquina y Queralt redujeron ese umbral a 5 o más palabras para el español [2014]. Cuando testamos el corpus de análisis, vimos que los entremeses de diferentes autores

compartían con frecuencia decreciente secuencias de 2, 3, 4 y hasta 5 ítems iguales, pero no secuencias de más ítems, salvo que se tratara de refranes o expresiones comunes de la lengua, compartidas por multitud de hablantes (véase apartado 5.4.3).

Si en *La cárcel de Sevilla* resultaba significativo el uso de la secuencia “eso haré yo de muy buena gana” (7 ítems), que aparecía hasta en 6 ocasiones en Cervantes (aunque también en otros dos autores entre 1547 y 1616), en el caso de *Los romances* no encontramos ningún *verbatim* de más de 5 ítems:

- Hay muchos *verbatim* de 3 ítems, sin fuerza idiosincrásica, del tipo “por lo que” (Dub_EV_romances_R, Cerv_EP_juez, Cerv_EP_viejo y Cerv_EV_rufian), “por tu vida” (Dub_EV_romances_R, Cerv_EP_cueva y Cerv_EP_vizcaino), “que me ha” (Dub_EV_romances_R, Cerv_EP_cueva, Cerv_EP_viejo y Cerv_EP_vizcaino), “es lo que” (Dub_EV_romances_R y Cerv_EV_eleccion), “no os enojéis” (Dub_EV_romances_R y Cerv_NE_gitanilla), “poco a poco” (Dub_EV_romances_R, Cerv_NE_casamiento y Cerv_EP_retablo), etc.
- También hay algún caso de *verbatim* de 4 ítems, como por ejemplo “qué es lo que” (Dub_EV_romances_R, Cerv_EP_guarda y Cerv_EP_vizcaino), “qué he de hacer” (Dub_EV_romances_R, Cerv_C_laberinto y Cerv_C_sultana) o “que me ha dado” (Dub_EV_romances_R, Cerv_EP_guarda), cuyo uso tampoco constituye un rasgo idiolectal.
- Hay un caso de *verbatim* de 5 ítems que afecta, precisamente, a una variación que se introduce en el romance “Ensílleme el potro rucio”. La secuencia a la que nos referimos es “mi mujer, que está presente”, que, según registra CORDE (1547-1617), sólo se da, además de en el texto dubitado, en Cervantes y en Lope de Vega (Tabla 51). Es preciso notar, no obstante, que la fórmula “mi + palabra que designe parentesco + que está presente” es altamente operativa en esos años: “mi compadre Fulano, que está presente” (Francisco de Luque Fajardo), “mi hermano, que está presente” (Anónimo), “mi ayo, que está presente” (Carlos Boil), “mi hija, que está presente” (Lope de Vega), “mi hermano, que está presente” (Cervantes), “mi sobrina, que está presente”, etc.

AUTOR	OBRA	FRAGMENTO
Autor dubitado	<i>Los romances</i>	Pondrasle el orillo azul que me dio para ponerme Teresa la del Villar, mi mujer, que está presente.
Cervantes	<i>El juez de los divorcios</i>	Por cuatro causas bien bastantes, vengo a pedir a vuesa merced, señor juez, haga divorcio entre mí y la señora doña Aldonza de Minjaca, mi mujer, que está presente.
Lope de Vega	<i>Comedia nueva del perseguido</i>	Leonora es mi mujer, que está presente, aquesto ya lo sabes porque ha sido a tu sosiego y honra conveniente decirte todo el caso sucedido y, jurando callarlo eternamente, de tu mujer injusta persiguido, sé yo que se lo has dicho, aunque Leonora, mi fe y sus celos engañada llora.

Tabla 51. Presencia del *verbatim* “mi mujer, que está presente” en CORDE entre 1547 y 1617

El hecho de que el texto dubitado real sea tan breve puede limitar la detección de los *verbatim* de más de 5 ítems; ahora bien, si un autor suele repetir frecuentemente, como hemos visto, cadenas de palabras iguales, ¿cuánto más no debiera hacerlo en obras literarias con una vinculación tan estrecha si realmente él fuera el autor de las dos?

7.3.13. Estudio de las palabras de función: orden de frecuencia, frecuencia relativa y PCA

En este apartado vamos a llevar a cabo varios análisis relacionados con el orden y la frecuencia de las palabras de función, siguiendo el método empleado para *La cárcel de Sevilla* [Burrows, 1988; Craig, 2009]. Como explicamos entonces, hemos deshecho en todos los textos las formas contractas (*al, del, deste, dello, etc.*) o enclíticas (*veisle, vámosle...*) para asegurarnos de que los resultados que ofrecemos sean exactos.

En la siguiente tabla mostramos las diez palabras de función más frecuentes en la parte real de *Los romances* —que, en orden de mayor a menor frecuencia, son *que*, *y*, *de*, *a*, *el*, *no*, *la*, *se*, *en* y *es*— y las comparamos con las diez que aparecen en el conjunto de entremeses indubitados. Para facilitar la interpretación, marcamos en verde las palabras que coinciden con el texto dubitado en la misma posición, en azul las que coinciden en distinta posición y en rojo las que no coinciden.

LISTADO DE LAS DIEZ PRIMERAS PALABRAS DE FUNCIÓN										
	1 ^a	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a	7 ^a	8 ^a	9 ^a	10 ^a
Dub_R	que	y	de	a	el	no	la	se	en	es
Av	que	de	y	el	a	la	en	no	por	es
Barr	que	y	de	a	no	v. m.	el	la	en	se
Belm	de	que	y	no	el	en	a	es	la	lo
Cast	de	el	que	y	en	a	no	un	con	la
Cerv	que	de	y	el	la	a	no	en	me	por
Hurt	de	que	el	y	en	a	no	la	es	que
Quev	de	y	que	el	la	a	no	en	lo	es
Quiñ	que	de	y	la	el	a	en	no	con	es
Salas	que	de	y	el	a	en	la	no	con	es
Vel	que	de	y	el	a	la	es	no	es	si

Tabla 52. Listado de las 10 primeras palabras de función
(Elaboración propia, a partir de los datos obtenidos en RStudio)

La Tabla 52 pone de manifiesto el uso de unas mismas palabras de función entre todos los entremesistas, pero, también, que no todos las utilizan en el mismo orden de frecuencia (claro predominio de las palabras marcadas en color azul). Si *La cárcel de Sevilla* compartía con Cervantes 9 de las 10 palabras de función más frecuentes, ocho de ellas en la misma posición, *Los romances* comparte 8, pero sólo una de ellas en la misma posición (*que*). El autor que más se aproxima en esta variable al texto dubitado es Barrionuevo (con 8 palabras coincidentes, 4 de ellas en la misma posición), aunque

esta similitud no es tan significativa como la que veíamos entre Cervantes y *La cárcel de Sevilla*²⁵⁶.

Calculamos, en segundo lugar, la frecuencia relativa de las diez palabras de función más utilizadas en la parte real de *Los romances* (Dub_EV_romances_R) y la comparamos con la que tienen en el resto de autores de nuestro corpus (Tabla 53). Acompañamos la tabla de un gráfico de barras para cada una de las diez palabras de función (Gráficos 83-92):

FRECUENCIA RELATIVA DE LAS PALABRAS DE FUNCIÓN											
	Dub_R	Av	Barr	Belm	Cast	Cerv	Hurt	Quev	Quiñ	Salas	Vel
que	0.0390	0.0565	0.0549	0.0436	0.0360	0.0506	0.0411	0.0422	0.0418	0.0456	0.0452
y	0.0390	0.0348	0.0469	0.0294	0.0338	0.0452	0.0330	0.0449	0.0341	0.0327	0.0344
de	0.0297	0.0469	0.0416	0.0488	0.0518	0.0502	0.0513	0.0505	0.0341	0.0445	0.0393
a	0.0212	0.0262	0.0258	0.0163	0.0198	0.0239	0.0216	0.0253	0.0236	0.0243	0.0214
el	0.0195	0.0322	0.0218	0.0220	0.0360	0.0262	0.0333	0.0320	0.0281	0.0316	0.0236
no	0.0195	0.0137	0.0249	0.0283	0.0151	0.0198	0.0214	0.0175	0.0151	0.0158	0.0184
la	0.0178	0.0176	0.0180	0.0157	0.0125	0.0241	0.0160	0.0260	0.0286	0.0209	0.0209
se	0.0169	0.0070	0.0139	0.0115	0.0103	0.0103	0.0064	0.0096	0.0125	0.0080	0.0095
en	0.0153	0.0154	0.0163	0.0199	0.0209	0.0193	0.0217	0.0166	0.0156	0.0213	0.0160
es	0.0127	0.0129	0.0131	0.0157	0.0118	0.0093	0.0137	0.0115	0.0096	0.0134	0.0190

Tabla 53. Frecuencia relativa de las 10 palabras de función más frecuentes en el texto dubitado (Elaboración propia, a partir de los datos obtenidos en RStudio)

²⁵⁶ Hemos recogido el orden que refleja el análisis con RStudio, si bien es cierto que, tal y como se desprende de la tabla de frecuencias relativas, la primera posición la comparten *que* e *y* y la quinta *el* y *no*. Si invertimos el orden de estos elementos, las conclusiones son las mismas.

CAPÍTULO 7

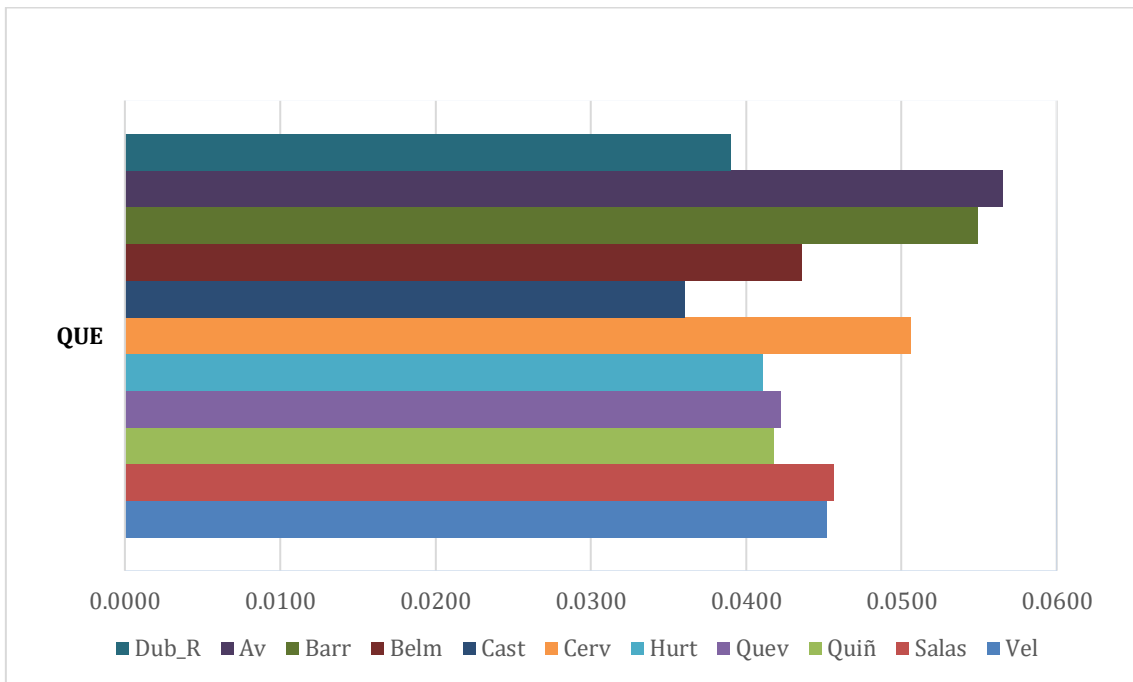


Gráfico 83. Análisis comparado de la frecuencia relativa de “que”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

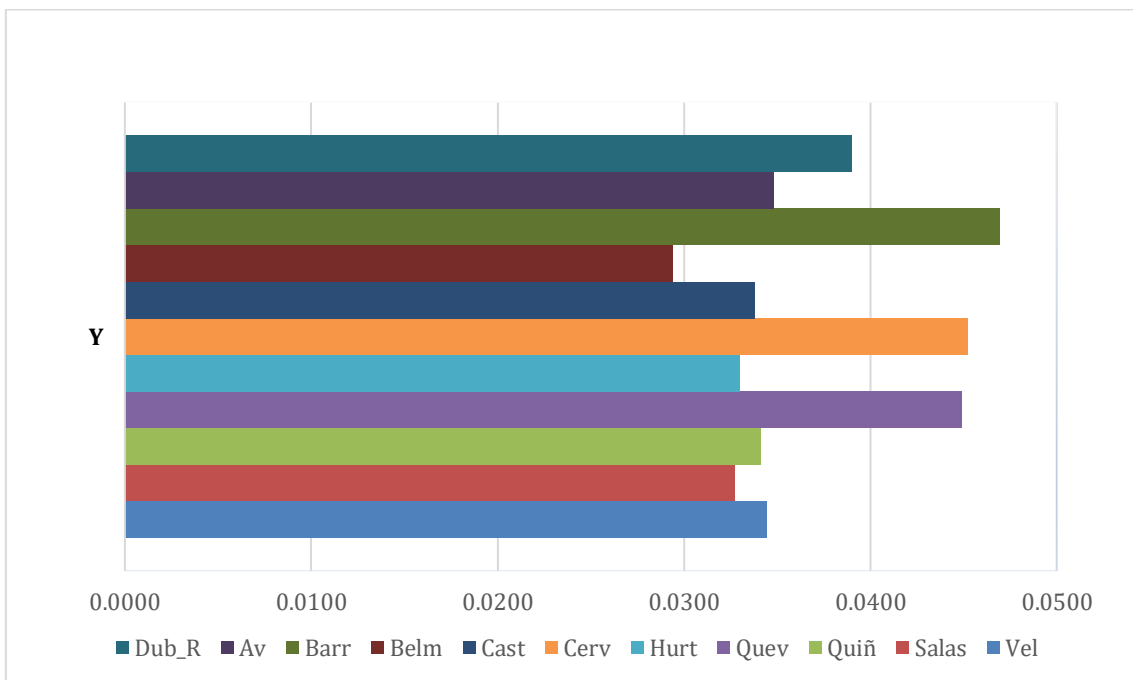


Gráfico 84. Análisis comparado de la frecuencia relativa de “y”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

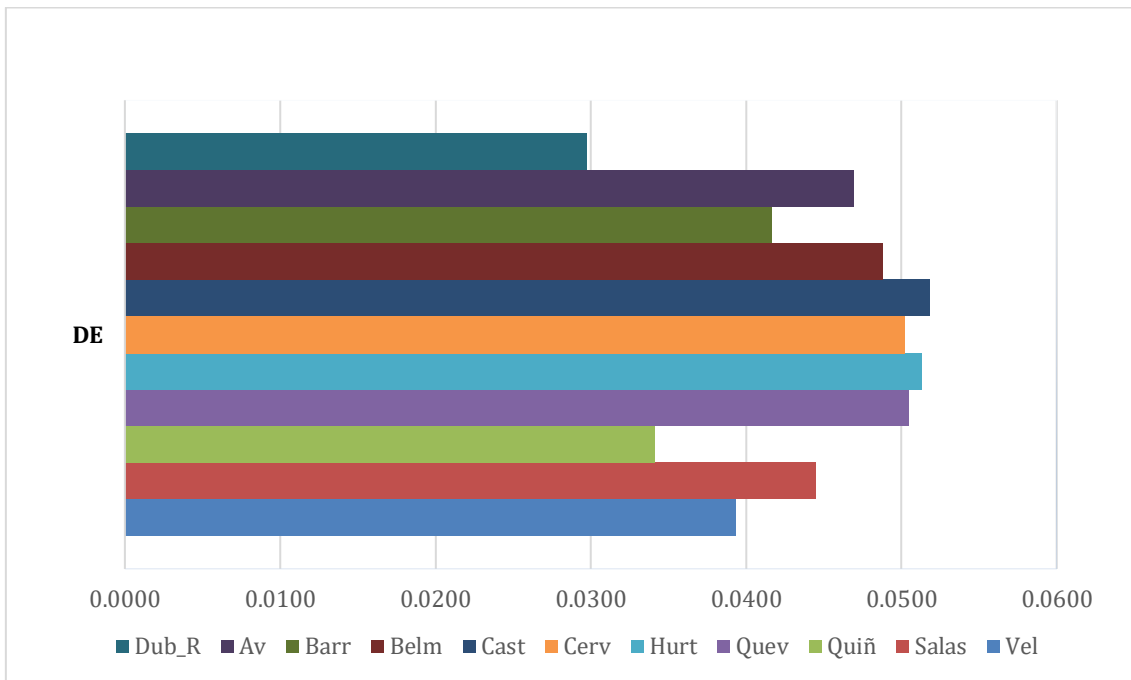


Gráfico 85. Análisis comparado de la frecuencia relativa de “de”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

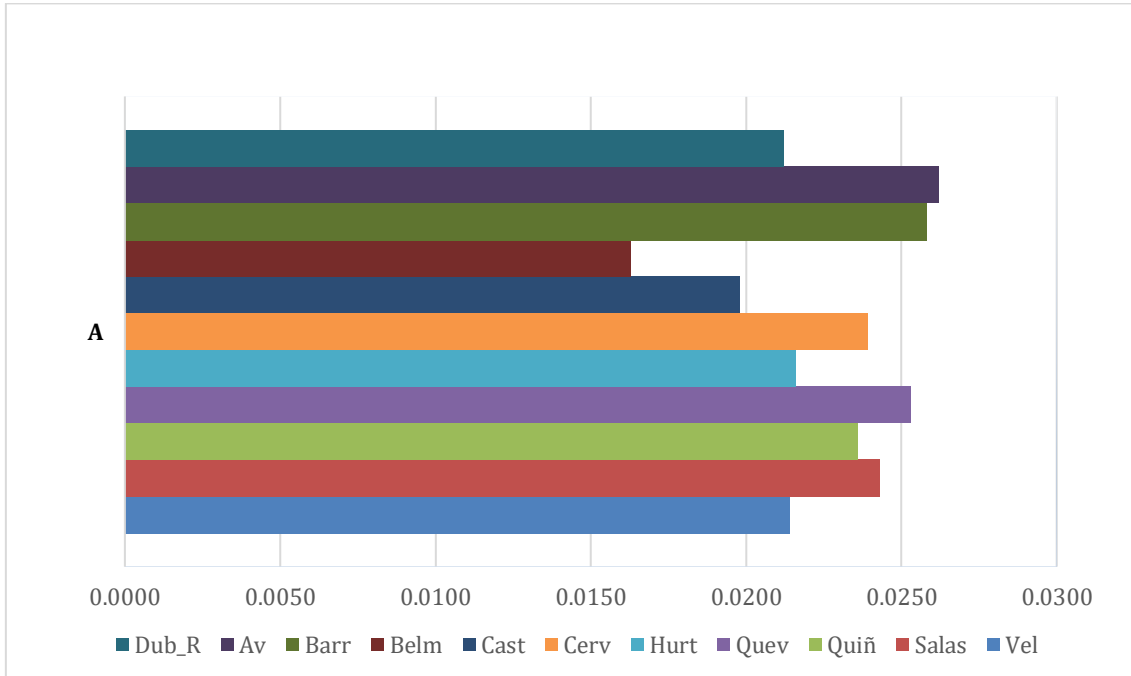


Gráfico 86. Análisis comparado de la frecuencia relativa de “a”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

CAPÍTULO 7

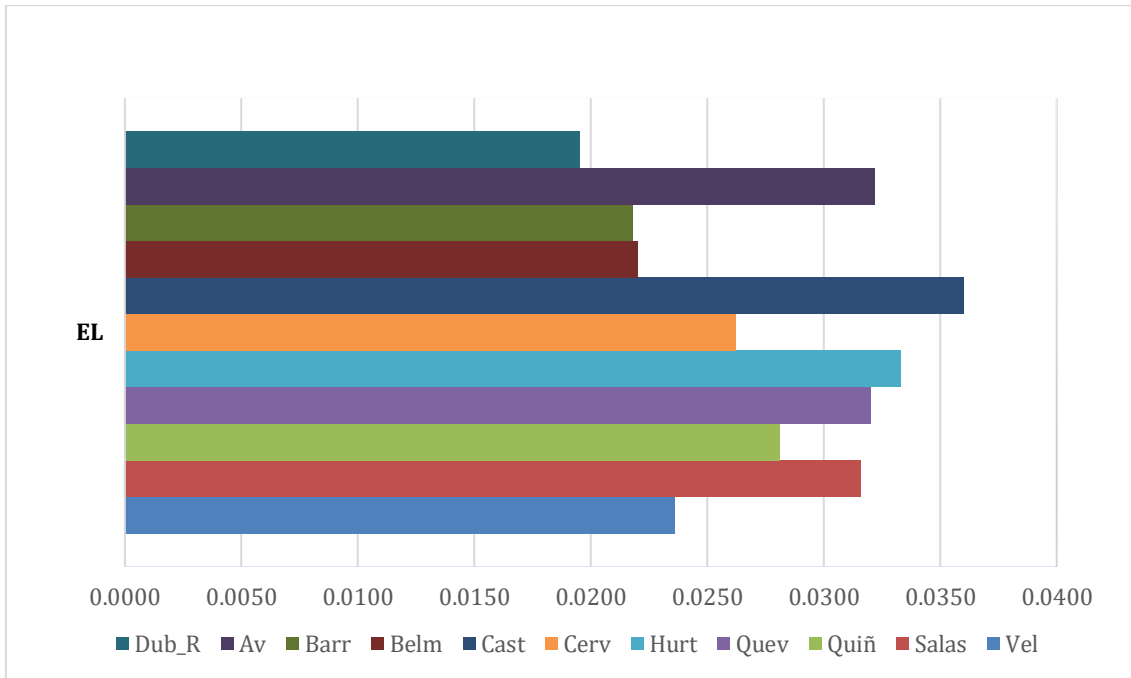


Gráfico 87 Análisis comparado de la frecuencia relativa de “el”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

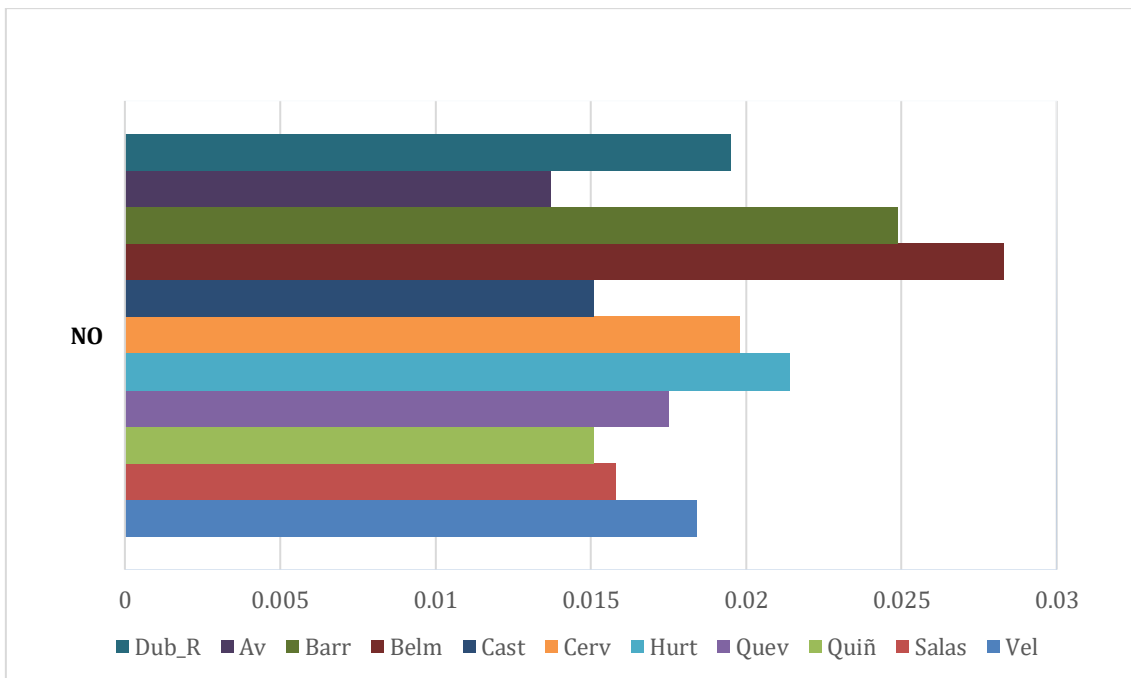


Gráfico 88. Análisis comparado de la frecuencia relativa de “no”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

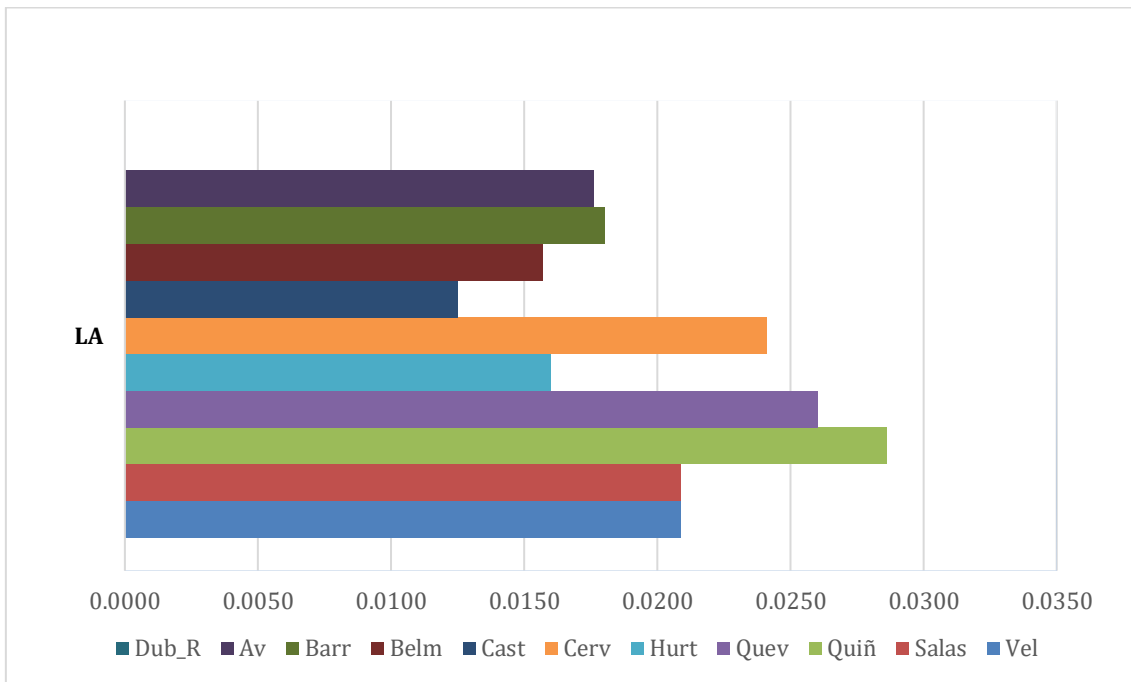


Gráfico 89. Análisis comparado de la frecuencia relativa de “la”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

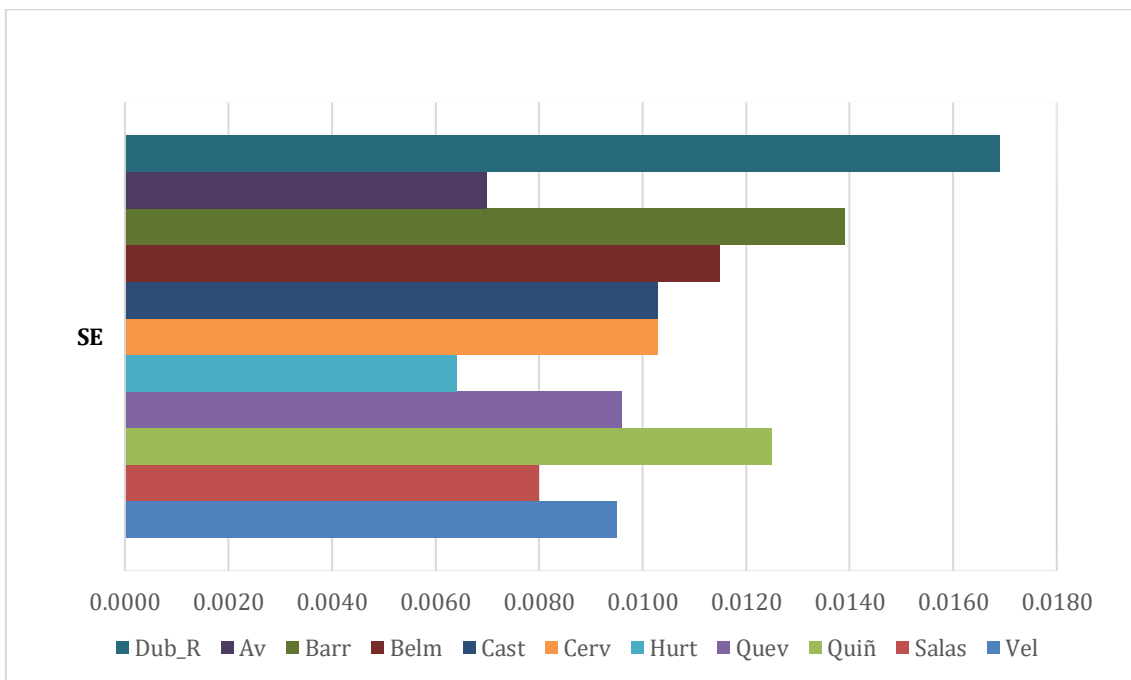


Gráfico 90. Análisis comparado de la frecuencia relativa de “se”
(Elaboración propia, a partir de los datos obtenidos en RStudio)

CAPÍTULO 7

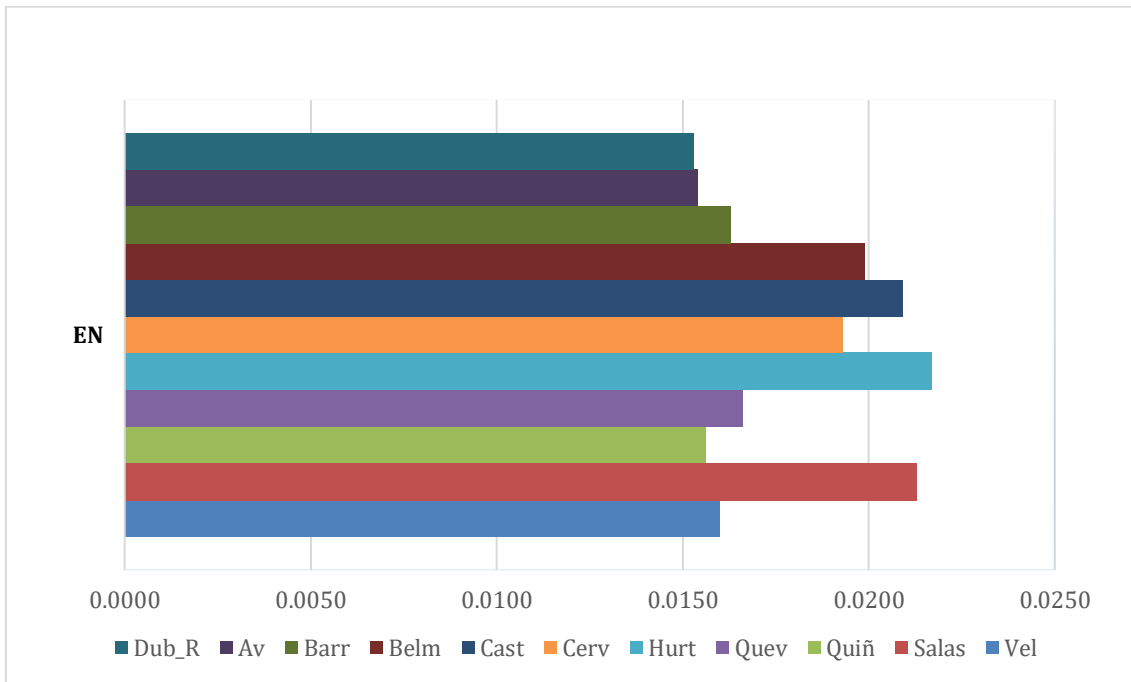


Gráfico 91. Análisis comparado de la frecuencia relativa de "en"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

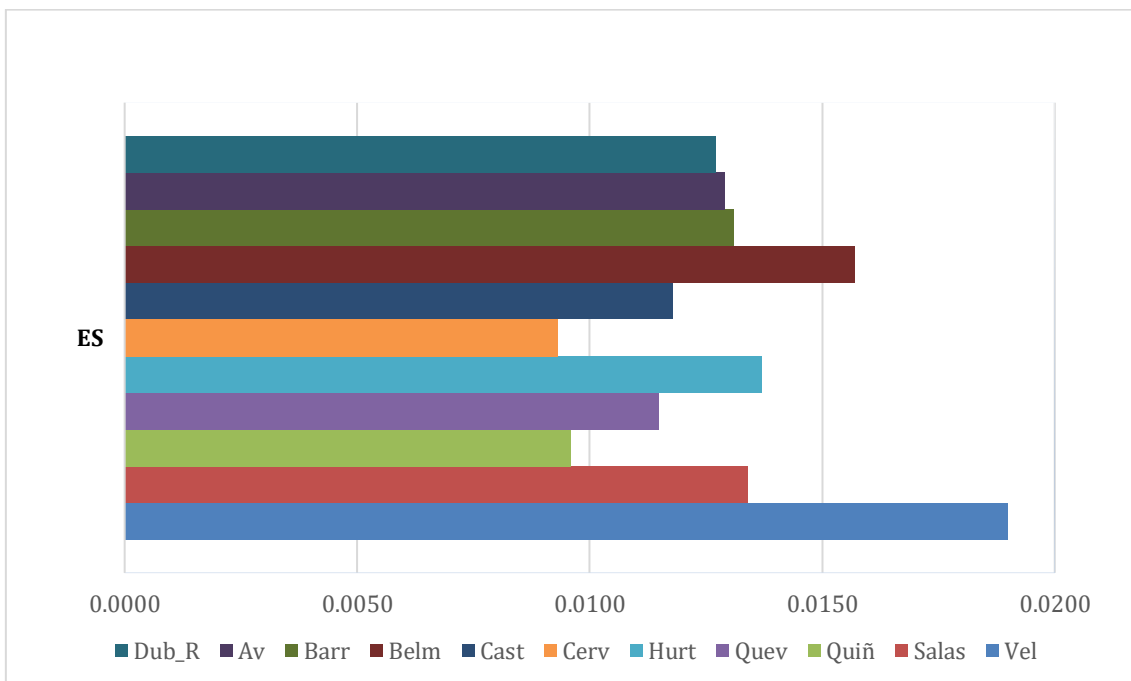


Gráfico 92. Análisis comparado de la frecuencia relativa de "es"
(Elaboración propia, a partir de los datos obtenidos en RStudio)

Los diagramas de barras reflejan un comportamiento absolutamente diferencial entre el autor anónimo y Cervantes en el uso de estas diez palabras de función. Las únicas frecuencias relativas que parecen ser acordes a ambos son “no” y, en menor medida, “a”. Así queda también de manifiesto en el siguiente gráfico, en donde mostramos la tendencia de las diez palabras de manera conjunta:

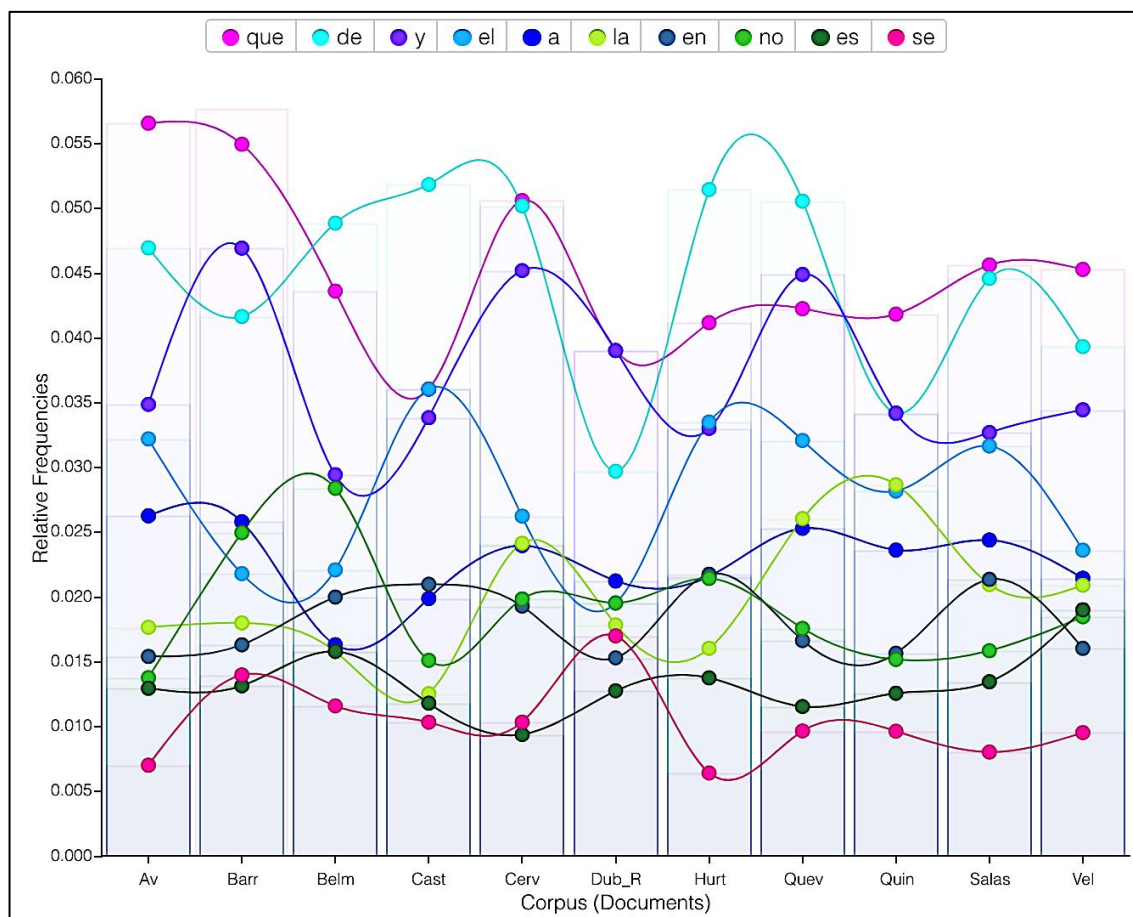


Gráfico 93. Análisis comparado de la frecuencia relativa de las diez primeras palabras de función en *Los romances* (Voyant Tools. Tendencias)

Vamos a realizar ahora el PCA de la tabla de frecuencias relativas (Tabla 53), con el objetivo de condensar la información en los componentes principales. Mostramos el gráfico de individuos (Gráfico 94), el de variables (Gráfico 95) y el biplot de individuos y variables (Gráfico 96). Las claves numéricas para cada autor son las siguientes:

CAPÍTULO 7

(1) Dub_R, (2) Av, (3) Barr, (4) Belm, (5) Cast, (6) Cerv, (7) Hurt, (8) Quev, (9) Quiñ, (10) Salas y (11) Vel.

El gráfico de individuos (autores) refleja la cercanía entre el autor del texto anónimo (1) y Hurtado de Mendoza (7), situados ambos en el tercer cuadrante, a escasa distancia. Cervantes (6) está muy alejado del centro, lo que significa que sus datos representan una variabilidad más extrema.

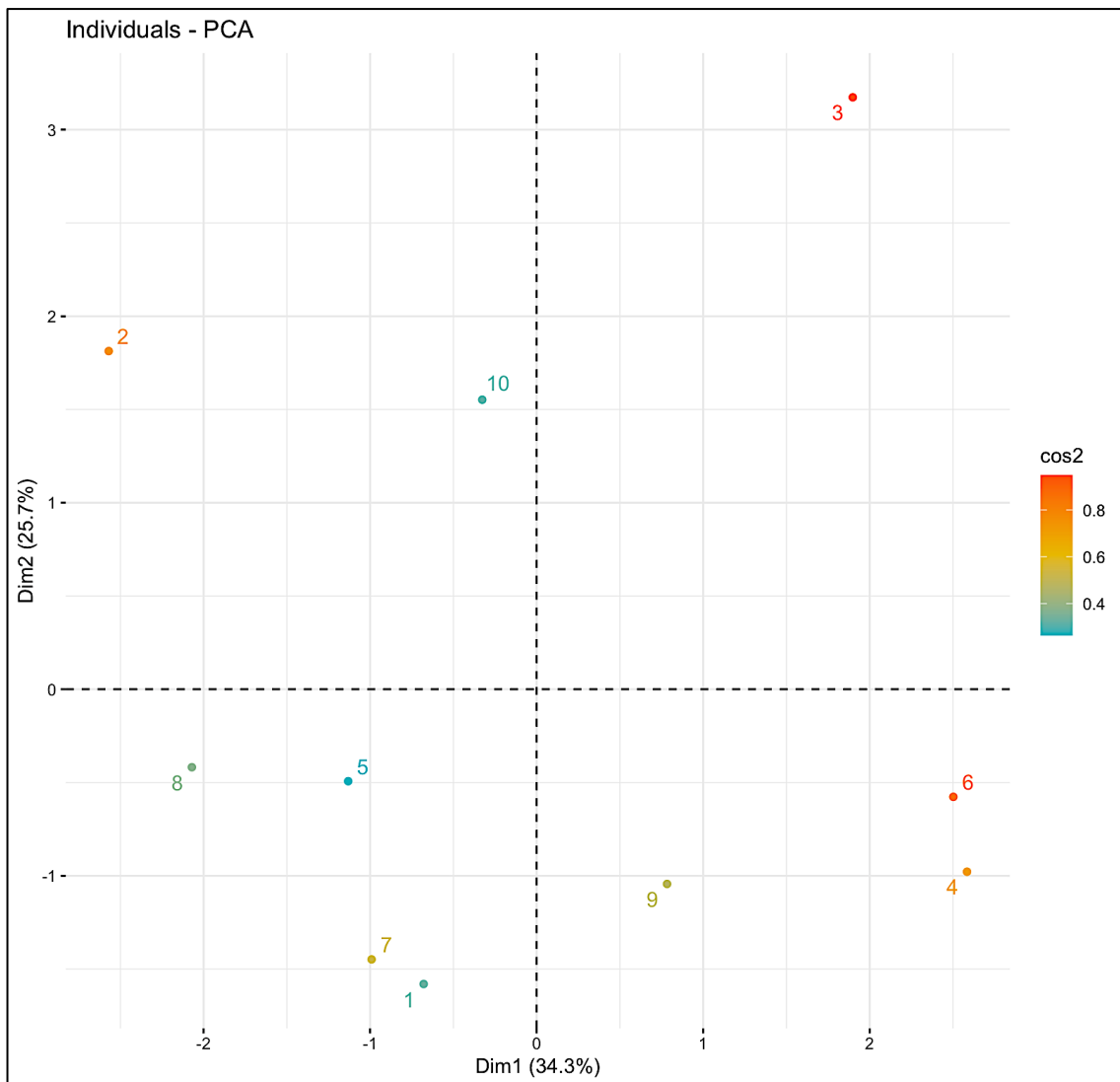


Gráfico 94. Visualización del PCA de las palabras de función. Gráfico de individuos. (RStudio. PCA. Graph of individuals)

El gráfico de variables ofrece la visualización de las diez palabras de función en vectores. Las variables correlacionadas positivas apuntan al mismo lado de la gráfica, mientras que las negativas lo hacen en lados opuestos:

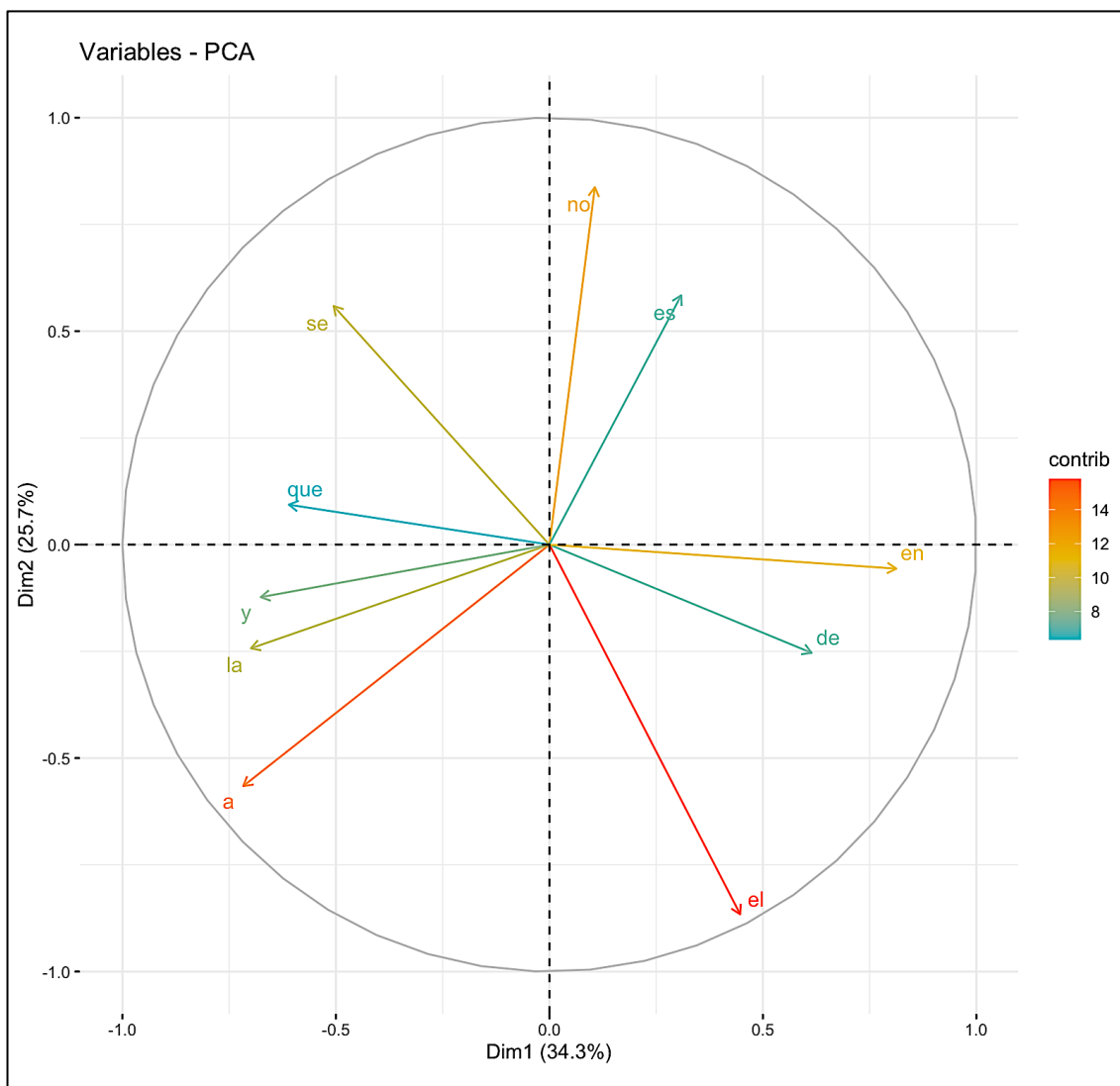


Gráfico 95. Visualización del PCA de las palabras de función. Gráfico de variables.
(RStudio. PCA. Graph of variables)

El biplot que relaciona a los individuos (autores) con las variables (palabras de función) muestra una evidente cercanía entre el autor de la parte real de *Los romances* (1) y Hurtado de Mendoza (7) y en el uso de “y”, “la” y “a”. Cervantes está en el cuarto cuadrante.

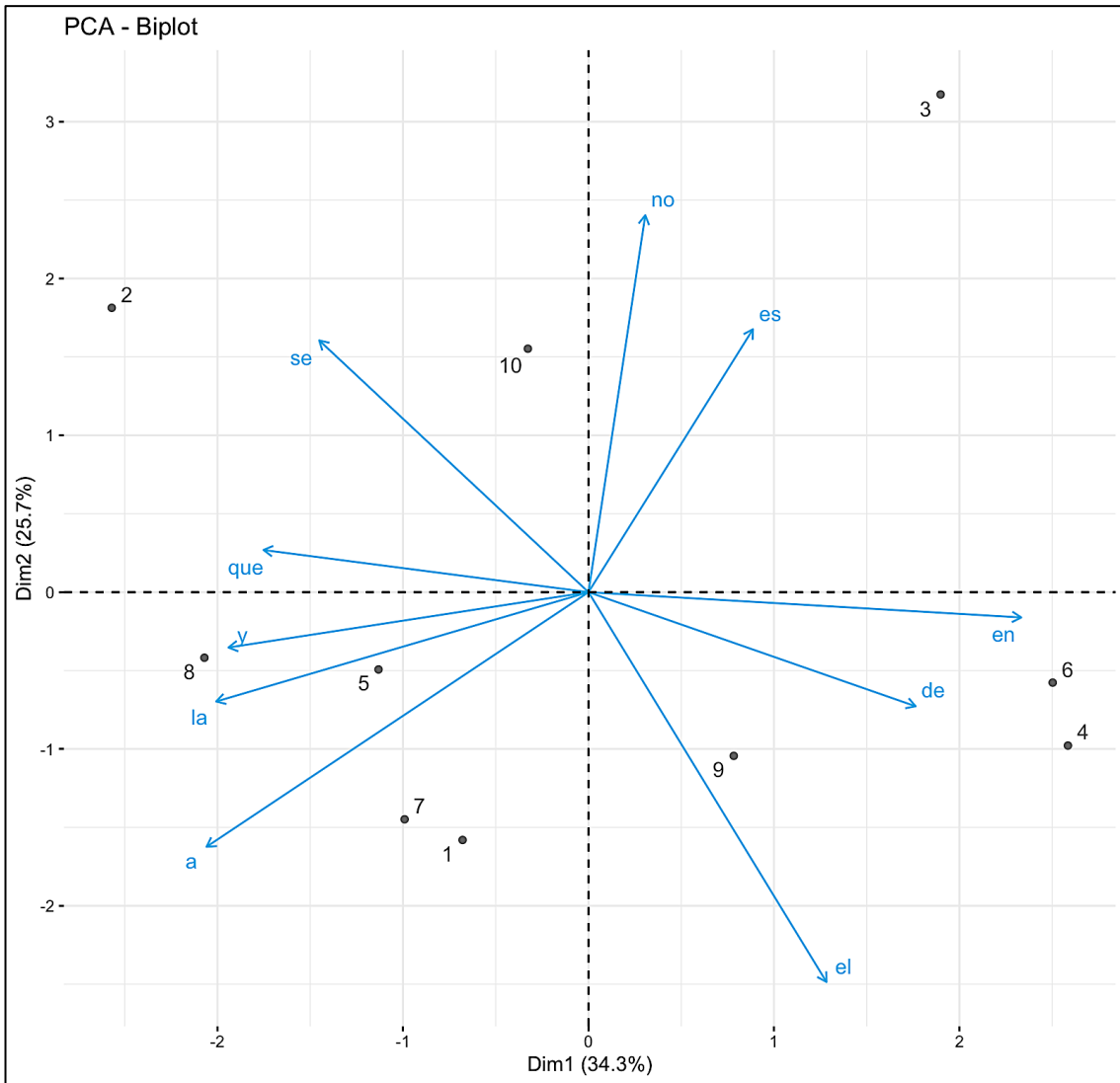


Gráfico 96. Visualización del PCA de las palabras de función. Biplot de individuos y variables. (RStudio. PCA. Biplot of individuals and variables)

El estudio del orden y la frecuencia de las palabras de función también contribuye a reforzar la hipótesis de que Cervantes no es el autor del *Entremés de los romances*.

7.3.14. Frecuencia de clases de palabras

En este apartado vamos a calcular la frecuencia de las distintas clases de palabras en *Los romances* y los dos entremeses en verso de Cervantes, mediante la ayuda del

etiquetador automático FreeLing. No analizamos, en esta ocasión, la frecuencia de los bigramas y trigramas de palabras, ya que en las obras en verso los patrones subconscientes pueden verse alterados por la necesidad de cuadrar el ritmo, el metro y la rima.

En la Tabla 54 mostramos el resultado final para cada categoría, tras la revisión de las etiquetas (POS-tagging); revisión muy necesaria en todos los casos, pero, más aún, como explicamos detalladamente en el apartado 6.3.15, en los textos del Siglo de Oro:

CLASES DE PALABRAS ²⁵⁷	DUB_EV_ROMANCES_R	CERV_EV
A. ADJETIVOS (ADJ)	39	128
CC. CONJUNCIONES COORDINADAS (CCONJ)	63	159
CS. CONJUNCIONES SUBORDINADAS (SCONJ)	48	94
D. DETERMINANTES (DET)	108	386
F. SIGNOS DE PUNTUACIÓN (PUNT)	380	669
I. INTERJECCIONES (INTJ)	20	24
M. NUMERALES (NUM)	2	32
NC. SUSTANTIVOS (SUST)	189	492
NP. NOMBRES PROPIOS (NPROP)	83	111
P. PRONOMBRES (PRON)	176	310
R. ADVERBIOS (ADV)	67	165
S. PREPOSICIONES (PREP)	106	327
VA. AUXILIARES (AUX)	40	89
VM. VERBOS (VERB)	241	466

Tabla 54. Frecuencias absolutas por clases de palabras en *Los romances* (Dub_EV_romances) y los entremeses en verso de Cervantes (RStudio. FreeLing)

²⁵⁷ Señalamos la nomenclatura de Eagles y también, entre paréntesis, la que utilizaremos nosotros: A (ADJ), CC (CCONJ), CS (SCONJ), D (DET), F (PUNT), I (INTJ), M (NUM), NC (SUST), NP (NPROP), P (PRON), R (ADV), S (PREP), VA (AUX), VM (VERB).

Dado que mostramos las frecuencias absolutas (no las relativas), vamos a reflejarlas en un gráfico para poder apreciar mejor las similitudes y diferencias entre ambos conjuntos:

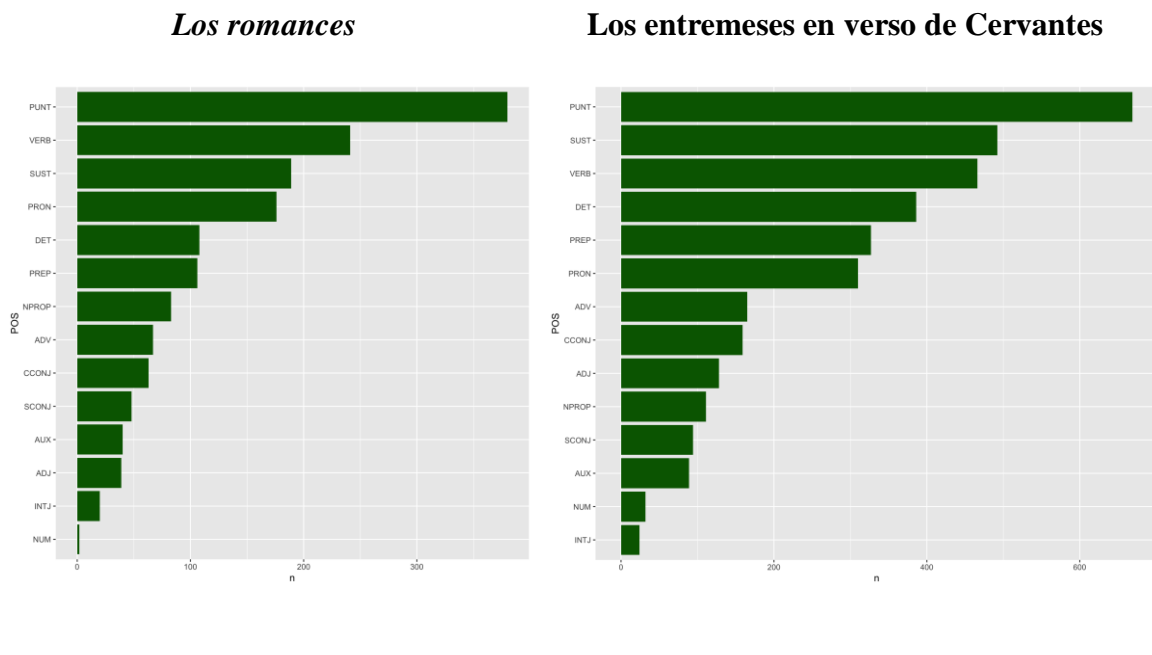


Gráfico 97. Comparativa de clases de palabras. Frecuencias absolutas.
(RStudio. FreeLing. Ggplot)

Dejando a un lado la puntuación, que no es marca propia de autoría en el Siglo de Oro, apreciamos algunas diferencias entre ambos conjuntos textuales. En *Los romances*, la clase de palabra más utilizada es el verbo, seguida a cierta distancia del nombre común; en los entremeses en verso de Cervantes, lo es el nombre común, seguido muy de cerca por el verbo. También hay un comportamiento desigual en el uso de los adjetivos (con mucha mayor presencia en Cervantes) y los nombres propios y los pronombres (mucho más habituales en el texto dubitado). En el caso de las conjunciones, hay más subordinación en *Los romances* que en Cervantes.

7.3.15. Uso y frecuencia relativa de las conjunciones coordinadas y subordinadas adverbiales

Como el texto dubitado real, además de no ser un texto completo, es de poca extensión, hemos optado por analizar el uso y la frecuencia relativa de las conjunciones que lo caracterizan, para ver si son las propias de Cervantes o no.

Analizamos, en primer lugar, las frecuencias relativas de las conjunciones coordinadas más habituales: la copulativas “y” e “ni”, la disyuntiva “o”, la explicativa “esto es” y las adversativas “pero”, “sino” y “mas” (Gráficos 98 y 99).

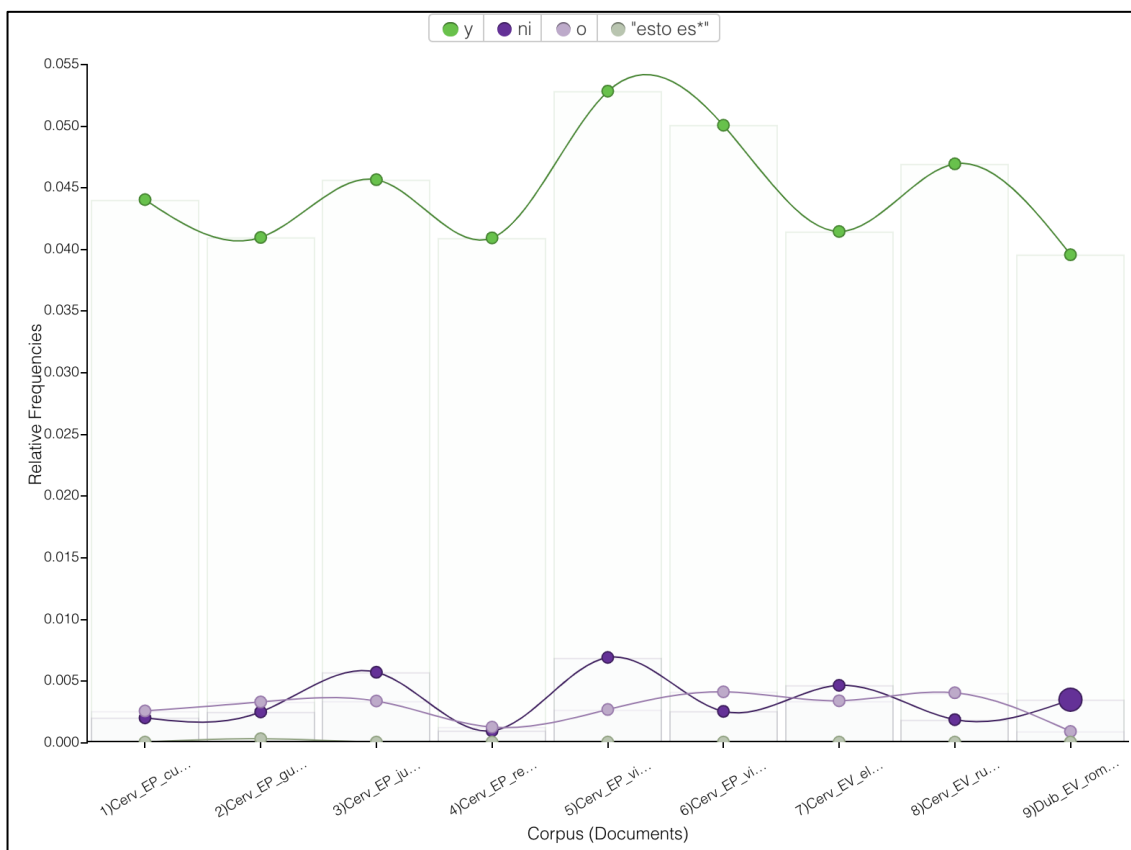


Gráfico 98. Comparativa de conjunciones coordinadas copulativas, disyuntivas y explicativas. Frecuencias relativas. (Voyant Tools. Tendencias)

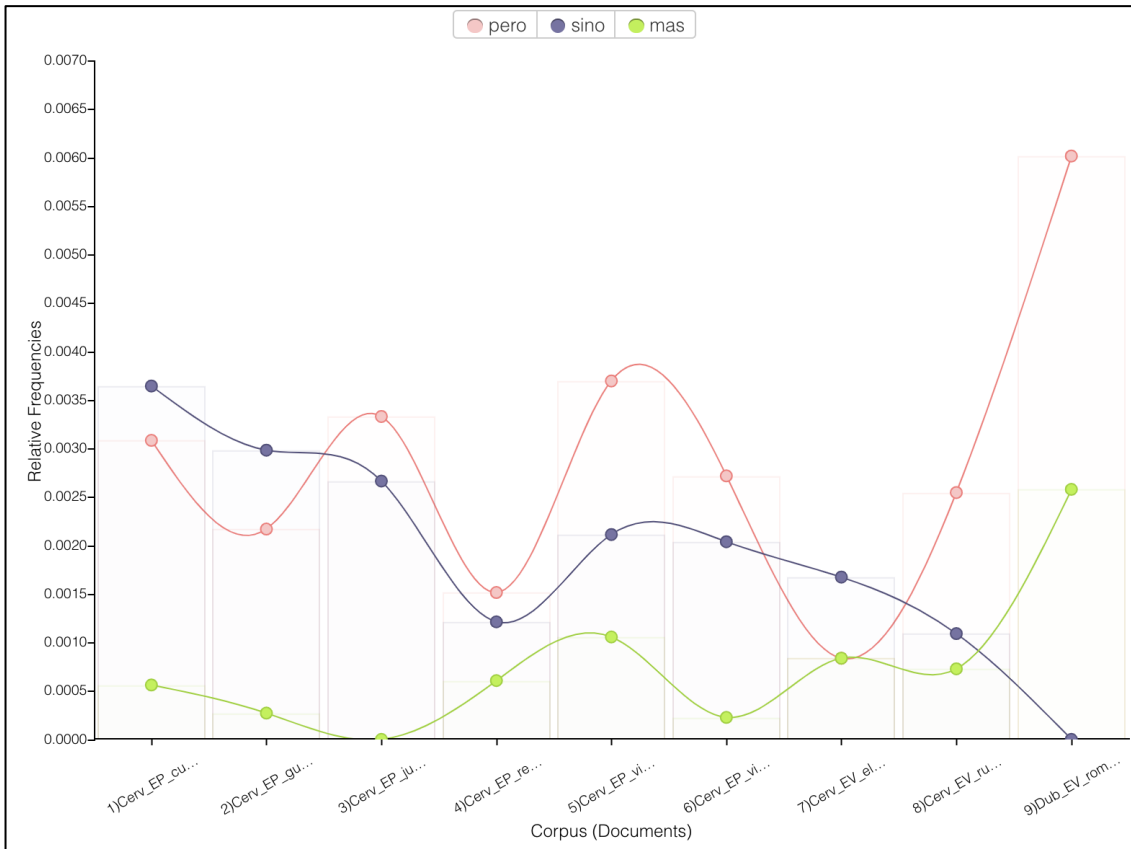


Gráfico 99. Comparativa de conjunciones coordinadas adversativas. Frecuencias relativas. (Voyant Tools. Tendencias)

Tal y como se desprende en el Gráfico 98, no hay grandes disonancias en el uso de las conjunciones coordinadas copulativas, disyuntivas y explicativas; las conjunciones “y” y “o” son algo menos usadas por el autor anónimo que por Cervantes, pero la diferencia es poco relevante. Mayores diferencias apreciamos en el uso de las conjunciones coordinadas adversativas (Gráfico 99), pues el texto dubitado muestra una clara predilección por “pero” y, en menor medida, por “mas”, mientras que en los ocho entremeses de Cervantes se da el comportamiento contrario: mayor frecuencia relativa de “sino” (que en el anónimo no aparece) y menor de “pero” y “mas”.

Pasamos ahora al análisis de las principales conjunciones subordinadas: “si” (condicional), “para que” (final), “pues” (causal), “porque” (causal), “ya que” (causal), “cuando / cuándo” (temporal), “donde / dónde” (de lugar), “aunque” (concesiva en la

mayoría de los casos), “según” (modal), “más que” (comparativa) y “menos que” (comparativa).

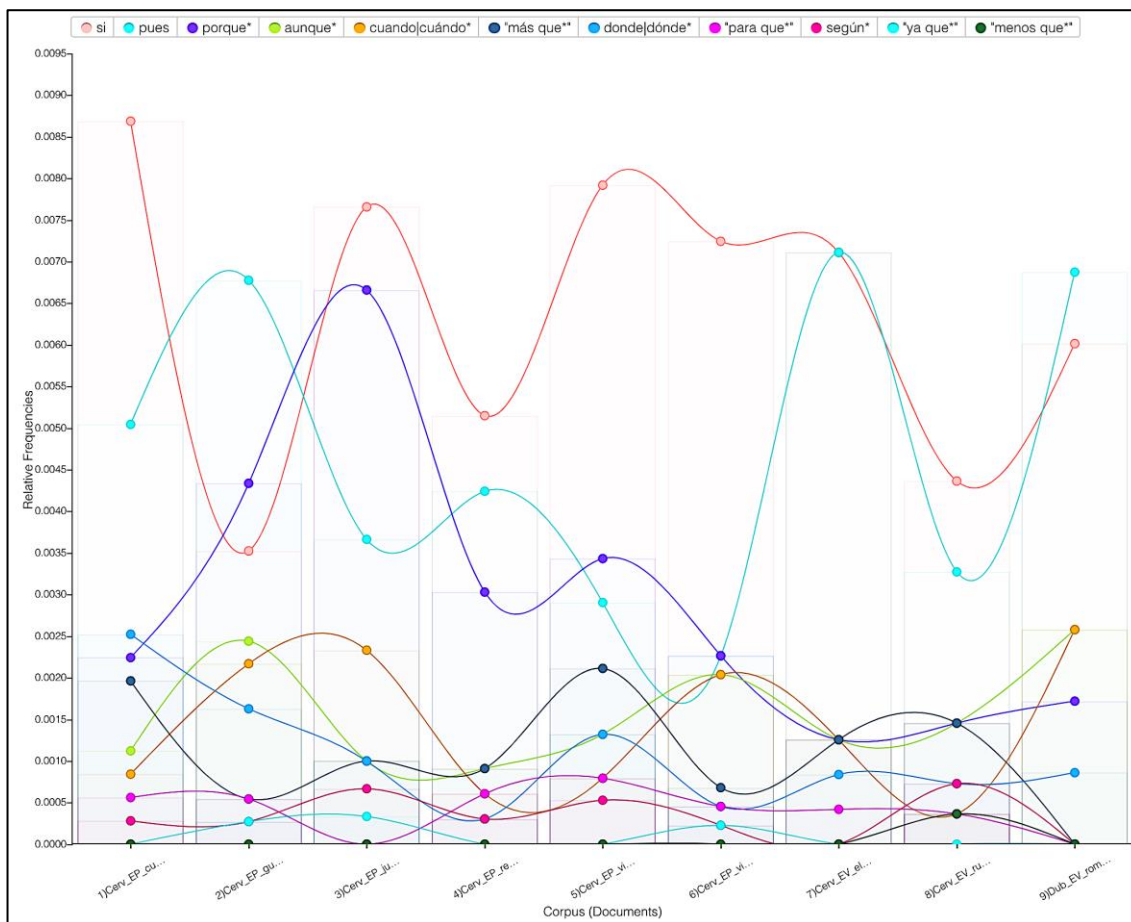


Gráfico 100. Comparativa de conjunciones coordinadas adversativas. Frecuencias relativas. (Voyant Tools. Tendencias)

Las líneas de tendencia del Gráfico 100 nos permiten hacer algunas consideraciones:

- En las ocho obras cervantinas, hay una alta variabilidad en las frecuencias relativas de ciertas conjunciones, como “si”, “pues” o “porque”, lo que impide establecer conclusiones fiables sobre ellas. En todo caso, el comportamiento del texto dubitado se ubica en los tres casos dentro de los intervalos cervantinos.

- La frecuencia de las conjunciones “donde / dónde” y “menos que” es similar en todos los textos.
- Las conjunciones “para que” y “según” aparecen en siete de los ocho entremeses de Cervantes, pero no en *Los romances*.
- La conjunción “más que” está presente en los ocho textos cervantinos, con una frecuencia de uso de [0,0005 , 0,0021], pero no en *Los romances*.
- La conjunción “aunque” aparece en todos los textos, tanto en los cervantinos como en *Los romances*, pero su frecuencia en el anónimo (0,0026) es ligeramente superior a la que presenta en los entremeses indubitados [0,009 , 0,0024]. Lo mismo ocurre con “cuando / cuándo”, con una presencia en el anónimo (0,0026) superior a la de Cervantes ([0,0004 , 0,0023]).

7.3.16. Búsqueda de palabras y expresiones del texto dubitado en corpus de referencia de la época

En este último apartado, y tratando de dar respuesta al trabajo de Rodríguez López-Vázquez [2020], vamos a realizar una serie de búsquedas de palabras y expresiones del texto dubitado “real”, a través de la base de datos del CORDE, para comprobar si su uso es idiosincrásico de Cervantes o si, por el contrario, es compartido por varios autores; para que el rastreo resulte significativo, delimitaremos cronológicamente la búsqueda al periodo temporal comprendido entre 1547, año de su nacimiento, y 1617, año en que se publica póstumamente *Los trabajos de Persiles y Segismunda*.

Como hemos explicado en varios puntos a lo largo de este trabajo, este método, por sí sólo, no es válido para determinar autorías, ya que CORDE es un repositorio limitado (y más si restringimos la búsqueda al periodo 1600-1613, en donde no está recogida, por ejemplo, ninguna obra dramática de Lasso de la Vega). Esta acotación temporal lleva a Rodríguez López-Vázquez a afirmar que “ninguno de los autores que hemos escrutado, ni Cervantes, ni Góngora, ni Gabriel Lobo Lasso de la Vega” usan la forma “dolo” [2020: 235], cuando, en realidad, está presente en todos ellos, a excepción

de Góngora: Cervantes la usa dos veces en el *Quijote* de 1615, Lasso de la Vega en *Manojuelo de romances y Mexicana* y Quevedo en sus *Poesías* y en el *Sueño del mundo por dentro*²⁵⁸.

Para dar respuesta a su análisis, hemos realizado varias decenas de búsquedas de palabras y expresiones del texto dubitado en el CORDE, aunque sólo mostraremos aquí aquéllas con una ocurrencia igual o inferior a 10 casos, por ser las más significativas (Tabla 55). Para no incurrir en el *cherry picking*, incluiremos también en esa tabla las diez búsquedas analizadas por Rodríguez López-Vázquez en su trabajo de atribución: “Pero Tanto/ pero tanto”, “¿Eso es verdad?”, “Mas me espanto”, “No os enojéis”, “no lo digo por tanto”, “Romancero”, “a pocos lances”, “tan fuera de sí”, “mi velado / su velado” y “Dolo” [2020: 234-235]:

CORDE (1547 – 1617)		
Secuencia [forma de búsqueda]	Casos / documentos	Autores
		Miguel de Cervantes (4)
		Juan Huarte de San Juan (4)
		San Juan Bautista de la Concepción (2)
		Jerónimo de Pasamonte (1)
		Diego Duque de Estrada (1)
<i>pero tanto</i> [pero tanto]	18 / 15	Lope de Vega Carpio (1)
		Juan de Mariana (1)
		Jerónimo Zurita (1)
		Luis Belmonte Bermúdez (1)
		Alonso de Ercilla (1)
		Anónimo (1)

²⁵⁸ Comete además graves incoherencias, ya que ofrece los resultados para el periodo 1600-1613, pero en algunos casos cita obras de Cervantes publicadas fuera de ese periodo, como *El trato de Argel* (1580) o la segunda parte del *Quijote* (1615). Véase este caso: “Parece claro que el personaje de Pero Tanto procede de la expresión «pero tanto». Cervantes la usa 4 veces: en las dos partes del *Quijote*, en *El trato de Argel* y en la novela ejemplar de *La ilustre fregona*. Extracto un pasaje de la primera parte del *Quijote*: «pero tanto cuanto tenía de esforzado y valeroso tenía de desdichado». Se podría pensar que se trata de una secuencia habitual, pero la consulta en el CORDE, entre la variante «pero tanto» y la variante «Pero tanto» nos da tan solo 15 casos entre 1600 y 1613. No aparece ni en Vélez de Guevara, ni en Quevedo, ni en Góngora, ni en Lasso de la Vega, posibles autores alternativos a Cervantes” [2020: 234].

CAPÍTULO 7

<p><i>¿Eso es verdad?</i> [?so es verdad]</p>	<p>28 / 23</p>	<p>Juan de Arce de Otárola (6) Lope de Vega Carpio (4) Miguel de Cervantes (3) San Juan Bautista de la Concepción (3) Fray Alonso de Cabrera (2) Tirso de Molina (1) Francisco Narváez de Velilla (1) Gregorio González (1) Melchor de Santa Cruz de Dueñas (1) Juan Méndez Nieto (1) Fray Luis de Granada (1) Juan de Pineda (1) Gonzalo Fernández de Oviedo (1) Fernán González de Eslava (1) Anónimo (1)</p>
<p><i>Mas me espanto</i> [?as me espanto]</p>	<p>3 / 3</p>	<p>Santa Teresa de Jesús (1) Juan Rufo (1) Luis Belmonte Bermúdez (1)</p>
<p><i>No os enojéis</i> [‘?o’ os enojéis] y [‘?o’ os enojeis]</p>	<p>10 / 10</p>	<p>Miguel de Cervantes (1) Juan de Arce de Otárola (1) Juan de Timoneda (1) Diego Hurtado de Mendoza (1) Pedro de Ribadeneira (1) Jaime Montañés (1) Jerónimo de Urrea (1) Sebastián de Horozco (1) Juan Ruiz de Alarcón (1) Anónimo (1)</p>
<p><i>no lo digo por tanto</i> [‘?o’ lo digo por tanto]</p>	<p>7 / 5</p>	<p>Miguel de Cervantes (2) Diego A. Velásques Velasco (2) Luis de Góngora (1) Juan de Arce de Otárola (1) Juan Pérez de Moya (1)</p>
<p><i>que no lo digo por tanto</i> [?ue ‘no’ lo digo por tanto]</p>	<p>1 / 1</p>	<p>Luis de Góngora (1)</p>
<p><i>Romancero</i></p>	<p>977 / 119</p>	<p>“No se puede Recuperar. Demasiados documentos”</p>

<i>partirse a la guerra</i> [parti* a la guerra]	10 / 7	Pedro de Luján (3) Juan Jerez (2) Diego López (1) Juan de Arce de Otárola (1), Francisco López de Gómara (1) Pedro Sarmiento de Gamboa (1) Anónimo (1)
<i>a pocos lances</i> [a pocos lances] y [A pocos lances]	26 / 14	Francisco de Luque Fajardo (9) Miguel de Cervantes (5) Diego Duque de Estrada (3) Fray Juan Márquez (2) Mateo Alemán (2) Cristóbal Suárez de Figueroa (1) José de Acosta (1) Antonio Hurtado de Mendoza (1) Fray José Sigüenza (1) Anónimo (1)
<i>tan fuera de sí</i> [tan fuera de sí]	42 / 28	Miguel de Cervantes (7) Jorge de Montemayor (4) Santa Teresa de Jesús (3) Jerónimo Fernández (3) Jerónimo de Urrea (2) Alonso Núñez de Reinoso (2) Diego Ortúñez de Calahorra (2) Fray Luis de Granada (2) Jerónimo de Urrea (2) Francisco López de Úbeda (1) Mateo Alemán (1) Francisco de Luque Fajardo (1) Lope de Vega (1) Fray Juan de los Ángeles (1) Pedro de Ribadeneira (1) Cristóbal Mosquera de Figueroa (1) Fray Martín Murúa (1) Fray Jerónimo Mendieta (1) Juan de Salinas (1) Luis Vélez de Guevara (1) Hernando de Ávila (1) Pedro de la Sierra (1) Diego Duque de Estrada (1) Antonio de Torquemada (1)

		Anónimos (28)
		Hugo de Celso (17)
		Juan de Castellanos (10)
		Fray Bartolomé de las Casas (5)
		Fray Juan de los Ángeles (5)
		Mateo Alemán (6)
		San Juan Bautista de la Concepción (4)
		Antonio de Torquemada (4)
		Francisco de Quevedo (4)
		Diego López (3)
		Fray Prudencio de Sandoval (3)
		Fray Diego de Hojeda (3)
		Juan Rufo (3)
		Miguel de Cervantes (2)
		Gabriel Lobo Lasso de la Vega (2)
		Lope de Vega Carpio (2)
		Francisco López de Úbeda (2)
		Luis Belmonte Bermúdez (2)
		Diego de Santiago (2)
		Fray Alonso de Cabrera (2)
		Juan Méndez Nieto (2)
		Luis Valle de la Cerda (2)
		Jerónimo Zurita (2)
		Juan de la Cueva (1)
		Fray Juan Márquez (1)
		Francisco Cascales (1)
		Luis de Toro (1)
		Pedro Hernández de Villaumbrales (1)
		Juan Valladares de Valdelomar (1)
		Hugo de Celso (1)
		Ginés Pérez de Hita (1)
		Fray Diego de Ocaña (1)
		Alonso de Santa Cruz (1)
		Hernando Arias de Saavedra (1)
		Bartolomé Cairasco de Figueroa (1)
		José de Villaviciosa (1)
		Luis Barahona de Soto (1)
		Guillén de Castro (1)
		Francisco de las Natas (1)
		Fray Prudencio de Sandoval (2)
		Alonso de Santa Cruz (1)
<i>Dolo</i>	131 / 66	
[dolo] y [Dolo]		
<i>más le persuadí</i>	3 / 2	Fray Prudencio de Sandoval (2)
[m?s le persuad*]		Alonso de Santa Cruz (1)

ESTUDIO DE AUTORÍA DEL *ENTREMÉS DE LOS ROMANCES*

<i>está tan fuera de sí</i> [est* tan fuera de s?]	6 / 5	Jorge de Montemayor (1) Lope de Vega (1) Santa Teresa de Jesús (4)
<i>mi velado</i> [mi velado]	3 / 3	Sebastián de Horozco (1) Fernán González de Eslava (1) Diego A. Velásques Velasco (1)
<i>tengo la culpa yo</i> [ten* la culpa yo]	3 / 3	Francisco López de Úbeda (1) Diego Hurtado de Mendoza (1) Cervantes de Salazar (1)
<i>¿Qué es aquesto en ..?</i> [qu? Es aquesto en]	1 / 1	Lope de Vega (1)
<i>a reñir con</i> [a reñir con]	9 / 9	Sebastián de Horozco (1) Antonio de Torquemada (1) Gonzalo Fernández de Oviedo (1) Juan de Jaúregui (1) Jerónimo de Urrea (1) San Juan Bautista (1) Fray Bernardino de Sahagún (1) Fray Bartolomé de las Casas (1) Tirso de Molina (1)
<i>sola y moza</i> [sola dist/4 moza]	6 / 5	Miguel de Cervantes (1) Lope de Vega (1) Juan Rufo (1) Juan de Pineda (1) Jerónimo de Urrea (2)
<i>tantas sinrazones en guerra y en paz</i> [sinrazones dist/5 guerra]	1 / 1	Gonzalo Fernández de Oviedo (1)
<i>y contado el caso</i> [‘y’ contando el caso]	4 / 3	Antonio de Torquemada (2) Gonzalo García (1) Alonso de Villegas (1)
<i>ir a embarcar</i>	9 / 5	Bernal Díaz del Castillo (3) Fernández de Oviedo (2) Gutiérrez de Santa Clara (2) Fray Bartolomé de las Casas (1) Alonso de Santa Cruz (1)
<i>senda alguna ni vereda</i> [senda dist/3 vereda]	1 / 1	San Juan Bautista (1)

CAPÍTULO 7

<i>tuya es la culpa</i> [?uya es la culpa]	2 / 2	Francisco López de Úbeda (1) Lope de Vega (1)
<i>te oigan los sordos</i> [oi* l?s sord*]	2 / 2	Gregorio González (1) Francisco de Quevedo (1)
<i>moro ni mora</i>	4 / 4	Anónimo (2) Anónimo (1) Luis de Mármol Carvajal (1)
<i>diablo que los aguarde</i> [diablo que l* aguar*]	2 / 2	Francisco de Quevedo (1) Anónimo (1)
<i>¡Santa María me valga!</i> [Santa Mar?a dist/5 val*]	1 / 1	Jerónimo de Urrea (1) ²⁵⁹
<i>Sin estoque ni puñal</i> [estoque dist/3 puñal]	1 / 1	Miguel de Cervantes (1)²⁶⁰
<i>Lleve el diablo el romancero</i> [¿leve el diablo el]	8 / 7	Lope de Vega (2) Miguel de Cervantes (1) Gabriel Lasso de la Vega (1) Juan Ruiz de Alarcón (1) Anónimo (1) Anónimos (2)
<i>Para alegrar la boda</i> [para aleg* la boda]	1 / 1	Luis Cabrera de Córdoba (1)
<i>Como un cochino roncaba</i> [cochin* dist/3ronc*]	1 / 1	Mateo Alemán (1)
<i>¡Maldita sea la casta!</i> [maldita dist/5 casta]	5 / 5	Felipe Guamán (1) Poma de Ayala (1) Luis de Mármol Carvajal (1) Pedro de Aguilar (1) Guillén de Castro (1) Anónimo (1)

Tabla 55. Frecuencias absolutas por clases de palabras en *Los romances* (Dub_EV_romances) y los entremeses en prosa de Cervantes (RStudio. FreeLing)

²⁵⁹ En este autor aparece una expresión muy similar: “¡Santa María, valme!”.

²⁶⁰ En el capítulo XV del *Quijote* de 1605 leemos: “tenía estoque, espada ni puñal”.

La Lingüística de corpus no permite atribuir el anónimo entremés a Cervantes:

- Las diez palabras/expresiones elegidas por Rodríguez López-Vázquez no son las más significativas, pues, de hecho, seis de ellas ofrecen más de 10 ocurrencias entre 1600 y 1613. De las otras cuatro (“más me espanto”, “no os enojéis”, “no lo digo por tanto” y “mi velado”), tres de ellas se dan en Cervantes, pero nunca en exclusividad.
- Por lo general, las palabras/expresiones más significativas (es decir, aquellas no superan las 10 ocurrencias) están presentes en multitud de autores. En varias ocasiones se repite el nombre de Cervantes, pero también el de Lope de Vega, Quevedo, Góngora, San Juan Bautista de la Concepción, Juan Rufo, Santa Teresa de Jesús o Gonzalo Fernández de Oviedo.
- Sólo hemos detectado siete resultados excluyentes:
 - o cuatro coocurrencias: “¡Santa María me valga!” (Jerónimo de Urrea), “sin estoque ni puñal” (Cervantes), “como un cochino roncaba” (Luis Cabrera de Córdoba), “tantas sinrazones en guerra y en paz” (Gonzalo Fernández de Oviedo) y “senda alguna ni vereda” (San Juan Bautista);
 - o una expresión cuasi coincidente de 4 ítems: “para alegrar la boda” (Luis Cabrera de Córdoba);
 - o un *verbatim* de 4 ítems: “¿Qué es aquesto en...?” (Lope de Vega);
 - o y un *verbatim* de 6 ítems, ciertamente relevante: “que no lo digo por tanto” (Góngora)
- Resulta interesante descubrir que existen concomitancias textuales bastante significativas entre la parte real del entremés dubitado y varios poemas anónimos del *Romancero general*, lo que parece confirmar que, incluso cuando no copia literalmente versos de romances conocidos, el ignoto entremesista está influenciado por esas composiciones (que, como hemos apuntado, posiblemente tenga delante mientras piensa, conforma y compone los versos de su entremés). Sirva a modo de ejemplo el caso de la búsqueda “lleve el diablo el romancero” [¿leve el diablo el], con la que el CORDE nos

remite a una *Sátira de romances moriscos* que comienza “Lleve el diablo el potro rucio / del alcaide de los Vélez”, poema que guarda similitud con otros versos de la pieza dubitada.

7.4. CONCLUSIONES SOBRE LA AUTORÍA CERVANTINA DE ESTE ENTREMÉS

El estudio de atribución de autoría de *El entremés de los romances* es un caso complejo, pues partimos de la certeza de que coexisten varios tipos de autoría: una autoría ejecutiva inicial que construye el cincuenta por ciento de la pieza sobre la autoría precursora de varios romancistas; una autoría ejecutiva posterior que modifica el texto a su conveniencia, en función de las peculiaridades de la compañía o la ciudad en que se representa; y una autoría final de revisión, que se encarga de componer, editar e imprimir el manuscrito (y que factiblemente dista en algún punto del original)²⁶¹.

Los análisis contrastivos realizados en este capítulo nos permiten establecer una serie de conclusiones sobre la posible autoría cervantina de la parte “real” de esta pieza, es decir, aquella ajena a los versos parodiados del Romancero:

1. Aunque presenta una idea temática análoga, un argumento coincidente y una estructura subyacente común a la de la primera salida de don Quijote, su estrategia paródica es más plana, débil y encierra una serie de aspectos incongruentes que la separan de la genialidad creadora cervantina. Además, la presencia de cuatro mujeres entre los personajes se aleja de lo habitual en los entremeses de Cervantes, en donde aparecen dos o tres mujeres como máximo [Agostini, 1964: 275].

²⁶¹ A este respecto, recordemos que Cervantes, en la dedicatoria de sus *Ocho comedias y ocho entremeses*, advierte que “si alguna cosa llevan de razonable, es que no van manoseados ni han salido al teatro, merced de los farsantes”.

2. La pieza adolece de varios de los rasgos que cualitativamente definen el lenguaje cervantino, como los superlativos en *-ísimo*, los diminutivos en *-ico*, los adverbios en *-mente* y los latinajos. También echamos en falta la composición de palabras nuevas con fines humorísticos o la caracterización de los personajes a través del lenguaje. Los únicos rasgos cervantinos que hemos detectado son el leísmo, la utilización pragmática de diálogos interrogativos y la presencia de reduplicaciones y sinónimos coordinados (aunque en mucha menor medida).
3. La pieza tiene un comportamiento métrico diametralmente opuesto al de los dos entremeses en verso de Cervantes, que se decantan por el endecasílabo suelto, rehúyen el empleo de la redondilla y la quintilla (ya sea suelta o en copla real) y sólo utilizan el romance en pasajes cantados. Ahora bien, estas diferencias podrían estar condicionadas por el tema de la pieza, que reclama, en cierto modo, la utilización de romances.
4. Tal y como muestran los distintos análisis de la distribución de la frecuencia de los primeros 1000 gramas de palabra y 5-gramas de caracteres, realizados con el algoritmo Eder's Delta (*Analysis Cluster* y *Bootstrap Consensus Tree*), el texto dubitado siempre se ubica en una rama diferente a la que acoge a los ocho entremeses de Cervantes.
5. El análisis de componentes principales de la distancia Delta (Eder's Delta) muestra una clara disparidad entre la parte real de *Los romances* y las ocho piezas cervantinas. El texto dubitado se sitúa en el cuarto cuadrante, en el que no figura ninguno de los ocho entremeses cervantinos.
6. El análisis de clasificaciones supervisadas por aprendizaje automático (Classify) asocia correctamente todos los textos indubitados a partir de las 700 palabras más frecuentes; el texto dubitado sólo se asocia con una obra de Cervantes en las 100 MFW, pero, a partir de esa cantidad, lo hace siempre con una obra de Francisco de Ávila (200-1000 MFW).
7. *Los romances* muestra un comportamiento léxico muy alejado del que tienen los dos entremeses en verso de Cervantes, ya que presenta menor porcentaje

de palabras de función, mayor densidad en la ratio *type/token* de palabras de contenido (TTR Content) y menor riqueza léxica (*Lexical richness*).

8. Cinco de los diez entremeses del corpus con mayores porcentajes de similitud textual con el texto dubitado son de Cervantes, si bien los porcentajes son muy inferiores a los que veíamos en *La cárcel de Sevilla* (pues en este caso nunca superan el 28%). Es muy significativo que sólo mantenga un 18% de similitud textual con los seis primeros capítulos del *Quijote* de 1605, teniendo en cuenta la vinculación genética que existe entre ambas obras (ya sea en una dirección u otra), lo que nos lleva a pensar en autorías independientes. Las mismas conclusiones obtenemos al advertir los bajos porcentajes de vocabulario léxico compartido una sola vez (*once only in both*) y los altos porcentajes de vocabulario léxico exclusivo (*only in*).
9. El análisis de verificación de autoría, realizado con el método General Imposters (GI), permite reforzar la hipótesis no sólo de que Cervantes no es el autor de *Los romances*, sino también de que probablemente no lo sea ninguno de los autores del corpus de respaldo (pues sólo en una ocasión se aproxima a Belmonte por encima del 0,63).
10. El texto dubitado comparte varios *verbatim* con Cervantes, pero ninguno de ellos de más de 5 ítems, que, como hemos explicado, son los realmente significativos. Cabe decir, no obstante, que CORDE (1547-1617) sólo registra el uso del *verbatim* “mi mujer, que está presente”, de 5 ítems, en Cervantes y Lope de Vega.
11. El estudio del orden de frecuencia, la frecuencia relativa y el análisis de componentes principales (PCA) de las palabras de función tampoco permite identificar al autor del texto anónimo con Cervantes.
12. *Los romances* y los dos entremeses en verso de Cervantes presentan diferencias en el uso de ciertas clases de palabras. El autor del texto dubitado utiliza más verbos, nombres propios, pronombres y conjunciones subordinadas que Cervantes y menos adjetivos y nombres comunes.

13. El autor del texto anónimo se aproxima bastante a Cervantes en el uso de las conjunciones coordinadas copulativas (“y” y “ni”), disyuntivas (“o”) y explicativas (“esto es”), pero no en el de las adversativas: mientras él prefiere el uso de “pero” y, en menor medida, de “mas”, el autor del *Quijote* se decanta por el uso de “sino” (que en el anónimo no aparece) y utiliza en menor proporción las formas “pero” y “mas”. En el caso de las conjunciones subordinadas adverbiales, hay ciertos usos similares (como ocurre con “donde / dónde”) pero también notables diferencias (“más que”, “para que”, “según” “aunque” y “cuando / cuándo”).
14. La búsqueda de palabras y expresiones del texto dubitado “real” en CORDE (1547-1617) no permite detectar rasgos exclusivos de Cervantes.

Aunque la breve extensión del dubitado “real” y el entramado de cruces de autoría dificultan en gran medida la identificación del texto, a tenor del análisis cualitativo y cuantitativo realizado previamente podemos concluir —siguiendo la escala de opinión verbal propuesta en el apartado 5.6.— que Miguel de Cervantes Saavedra no es el autor del *Entremés de los romances*, con un grado de probabilidad (-1) bastante alto, ya que el idiolecto que se desprende del estudio de sus textos indubitados presenta muy pocos rasgos coincidentes con el del idiolecto del texto dubitado.

CONCLUSIONES

Tal y como hemos dejado patente a lo largo de estas páginas, la autoría es una realidad compleja, heterogénea y cambiante, que afecta a muy diferentes ámbitos del conocimiento. En sentido teórico, un autor literario es aquél que crea o inventa una obra; pero la práctica nos demuestra, como ha estudiado Love, que, junto a la autoría puramente ejecutiva, hay autorías de revisión, colaborativas, declarativas o precursoras, que complican enormemente la definición de este concepto [2002]. Además, más allá de la complejidad terminológica, la autoría lleva implícitos ciertos condicionantes relacionados con la función social del sujeto, su economía o sus derechos y responsabilidades sobre el objeto creado, lo que ha influido, a lo largo de la historia, en la proliferación de textos anónimos, apócrifos, con pseudónimo, mal atribuidos o censurados.

Al margen de todas estas cuestiones, lo que no se puede negar es la importancia que las sociedades, desde las más primitivas a las actuales, confieren al asunto de la autoría. Los escritores se han preocupado desde antiguo por ella, ya sea para fijarla u ocultarla, en función del beneficio o perjuicio que ese acto llevara aparejado. También se han interesado por ella los estudiosos del hecho literario; basta pensar, por ejemplo, en los miles de páginas que se han escrito sobre los autores de la Biblia hebrea o sobre la existencia real de Homero. Incluso Foucault, que cuestionó la función-autor en la mayoría de los textos, concede a la Literatura un estatus especial, al defender que su anonimato —ya sea accidental o voluntario— no es socialmente soportable.

CONCLUSIONES

A pesar de todo ello, los estudios de autoría en el ámbito literario siempre han sido vistos con cierto recelo, algo así como “investigar la prostitución o algún otro tema supuestamente «indecente»” [Eisenberg, 1990]. La mayor parte de la crítica acepta sin ningún recelo las atribuciones basadas en evidencias externas, es decir, aquéllas que provienen del mundo social dentro del cual se crea, promulga y lee la obra; pero sienten mucha mayor reticencia a la hora de avalar las evidencias internas, al entender —no sin cierta razón— que los métodos utilizados hasta ahora adolecían de rigor científico. En la última década, gracias principalmente a la solvencia que están demostrando los análisis cuantitativos desarrollados por la Estilometría, las evidencias internas han empezado a ganar adeptos; estos estudios, de gran rigor metodológico, no han logrado, sin embargo, establecer un sistema consensuado para expresar las conclusiones alcanzadas.

La falta de atención que reciben los estudios de atribución de autoría en el mundo de la Filología contrasta con su auge en el ámbito forense²⁶². La Lingüística forense se ha convertido en una disciplina clave en el entorno de la Justicia y de los Cuerpos y Fuerzas de Seguridad del Estado, ayudando a la investigación y el esclarecimiento de casos claves de nuestra historia criminal. Es en este contexto en el que surge la idea de este proyecto de tesis doctoral, que plantea, primero, las ventajas de emplear los fundamentos de esta disciplina a los estudios de autoría literaria, para, posteriormente, aplicarlos al caso concreto —y complejo— de los diez entremeses que alguna vez se han atribuido a Miguel de Cervantes Saavedra (*Los habladores*, *El hospital de los podridos*, *La cárcel de Sevilla*, *Los refranes*, *Los mirones*, *Los romances*, *Melisendra*, *Durandarte y Belerma*, *Doña Justina y Calahorra* y *Ginetilla ladrón*, hoy perdido).

²⁶² Recordamos aquí que, a día de hoy, la etiqueta “atribución de autoría” se utiliza de manera genérica para referirnos a cualquier proceso mediante el cual se analizan las características lingüísticas de unos determinados textos para extraer conclusiones sobre la identidad de su autor, con independencia de la tipología textual (sólo dubitados o dubitados e indubitados) y del número de candidatos a autor (ninguno, uno o varios) [Zheng, Qin, Huang y Chen, 2003]. Para profundizar en esta cuestión, véase la nota 51.

CONCLUSIONES

Las conclusiones alcanzadas pueden resumirse en los siguientes puntos:

Primero. Los estudios de autoría en el ámbito literario deben asumir los fundamentos teóricos, metodológicos y conclusivos de la Lingüística forense, para lograr ser más objetivos, fiables y precisos:

- El principio teórico fundamental que sustenta los estudios de autoría en el ámbito forense es la existencia del idiolecto, es decir, una versión personal, marcada y distintiva de la lengua que hablamos y escribimos, que nos define y diferencia de los demás. Dicha versión, conformada por multitud de rasgos lingüísticos, puede cambiar ligeramente a lo largo del tiempo y adaptarse a las diferentes circunstancias comunicativas, pero, aún así, la variación intrapersonal será menor que la variación interpersonal. Es por ello por lo que el idiolecto puede usarse como evidencia lingüística para discriminar autorías (*Language as evidence*).
- La metodología desarrollada por la Lingüística forense, basada en el principio de certeza, comprende el análisis comparado de las distintas muestras lingüísticas, teniendo en consideración las posibles limitaciones derivadas de su calidad, extensión, datación, temática o tipología textual. Esta metodología, que abarca los distintos niveles de la lengua, recomienda la combinación de técnicas cualitativas y cuantitativas, para poder ofrecer resultados objetivos y replicables sin renunciar por ello a la riqueza interpretativa y la profundidad de significados que aporta la observación directa (metodología mixta). Nunca la coincidencia o disimilitud de un solo rasgo lingüístico podrá usarse como prueba para atribuir o *desatribuir* una obra a un determinado autor.
- La expresión de las conclusiones es otro de los elementos clave de la Lingüística forense (del que, hasta el momento, adolecían los estudios literarios). Todo esfuerzo teórico y metodológico es insuficiente si el investigador no es capaz de ofrecer unos resultados ecuanímenes, que todos los receptores sean capaces de interpretar de igual modo, en base al principio de probabilidad de la criminalística moderna. En el ámbito forense de la atribución de autoría se trabaja

CONCLUSIONES

con escalas de opinión verbal, en las que se recogen los distintos niveles de probabilidad dentro de los que puede enmarcarse el resultado global; además, de un tiempo a esta parte, se están empezando a desarrollar escalas bayesianas, aunque la falta de estudios poblacionales, indispensables para el cálculo de la razón de verosimilitud, hace aún muy difícil su aplicación práctica. Nosotros proponemos, por tanto, una escala de opinión verbal que represente los seis grados de probabilidad de las dos hipótesis posibles (sí / no), así como la imposibilidad de llegar a alguna de ellas (resultado no concluyente) (Tabla 14).

Segundo. A pesar de que el entremés es un género breve, repetitivo y cambiante, cuya impresión se produce, por lo general, tras un largo proceso de contaminación textual y/o autorial, las pruebas realizadas para testar la operatividad de ciertos análisis cuantitativos en el corpus indubitado ofrecen resultados altamente satisfactorios:

- El análisis de la distribución de la frecuencia de las 800 y 1000 palabras más empleadas en los textos, realizado con el paquete Stylo de RStudio (Eder's Delta), muestra grupos de asociación diferenciados para las obras de Cervantes, Ávila, Salas Barbadillo, Hurtado de Mendoza, Castillo de Solórzano, Quiñones de Benavente, Quevedo y Belmonte; lo que supone un porcentaje de fiabilidad superior al 90% (Gráficos 3 y 4). Los únicos errores de agrupación se producen entre las tres obras de Vélez de Guevara, cuyas causas, posiblemente, estén relacionadas con las prácticas pseudoplagiarias atribuidas a este autor o la baja calidad de las muestras seleccionadas. Mención especial merece el caso de Barrionuevo, cuyo único entremés se asocia siempre con los de Cervantes (nunca de manera directa), pero del que no poseemos ninguna otra obra indubitada que permita certificar, como sería lo esperable, que la distancia entre ellas es menor que la que mantiene con cada una de las cervantinas (*nearest neighbor problem*). Estos resultados son ciertamente relevantes, pues algunos especialistas habían apuntado la imposibilidad de aplicar la distancia Delta a textos del teatro español del Siglo de Oro de menos de 2000-2500 palabras, que es la extensión que tienen varios de los textos que conforman nuestro corpus de análisis [Campión Larumbe y Cuéllar, 2021: 59].

CONCLUSIONES

- El análisis de la distribución de la frecuencia de las 800 y 1000 cadenas de 5 caracteres más frecuentes, realizado de nuevo con el paquete Stylo de RStudio (Eder's Delta), también ofrece resultados fiables para discriminar entre los diferentes autores, aunque en menor grado que con palabras independientes. Este resultado era el esperado, pues debemos tener en cuenta que ni las variantes gráficas ni los signos de puntuación contenidos en los textos son los realmente utilizados por los autores primigenios. Aún así, todas las asociaciones directas se producen entre obras de un mismo autor (Gráficos 16 y 17).
- El análisis del porcentaje de similitud textual, llevado a cabo con CopyCatch Gold tras aplicar una *stopword list* de palabras de función, revela que los ocho entremeses de Cervantes guardan entre sí una similitud del [26 , 44] %, superior a la que mantienen con los del resto de autores [15 , 35] %, a excepción de Barrionuevo, con cuyo entremés sí supera el controvertido umbral del 40% (Gráficos 5-14). De nuevo, el hecho de no poseer otras obras indubitadas de este autor imposibilita profundizar en el resultado; pero, cuanto menos, resultan llamativos los porcentajes de similitud detectados entre *El triunfo de los coches* y *El juez de los divorcios* (41%), *El viejo celoso* (42%), *La cueva de Salamanca* (42%) y *El vizcaíno fingido* (46%).
- El estudio de *verbatim*, es decir, de las secuencias de palabras iguales y en el mismo orden, demuestran que —como establece la Lingüística forense— los casos compartidos de más de 5 ítems sólo se dan entre obras de unos mismos autores. Hemos detectado, por ejemplo, *verbatim* de 6 ítems entre obras de Vélez de Guevara, de 8 entre obras de Cervantes y hasta de 23 entre obras de Salas Barbadillo (Tabla 11).

Tercero. El repaso de la historia crítica de los diez entremeses pro hijados a Cervantes pone de manifiesto la existencia de dos posturas enfrentadas: la primera de ellas provoca la atribución masiva de obras a este autor durante todo el siglo XIX, como consecuencia de la idolatría que producía el éxito alcanzado por el *Quijote* a nivel mundial; la segunda, que surge a raíz del descubrimiento de la falsificación del *Buscapié*, genera un rechazo total hacia todas ellas. Los argumentos esgrimidos en

CONCLUSIONES

ambos casos —basados en meras coincidencias temáticas o estilísticas— han contribuido, en parte, al desprestigio ligado a este tipo de estudios. En la última década, algunos investigadores han vuelto a reivindicar la autoría cervantina de varias de estas obras (*Los habladores*, *Los romances*, *Los mirones*, *Doña Justina* y *Calahorra* y *La cárcel de Sevilla*), en nuestra opinión, como hemos explicado detalladamente en varios puntos, con métodos insuficientes para alcanzar conclusiones concluyentes [Rodríguez López-Vázquez, 2011a; 2011b; 2013a; 2019; y Peña Muñoz, 2010].

Cuarto. Aunque nuestro trabajo se limita al estudio de la posible autoría cervantina de dos de los nueve entremeses conservados, uno en prosa (*La cárcel de Sevilla*) y otro en verso (*Los romances*), los análisis realizados con todos ellos (*PCA*, *Delta distance*, *similarity* y *verbatim*) nos permiten plantear las siguientes hipótesis de trabajo, sobre las que esperamos vuelvan la vista los estudiosos de la literatura aurea o nosotros mismos en un futuro:

- El único entremés en el que se aprecian coincidencias significativas con el idiolecto de los ocho entremeses cervantinos es el de *La cárcel de Sevilla*;
- *Los habladores*, *El hospital de los podridos*, *Los refranes*, *Los romances*, *Melisendra* y *Doña Justina* y *Calahorra* no presentan tales coincidencias, por lo que, al menos *a priori*, no parecen ser de Cervantes.
- No podemos por el momento establecer hipótesis concluyentes sobre la posible autoría cervantina de *Los mirones* y *Durandarte* y *Belerma*, pues, aunque en el código del setecientos conservado en la Biblioteca Colombina de Sevilla (BC, Ms. 82-3-40) llevan el título de “entremés”, lo cierto es que su extensión y configuración ponen en duda esta clasificación genérica. El primero es una especie de coloquio o novela ejemplar dialogada (del tipo de *El coloquio de los perros*) y el segundo una comedia burlesca; siendo así, los análisis contrastivos, realizados dentro de un corpus integrado únicamente por entremeses, podrían desvirtuar los resultados.

CONCLUSIONES

Quinto. El primer componente (PC1) del análisis de componentes principales (PCA) aplicado a la distancia Delta de las palabras más empleadas en los textos tiende a diferenciar entre entremeses en verso (mitad derecha) y en prosa (mitad izquierda) (Gráfico 20). Aunque no hemos profundizado en este análisis, por no ser el objeto de nuestro estudio, dicha diferenciación podría corresponderse, en realidad, con una ordenación temporal, tal y como se ha demostrado que ocurre en obras de unos mismos autores [Celma Valero y Ruiz Urbón, 2021], pues es sabido que la moda del entremés en prosa es anterior a la del entremés en verso. De ser así, podríamos proponer una nueva teoría de la fecha de redacción de los ocho entremeses cervantinos basándonos en su distribución en los cuatro cuadrantes de una gráfica de puntuaciones.

Sexto. El estudio del *Entremés de la cárcel de Sevilla* nos lleva a las siguientes consideraciones:

- Tal y como consta en un apógrafo dieciochesco de la comedia *La viuda, casada y doncella* de Lope de Vega (BNE, Ms. 22 423), este entremés obtuvo licencia de representación en el año 1600 a cargo de Tomás Gracián Dantisco; su fecha de composición, por tanto, es muy anterior a la que le ha otorgado la crítica tradicional (1610-1612), basándose, exclusivamente, en la presencia del personaje de Escarramán [Asensio, 1971: 90-91].
- Las coincidencias entre este entremés y las dos primeras partes de la *Relación* de Cristóbal de Chaves ponen de manifiesto que su autor no sólo había leído estos textos, sino que, además, hubo de tenerlos delante al componer su obra; sólo así pueden explicarse las coincidencias textuales que se aprecian entre ellos (Tabla 15). Pero más allá de todas estas coincidencias —propias de la reescritura creativa—, el idiolecto de Chaves está muy alejado del que aparece en el dubitado (Gráficos 41, 45 y 46 y Tablas 22 y 26).
- Los diferentes análisis realizados (léxicos, morfológicos, sintácticos, estructurales, etc.), cuyas conclusiones hemos detallado en el apartado 6.4., permiten determinar que Cervantes sí es el autor de este entremés con un grado de probabilidad “bastante alto” (nivel 4 en una escala del 0 al 5).

CONCLUSIONES

Séptimo. El estudio detallado del *Entremés de los romances* nos lleva a las siguientes consideraciones:

- Las similitudes temático-argumentales entre esta pieza y la primera salida de don Quijote (*Quijote*, I, 1-6) invitan a pensar en una relación inter-textual entre ambos escritos; esta relación, sin embargo, no es directa, pues no se detectan casos de *verbatim* entre ambas obras (como sí veíamos entre *La cárcel de Sevilla* y la *Relación de Chaves*), al margen de la citación de unos mismos versos (con variantes) del *Romance del Marqués de Mantua*.
- Aunque la direccionalidad de la copia entre ambos textos no afecta al asunto de la autoría, consideramos, en contra de lo que piensa la tendencia mayoritaria, que el entremés es posterior. Nuestros argumentos se basan en (i) la presencia del escudero; (ii) la falta de coherencia en la caracterización del habla de los personajes cuerdos; (iii) el empleo de la forma versificada y del romance dialogado, propios de principios del XVII; (iv) y el empleo de la fórmula sintética “vuesastedes”, frente al uso analítico anterior “vuesa/s merced/es”, que es el que encontramos en el *Quijote*.
- El estudio de autoría de este entremés supone todo un reto, pues el texto original, es decir, aquél resultante de eliminar los 32 fragmentos de conocidos romances insertados en él, apenas supera las 1150 palabras. Los diferentes análisis realizados (léxicos, morfológicos, sintácticos, estructurales, métricos, etc.), cuyas conclusiones hemos detallado en el apartado 7.4., permiten determinar que Cervantes no es su autor, con un grado de probabilidad “bastante alto” (nivel -1, en una escala del -5 al 0).

Es cierto que en el rompecabezas literario del Siglo de Oro español aún quedan muchas piezas por encajar y que, muy probablemente, algunas de ellas pertenezcan a Cervantes; pero no es menos cierto que los criterios con los que se le han ido adjudicando cerca de un centenar de obras a lo largo de los tiempos (incluidas *El cerco de Numancia* y *El trato de Argel*) son a día de hoy insuficientes. Nos parece inadmisibile seguir manteniendo la atribución de obras a este autor por su calidad literaria, pues ni

CONCLUSIONES

todos los buenos textos fueron escritos por él, ni él escribió siempre buenos textos; o por la semejanza de argumento con alguna de sus obras reconocidas, pues bien es sabido que los autores de nuestra literatura barroca bebieron de fuentes comunes y emplearon la refundición temático-argumental en gran parte de sus creaciones. Por eso, aunque útiles en su momento y dentro de su contexto, todo este *maremágnum* de filiaciones cervantinas ha quedado totalmente obsoleto, siendo preciso revisarlo y reelaborarlo desde la óptica del siglo XXI y las posibilidades metodológicas que ofrece la Lingüística forense como disciplina científica.

Confiamos en que nuestro trabajo contribuya a esta gran empresa. Pero, sobre todo, que permita empezar a sentar las bases de lo que deberían ser —y lo que no— los estudios de autoría literaria.

CONCLUSIONS

As we have made clear throughout these pages, authorship is a complex, heterogeneous and changing reality that affects very different areas of knowledge. In a theoretical sense, a literary author is someone who creates or invents a work; but as Love has concluded from his studies, practice shows us that along with purely executive authorship there are also revisionary, collaborative, declarative and precursory authorships, which hugely complicate the definition of this concept [2002]. Also, beyond the complexity of the terminology, authorship implicitly includes certain conditioners related to the social function of the subject, the author's economic situation or rights and responsibilities over the created object. Throughout history, such factors have had an influence on the proliferation of anonymous, apocryphal, pseudonymous, misattributed or censored texts.

Aside from all these questions, what cannot be denied is the importance that societies, from the most primitive to those of today, attach to the issue of authorship. Since ancient times, writers have been concerned with either establishing it or concealing it, depending on the benefit or harm this could entail. Literary scholars have also been interested in it. One need only think, for example, of the thousands of pages that have been written about the authors of the Hebrew Bible or about whether Homer ever actually existed. Even Foucault, who questioned the author function in most texts, grants literature a special status, arguing that its anonymity—whether accidental or deliberate—is no longer acceptable today.

CONCLUSIONS

Despite all this, studies of authorship in the field of literature have always been regarded with a certain amount of suspicion, rather like “researching prostitution or some other supposedly «indecent» subject” [Eisenberg, 1990]. Most critics readily accept attributions based on external evidences (that which comes from the social world within which the work is created, promulgated and read). But they are much more reluctant to endorse internal evidences, believing —not entirely unreasonably— that the methods used up to now lacked scientific rigour. In the last decade, thanks mainly to the solvency being demonstrated by the quantitative analyses developed by stylometry, internal evidence has grown in popularity. However, although of great methodological rigour, these studies have not managed to establish a consensual system for expressing the conclusions reached.

The lack of attention given to authorship attribution studies in the world of philology contrasts with its boom in the forensic field²⁶³. Forensic linguistics has become a key discipline in the area of justice and for the security agencies and forces of the State, helping in the investigation and clarification of key cases in our criminal history. It is in this context that the idea of this doctoral thesis arose, which first of all proposes the advantages of using the basic principles of this discipline in literary authorship studies in order to subsequently apply them to the concrete —and complex— case, of the ten *entremeses* that have been attributed to Miguel de Cervantes Saavedra (*Los habladores, El hospital de los podridos, La cárcel de Sevilla, Los refranes, Los mirones, Los romances, Melisendra, Durandarte y Belerma, Doña Justina y Calahorra* and the now lost *Ginetilla ladrón*).

The conclusions of the thesis can be summarised as follows:

²⁶³ We recall here that, nowadays, the label "authorship attribution" is used generically to refer to any process by which the linguistic characteristics of certain texts are analysed to draw conclusions about the identity of their author, regardless of the textual typology (only disputed or disputed and known) and the number of authorial candidates (none, one or several) [Zheng, Qin, Huang and Chen, 2003]. For more on this issue, see note 52.

First. Authorship studies in the field of literature must take on the theoretical, methodological and conclusive fundamentals of forensic linguistics if they are to succeed in being more objective, reliable and precise:

- The fundamental theoretical principle that underpins forensic authorship studies is the existence of the *idiolect*: a personal, marked and distinctive version of the language we speak and write, which defines us and differentiates us from others. This idiolect, which comprises a multitude of linguistic features, may change slightly over time and adapt to different circumstances of communication; but even so, intrapersonal variation will be less than interpersonal variation. This is the reason why idiolect can be used as linguistic evidence to discriminate authorship (*Language as evidence*).

- The methodology developed by forensic linguistics, based on the principle of certainty, involves the comparative analysis of different linguistic samples, taking into consideration the possible limitations due to their quality, length, dating, subject matter or textual typology. This methodology, which covers the different levels of the language, recommends combining qualitative and quantitative techniques in order to be able to obtain objective and replicable results without thereby renouncing the interpretative richness and depth of meaning obtained through direct observation (mixed methodology). The coincidence or dissimilarity of a single linguistic feature can never be used as evidence to attribute or not to attribute a work to a particular author.

- The expression of conclusions is another key element of forensic linguistics (which, up to now, literary studies have lacked). No theoretical and methodological effort is sufficient if the researcher is not capable of offering fair and balanced results that all recipients can interpret in the same way, based on the principle of probability in modern forensics. The forensic field of attribution of authorship works with verbal opinion scales, in which the different levels of probability within which the overall result can be framed are collected. Furthermore, Bayesian scales are beginning to be developed; although the lack of population studies, which are essential for calculating the likelihood ratio, still

CONCLUSIONS

makes the practical application of these scales very difficult. We therefore propose a verbal opinion scale representing the six degrees of probability of the two possible hypotheses (yes / no), as well as the impossibility of arriving at either (inconclusive result) (Table 14).

Second. Despite the fact that the *entremés* is a short, repetitive and changing genre, which is generally printed after a long process of textual and/or authorial contamination, the tests carried out to test the operability of certain quantitative analyses in the corpus in question offer highly satisfactory results:

- The analysis of the frequency distribution of the 800 and 1,000 most used words in the texts, carried out with RStudio's Stylo package (Eder's Delta), shows differentiated association groups for the works of Cervantes, Ávila, Salas Barbadillo, Hurtado de Mendoza, Castillo de Solórzano, Quiñones de Benavente, Quevedo and Belmonte (indicating a reliability rate greater than 90%) (Figures 3 and 4). The only grouping errors that occur are in the three works by Vélez de Guevara, the causes of which may be related to the pseudo-plagiaristic practices attributed to this author or the low quality of the selected samples. Special mention should be made of the case of Barrionuevo, whose sole *entremés* is always associated with those of Cervantes (never directly) but of whom we do not possess any other indisputable work that would allow us to certify, as would be desirable, that the distance between those two works is less than that from any of the works of Cervantes (*nearest neighbor problem*). These results are definitely relevant, as some specialists had pointed out the impossibility of applying Delta distance to Spanish Golden Age theatre texts of less than 2,000 – 2,500 words, which is the length of several of the texts that make up our corpus of analysis [Camióñ Larumbe y Cuéllar, 2021: 59].

- Analysis of the frequency distribution of the 800 and 1,000 most frequent 5-character strings, again carried out with RStudio's Stylo package (Eder's Delta), also provides reliable results for discriminating between the different authors, although to a lesser degree than with independent words. This result was to be expected, since we must bear in mind that neither the graphic variants nor the

punctuation marks contained in the texts are those that were actually used by the original authors. Even so, all the direct associations occur between works by the same author (Figures 16 and 17).

- Analysis of the percentage of textual similarity, carried out using CopyCatch Gold after applying a *stopword list* of function words, reveals that the eight *entremeses* by Cervantes have a similarity between them of [26 , 44]%, higher than that of the other authors [15 , 35]%, with the exception of Barrionuevo, whose *entremés* does exceed the controversial threshold of 40% (Figures 5-14). Again, the fact that we have no other undisputed works by this author makes it impossible to go into the result in depth; but at least the percentages of similarity detected between *El triunfo de los coches* and *El juez de los divorcios* (41%), *El viejo celoso* (42%), *La cueva de Salamanca* (42%) and *El vizcaíno fingido* (46%) are striking.

- The study of *verbatim*, that is, of sequences of the same words and in the same order, shows that —as established by forensic linguistics— shared cases of more than 5 items only occur between works by the same authors. We have detected, for example, *verbatim* of 6 items between works by Vélez de Guevara, 8 between works by Cervantes and up to 23 between works by Salas Barbadillo (Table 11).

Third. A review of the critical history of the ten Cervantes *entremeses* reveals two opposing positions: the first of these leads to the massive attribution of works to this author throughout the 19th century as a result of the idolatry produced by the worldwide success of *Don Quixote*; the second, which arose as a result of the discovery of the forgery of the *Buscapié*, causes them all to be rejected. The arguments put forward in both cases —based on mere thematic or stylistic coincidences— have partly contributed to the discredit attached to studies of this type. In the last decade, some researchers have once again claimed Cervantes's authorship of several of these works (*Los habladores*, *Los romances*, *Los mirones*, *Doña Justina y Calahorra* and *La cárcel de Sevilla*). In our opinion, as we have explained in detail at various stages, this has been done using insufficient methods to reach definite conclusions [Rodríguez López-Vázquez, 2011a; 2011b; 2013a; 2019; and Peña Muñoz, 2010].

CONCLUSIONS

Fourth. Although we limited our work to the study of the possible Cervantine authorship of two of the nine preserved *entremeses* —one in prose (*La cárcel de Sevilla*) and the other in verse (*Los romances*)— the analyses carried out on all of them (*PCA*, *Delta distance*, *similarity* and *verbatim*) allow us to propose the following working hypotheses, to which we hope scholars of the literature of the Golden Age or we ourselves will turn our attention in the future:

- The only *entremés* in which one notes significant coincidences with the idiolect of Cervantes's eight *entremeses* is that of *La cárcel de Sevilla*;
- *Los habladores*, *El hospital de los podridos*, *Los refranes*, *Los romances*, *Melisendra* and *Doña Justina y Calahorra* do not show such coincidences, so, at least *a priori*, they do not seem to be by Cervantes.
- We cannot for the moment establish any conclusive hypotheses about the possible Cervantine authorship of *Los mirones* and *Durandarte y Belerma*, since, although in the seventeenth-century codex preserved in the Biblioteca Colombina in Seville (BC, Ms. 82-3-40) they bear the title of "*entremés*", their length and configuration cast doubt on this generic classification. The former is a kind of colloquy or exemplary dialogue novel (of the type of *El coloquio de los perros*); the latter a burlesque comedy. Contrastive analyses, therefore, carried out within a corpus made up solely of *entremeses* could distort the results.

Fifth. The first component (PC1) of the principal component analysis (PCA) applied to the Delta distance of the most used words in the texts tends to differentiate between *entremeses* in verse (right half) and in prose (left half) (Figure 20). Although we have not analysed this in greater depth, it not being the object of our study, this differentiation could actually correspond to a temporal arrangement, such as has been shown to occur in works by the same authors [Celma Valero and Ruiz Urbón, 2021], since it is known that the fashion for the prose *entremés* predates that of the verse *entremés*. Should that be so, we could propose a new theory of the date of writing of Cervantes's eight *entremeses* based on their distribution in the four quadrants of a score chart.

Sixth. The study of the *Entremés de la cárcel de Sevilla* leads us to the following considerations:

- According to an eighteenth-century apograph of the comedy *La viuda, casada y doncella* by Lope de Vega (BNE, Ms. 22 423), this *entremés* was licensed for performance in 1600 by Tomás Gracián Dantisco. Its date of composition, therefore, is much earlier than that given by traditional critics (1610 – 1612), had been based exclusively on the presence of the character Escarramán [Asensio, 1971: 90 – 1].
- The coincidences between this *entremés* and the first two parts of the *Relación* by Cristóbal de Chaves show that not only had the author read these texts but also that he must have had them before him when composing his work. This is the only explanation for the textual coincidences that can be seen between them (Table 15). But apart from all these coincidences —typical of creative rewriting— the idiolect of Chaves is very different from that which appears in the disputed one (Figures 41, 45 and 46 and Tables 22 and 26).
- The different analyses carried out (lexical, morphological, syntactic, structural, etc.), the conclusions of which we have examined in detail in section 6.4, allow us to determine with a "fairly high" degree of probability (level 4 on a scale of 0 to 5) that Cervantes is indeed the author of this *entremés*.

Seventh. A detailed study of the *Entremés de los romances* leads us to the following considerations:

- The theme argument similarities between this piece and Don Quixote's first excursion (*Quixote*, I, 1-6) lead us to consider an inter-textual relationship between both writings. This relationship, however, is not direct, as no cases of *verbatim* are detected between both works (such as we saw between *La cárcel de Sevilla* and Chaves's *Relación*), apart from the citation of the same verses (with variants) from the *Romance del Marqués de Mantua*.
- Although the directionality of the copying between the two texts does not affect the question of authorship, contrary to the majority opinion we are of the

CONCLUSIONS

view that the *entremés* came later. Our arguments are based on (i) the presence of the squire; (ii) the lack of coherence in the characterisation of the speech of the sane characters; (iii) the use of versified form and dialogue romance, typical of the early seventeenth century; (iv) and the use of the synthetic formula "vuesastedes", rather than the earlier analytical use of "vuesa/s merced/es", which is what we find in *Don Quixote*.

- The study of the authorship of this *entremés* is quite a challenge, as the original text —the one resulting from eliminating the 32 fragments of well-known romances inserted in it— barely exceeds 1,150 words. The different analyses carried out (lexical, morphological, syntactic, structural, metrical, etc.), the conclusions of which we have set out in detail in section 7.4., allow us to determine with a "fairly high" degree of probability (level -1, on a scale from -5 to 0) that Cervantes is not its author.

It is true that in the literary puzzle of the Spanish Golden Age there are still many pieces to be fitted together and that, very probably, some of them belong to Cervantes. But it is no less true that the criteria by which around a hundred works have been attributed to him over the years (including *El cerco de Numancia* and *El trato de Argel*) are insufficient today. It seems inadmissible to us to continue to maintain the attribution of works to this author because of their literary quality, (i) since not all good texts were written by him, nor did he himself always write good texts or (ii) because of the similarity of the plot to any of his recognised works, since it is well known that the authors of our Baroque literature drew from a common well and used theme argument recasting in a large part of their creations. Therefore, although useful at the time and within its context, this entire *welter* of Cervantine affiliations has become completely obsolete, and needs to be reviewed and reworked from a 21st century perspective and in the light of the scientific methodological possibilities offered by forensic linguistics.

We hope that our work will contribute to this great endeavour; but, above all, that it will help lay the foundations of what literary authorship studies should and should not be.

SIGLAS

BC	Biblioteca Colombina.
BNE	Biblioteca Nacional de España.
CORDE	Corpus Diacrónico del Español.
GI	<i>General Imposters</i>
IAFL	<i>International Association of Forensic Linguistics</i> . En español, Asociación Internacional de Lingüistas forenses.
IAFLL	<i>International Association for Forensic and Legal Linguistics</i> . En español, Asociación Internacional de Lingüística legal y forense.
LR	<i>Likelihood Ratio</i> . En español, ratio de verosimilitud.
MFW	<i>Most Frequent Words</i> . En español, palabras más frecuentes.
PCA	<i>Principal Component Analysis</i> . En español, análisis de componentes principales.
POST	<i>Part-Of-Speech Tagging</i> o <i>Pos tagging</i> . En español, etiquetado gramatical.
RAE	Real Academia Española de la Lengua.
RTT	<i>Type/token ratio</i> . En español, ratio type/token.
Tf-idf:	<i>Term frequency - inverse document frequency</i> . En español, frecuencia de término - frecuencia inversa de documento.
OMPI	Organización Mundial de la Propiedad Intelectual (OMPI).

GLOSARIO

ANÁLISIS DE CONGLOMERADOS (*CLUSTER ANALYSIS*): Técnica estadística multivariante cuya finalidad es dividir un conjunto de objetos en grupos (*cluster*).

ANÁLISIS DE COMPONENTES PRINCIPALES (PCA): Técnica estadística de síntesis de la información que consiste en describir un determinado conjunto de datos en términos de nuevas variables (o componentes) no correlacionadas; las nuevas variables se ordenan por la cantidad de varianza original que describen, lo que permite reducir la dimensionalidad del conjunto [Jolliffe, 2002: 2].

BILOT: Tipo de gráfico exploratorio usado en estadística para representar tres o más variables en un espacio de dimensión reducida, normalmente de dimensión dos. Un biplot permite mostrar gráficamente la información de las filas (individuos, obras, etc.) y las columnas (variables) de una matriz de datos multivariantes.

BOOTSTRAP O BOOTSTRAPPING: Método de remuestreo, propuesto por Bradley Efron en 1979, que emplea un gran número de cálculos repetitivos para obtener la forma más estable de la distribución de las muestras.

CANONICALIZACIÓN: Conjunto de operaciones realizadas para evitar aquellos elementos que podrían distorsionar el análisis. Estas operaciones deben hacerse sistemáticamente, siguiendo las pautas establecidas por la comunidad científica especializada, y deben aplicarse a todos los textos del corpus por igual.

CHERRY PICKING: Término utilizado en investigación para hacer referencia a la falacia de evidencia incompleta o selectiva. Consiste, básicamente, en elegir los datos que apoyan nuestra hipótesis y desechar u ocultar aquellos que la cuestionan.

CLOSE READING / DISTANT READING: La lectura de cerca (*close reading*) es el método tradicional de análisis de textos literarios, basado en la lectura minuciosa y en la observación cualitativa de los datos. La lectura de lejos (*distant reading*), por el contrario, está relacionada con el análisis cuantitativo y consiste en la aplicación de métodos computacionales a los textos literarios.

COEFICIENTE DE CORRELACIÓN O COEFICIENTE PHI. Medida de asociación, comprendida entre el -1 y el +1, que permite medir la fuerza y la dirección de la relación lineal entre dos variables binarias. Ambas variables pueden intercambiarse sin que la correlación varíe.

COOCURRENCIA: Palabra que coaparece a una cierta distancia de otra, sin un orden exacto y estable.

CORPUS: “Conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etc., que pueden servir de base a una investigación” (*Diccionario de la lengua española*).

CORRELACIÓN: Término usado en probabilidad y estadística para indicar la fuerza y dirección de una relación lineal y la proporcionalidad entre dos variables estadísticas.

DENDROGRAMA: Representación visual, en forma de árbol (gr. δένδρον 'árbol'), que organiza los conjuntos de datos en subcategorías. Este tipo de representación permite evaluar la cohesión de los conglomerados o ramas que se forman.

DISTANCIA DELTA O DELTA DISTANCE: Algoritmo que parte de la idea de que la comparación en las frecuencias de las palabras más empleadas en un conjunto de textos permite llegar a conclusiones fiables en relación con su autoría, datación o género textual.

DOCUMENTO DUBITADO: Aquel sobre cuya veracidad existen dudas. En el contexto de los estudios de atribución de autoría, un texto es dubitado cuando no se sabe con certeza quién lo ha escrito.

DOCUMENTO INDUBITADO: Aquel sobre cuya veracidad no cabe duda alguna (Artículo 608 de la Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil). En el contexto de los estudios de atribución de autoría, un texto es indubitado cuando no existen dudas sobre quién lo ha escrito.

ESCALA DE OPINIÓN VERBAL: Representa los niveles de certeza en que puede expresarse la respuesta de una investigación, desde la probabilidad nula al grado máximo de probabilidad.

ESTILOMETRÍA: Disciplina que tiene por objeto la cuantificación, con ayuda de herramientas informáticas, de diferentes aspectos relacionados con el estilo de un determinado texto. La Estilometría se basa en el concurso de otras disciplinas auxiliares; sobre todo, de la Informática, la Estadística y la Filología.

HAPAX DISLEGOMENA: Número de palabras usadas dos veces en un texto, es decir, aquellas cuya presencia de aparición es igual a 2.

HAPAX LEGOMENA: Número de palabras usadas una sola vez en un texto, es decir, aquellas cuya presencia de aparición es igual a 1.

IDIOLECTO O ESTILO IDIOLECTAL: Variante lingüística característica de un determinado individuo.

ITEM O ÍTEM: “Cada uno de los artículos o capítulos de que consta una escritura u otro instrumento” (*Diccionario de la lengua española*). En nuestro trabajo, hablamos de “ítems” para referirnos al número de elementos que componen un *verbatim*.

LEY DE ZIPF: Esta ley, formulada en la década de 1930 por el lingüista norteamericano George Kingsley Zipf, hace referencia a la relación constante e inversa entre el número de veces que aparece una palabra en un texto y su posición relativa

(rango) en la lista de frecuencia de aparición. Según esta ley, la frecuencia de aparición de una palabra es proporcional al inverso de la posición que ocupa en dicha lista.

LINGÜÍSTICA DE CORPUS: Es una rama de la Lingüística que basa sus investigaciones en datos obtenidos a partir de un corpus, esto es, muestras reales de uso de la lengua, lo que permite “llevar a cabo investigaciones empíricas en contextos auténticos” [Parodi, 2008]. En rigor, el término no define una disciplina lingüística, como lo pueden ser la Morfología, la Sintaxis o la Pragmática, sino un enfoque metodológico que es posible adoptar desde disciplinas diversas, que se contraponen a una metodología basada fundamentalmente en la introspección.

LINGÜÍSTICA FORENSE: Puede definirse, de manera global, como la interfaz entre Lengua y Derecho. Es una ciencia multifacética e interdisciplinaria que pone en conexión los campos de la Lingüística y el Derecho con el fin de dar respuesta a las cuestiones legales que, de un modo u otro, implican la lengua.

MOST FREQUENT WORDS (MFW): Palabras más empleadas.

N-GRAMS O NGRAMAS: Secuencia de n elementos en un texto. Puede hacer referencia a una secuencia de fonemas, grafemas, caracteres, palabras, sintagmas...

PALABRAS FUNCIONALES, VACÍAS O SINSEMÁNTICAS: Son aquellas con significado gramatical y sin contenido léxico; no son significativas por sí mismas, por lo que se utilizan para indicar la relación existente entre las unidades léxicas.

PALABRAS LEXICALES, LLENAS, PLENAS, SEMÁNTICAS O DE CONTENIDO: Son aquellas que poseen significado léxico y contenido conceptual; son significativas en sí mismas y representan de manera inmediata la realidad extralingüística, por lo que pueden aparecer aisladamente en el discurso.

PART-OF-SPEECH TAGGER (POST o POS TAGGING): Secuencia de etiquetas gramaticales.

RAZÓN DE VEROSIMILITUD O LIKELIHOOD RATIO (LR): Indicador estadístico que permite cuantificar la validez o certeza de un resultado entre dos hipótesis excluyentes.

RATIO TYPE/TOKEN (TTR): Relación que existe entre el número de palabras distintas de un texto (*types*) y el número total de palabras que lo conforman (*tokens*). Sirve para calcular el nivel de densidad léxica de un texto: si contiene muchas palabras distintas que se repiten pocas veces, será alto, mientras que si contiene pocas palabras que se repiten muchas veces, será bajo. Véanse *TOKENS* y *TYPES*.

STOPWORD LIST O STOP LIST: Lista de palabras funcionales (sin significado léxico) que deben usarse para calcular variables como la riqueza léxica, las ratios *type/token* de las palabras de contenido o el nivel de similitud textual entre distintos pares textuales.

TOKENS: Número total de palabras de un texto. En español, *formas*.

TYPES: Número total de palabras distintas de un texto. En español, *tipos*.

VARIABLES Y VARIANTES LINGÜÍSTICAS: Una variable es el conjunto de manifestaciones o expresiones de un mismo elemento, y cada una de esas manifestaciones o expresiones recibe el nombre de variante lingüística [Moreno Fernández, 1998: 18]. Ejemplos de variantes lingüísticas en español pueden ser las distintas realizaciones del fonema /s/ implosivo ([s], [h] o [Ø]); la elección del morfema *-ra* o *-se* en el imperfecto de subjuntivo (*cantara* o *cantase*); la presencia o ausencia del sujeto pronominal; o la manera de escribir los años (2020 o 2.020).

VARIANZA: Medida de dispersión que representa la variabilidad de una serie de datos respecto de la media aritmética de los mismos. Si se comparan datos de un mismo tipo, una varianza elevada significaría que los datos están más dispersos, mientras que una varianza baja indicaría que los valores están, por lo general, más próximos a la media. Una varianza cero supondría que todos los datos coinciden con la media, es decir, que todos los datos son iguales.

GLOSARIO

VERBATIM: Adverbio latino que significa ‘textualmente’, es decir, ‘palabra a palabra’. Es la reproducción exacta de una determinada secuencia de palabras, que deben aparecer en el mismo orden, sin paráfrasis, sustitución o abreviación de ningún tipo.

BIBLIOGRAFÍA

- ABBASI, Ahmen y Hsinchun CHEN (2008), “Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace”, *ACM Transactions on Information Systems*, 2008, 26 (2), pp. 1-29.
- AGOSTINI DE DEL RÍO, Amelia (1964), “El teatro cómico de Cervantes (continuará)”, *Boletín de la Real Academia Española (BRAE)*, 44 (173), pp. 475-540 y 475-539.
- AGOSTINI DE DEL RÍO, Amelia (1965), “El teatro cómico de Cervantes (conclusión)”, *Boletín de la Real Academia Española (BRAE)*, 45 (174-175), pp. 65-116.
- ALBOREDA, Carlos A. (1991), *Teoría y formas del metateatro en Cervantes*, Salamanca, Universidad de Salamanca.
- ALCALÁ GALÁN, Mercedes (2002), “«Dios te dé salud y a mí paciencia»: teoría del teatro en Cervantes”, en José Manuel Martín Morán (ed.), *La media semana del jardincito. Cervantes y la reescritura de los códigos*, Padova, Unipress, pp. 255-275.
- ALONSO CORTÉS, Narciso (1903), *Un pleito de Lope de Rueda. Nuevas noticias para su biografía*, Valladolid, Imprenta de Juan R. Hernando.
- ALONSO CORTÉS, Narciso (1917a), “Lope de Rueda en Valladolid”, *Boletín de la Real Academia Española (BRAE)*, 3, pp. 219-220.
- ALONSO CORTÉS, Narciso (1917b), “El teatro en Valladolid”, *Boletín de la Real Academia Española (BRAE)*, 4, pp. 598-611.

BIBLIOGRAFÍA

- ALONSO CORTÉS, Narciso (2005), *Volumen cervantino*, Valladolid, Ayuntamiento de Valladolid.
- ALONSO, Dámaso (1936), *El hospital de los podridos y otros entremeses alguna vez atribuidos a Cervantes*, Madrid, Signo.
- ALVAR EZQUERRA, Alfredo (2014), *Un maestro en tiempos de Felipe II: Juan López de Hoyos y la enseñanza humanista del siglo XVI*, Madrid, La esfera de libros.
- ALVAR EZQUERRA, Carlos, Florencio SEVILLA ARROYO y Alfredo ALVAR EZQUERRA (dirs.) (2005), *Gran enciclopedia cervantina*, Madrid, Castalia.
- ÁLVAREZ RAMOS, Eva y Javier BLASCO PASCUAL (coords.) (2018), *Humanidades digitales. Retos, recursos y nuevas propuestas*, Valladolid, Agilice Digital.
- ANÓNIMO (1820), “Propiedad literaria”, *El censor. Periódico político y literario*, 15, pp. 223-231.
- ANTHONY, Laurence (2013), “A critical look at software tools in Corpus Linguistics”, *Linguistic Research*, 30 (2), pp. 141-161.
- ARATA, Stefano (1992), “La Conquista de Jerusalén, Cervantes y la generación teatral de 1580”, *Criticón*, 54, pp. 9-112.
- ARATA, Stefano (1997), “Notas sobre *La Conquista de Jerusalén* y la transmisión manuscrita del primer teatro cervantino”, *Edad de oro*, 16, pp. 53-66.
- ARCO Y GARAY, Ricardo del (1951), “Cervantes y la farándula”, *Boletín de la Real Academia Española (BRAE)*, 31 (133), pp. 311-330.
- ARELLANO-TORRES, Ignacio D. (2018), “El entremés de *Los habladores*, atribuido a Cervantes”, *Anales cervantinos*, 50, pp. 299-323.
- ARELLANO, Ignacio (2002), *Comedias burlescas del Siglo de Oro. Tomo III*, Madrid, Iberoamericana / Frankfurt am Main, Vervuert.
- ARELLANO, Ignacio y Celsa Carmen GARCÍA-VALDÉS (1997), “El Entremés *el marido fantasma*, de Quevedo”, *La Perinola*, 1, pp. 41-68.

BIBLIOGRAFÍA

- ARGAMON, Shlomo (2008), "Interpreting Burrows's Delta: Geometric and probabilistic foundations", *Literary and Linguistic Computing*, 23 (2), pp. 131-147.
- ARGAMON, Shlomo y Shlomo LEVITAN (2005), "Measuring the usefulness of function words for authorship attribution", en *Proceedings of the 2005 ACH/ALLC Conference*, Victoria, BC, pp. 1-3.
- Ariza Viguera, Manuel (1998), "El sufijo -ICO", en Claudio García Turza, Fabián González Bachiller y José Javier Mangado Martínez (coords), *Actas del IV Congreso Internacional de Historia de la Lengua Española: La Rioja, 1-5 de abril de 1997*, Logroño, Universidad de La Rioja, vol. I, pp. 351-359.
- ASENSIO, Eugenio (1970), "Introducción crítica: Los entremeses de Cervantes", en Miguel de Cervantes, *Entremeses*, Madrid, Castalia, pp. 4-49.
- ASENSIO, Eugenio (1971), *Itinerario del entremés, desde Lope de Rueda a Quiñones de Benavente con cinco entremeses inéditos de D. Francisco de Quevedo*, Madrid, Gredos, 2ª edición revisada.
- ASENSIO, Eugenio (1973), "Entremeses", en J. B. Avallé-Arce y E.C. Riley (eds.), *Suma cervantina*, London, Tamesis books, pp. 171-197.
- ASENSIO, José María (1874), "Sol y sombra. Cartas á los insignes cervantistas, D. José de Palacio Vitery y D. Mariano Pardo de Figueroa, sobre Asuntos y Zarandajas de crónica escandalosa cervantina", *Revista Europea*, año I, tomo II, 35, pp. 555-562.
- ASENSIO, José María (1902), *Cervantes y sus obras*. Barcelona, Seix Editor.
- ASTRANA MARÍN, Luis (1948-1958), "Escritos probables, atribuidos, dudosos, apócrifos y falsos", en *Vida ejemplar y heroica de Miguel de Cervantes Saavedra*, Madrid, Instituto Editorial Reus, vol. VII, pp. 751-767.
- AVALLE-ARCE, Juan Bautista (1973), "Atribuciones y supercherías", en J. B. Avallé-Arce y E. C. Riley (eds.), *Suma cervantina*, Londres, Tamesis, pp. 399-408.
- AVELLANEDA, Alonso Fernández de (2000), *El ingenioso hidalgo Don Quijote de la Mancha*, ed. de Luis Gómez Canseco, Madrid, Biblioteca Nueva.

BIBLIOGRAFÍA

- AYALA, Francisco (1983), “Cervantes no sólo escribió el *Quijote*”, en *Palabras y plumas*, Edhasa, Barcelona, pp. 234-136.
- AYLWARD, Edward T. (1982), *Cervantes: Pioneer and Plagiarist*, Londres, Tamesis.
- AZCUE CASTILLÓN, Verónica (2002), “La disputa del baciuelmo y «El retablo de las maravillas»: sobre el carácter dramático de los capítulos 44 y 45 de la primera parte de don Quijote”, *Cervantes: Bulletin of the Cervantes Society of America*, 22 (1), pp. 71-81.
- AZCUNE, Valentín y Manuel FERNÁNDEZ NIETO (2004), “Cervantes no imitó el entremés de los romances”, *Cuadernos para investigación de la literatura hispánica*, 29, pp. 103-118.
- BAAYEN, R. Harald (2009), *Analyzing Linguistic Data*, Cambridge, Cambridge University Press.
- BAAYEN, R. Harald, Hans VAN HALTEREN y Fiona J. TWEEDIE (1996), “Outside the cave of shadows: Using syntactic annotation to enhance authorship attribution”, *Literary and Linguistic Computing*, 11 (3), pp. 121-131.
- BAILEY, Richard W. (1979), “Authorship attribution in a forensic setting”, en D. E. Ager, F. E. Knowles and J. Smith (eds.), *Advances in Computer-aided Literary and Linguistic Research*, Birmingham, AMLC, pp.1-20.
- BALDWIN, John (1979), “Phonetics and speaker identification”, *Medicine, Science and the Law*, 19, pp. 231-232.
- BANKS, J. (ed.) (1856), *The Works of Hesiod, Callimachus and Theognis*, London, Henry G. Bohn.
- BAQUERO ESCUDERO, Ana Luisa (2005), “Las novelas sueltas, pegadizas y pegadas en el *Quijote*”, en Eva Reichenberger y Kurt Reichenberger (eds.), *Cervantes y su mundo II*, Kassel, Reichenberger, pp. 23-52.
- BARAS ESCOLÁ, Alfredo (1993), “El *Entremés de los romances* y la novela corta del *Quijote*”, en *Actas del III Coloquio Internacional de la Asociación de*

BIBLIOGRAFÍA

- Cervantistas: Alcalá de Henares, 12-16 de noviembre, 1990*, Barcelona, Anthropos / Madrid, Ministerio de Asuntos Exteriores, pp. 331-335.
- BARRERA Y LEIRADO, Cayetano Alberto de la (1860), *Catálogo bibliográfico y biográfico del teatro antiguo español, desde sus orígenes hasta mediados del siglo XVIII*, Madrid, Rivadeneyra.
- BARTHES, Roland (1968), “La mort de l’auteur”, *Manteia*, 5, pp. 12-17.
- BARTHES, Roland (1977), *Elements of Semiology*, New York, Hill and Wang.
- BENNETT, Andrew (2005), *The Author. The New critical Idiom*, London and New York, Routledge.
- BERGAMIN, José (1959), *Lázaro, Don Juan y Segismundo*, Madrid, Espasa-Calpe.
- BIBER, Douglas (1993), “Representativeness in Corpus Design”, *Literary and Linguistic Computing*, 8 (4), pp. 243-257.
- BINGHAM KIRBY, Carol (1986), “La verdadera edición crítica de un texto dramático del Siglo de Oro: teoría, metodología y aplicación”, *Incipit*, 6, pp. 71-99.
- BLASCO, Francisco Javier (2005a), *Baltasar Navarrete, posible autor del “Quijote” apócrifo*, Segovia, Instituto Castellano y Leonés de la Lengua,
- BLASCO, Javier (2005b), *Miguel de Cervantes Saavedra: regocijo de las musas*, Valladolid, Universidad de Valladolid / Secretariado de Publicaciones e Intercambio Editorial.
- BLASCO, Javier (2006), *Cervantes. Un hombre que escribe*, Valladolid, Difácil.
- BLASCO, Javier (2010), “La cuestionada autoría del *Diálogo entre Cilenia y Selanio*”, en Javier Blasco, Patricia Marín Cepeda y Cristina Ruiz Urbón (eds.). *Hos ergo versículos feci... Estudios de atribución y plagio*, Madrid, Iberoamericana, 2010, pp. 19-74.
- BLASCO, Javier (2013), “Más allá del romancero: *Entremés de los romances*”, *Edad de oro*, 32, pp. 31-46.

BIBLIOGRAFÍA

- BLASCO, Javier (2016), “Avellaneda desde la estilometría”, en Pedro Ruiz Pérez (ed.), *Cervantes: los viajes y los días*, Madrid, Sial Ediciones, pp. 97-116.
- BLASCO, Javier (2019a), “Atribuciones cervantinas desde la estilometría. *El entremés de Los mirones*”, en Guillermo Laín Corona, Rocío Santiago Nogales y José Romera Castillo (coords.), *Cartografía teatral en homenaje al profesor José Romera Castillo*, Madrid, Visor Libros, pp. 151-168.
- BLASCO, Javier (2019b), “La graciosa y gratuita disputa sobre la autoría de la *Historia verdadera del inconfundible Bernal Díaz del Castillo*”, *Boletín de la Real Academia Española (BRAE)*, 99 (319), pp. 5-44.
- BLASCO, Javier y Cristina RUIZ URBÓN (2009), “Evaluación y cuantificación de algunas técnicas de ‘Atribución de autoría’ en textos españoles”, *Castilla. Estudios de literatura*, 0, pp. 27-47.
- BLECUA, José Manuel (1979), *Lope de Rueda, Pasos. Miguel de Cervantes, Entremeses*, Zaragoza, Ebro.
- BLECUA, José Manuel (2004), “El *Quijote* en la historia de la lengua española”, en Miguel de Cervantes, *Don Quijote de la Mancha*, Madrid, Real Academia Española / Asociación de academias de la lengua española, pp. 1115-1122.
- BONDÍA ROMÁN, Fernando (1988), *La Propiedad intelectual: su significado en la sociedad de la información (la nueva ley de 11 de noviembre de 1987)*, Madrid, Trivium.
- BONILLA Y SAN MARTÍN, Adolfo (1916), “Introducción”, en Miguel de Cervantes, *Entremeses*, Madrid, Asociación de la librería de España, pp. 1-39.
- BOURDIEU, Pierre (1989-1990), “El campo literario. Prerrequisitos críticos y principios de método”, *Criterios*, 25-28, pp. 20-42.
- BOURDIEU, Pierre (1995), *Las reglas del arte. Génesis y estructura del campo literario*, Barcelona, Anagrama.
- BREUIL, Henri y Lilo BERGER-KIRCHNER (1962), “Arte rupestre Franco-cantábrico”, en *El arte de los pueblos. Edad de Piedra*, Barcelona, Praxis/Seix Barral, pp. 11-70.

BIBLIOGRAFÍA

- BRIOSO SANTOS, Héctor (1996), “Vélez de Guevara y la sátira barroca: el tema de los encochados”, en Piedad Bolaños Donoso y Marina Martín Ojeda (eds.), *Luis Vélez de Guevara y su época. IV Congreso de Historia de Écija (Écija, 20-23 de octubre de 1994)*, Sevilla, Fundación El Monte, pp. 227-236.
- BRIOSO SANTOS, Héctor (coord.) (2007), *Cervantes y el mundo del teatro*, Kassel, Reichenberger.
- BRIOSO SANTOS, Héctor (2018), “Cervantes frente a la Comedia nueva: *El retablo de las maravillas* y las maravillas del nuevo teatro”, *eHumanista*, 38, pp. 723-745.
- BROEDERS, A. P. A. (1999), “Some observations on the use of probability scales in forensic identification”, *Forensic Linguistics*, 6 (2), pp. 228-41.
- BROWN, Roger (1982), “What is speaker recognition”, *Journal of the International Phonetics Association*, 12 (1), pp. 13-24.
- BURRIDGE, Kate y Jean MULDER (1988), *English in Australia and New Zealand: An introduction to its structure, history and use*, Melbourne, Oxford University Press.
- BURROWS, John Frederick (1989), “«An ocean where each kind»: Statistical analysis and some major determinants of literary style,” *Computers and the Humanities*, 23 (4-5), pp. 309-21.
- BURROWS, John Frederick (2002), “«Delta»: A measure of stylistic difference and a guide to likely authorship”, *Literary and Linguistic Computing*, 17 (3), pp. 267-287.
- BURROWS, John Frederick (2003), “Questions of Authorship: Attribution and Beyond”, *Computers and the Humanities*, 37 (1), pp. 5-32.
- BURROWS, John Frederick (2007), “All the Way Through: Testing for Authorship in Different Frequency Strata”, *Literary and Linguistic Computing*, 22 (1), pp. 27-47.
- BURROWS, John Frederick y Anthony J. HASSALL (1988), “Anna Boleyn and the authenticity of Fielding’s feminine narratives”, *Eighteenth Century Studies*, 21, pp. 427-53.

BIBLIOGRAFÍA

- CALVO TELLO, José (2016), “Entendiendo Delta desde las Humanidades”, *Caracteres. Estudios culturales y críticos de la esfera digital*, 5 (1), pp. 140-176.
- CALVO TELLO, José (2019), “Diseño de corpus literario para análisis cuantitativos”, *Revista de Humanidades Digitales*, 4, pp. 115-135.
- CAMPILLO, Antonio (1992), "El autor, la ficción, la verdad", *Daimon*, 5, pp. 25-46
- CAMPIÓN LARUMBE, Miguel y Álvaro CUÉLLAR (2021), “Discernir entre original y refundición en el teatro del Siglo de Oro a través de la estilometría: el caso de *El mejor amigo, el muerto*”, *Talía. Revista de estudios teatrales*, 3, pp. 59-69.
- CANAVAGGIO, Jean (1977), *Cervantès dramaturge. Un théâtre à naître*, Paris, Presses Universitaires de France.
- CANAVAGGIO, Jean (2000), *Miguel de Cervantes Saavedra entre vida y creación*, Alcalá de Henares, Centro de Estudios Cervantinos.
- CANAVAGGIO, Jean (2003), *Cervantes*, Madrid, Espasa-Calpe.
- CANAVAGGIO, Jean (2014), “*Cervantès dramaturge, veinticinco años después*”, en *Retornos a Cervantes*, New York, Idea, pp. 121-135.
- CANTER, David y Joanne CHESTER (1997), “Investigation into the claim of weighted Cusum in authorship attribution studies”, *Forensic Linguistics. The International Journal of Speech, language and the Law*, 4 (2), pp. 252-261.
- CAPARRÓS, José Domínguez (2002), *Métrica de Cervantes*, Alcalá de Henares, Centro de Estudios Cervantinos.
- CASALDUERO, Joaquín (1951), *Sentido y forma del teatro de Cervantes*, Madrid, Aguilar.
- CASCÓN, Miguel (1947), “Los jesuitas según Cervantes”, *Boletín de la Biblioteca de Menéndez Pelayo*, 23, pp. 179-211.
- CÁSEDA TERESA, Jesús Fernando (2019), “En tiempos de Lope de Vega: El memorió Luis Ramírez de Arellano”, *Studia Aurea*, 13, pp. 297-317.
- CASTRO DÍAZ, Antonio (1983), “Acerca de una nueva edición de los entremeses de Cervantes”, *Cuadernos Hispanoamericanos*, 339, pp. 137-142.

BIBLIOGRAFÍA

- CASTRO, Adolfo de (1874), *Varias obras inéditas de Cervantes*, Madrid, A. de Castro e hijos editores.
- CASTRO, Adolfo de (1957), *Cuatro entremeses atribuidos a Miguel de Cervantes*, Barcelona, Gráficas Aymamí.
- CASTRO, Américo (1969), “La palabra escrita y el *Quijote*”, en *Hacia Cervantes*, Madrid, Taurus, pp. 292-324.
- CASTRO, Edgardo (2018), *Diccionario Foucault: Temas, conceptos y autores*, Buenos Aires, Siglo XXI.
- CELMA VALERO, María Pilar y Cristina RUIZ URBÓN (2021), “Una aproximación a la escritura de Miguel Delibes desde la Estilometría”, *Tonos digital*, 41, pp. 1-13.
- CEREZO SOLER, Juan y José CALVO TELLO (2019), “Autoría y estilo. Una atribución cervantina desde las humanidades digitales. El caso de *La conquista de Jerusalén*”, *Anales Cervantinos*, 51, pp. 231-250.
- CERVANTES, Miguel de (1914-1941), *Obras completas*, edición de Rudolph Schevill y Adolfo Bonilla y San Martín, Madrid, Gráficas reunidas.
- CERVANTES, Miguel de (1948), *Entremeses*, edición, prólogo y notas de Agustín del Campo, Madrid, Clásicos Castilla.
- CERVANTES, Miguel de (1970), *Entremeses*, edición de Eugenio Asensio, Madrid, Castalia.
- CERVANTES, Miguel de (1988), *Don Quijote de la Mancha*, ed. de Francisco Rico, Barcelona, Instituto Cervantes / Editorial Crítica.
- CERVANTES, Miguel de (1998), *Obras completas*, ed. de Florencio Sevilla Arroyo, Madrid, Castalia.
- CERVANTES, Miguel de (1999), *Obras completas*, ed. de Florencio Sevilla Arroyo y Antonio Rey Hazas, Madrid, Alianza.
- CERVANTES, Miguel de (2012), *Entremeses*, edición, estudio y notas de Alfredo Baras Escolá, Madrid, Real Academia Española / Barcelona, Galaxia Gutenberg y Círculo de Lectores.

BIBLIOGRAFÍA

- CERVANTES, Miguel de (2018), *La tía fingida*, edición de Adrián J. Sáez, Madrid, Cátedra.
- CERVANTES, Miguel de (2020), *Entremeses*, edición, introducción y notas de Adrián J. Sáez, Madrid, Cátedra.
- CÉSAR ÁLVAREZ, José (2005), *La disputada cuna de Cervantes*, Alcalá de Henares, Ediciones Bornova.
- CHAMPOD, Christophe e Ian W. EVETT (2000), “Commentary on A. P. A. Broeders (1999) ‘Some observations on the use of probability scales in forensic identification’”, *Forensic Linguistics* 6 (2): 228-41”, *Forensic Linguistics*, 7 (2), pp. 238-234.
- CHASKI, Carole Elisabeth (1996), “Linguistic methods of determining authorship”, en *National Institute of Justice Research Seminar. 49th American Academy of Forensic Sciences Meeting*, Nashville, TN.
- CHASKI, Carole Elisabeth (2001), “Empirical evaluations of language-based author identification techniques”, *Forensic Linguistics. The International Journal of Speech, Language and the Law*, 8 (1), pp. 26-43.
- CHAUCHADIS, Claude (1980), “Risa y honra conyugal en los entremeses”, en *Risa y sociedad en el teatro español del siglo de Oro. Actes du 3^e colloque du Groupe d’Estudes Sur le Théâtre Espagnol, Toulouse 31 janvier – 2 février 1980*, Toulouse / Paris, C.N.R.S., pp. 165-185.
- CHEVALIER, Maxime (1976), *Lectura y lectores en la España del siglo XVI y XVII*, Madrid, ediciones Turner.
- CICRES, Jordi (2014), “Comparación forense de voces mediante el análisis multidimensional de las pausas llenas”, *Revista Signos. Estudios de Lingüística*, 47 (86), pp. 365-384.
- CICRES, Jordi y M. Teresa TURELL (2010), "Investigación, docencia y práctica profesional en Lingüística Forense", en E. Garayzábal Heinze, M. Jiménez Bernal y M. Reigosa Riveiros (eds.), *Panorama actual de la Lingüística Forense en el*

BIBLIOGRAFÍA

ámbito legal y policial. Teoría y práctica. Jornadas (in)formativas de Lingüística Forense, Madrid, Euphoria Ediciones, S.L., pp. 32-44.

CICRES, Jordi y NURIA GAVALDÀ (2014), “La Lingüística Forense: La Llengua com a Evidència”, *Revista de Llengua i Dret*, núm. 61, pp. 60-71.

CLARKE, Dorothy Clotelle (1952), *A chronological sketch of Castilian versification together with a list of its metric terms*, Berkeley and Los Ángeles, University of California Press.

CORREAS, Gonzalo (1924), *Vocabulario de refranes y frases proverbiales y otras fórmulas comunes de la lengua castellana*, Madrid, Tipografía de la Revista de Archivos, Bibliotecas y Museos.

CORTIJO OCAÑA, Antonio (2006), “El amor como burla y la parodia erótica en la comedia burlesca áurea. A propósito de *El amor más verdadero*, *Durandarte* y *Belerma* de Guillén Pierres”, en J. Ignacio Díez y Adrienne L. Martín (eds.), *Venus venerada. Tradiciones eróticas de la literatura española*, Madrid, Editorial Complutense, pp. 165-189.

CORTIJO OCAÑA, Antonio y Adelaida CORTIJO OCAÑA (2002), “Introducción” a la edición de *El amor más verdadero*, *Durandarte* y *Belerma*, en I. Arellano (dir.), *Comedias burlescas del Siglo de Oro*, Madrid, Iberoamericana / Frankfurt am Main, Vervuert, tomo III, pp. 73-92.

CORTIJO OCAÑA, Antonio y Adelaida CORTIJO OCAÑA (2006), “*Sierra Morena de las mujeres*. Un entremés de Luis de Belmonte Bermúdez”, *Dicenda. Cuadernos de Filología Hispánica*, 24, pp. 61-80.

COTARELO VALLEDOR, Armando (1915), *El teatro de Cervantes*, Madrid, Tipografía de la Revista de Archivos, Bibliotecas y Museos.

COTARELO VALLEDOR, Armando (1943), *Cervantes lector*, Madrid, Publicaciones del Instituto de España.

COTARELO VALLEDOR, Armando (1947-1948), “Obras perdidas de Cervantes que no se han perdido”, *Boletín de la Real Academia Española (BRAE)*, 27, pp. 61-77.

BIBLIOGRAFÍA

- COTARELO Y MORI, Emilio (1901), *Lope de Rueda y el teatro de su tiempo*, Madrid, Imprenta de “La revista española”.
- COTARELO Y MORI, Emilio (1904), *Bibliografía de las controversias sobre la licitud del teatro en España*, Madrid, Revista de Archivos, Bibliotecas y Museos.
- COTARELO Y MORI, Emilio (1911), *Colección de entremeses. Loas, bailes, jácaras y mojigangas desde fines del siglo XVI á mediados del XVIII*, Madrid, Casa Bailly Bailliére, tomo I, vol. I.
- COTARELO Y MORI, Emilio (1920), “Sobre los orígenes y formación del *Quijote*”, en *Últimos estudios cervantinos*, Madrid, Tipografía de la Revista de Archivos, Bibliotecas y Museos, pp. 45-66.
- COUDERC, Christophe (2009), “El autor ante la edición de sus obras. Los prólogos de las Partes de comedias”, en Soledad Arredondo Sirodey, Pierre Civil y Michel Moner, *Paratextos en la literatura española (siglos XV-XVIII)*, Madrid, Casa de Velázquez, pp.119-134.
- COULTHARD, Malcolm (1994), “On the use of corpora in the analysis of forensic texts”, *Forensic linguistics: The International Journal of Speech Language and the Law*, 1 (1), pp. 26-43.
- COULTHARD, Malcolm (1998), “Identifying the Author”, *Cahiers de Linguistique Française*, 20, pp. 139-161.
- COULTHARD, Malcolm (2004), “Author identification, idiolect and linguistic uniqueness”, *Applied Linguistics*, 25 (4): pp. 431-447.
- COULTHARD, Malcolm (2005a), “Algunas aplicaciones forenses de la Lingüística descriptiva”, en M. Teresa Turell (ed.), *Lingüística forense, lengua y derecho. Conceptos, métodos y aplicaciones*, Barcelona, Institut Universitari de Lingüística Aplicada, pp. 249-266.
- COULTHARD, Malcolm (2005b), “The Linguistic as Expert Witness”, *Linguistics and the Human Sciences*, 1 (1), pp. 39-58.

BIBLIOGRAFÍA

- COULTHARD, Malcolm (2005c), "Some forensic applications of descriptive linguistics", *Veredas*, 9, pp. 9-28.
- COULTHARD, Malcolm (2006), "...and then... Language description and author attribution". Disponible en https://www2.aston.ac.uk/migrated-assets/applicationpdf/lss/10186-Andthen_Coulthard.pdf [Fecha de consulta: 08.12.2021].
- COULTHARD, Malcolm y Alison JOHNSON (2010), *The Routledge Handbook of Forensic Linguistics*, London / New York, Routledge.
- COULTHARD, Malcolm, Alison JOHNSON y David WRIGHT (2017), *An introduction to Forensic Linguistics: Language in Evidence*, London / New York, Routledge.
- CRAIG, Hugh y Arthur F. KINNEY, *Shakespeare, Computers, and the Mystery of Authorship*, Cambridge University Press, 2009.
- CRUICKSHANK, Don W. (2000), "Los «hurtos de la prensa» en las obras dramáticas", en Francisco Rico (coord.), *Imprenta y crítica textual en el Siglo de Oro*, Valladolid, Universidad de Valladolid, pp. 129-150.
- CRUZ CASADO, Antonio (2005), "Miguel de Cervantes, 'natural de la ciudad de Córdoba'. El proceso del cómico Tomás Gutiérrez contra la Cofradía del Santísimo Sacramento (Sevilla, 1593)", *Boletín de la Real Academia de Córdoba de Ciencias, Bellas Letras y Nobles Artes*, 84 (149), pp. 145-180.
- CRYSTAL, David, (2008), *'Think on my Words' Exploring Shakespeare's Language*, Cambridge, Cambridge University Press, 2008.
- CUARTERO SÁNCHEZ, Juan Manuel (2002), "«Significado léxico» y «Significado gramatical» en las gramáticas de español moderno", *Boletín de la Sociedad Española de Historiografía Lingüística*, 3, pp. 43-78.
- CUENCA MUÑOZ, Paloma (1999), "Estudio paleográfico de algunos autógrafos teatrales de Lope de Vega y edición del entremés de *Melisendra* (Res: 88)", *Cuadernos para investigación de la literatura hispánica*, pp. 149-194.

BIBLIOGRAFÍA

- DE PERETTI DELLA ROCCA, Cristina (1989), *Jacques Derrida: texto y deconstrucción*, Barcelona Anthropos.
- DELGADO ROMERO, Carlos (2001), *La identificación de locutores en el ámbito forense*, tesis doctoral, Universidad Complutense de Madrid. Disponible en <https://eprints.ucm.es/id/eprint/4488/> [Fecha de consulta: 08.12.2021].
- DELGADO ROMERO, Carlos (2005), “Comentarios sobre el contexto actual de la identificación forense de locutores”, en M. Teresa Turell (ed.), *Lingüística forense. lengua y derecho. Conceptos, métodos y aplicaciones*, Barcelona, Institut Universitari de Lingüística Aplicada, pp. 113-129.
- DEMATTE, Claudia (2021), “El Segundo tomo de comedias de Juan Pérez de Montalbán: hacia un mapa estilométrico de las comedias auténticas, dudosas y ajenas”, *Talía. Revista de estudios teatrales*, 3, pp. 71-78.
- DI PINTO, Elena (2005), *La tradición escarramanesca en el teatro del Siglo de Oro*, Madrid-Frankfurt Am Main, Iberoamericana-Vervuert.
- DÍAZ DE ESCOVAR, Narciso (1905), *Apuntes escénicos cervantinos: ó sea, Un estudio histórico, bibliográfico y biográfico de las comedias y entremeses escritos por Miguel de Cervantes Saavedra, con varias de sus opiniones sobre las comedias y los cómicos y noticias de los comediantes que debió conocer ó mencionó en sus libros inmortales*, Madrid, Viuda de Rico.
- DÍAZ PLAJA, Guillermo (1963a), “El *Quijote* como situación teatral”, en *Cuestión de límites. Cuatro ejemplos de estéticas fronterizas (Cervantes, Velázquez, Goya el cine)*, Madrid, Revista de Occidente, pp. 13-136.
- DÍAZ PLAJA, Guillermo (1963b), “El retablo de maese Pedro”, *Ínsula*, 18 (204), pp. 1-12.
- DÍAZ PLAJA, Guillermo (1968), *Historia general de las literaturas hispánicas. Pre-Renacimiento y Renacimiento*, Barcelona, Editorial Vergara, tomo II.
- DIDEROT, Denis (1997), *Carta sobre el comercio de librería*, Salamanca, Junta de Castilla y León / CEGAL.

BIBLIOGRAFÍA

- DÍEZ BORQUE, José María (1978), *Sociedad y teatro en la España de Lope de Vega*, Barcelona, Bosch.
- DÍEZ BORQUE, José María (2007), “Cervantes y la vida teatral en el Siglo de Oro”, en *En torno al Teatro del Siglo de Oro. Jornadas XXI-XXIII. 2004, 2005 y 2006 (Almería)*, coord. Antonio Serrano, Almería, Instituto de Estudios Almerienses / Diputación de Almería, pp. 163-184.
- DÍEZ BORQUE, José María (2016), «Prólogo», en Ignacio Arellano y Gonzalo Santonja, Nueva York, Idea, pp. 9-16.
- DOCK, Marie-Claude (1974), “Génesis y evolución de la noción de propiedad literaria”, traducido al español por Juana Martínez-Arretz, *Revue Internationale du Droit d’Auteur (RIDA)*, 79, pp. 130-154.
- DOMÍNGUEZ CAPARRÓS, José (2002), *Métrica de Cervantes*, Alcalá de Henares, Centro de Estudios Cervantinos.
- EDER, Maciej (2013a), “Computational stylistics and Biblical translation: How reliable can a dendrogram be”, en Tadeusz Piotrowski y Łukasz Grabowski (eds.), *The Translator and the Computer*, Wrocław, WSF Press.
- EDER, Maciej (2013b), “Bootstrapping Delta: A safety net in open-set authorship attribution”, *Digital Humanities 2013. Conference abstracts. University of Nebraska/Lincoln, USA. 16-19 July 2013*, pp. 169-172.
- EDER, Maciej (2016), “Rolling stylometry”, *Digital Scholarship in the Humanities*, 31 (3), pp. 457-469.
- EDER, Maciej (2017), “Visualization in stylometry: Cluster analysis using networks. Digital”, *Scholarship in the Humanities*, 32 (1), pp. 50-64.
- EDER, Maciej (2018), “Elena Ferrante: A virtual author”, en Arjuna Tuzzi y Michele A. Cortelazzo (eds.), *Drawing Elena Ferrante profile*, Padydova, Padova University Press, pp. 31-45.
- EDER, Maciej, Mike KESTEMONT y Jan RYBICKI (2016), “Stylometry with R: A Package for Computational Text Analysis”, *The R Journal*, 8 (1), pp. 107-121.

BIBLIOGRAFÍA

- EISENBERG, Daniel (1987), “La biblioteca de Cervantes”, en *Studia in Honorem prof. Martín de Riquer*, II, Barcelona, Quaderns Crema, pp. 271-328.
- EISENBERG, Daniel (1990), “Repaso crítico de las atribuciones cervantinas”, *Nueva Revista de Filología Hispánica*, 38 (2), pp. 477-492.
- EISENBERG, Daniel (1991), *Estudios cervantinos*, Barcelona, Sirmio.
- EISENBERG, Daniel (1994), “Cervantes, autor de la *Topografía e Historia de Argel* publicada por D. de Haedo”, *Aljamía. Revista de la Consejería de Educación en Marruecos*, 6, 1994, pp. 19-27.
- ESPÍN CÁNOVAS, Diego (1995), *Fundamento y protección del derecho de autor en las obras de arte, real academia de Jurisprudencia y Legislación*, Madrid, Editorial Civitas.
- EVERT, Stefan, Thomas PROISL, Fotis JANNIDIS, Isabella REGER, Steffen PIELSTRÖM, Christof SCHÖCH, y Thorsten VITT (2017), “Understanding and explaining Delta measures for authorship attribution”, *Digital Scholarship in the Humanities*, 32 (2), pp. 4-16.
- EVETT, Ian W. y Bruce S. Weir (1998), *Interpreting DNA Evidence. Statistical Genetics for Forensic Scientists*, Sunderland, Sinauer Associates, Inc.
- FARRINGDON, Jill (1996), *Analysing for authorship: A guide to the Cusum technique*, Cardiff, University of Wales Press.
- FEIGUINA, Olga y Graeme HIRST (2007), “Authorship attribution for small texts: Literary and forensic experiments”. Conferencia presentada al *International Workshop on Plagiarism Analysis, Authorship Identification, and Near-Duplicate Detection*, 30th Annual International ACM SIGIR Conference (SIGIR '07). Disponible en <http://ceur-ws.org/Vol-276/paper3.pdf> [Fecha de consulta: 08.12.2021].
- FERNÁNDEZ FERREIRO, María (2012), “Adaptaciones teatrales del *Quijote* (siglos XX-XXI). Selección de un corpus”, en Carlos Mata Induráin y Adrián J. Sáez (coords.), “*Scripta manent*”. *Actas del I Congreso Internacional Jóvenes*

BIBLIOGRAFÍA

Investigadores Siglo de Oro, Navarra, Servicio de Publicaciones de la Universidad de Navarra, pp. 185-194.

FERNÁNDEZ GUERRA, Aureliano (1863), “Noticia de un precioso códice de la biblioteca colombina, con varios rasgos inéditos de Cetina, Cervantes y Quevedo”, en Bartolomé José Gallardo, *Ensayo de una biblioteca española de libros raros y curiosos*, apéndice al tomo I, Madrid, Imprenta y estereotipia de M. Rivadeneyra, pp. 1245-1326.

FERNÁNDEZ OBLANCA, Justo (1991), “La pasión por los coches en el siglo XVII y su reflejo cómico en los entremeses barrocos”, *Archivum: Revista de la Facultad de Filosofía y Letras*, tomo 41-42, pp. 105-124.

FERNÁNDEZ OBLANCA, Justo (1991), *Aproximación sociológica a los entremeses del siglo XVII*, Oviedo, Universidad de Oviedo.

FERNÁNDEZ OBLANCA, Justo (1992), *Literatura y sociedad en los entremeses del siglo XVII*, Oviedo, Universidad de Oviedo.

FERNÁNDEZ TRILLO, Juan Jesús (2015), *La problemática autoría de Carlos Castaneda. Un estudio estilométrico de “candidato único”*, tesis doctoral, Universidad Autónoma de Madrid. Disponible en <https://repositorio.uam.es/handle/10486/666725> [Fecha de consulta: 28.12.2021].

FERNÁNDEZ VALBUENA, Ana Isabel (ed.) (2006), *La Comedia del Arte: materiales escénicos*, Madrid, Editorial Fundamentos.

FERRER VALLS, Teresa (1991), “Lope de Vega y el teatro por encargo: plan de dos comedias” en M. V. DIAGO y T. FERRER (eds), *Comedias y Comediantes: estudios sobre el teatro clásico español*, Valencia, Universitat de València, pp. 189-202.

FITZGERALD, James R. (2004) “Using a Forensic linguistic approach to track the Unabomber”, en John H. Campbell y Don DeNevi (eds.), *Profilers. Leading Investigators Take You Inside the Criminal Mind*, Amherst, N.Y., Prometheus Books, pp. 193-222.

FLORIT DURÁN, Francisco (2009), “En torno a la teatralización de la acción narrativa en el *Quijote*”, en Joaquín Álvarez Barrientos, Oscar Cornago Bernal, Abraham

BIBLIOGRAFÍA

- Madroñal Durán, Carmen Menéndez Onrubia (coords.), *En buena compañía: estudios en honor de Luciano García Lorenzo*, pp. 287-298.
- FORD, Jeremiah D. M. y Ruth LANSING (1931), *Cervantes. A tentative Bibliography of his Works and the Biographical and Critical Material Concerning Him*, Cambridge, Harvard University Press.
- FOSTER, Don (2000), *Author unknown. On the trail of anonymous*, New York, Henry Holt and Company.
- FOUCAULT, Michel (1969), “¿Qu'est-ce qu'un auteur?”, *Bulletin de la Société française de Philosophie*, 34, pp. 73-104.
- FRADEJAS RUEDA, José Manuel (2009-2010), “La codificación XML/TEI de textos medievales”, *Memorabilia: boletín de literatura sapiencial*, 12, pp. 219-247.
- FRADEJAS RUEDA, José Manuel (2016a), “El análisis estilométrico aplicado a la literatura española: las novelas policíacas e históricas”, *Caracteres: estudios culturales y críticos de la esfera digital*, 5 (2), pp. 196-245.
- FRADEJAS RUEDA, José Manuel (2016b), “Lingüística forense y crítica textual. El caso Ayala–Cervantes”, *Signa: Revista de la Asociación Española de Semiótica*, 25, pp. 193-220.
- FRADEJAS RUEDA, José Manuel (2019a), “El Ministerio del Tiempo: minería textual de un éxito transmedia”, en Teresa Gómez Trueba y María Martínez Deyros (coords.), *Página y pantalla: interferencias metaficcionales*, Gijón, Ediciones Trea, pp. 319-334.
- FRADEJAS RUEDA, José Manuel (2019b), “Estilometría y Edad Media castellana”, *Romanische Studien*, 6, pp. 49-74.
- FRADEJAS RUEDA, José Manuel (2020), *Cuentapalabras. Estilometría y análisis de texto con R para filólogos*. Disponible en <http://www.aic.uva.es/cuentapalabras/> [Fecha de consulta: 08.12.2021].
- FRASER, B. (1998), “Threatening”, *Forensic Linguistics. The International Journal of Speech, Language and the Law*, 5 (2), pp. 159-173.

BIBLIOGRAFÍA

- FRENCH, Peter y Malcolm COULTHARD (1994), “Editorial introduction”, *Forensic Linguistics*, 1 (1), London, Routledge, pp. 7-8.
- FRENCH, Peter y Philip HARRISON (2007), “Position Statement concerning use of impressionistic likelihood terms in forensic speaker comparison cases”, *The International Journal of Speech, Language and the Law*, 14 (1), pp. 137-144.
- GAINS, Jonathan (1999), “Electronic mail. A new style of communication or just a new medium? An investigation into the text features of e-mail”, *English for Specific Purposes*, 18 (1), pp. 81-101.
- GAMBIN, Felice (2002), “Intermezzi della follia: malinconia e meditazione nei discorsi di don Quijote”, en José Manuel Martín Morán (ed.), *La media semana del jardincito. Cervantes y la reescritura de los códigos*, Padova, Unipress, pp. 151-166.
- GARAYZÁBAL HEINZE, Elena, Miriam JIMÉNEZ BERNAL y Mercedes REIGOSA RIVEIROS (2014), *Lingüística Forense: La Lingüística en el ámbito legal y policial*, Madrid, Euphoria Ediciones.
- GARAYZÁBAL HEINZE, Sheila QUERALT y Mercedes REIGOSA RIVEIROS (2019), *Fundamentos de la Lingüística forense*, Madrid, Síntesis.
- GARCÍA AGUILAR, Ignacio, Luis GÓMEZ CANSECO y Adrián J. SÁEZ (2016), *El teatro de Miguel de Cervantes*, Madrid. Visor Libros.
- GARCÍA LORENZO, Luciano (ed.) (1988), *Los géneros menores en el teatro español del siglo de oro (Jornadas de Almagro 1987)*, Madrid, Ministerio de cultura, Instituto nacional de las Artes escénicas y de la Música.
- GARCÍA LORENZO, Luciano y J. E. VAREY (eds.) (1991), *Teatros y vida teatral en el siglo de oro a través de las fuentes documentales*, London, Tamesis Books.
- GARCÍA MARTÍN, Manuel (1991), *Cervantes y la comedia española en el siglo XVII*, Universidad de Salamanca.
- GARCÍA SORIANO, Justo (1945), *El teatro universitario y humanístico en España*, Toledo, Talleres tipográficos de don Rafael Gómez Méndez.

BIBLIOGRAFÍA

- GARCÍA-REIDY, Alejandro (2008), “Una comedia inédita de Andrés de Claramonte: *San Carlos o las dos columnas de Carlos*”, *Criticón*, 102, pp. 177-193.
- GARCÍA-REIDY, Alejandro (2013a), “*Mujeres y criados*, una comedia recuperada de Lope de Vega”, *Revista de Literatura*, 75 (150), pp. 417-438.
- GARCÍA-REIDY, Alejandro (2013b), *Las musas ramera. Oficio dramático y conciencia profesional en Lope de Vega*, Madrid, Iberoamericana - Frankfurt am Main, Vervuert.
- GARCÍA-REIDY, Alejandro (2019), “Deconstructing the Authorship of *Siempre ayuda la verdad*: A Play by Lope de Vega?”, *Neophilologus*, 103, pp. 493-510.
- GARCÍA-REIDY, Alejandro, Ramón VALDÉS GÁZQUEZ y Germán VEGA GARCÍA-LUENGOS (2021), “Una nueva edición (¿*princeps*?) de *El castigo sin venganza*”, *Anuario de Lope de Vega. Texto, literatura, cultura*, 27, pp. 270-329.
- GIBBONS, John y M. Teresa TURELL (2008), *Dimensions of Forensic Linguistics*, Amsterdam / Philadelphia, John Benjamins Publishing Company.
- GIES, David T. (2004), “El reto imposible de la historia literaria: el caso Cambridge”, en Leonardo Romero Tobar (ed.), *Historia literaria / Historia de la literatura*, Zaragoza, Prensas Universitarias de Zaragoza, pp. 273-283.
- GLASER, Edward (1954), “An addition to the Cervantes Canon?”, *Harvard Library Bulletin*, 3 (1), pp. 86-96.
- GOLD, Erica y Peter FRENCH (2011), “International practices in forensic speaker comparison”, *The International Journal of Speech, Language and the Law*, 18 (2), pp. 293-307.
- GÓMEZ, Jesús (2001), “El entremés de *Melisendra* atribuido a Lope de Vega y los orígenes de la comedia burlesca”, *Boletín de la Real Academia Española (BRAE)*, 81, pp. 205-221.
- GÓNGORA, Luis de (1998), *Romances*, edición crítica de Antonio Carreira, Barcelona, Quaderns Cremá, 4 vols.

BIBLIOGRAFÍA

- GONZÁLEZ GUTIÉRREZ, Cayo (1997), *El teatro escolar de los jesuitas (1555-1640) (Su influencia en el teatro del siglo de Oro)*. Edición de la «Tragedia de San Hermenegildo», Oviedo, Universidad de Oviedo.
- GONZÁLEZ, Aurelio (1993), “Cervantes y los temas del Romancero nuevo”, *Actas del III Coloquio Internacional de la Asociación de Cervantistas*, Barcelona, Anthropos, pp. 609-616.
- GONZÁLEZ, Aurelio (1997), “El proyecto teatral de las comedias de Cervantes”, en *El escritor y la escena V. Estudios sobre teatro español y novohispano de los Siglos de Oro: homenaje a Marc Vitse*, Ciudad Juárez, Universidad Autónoma de Ciudad Juárez, pp. 79-85
- GONZÁLEZ, Aurelio (2015), “Creación de atmósfera narrativa (clima) en *las Novelas ejemplares*”, en Aurelio González y Nieves Rodríguez Valle (eds.), *Las Novelas ejemplares: texto y contexto (1613-2013)*, Méjico, El Colegio de Méjico, Centro de Estudios Lingüísticos y Literarios, pp. 33-60.
- GONZALO SÁNCHEZ-MOLERO, José Luis (2007), “La ‘Epístola a Mateo Vázquez’, redescubierta y reivindicada”, *Cervantes: Bulletin of the Cervantes Society of America*, Fall 2007, 27 (2), pp. 181-211.
- GONZALO SÁNCHEZ-MOLERO, José Luis (2010), *La epístola a Mateo Vázquez: Historia de una polémica literaria en torno a Cervantes*, Alcalá de Henares, Centro de Estudio cervantinos.
- GRACIA, Jordi (2016), *Miguel de Cervantes: La conquista de la ironía*. Madrid, Taurus.
- GRANJA, Agustín de la (1987), “Hacia una bibliografía general del teatro breve del siglo de Oro. Primera parte: Estudios I”, *Criticón*, 37, pp. 227-246.
- GRANJA, Agustín de la (1990), “Hacia una bibliografía general del teatro breve del siglo de Oro. Primera parte: Estudios II”, *Criticón*, 50, pp. 113-124.
- GRANJA, Agustín de la (1994), “El entremés: la larga risa de un teatro breve”, en I. Arellano, V. García Ruiz y M. Vitse (eds.), *Del horror a la risa. Los géneros teatrales clásicos. Homenaje a Christiane Faliu-Lacourt*, Kassel, Reichenberger, pp. 161-189.

BIBLIOGRAFÍA

- GRANJA, Agustín de la (1995), “Apogeo, decadencia y estimación de las comedias de Cervantes”, *Cervantes*, 10 (2), pp. 225-254.
- GRANJA, Agustín de la y María Luisa LOBATO (1999), *Bibliografía descriptiva del teatro breve español (siglos XV-XX)*, Universidad de Navarra.
- GRANT, Tim (2007), “Quantifying evidence in forensic authorship analysis”, *The International Journal of Speech, Language and the Law*, 14 (1), pp. 1-25.
- GRANT, Tim (2008), “Approaching questions in Forensic authorship analysis”, en J. Gibbons y M. T. Turell, *Dimensions of Forensic Linguistics*, Amsterdam / Philadelphia, John Benjamins Publishing Company, pp. 216-229.
- GRANT, Tim y K. BAKER (2001), “Identifying reliable, valid markers of authorship: a response to Chaski”, *Forensic Linguistics. The International Journal of Speech, Language and the Law*, 8 (1), pp. 66-79.
- GRIEVE, Jack (2007), “Quantitative Authorship Attribution: An Evaluation of Techniques”, *Literary and Linguistic Computing*, 22 (3), pp. 251-270.
- GRIEVE, Jack, Isobelle CLARKE, Emily CHIANG, Hannah GIDEON, Annina HEINI, Andrea NINI y Emily WAIBEL (2018), “Attributing the Bixby Letter using n-gram tracing”, *Digital Scholarship in the Humanities*, 34 (3), pp. 493-512.
- GUTIÉRREZ CUADRADO, Juan (1988), “La lengua del *Quijote*: rasgos generales”, en Francisco Rico (ed.), *Don Quijote de la Mancha*, vol. complementario, Barcelona, Instituto Cervantes / Editorial Crítica, pp. 819-856.
- GUTIÉRREZ, Carlos M. (2004), “Cervantes o la relocalización del sujeto teatral”, *Bulletin of the comediantes*, 56 (2), pp. 289-310.
- GUTIÉRREZ, Carlos M. (2005a), *La espada, el rayo y la pluma. Quevedo y los campos literario y de poder*. West Lafayette, Purdue University Press.
- GUTIÉRREZ, Carlos M. (2005b), “Narrador, autor y personaje: facetas de la autorrepresentación literaria en Góngora, Lope, Cervantes y Quevedo”, *Espéculo: Revista de estudios literarios*, 31.

BIBLIOGRAFÍA

- HANSEN, Federico (1945), *Gramática histórica de la lengua castellana*. Buenos Aires: El Ateneo.
- HEIDEGGER, Martin (1958), *Arte y Poesía*, México, Fondo de Cultura Económica.
- HERNÁNDEZ ALONSO, César y Beatriz SANZ ALONSO (1999), *Alemania y sociedad en los Siglos de Oro*, Valladolid, Secretariado de Publicaciones e Intercambio Editorial de la Universidad de Valladolid.
- HERNÁNDEZ LORENZO, Laura (2019a), *Los textos poéticos de Fernando de Herrera: Aproximaciones desde la Estilística de corpus y la Estilometría*, tesis doctoral, Universidad de Sevilla. Disponible en <https://idus.us.es/handle/11441/93465>. [Fecha de consulta: 28.12.2021].
- HERNÁNDEZ LORENZO, Laura (2019b), “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), pp. 189-228.
- HERRERO GARCÍA, Miguel (1950), “Una hipótesis sobre las *Novelas ejemplares*”, *Revista Nacional de Educación*, 10, pp. 33-37.
- HOLMES, David I. (1998), “The Evolution of Stylometry in Humanities Scholarship”, *Literary and Linguistic Computing*, 13 (3), pp. 111-117.
- HOLMES, David I. y Richard. S. FORSYTH (1995), “The Federalist revisited: New directions in authorship attribution”, *Literary and Linguistic Computing*, 10 (2), pp. 111-127.
- HOLMES, David I., Michael ROBERTSON y Roxanna PAEZ (2001), “Stephen Crane and the *New-York Tribune*: A Case Study in Traditional and Non-Traditional Authorship Attribution”, *Computers and the Humanities*, 35, pp. 315-331.
- HOOVER, David L. (2004a), “Frequent collocations and authorial style”, *Literary and Linguistic Computing*, 18 (3), pp: 261-228.
- HOOVER, David L. (2004b), “Testing Burrows’s Delta”, *Literary and Linguistic Computing*, 19 (4), pp. 453-475.

BIBLIOGRAFÍA

- HOUVARDAS, John y Efstathios STAMATATOS (2006), “N-gram feature selection for authorship identification”, en J. Euzenat y J. Domingue (eds.), *Proceedings of Artificial Intelligence: Methodologies, Systems, and Applications*, Springer Verlag, pp. 77-86.
- HUERTA CALVO, Javier (1990), *Risa y Eros: del erotismo en los entremeses*, Madrid.
- HUERTA CALVO, Javier (1995), *El nuevo mundo de la risa. Estudios sobre el teatro breve y la comicidad en los siglos de Oro*, Barcelona, José J. de Olañeta, Editor.
- HUERTA CALVO, Javier (2001), *El teatro breve en la Edad de Oro*, Madrid, Laberinto.
- HUERTA CALVO, Javier (coord.) (2008), *Historia del teatro breve en España*, Madrid/Frankfurt am Main, Iberoamericana/Vervuert.
- HUERTA CALVO, Javier (dir.), Abraham MADROÑAL y Héctor URZÁIZ TORTAJADA (coords.) (2003), *Historia del teatro español. I. De la Edad Media a los siglos de Oro*, Madrid, Gredos.
- IGLESIAS CASAL, Isabel (1996), *Los relativos en la prosa renacentista castellana*, Oviedo, Universidad de Oviedo.
- INAMOTO, Kenji (1990), “Melisendra, entremés una vez atribuido a Cervantes”, en *Actas del II Coloquio Internacional de la Asociación de Cervantistas*, Barcelona, Anthropos, pp. 691-694.
- INAMOTO, Kenji (2004), “Sobre la atribución cervantina de *La tía fingida* o/y sobre el hiato delante del acento rítmico interno en el verso cervantino”, en Alicia Villar Lecumberri (coord.), *Peregrinamente peregrinos: Actas del V Congreso Internacional de la Asociación de Cervantistas*, vol. 2, pp. 1401-1418.
- INFANTES DE MIGUEL, Víctor (2015), *La librería Don Quijote y los libros de Cervantes (I, VI)*, Madrid, Turpin, D.L.
- JAKOBSON, Roman (1975), *Ensayos de Lingüística General*, Barcelona, Seix Barral.
- JANNIDIS, Fotis, Steffen PIELSTRÖM, Christof SCHÖCH y Thorsen VITT (2015), “Improving Burrows Delta. An empirical evaluation of text distance measures. *Digital Humanities*, 32 (2), pp. 4-16.

BIBLIOGRAFÍA

- JEFFRIES, Lesley y Daniel MCINTYRE (2010), *Stylistics*, Cambridge, Cambridge University Press.
- JOCKERS, Matthew y Daniela M. WITTEN (2010), “A comparative study of machine learning methods for authorship attribution”, *Literary and Linguistic Computing*, 25 (2), pp. 215-523.
- JOHNSON, Alison (1998), “Textual kidnapping. A case of plagiarism among three student texts?”, *Forensic Linguistics*, 4 (2), Birmingham: The University of Birmingham Press, pp. 210-225.
- JOHNSON, Carroll B. (1981), “El arte viejo de hacer teatro: Lope de Rueda, Lope de Vega y Cervantes”, *Cuadernos de Filología*, III, 1-2, pp. 247-259.
- JOHNSTONE, Barbara (1996), *The Linguistic Individual. Self-Expression in Language and Linguistics*, New York, Oxford University Press.
- JOLLIFFE, I. T. (2002), *Principal Component Analysis*, New York, Springer, 2ª edición.
- JONES, H. G. (1976), “Texto crítico de ‘Alta mar esquiva’, poema germanesco”, *Boletín de la Biblioteca Menéndez Pelayo*, 52, pp. 223-227.
- JULIÁ MARTÍNEZ, Eduardo (1948), “Estudio y técnica de las comedias de Cervantes”, *Revista de Filología Española*, XXXII, pp. 339-365.
- JUOLA, Patrick (2006), “Authorship Attribution”, *Information Retrieval*, 1.3, pp. 233-334.
- JUOLA, Patrick (2008), *Authorship Attribution*, Hanover, Now Publishers.
- JUOLA, Patrick (2013), “How a computer program helped reveal J. K. Rowling as author of *A Cuckoo’s Calling*”, *Scientific American*. Disponible en <https://www.scientificamerican.com/article/how-a-computer-program-helped-show-jk-rowling-write-a-cuckoos-calling/>. [Fecha de consulta: 28.12.2021].
- JUOLA, Patrick (2015), “The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions”, *Digital Scholarship in the Humanities*, 30 (1), pp. 100-113.

BIBLIOGRAFÍA

- JURADO SANTOS, Agapita (2005), *Obras teatrales derivadas de las novelas cervantinas (siglo XVII). Para una bibliografía*, Kassel, Reichenberger.
- KARSDORP, Folgert, Mike KESTEMONT y Allen RIDDELL (2021), *Humanities Data Analysis: Case Studies with Python*, Princeton, Princeton University Press.
- KESTEMONT, Mike, Justin STOVER, Moshe KOPPEL, Folgert KARSDORP y Walter DAELEMANS (2016a), “Authenticating the writings of Julius Caesar”, *Expert Systems with Applications*, 63, pp. 86-96.
- KESTEMONT, Mike, Justin STOVER, Moshe KOPPEL, Folgert KARSDORP y Walter DAELEMANS (2016b), “Authorship verification with the Ruzicka metric”, en *Digital Humanities 2016: Conference Abstracts*, Kraków, Jagiellonian University & Pedagogical University, pp. 246-49.
- KOPPEL, Moshe y Yaron WINTER (2014), “Determining if two documents are written by the same author”, *Journal of the Association for Information Science and Technology*, 65 (1), pp. 178-87.
- KOPPEL, Moshe, Jonathan SCHLER y Shlomo ARGAMON (2009), “Computational methods in authorship attribution”, *Journal of the American Society for Information Science and Technology*, 60 (1), pp. 9-26.
- LABOV, William (1966), *The social stratification of English in New York City*, Washington D.C., Center for Applied Linguistics.
- LABOV, William (1972), *Sociolinguistic Patterns*. Philadelphia, Pennsylvania University Press.
- LABOV, William (1994), *Principles of Linguistic Change, I: Internal Factors*, Oxford, Blackwell.
- LABOV, William (2001), *Principles of Linguistic Change, II: External Factors*, Oxford, Blackwell.
- LADRA, David (1994), “Algunas claves para la comprensión del teatro de Miguel de Cervantes”, *Primer acto. Cuadernos de investigación teatral*, 252, pp. 18-25.

BIBLIOGRAFÍA

- LANCASHIRE, Ian (1999), "Probing Shakespeare's Idiolect in *Troilus and Cressida* 1.3.1-29", *University of Toronto Quarterly*, 68.3, pp. 728-767.
- LARRA, Mariano José (seud. Bachiller don Juan Pérez de Munguía) (1832), "¿Quién es por acá el autor de una comedia?", *El pobrecito hablador. Revista satírica de costumbres*, 5, pp. 15-21.
- LÁZARO CARRETER, Fernando (1965), "El *Arte nuevo* y el término *entremés*", *Anuario de las letras*, 5, pp. 77-92.
- LEÓN MAÍNEZ, Ramón (1876), *Vida de Miguel de Cervantes Saavedra*, Cádiz, Tipografía la Mercantil.
- LEONARD, Robert Andrew (2006), "Forensic Linguistics: applying the scientific principles of language analysis to issues of the law", *International Journal of the Humanities*, 3 (7), Melbourne, Common Ground Publishing Pty. LTD., pp. 65-70.
- LEVI, Judith N. (1994), Language as evidence: the linguist as expert witness in North American Courts, *Forensic Linguistics. The International Journal of Speech, Language and the Law*, 1 (1), pp. 1-26.
- LOBATO, María Luisa (1991), "Tres calas en la métrica del teatro breve español del Siglo de Oro: Quiñones de Benavente, Calderón y Moreto", en Karl-Hermann Körner (ed.), Günther Zimmermann (ed.) y Rafael Lapesa Melgar (introd.), *Homenaje a Hans Flasche: Festschrift zum 80. Geburtstag am 25. November 1991*, Stuttgart, Franz Steiner, pp. 113-154.
- LOBATO, María Luisa (2013), "Maladros 'padre fundador' de la germanía: presencia literaria y edición del entremés inédito *Los valientes*, de Juan Vélez de Guevara", *La Perinola*, 17, pp. 229-258.
- LOEWENSTEIN, Joseph (1985), "The script in the marketplace", *Representations*, 12, pp. 101-114.
- LÓPEZ MARTÍN, Ismael (2012), "Las colecciones de entremeses en el Barroco español", *Castilla. Estudios de Literatura*, 3, pp. 1-17.

BIBLIOGRAFÍA

- LÓPEZ NAVÍO, José (1959-1960), “*El entremés de los romances, sátira contra Lope de Vega, fuente de inspiración de los primeros capítulos del Quijote*”, *Anales cervantinos*, VIII, pp. 151-239.
- LOVE, Harold (2002), *Attributing Authorship: An introduction*, Cambridge, Cambridge University Press.
- LUDWIG PFANDL (1952), *Historia de la literatura nacional española en la edad de oro*, Barcelona, Gili.
- LUDWIG PFANDL, (1933), *Historia de la literatura nacional española en la edad de oro*, Barcelona, Sucesores de J. Gili, S.A.
- LUYCKX, Kim (2010), *Scalability Issues in Authorship Attribution*, Brussels, University Press Antwerp.
- MACHADO, Manuel (1925), “La palabra vengada. Plan inédito de una comedia perdida de Lope de Vega”, *Revista de Bibliotecas, Archivos y Museos del Ayuntamiento de Madrid*, 2, pp. 302-306.
- MADROÑAL DURÁN, Abraham (1993), “El contador Gaspar de Barrionuevo (1562-c.1624?), poeta y dramaturgo toledano amigo de Lope de Vega”, *Voz y Letra: revista de literatura*, 4 (2), pp. 105-128.
- MADROÑAL DURÁN, Abraham (1994), “En torno a la *Jocoseria* de Luis Quiñones de Benavente”, en Francis Cerdán (coord.), *Hommage à Robert Jammes*, vol. 2, Toulouse, Presses Universitaires du Mirail, pp. 773-786.
- MADROÑAL DURÁN, Abraham (1995), “Catálogo de entremeses de la Biblioteca de la Real Academia Española”, *Boletín de la Real Academia Española (BRAE)*, 75 (246), pp. 523-568.
- MADROÑAL DURÁN, Abraham (1996), *Nuevos entremeses atribuidos a Luis Quiñones de Benavente*, Kassel, Reichenberger.
- MADROÑAL DURÁN, Abraham (1998), “Suplemento al *Catálogo de entremeses de la Biblioteca de la Real Academia Española*”, *Boletín de la Real Academia Española (BRAE)*, 78 (273), pp. 131-139.

BIBLIOGRAFÍA

- MADROÑAL DURÁN, Abraham (2004), “«Una acción entre plebeya gente»: los entremeses impresos por Velasco en Cádiz a mediados del XVII”, en Odette Gorsse y Frédéric Serralta (coords.), *El Siglo de Oro a escena. Homenaje a Marc Vitse*, Toulouse, PUM/Consejería de Educación de la Embajada de España en Francia, pp. 563-575.
- MADROÑAL DURÁN, Abraham (2004), “Estado actual de los estudios sobre teatro breve del Siglo de Oro”, *Arbor*, 699-700, pp. 455-474.
- MADROÑAL DURÁN, Abraham (2008), “Entremeses intercalados en el *Quijote*”, en *El Quijote y el pensamiento teórico literario. Actas del Congreso Internacional celebrado en Madrid los días del 20 al 24 de junio de 2005*, Miguel Ángel Garrido Gallardo y Luis Alburquerque García (coords.), Madrid, CSIC, pp. 265-277.
- MADROÑAL DURÁN, Abraham (2011), “Disidentes andaluces en el entremés barroco”, en Elisa García Lara y Antonio Serrano Aguiló (coords.), *Dramaturgos y espacios teatrales andaluces de los siglos XVI-XVII*, Almería, Instituto de Estudios Almerienses, pp. 69-82.
- MADROÑAL DURÁN, Abraham (2013), “De nuevo sobre la autoría de *Los refranes del viejo celoso*, entremés atribuido a Quevedo”, *La Perinola*, 17, pp. 155-177.
- MAESTRO, Jesús G. (1998), “Construcción e interpretación del diálogo en los *Entremeses* de Miguel de Cervantes”, en A. Bernat Vistarini (ed.), *Actas del tercer Congreso Internacional de la Asociación de Cervantistas*, Palma de Mallorca, Universitat de les Illes Balears, pp. 591-610.
- MAESTRO, Jesús G. (1998), “La *commedia dell'arte* y el entremés cervantino. Sobre el diálogo dramático”, *El Extramundi y los Papeles de Iria Flavia*, 14, pp. 15-44.
- MAESTRO, Jesús G. (2000), *La escena imaginaria. Poética del teatro de Miguel de Cervantes*, Madrid, Iberoamericana.
- MAESTRO, Jesús G. (2004), “Cervantes y Shakespeare: el nacimiento de la literatura metateatral”, *Bulletin of Spanish Studies. Essays in Memory of E. C. Riley on the Quatercentenary of Don Quijote*, 48 (4-5), pp. 599-611.

BIBLIOGRAFÍA

- MAESTRO, Jesús G. (2005), “Nuevo itinerario del entremés”, *Cervantes. Bulletin of the Cervantes Society of America*, 25 (1), pp. 201-203.
- MAESTRO, Jesús G. (2007), “De la teatralidad en el “Quijote”: Sancho en Barataria o la subversión de la preceptiva de lo cómico”, en Emilio Martínez Mata (coord.), *Cervantes y el Quijote: Actas del coloquio internacional organizado por la cátedra Emilio Alarcos*, Madrid, Arco Libros, pp. 97-112.
- MAESTRO, Jesús G. (2013), “Atribuciones teatrales cervantinas: el caso de los entremeses”, en *Calipso eclipsada. El teatro de Cervantes más allá del Siglo de Oro*, Madrid, Verbum, pp. 75-90.
- MANGUEL, Alberto (1998), *Una historia de la lectura*, Madrid, Alianza Editorial / Fundación Germán Sánchez Ruipérez.
- MANNION, David y Peter DIXON (1997), “Authorship attribution: the case of Oliver Goldsmith.”, *The Statistician*, 46 (1), pp. 1-18.
- MARAVALL, José Antonio (1975), *La cultura del Barroco: Análisis de una estructura histórica*, Barcelona, Ariel.
- MARCIAL, Marco Valerio (1997), *Epigramas, I*. Madrid, Gredos.
- MARCO MOLINA, Juana (1994), “Bases históricas y filosóficas y precedentes legislativos del derecho de autor”, *Anuario de derecho civil*, 47 (1), pp. 121-208
- MARCO MOLINA, Juana (1995), *La propiedad intelectual en la legislación española*, Madrid, Marcial Pons.
- MARCOS RODRÍGUEZ, Emma M^a (2021), “Texto, atribución y censura de *Próspera y Adversa fortuna de don Álvaro de Luna*”, *Talía. Revista de estudios teatrales*, 3, pp. 79-89.
- MARÍN CEPEDA, Patricia (ed.) (2006), “Miguel de Cervantes (atrib.). Relación de lo sucedido en la ciudad de Valladolid, desde el punto del felicísimo nacimiento del príncipe don Felipe Dominico Víctor nuestro señor, hasta que se acabaron las demostraciones de alegría que por él se hicieron”, *Cervantes: Bulletin of the Cervantes Society of America*, 25 (2), pp. 195-270.

BIBLIOGRAFÍA

- MARÍN CEPEDA, Patricia (2010), “Cuatro personajes en busca de autor para la *Topografía e historia general de Argel*: Haedo (arzobispo de Sicilia), Haedo (abad de Frómista), Sosa y Cervantes”, en Javier Blasco, Patricia Marín Cepeda y Cristina Ruiz Urbón (eds.), *Hos ergo versículos feci... Estudios de atribución y plagio*, Madrid, Iberoamericana, 2010, pp. 19-74.
- MARÍN CEPEDA, Patricia (2015), *Cervantes y la corte de Felipe II: escritores en el entorno de Ascanio Colonna, (1560-1608)*, Madrid, Polifemo, 2015.
- MAROTTI, Arthur F. (1995), *Manuscript Print, and the English Renaissance Lyric*, Ithaca Cornell University Press.
- MÁRQUEZ VILLANUEVA, FRANCISCO (1983), “Lope, infamado de morisco”, *Anuario de Letras*, 21, pp. 147-182.
- MARQUINA, Montse y Sheila QUERALT (2014), “Similarity threshold to detect plagiarism in Spanish”, *RAEL: revista electrónica de lingüística aplicada*, 13 (1), pp. 79-95.
- MARRAST, Robert (1957), *Cervantes dramaturge*, Paris, L'Arche.
- MARRAST, Robert (1961), “Introducción”, en Miguel de Cervantes, *El cerco de Numancia*, salamanca, Anaya, pp. 5-25.
- MARTÍN JIMÉNEZ, Alfonso (2005), *Cervantes y Pasamonte: la réplica cervantina al “Quijote” de Avellaneda*, Madrid: Biblioteca Nueva.
- MARTÍN JIMÉNEZ, Alfonso (2006), “El manuscrito de la primera parte del *Quijote* y la disputa entre Cervantes y Lope de Vega”, *Etiópicas*, 2, pp. 255-344.
- MARTINEZ LOPEZ, María José (1997), *El entremés: radiografía de un género*, Toulouse, Presses Universitaires du Mirail.
- MATTZA, Carmela V. (2013), “Écfrasis discursiva y metateatro: Rodamonte en el *Entremés del viejo celoso*”, en Alain Bègue y Emma Herrán Alonso (coords.), *Pictavia aurea. Actas del IX Congreso de la AISO (Poitiers, 11-15 de julio de 2011)*, Toulouse, Presses Universitaires du Mirail, pp. 973-980.

BIBLIOGRAFÍA

- MATTZA, Carmela V. (2016), “La historia de Belerma y Durandarte según Lope de Vega y Cervantes: apuntes para una estética de la emotividad en Don Quijote”, *Anuario de estudios cervantinos*, 12, pp. 229-242.
- MCENERY, Tony y Andrew HARDIE (2012), *Corpus linguistics: Method, theory and practice*, Cambridge, Cambridge University Press.
- MCMENAMIN, Gerald R. (2001), “Style markers in authorship studies”, *Forensic Linguistics. The International Journal of Speech, Language and the Law* 8 (2), pp. 93-97.
- MCMENAMIN, Gerald R. (2002), *Forensic Linguistics: Advances in Forensic Stylistics*, London/New York, CRC Press.
- MCMENAMIN, Gerald R. (2004), “Disputed authorship in US Law”, *The International Journal of Speech, Language and the Law* 11 (1), pp. 73-82.
- MCMENAMIN, Gerald R. (2017), *Introducción a la Lingüística forense. Un libro de curso*, Fresno, The Press at California State University.
- MEDINA MORALES, Francisca (2005), *La lengua del siglo de Oro. Un estudio de variación lingüística*, Granada, Universidad de Granada.
- MENÉNDEZ PELAYO, Marcelino (1884), *Historia de las ideas estéticas en España*, Madrid, Imprenta de A. Pérez Dubrull, tomo II, 2 vols.
- MENÉNDEZ PELAYO, Marcelino (1941), “Obras inéditas de Cervantes”, en *Estudios y discursos de crítica histórica y literaria*, Madrid, CSIC, tomo I, pp. 269-302.
- MENÉNDEZ PELAYO, Marcelino (ed.) (1903), *Flor de entremeses y sainetes de diferentes autores (1657)*, Madrid, Imprenta de Fortanet, segunda edición corregida.
- MENÉNDEZ PIDAL, Ramón (1920), “«Nota bibliográfica» de *Últimos estudios cervantinos de E. Cotarelo y Mori*”, *Revista de Filología Española*, 7, 1920, pp. 389-192.
- MENÉNDEZ PIDAL, Ramón (1964), “Un aspecto en la elaboración del *Quijote*”, *De Cervantes y Lope de Vega*, Madrid, Espasa-Calpe, pp. 9-60.

BIBLIOGRAFÍA

- MIGUEL MAGRO, Tania de (2020), “De *El juez de los divorcios* a *El descasamentero* de Salas Barbadillo”, *eHumanista: Journal of Iberian Studies*, 44, pp. 89-108.
- MILLÉ y GIMÉNEZ, Juan (1930), *Sobre la génesis del Quijote: Cervantes, Lope, Góngora, el "Romancero general", el "Entremés de los Romances", etc.*, Barcelona, Araluce.
- MIRÓ LLINARES, Fernando (2007), “El futuro de la propiedad intelectual desde su pasado: La historia de los derechos de autor y su porvenir ante la revolución de Internet”, *Revista de la Facultad de Ciencias Sociales y Jurídicas de la Universidad Miguel Hernández*, 2, pp. 103-155.
- MOLL, Jaime (1979), “Problemas bibliográficos del libro del Siglo de Oro español”, *Boletín de la Real Academia Española (BRAE)*, 59 (216), pp. 49-108.
- MOLL, Jaime (2003), “El impresor, el editor y el librero”, en Víctor Infantes, François Lopez, Jean-François Botrel (eds.), *Historia de la edición y de la lectura en España, 1472-1914*, Madrid, Fundación Germán Sánchez Ruipérez, pp. 77-84.
- MONTERO REGUERA, José (1995), “La obra literaria de Miguel de Cervantes (Ensayo de un catálogo)”, en AA.VV., *Cervantes*, Alcalá de Henares, Centro de Estudios cervantinos, pp. 43-74.
- MONTERO REGUERA, José (1997), *El Quijote y la crítica contemporánea*, Centro de Estudios Cervantinos.
- MONTERO REGUERA, José (1998), “Hacia una revisión de las atribuciones teatrales cervantinas: *La soberana de la Virgen de Guadalupe, y sus milagros, y grandezas de España*”, en *Actas del tercer Congreso internacional de la Asociación de cervantistas (III-CINDAC)*, Palma de Mallorca, Universidad de las Islas Baleares, pp. 661.617.
- MORÍNIGO, Marcos Antonio (1957), “El teatro como sustituto de la comedia en el siglo de oro”, *Revista de la Universidad de Buenos Aires*, II época, 2.1, pp. 41-61.
- MOSTELLER, Frederick y David L. WALLACE (1964), *Inference and disputed authorship: The Federalist*, New York, Springer-Verlag.

BIBLIOGRAFÍA

- MUÑOZ, Liliana (2011), “A propósito de *Doña Justina y Calahorra*, un entremés atribuido a Cervantes”, *Anales cervantinos*, 43, pp. 299-323.
- MURILLO, Luis Andrés (1986), “Cervantes y *El entremés de los romances*”, en A. David Kossoff, Ruth H. Kossoff, Geoffrey Ribbans, José Amor y Vázquez (coords.), *Actas del VIII Congreso de la Asociación Internacional de Hispanistas: 22-27 agosto 1983*, vol. II, pp. 353-357.
- NAVARRO COLORADO, Borja (2016), “Hacia un análisis distante del endecasílabo áureo: patrones métricos, frecuencias y evolución histórica”, *Rhythmica: revista española de métrica comparada*, 14, pp. 89-118.
- NAVARRO Y LEDESMA, Francisco (1944), *El ingenioso hidalgo Miguel de Cervantes Saavedra*, Madrid, Imp. Alemana, 1905; Madrid, Espasa-Calpe.
- NAZAR, Rogelio y Marta SANCHEZ POL (2007), “An Extremely Simple Authorship Attribution System”, en M. T. Turell, M. S. Spassova y J. Cicres (eds.), *Proceedings of the Second European IAFL Conference on Forensic Linguistics/Language and the Law*, Barcelona: Publicacions de l’IULA, pp. 197-203.
- NIDER, Valentina (2011), “La censura del disparate: *l’Entremés de la Infanta Palancona* (Pisa, 1616) e la commedia burlesca *Durandarte y Belerma*”, en Maria Grazia Profeti y Donatella Pini (eds.), *Leyendas negras e leggende auree*, Firenze, Alinea, pp. 157-182.
- NIDER, Valentina (2020), “Estudio preliminar del *Entremés de Durandarte y Belerma*. Versión no censurada de *El amor más verdadero, Durandarte y Belerma*, comedia burlesca de Mosén Guillén Pierres”, en Carlos Mata Indurain (coord.), *Antología de la literatura burlesca del Siglo de Oro*, Pamplona, Servicio de Publicaciones de la Universidad de Navarra, vol. 8, pp. 267-303.
- NORTHUP, George T. (1922a), *Ten Spanish Farces of the 16th, 17th and 18th Centuries*, Boston-Nueva York-Chicago, D.C. Heath & Co.
- NORTHUP, George T. (1922b), “Review” de “Un aspecto en la elaboración del *Quijote*” de Menéndez Pidal, *Modern Philology*, 19 (4), pp. 435-436.

BIBLIOGRAFÍA

- OLSSON, John (2008), *Forensic Linguistics*, London / New York, Continuum International Publishing Group.
- OLSSON, John (2009), *Wordcrime. Solving Crime Through Forensic Linguistics*, London / New York, Continuum International Publishing Group.
- ORTÉS, Federico (2002), *El triunfo de don Quijote. Cervantes y la Compañía de Jesús: un mensaje cifrado*, Brenes (Sevilla), Muñoz Moya editores.
- PALACÍN, G. B. (1952), “¿En dónde oyó Cervantes recitar a Lope de Rueda?”, *Hispanic review*, 20, pp. 240-243.
- PARAÍSO, Isabel (2000), *La métrica española en su contexto románico*, Madrid, Arco/Libros.
- PARODI, Giovanni (2008), “Lingüística de corpus: una introducción al ámbito”, *Revista de Lingüística Teórica y Aplicada*, 46 (1), pp. 93-119.
- PASCUAL, Pedro (1998), “Las Pragmáticas y la industria editorial española en el reinado de Felipe II”, en José Martínez Millán (dir. congr.), *Europa y la monarquía católica: Congreso Internacional "Felipe II (1598-1998), Europa dividida, la monarquía católica de Felipe II (Universidad Autónoma de Madrid, 20-23 abril 1998)*, vol. 4, pp. 403-423.
- PAZ GAGO, José María (1993), “Texto y representación en el teatro español del último cuarto del siglo XVI (Cervantes y Lope: una perspectiva comparada)”, *Bulletin of the Comediantes*, XLV, pp. 225-275.
- PAZ Y MELIÁ, Antonio, (1934), *Catálogo de las piezas de teatro que se conservan en el Departamento de manuscritos de la Biblioteca Nacional*, tomo I, Madrid, Biblioteca Nacional.
- PEDRAZA JIMÉNEZ, Felipe B. (1998), “El teatro mayor de Cervantes: Comentarios a contrapelo”, en José Ramón Fernández de Cano y Martín (coord.), *Actas del VIII Coloquio Internacional de la Asociación de Cervantistas*, Toledo, Ayuntamiento del Toboso, pp. 19-38.

BIBLIOGRAFÍA

- PEDRAZA JIMÉNEZ, Felipe B. (2006), *Cervantes y Lope de Vega: Historia de una enemistad y otros estudios cervantinos*, Barcelona, Octaedro.
- PEDRAZA Jiménez, Felipe B., Rafael GONZÁLEZ CAÑAL, Elena E. MARCELLO (eds.) (2017), *El entremés y sus intérpretes: XXXVIII jornadas de teatro clásico (9, 10 y 11 de julio de 2015: Almagro)*, Cuenca, Universidad de Castilla-La Mancha.
- Pedrosa, José Manuel (2006), “Reseña. Elena Di Pinto, *La tradición escarramanesca en el teatro del Siglo de Oro*. Madrid/Frankfurt am Main, Iberoamericana/Vervuert, 2005. 615 p.”, *Criticón*, 97-98, pp. 326-328.
- PEÑA, Margarita (2010), “Cervantes y el entremés de *La cárcel de Sevilla*: autobiografía y autoría”, en Pierre Civil y Françoise Crémoux (coords.), *Nuevos caminos del hispanismo. Actas del XVI Congreso de la Asociación Internacional de Hispanistas (París, julio de 2007)*, Madrid, Iberoamericana, vol. 2 [CD-ROM].
- PERCAS DE PONSETI, Helena (2003), “Cervantes y Lope de Vega: Postrimerías de un duelo literario y una hipótesis”, *Cervantes*, 23 (1), pp. 63-115.
- PÉREZ DE LEÓN, Vicente (1998), *Humor, reforma y contrautopía en los entremeses de Cervantes*, Am Arbor, Michigan, UMI Dissertations Service.
- PÉREZ DE LEÓN, Vicente (2005), *Tablas destempladas. Los entremeses de Cervantes a examen*, Madrid, Biblioteca de estudios cervantinos.
- PÉREZ DE LEÓN, Vicente (2009), “Entremeses de Miguel de Cervantes”, en Pablo Jauralde Pou (dir.), *Diccionario filológico de literatura española. Siglo XVI*, vol. I, Madrid, Castalia, 2009.
- PÉREZ LASHERAS, Antonio (1988), “El *Entremés de los romances* y los romances del *Entremés*”, en Jean-Pierre Etievre y Leonardo Romero Tobar (coord.), *La recepción del texto literario (Coloquio Casa de Velázquez-Departamento de Filología Española de la Universidad de Zaragoza. Jaca, abril de 1986)*, Secretariado de publicaciones de la Universidad de Zaragoza, pp. 61-76.
- PÉREZ LÓPEZ, José Luis (2009), “Los romances del realismo bucólico de Liñán de Rianza y de Lope de Vega, el *Entremés de los romances* y el *Quijote*”, *Anuario Lope de Vega*, 15, pp. 169-201.

BIBLIOGRAFÍA

- PÉREZ LÓPEZ, José Luis (2012), “El romance morisco *Ensíllenme el potro rucio*, atribuido a Liñán, y su parodia”, *Revista de Filología Española*, 92 (1), pp. 101-116.
- PÉREZ PASTOR, Cristóbal (1897), *Documentos cervantinos hasta ahora inéditos*, Madrid, Establecimiento Tipográfico de Fortanet.
- PÉREZ PASTOR, Cristóbal (1901), *Nuevos datos acerca del histrionismo español en los siglos XVI y XVII. Primera serie*. Madrid, Imprenta de la revista española.
- PÉREZ PASTOR, Cristóbal, Nicomedes SANZ Y RUIZ DE LA PEÑA y Amalia PRIETO CANTERO (1992), *Cervantes en Valladolid*, ed. facsímil, Valladolid, Grupo Pinciano / Caja España.
- PERKINS, Ria, Isabel PICORNELL y Malcolm COULTHARD (eds.) (2022), *Methodologies and Challenges in Forensic Linguistic Casework*, Hoboken (NJ), Wiley-Blackwell.
- PERROMAT AUGUSTÍN, Kevin (2010), *El plagio en las literaturas hispánicas: Historia, Teoría y Práctica*, tesis doctoral, Université Paris-Sorbonne / Paris IV. Disponible en http://www.ugr.es/~plagio_hum/Documentacion/06Publicaciones/LIB003.pdf [Fecha de consulta: 08.12.2021].
- PHILBRICK, Frederick A. (1949), *Language and the Law. The Semantics of Forensic English*, New York, The MacMillan Company.
- PLECHÁČ, Petr, Klemens BOBENHAUSEN y Benjamin HAMMERICH (2018), “Versification and authorship attribution. A pilot study on Czech, German, Spanish, and English poetry”, *Studia Metrica et Poetica*, 5 (2), pp. 29-54.
- PONTÓN, Gonzalo y Agustín SÁNCHEZ AGUILAR (1997), “Prólogo a la edición de los Entremeses”, en *Comedias de Lope de Vega*, Lleida, Editorial Milenio, parte I, vol. 3, pp. 1809-1831.
- PRESOTO, Marco (1997), “Hacia la producción del texto-espectáculo en las comedias autógrafas d Lope”, *Anuario de Lope de Vega*, 3, pp. 153-168.
- PROFETI, Maria Grazia (1996), “Editar el teatro del fénix de los ingenios”, *Anuario de Lope de Vega*, 2, pp. 129-152.

BIBLIOGRAFÍA

- QUERALT, Sheila (2014), “Acerca de la prueba lingüística en atribución de autoría hoy”, *Revista de Llengua i Dret*, 62, pp. 35-48.
- QUERALT, Sheila (2019), “El uso de recursos tecnológicos en lingüística forense”, *Pragmalingüística*, 28, pp. 212-237.
- QUERALT, Sheila y Roser GIMÉNEZ GARCÍA (2018), “La imitación como contraargumento en peritajes de atribución de autoría: estudio de un caso”. *Estudios de Lingüística Aplicada*, 68, pp. 131-164.
- REAL ACADEMIA ESPAÑOLA (2005), *Corpus diacrónico del español (CORDE)*, Disponible en <http://corpus.rae.es/cordenet.html> [Fecha de consulta: 08.12.2021].
- REAL ACADEMIA ESPAÑOLA (2009), *Nuevo Tesoro Lexicográfico de la lengua española (NTLLE)*. Disponible en <https://apps.rae.es/ntlle/SrvltGUISalirNtlle> [Fecha de consulta: 08.12.2021].
- RECOULES, Henri (1973), *Les intermèdes des collections imprimées. Vision caricaturale de la société espagnole au XVIIème siècle*, tesis doctoral, Université de Lille III, 2 tomos.
- REDONDO, Agustín (1997), *Otra manera de leer el «Quijote»*. Historia, tradiciones culturales y literatura, Madrid, Castalia.
- REICHENBERGER, Kurt (1990), “Ediciones críticas de textos dramáticos. Problemas antiguos y recientes”, en Jesús Cañedo Fernández e Ignacio Arellano Ayuso (coords.), *Crítica textual y anotación filológica en obras del Siglo de Oro. Actas del Seminario Internacional para la Edición y Anotación de Textos del Siglo de Oro*, Madrid, Castalia, pp. 417-430.
- REINA, Manuel Francisco (2012), *El plagio como una de las bellas artes*, Barcelona, Ediciones B.
- RESTA, Ilaria (2016), *Fuentes, reescrituras e intertextos: la “novella” italiana en el entremés de Siglo de Oro*, Madrid / Frankfurt am Main, Iberoamericana / Vervuert.
- REY HAZAS, Antonio (1999), “Cervantes se reescribe: teatro y *Novelas ejemplares*”, *Criticón*, 76, pp. 119-164.

BIBLIOGRAFÍA

- REY HAZAS, Antonio (2006), “Cervantes, Lope, Góngora, el *Entremés de los Romances* y los primeros capítulos del *Quijote*”, *Edad de oro*, 25, pp. 473-502.
- REY HAZAS, Antonio (2007), “Estudio del *Entremés de los romances*”, *Revista de estudios cervantinos*, 1, pp. 1-57.
- REY HAZAS, Antonio (coord.) (2002), *Teatro breve del Siglo de Oro*, Alianza Editorial.
- REYES, Alfonso, *Libros y libreros en la Antigüedad*, Madrid, Fórcola, 2011.
- RIBERA BLANES, Begoña (2002), *El derecho de reproducción en la propiedad intelectual*, Madrid, Dykinson.
- RICO, Francisco (2005), “Sobre la cronología de las novelas de Cervantes”, en *Por discreto y por amigo: Mélanges offerts à Jean Canavaggio*, Madrid, Casa de Velázquez, pp. 159-165.
- RICO, Francisco (coord.) (2000), *Imprenta y crítica textual en el siglo de oro*, Valladolid, Universidad de Valladolid- Secretariado de Publicaciones e Intercambio editorial.
- RILEY, Edward C. (1971), *Teoría de la novela en Cervantes*, Madrid, Taurus, 1981.
- RILEY, Edward C. (1973), “Teoría literaria”, en J. B. Avalle-Arce y E. C. Riley (eds.), *Suma cervantina*, Londres, Támesis, pp. 293-322.
- RIQUER, Martín de (1988), *Cervantes, Passamonte y Avellaneda*. Barcelona: Sirmio.
- RIBLER-PIPKA, Nanete (2016), “Der falsche *Quijote*? Autorschaftsattribuion für spanische Prosa der frühen Neuzeit”, en *DHd 2016 Modellierung, Vernetzung, Visualisierung*, Leipzig, DHd, pp. 212-217.
- RIUS, Leopoldo (1895-1905), *Bibliografía crítica de las obras de Miguel de Cervantes Saavedra*, 3 vols., Madrid, Librería de M. Murillo.
- RODRÍGUEZ CUADROS, Evangelina (1991), “Sobre los entremeses atribuidos a Lope de Vega”, en Ch. Davis y A. Deyermond (eds.), *Golden Age Spanish Literature. Studies in Honour of John Varey by his Colleagues and pupils*, Londres, Queen Mary and Westfield College, pp. 179-189.

BIBLIOGRAFÍA

- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2011), “El *Quijote* de Avellaneda: nuevos índices de atribución a José de Villaviciosa”, *Lemir*, 15, pp. 9-22
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2011a), “Un tríptico presuntamente cervantino: la *Jerusalén*, *Los habladores* y el *Auto de la Virgen de Guadalupe*. *Rivista di filologia e letteratura ispaniche*, 14, pp. 57-76.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2011b), “Cervantes y el entremés de *Los mirones*. Bases objetivas para su atribución”, *Etiópicas*, 7, pp. 57-63.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2011c), “La *Jerusalén de Cervantes*: nuevas pruebas de su autoría”, *Artifara: Revista de lenguas y literaturas ibéricas y latinoamericanas*, 11, pp. 1-6.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2013a), “Dos entremeses de Cervantes: “*Doña Justina y Calahorra*” y “*La cárcel de Sevilla*”, *Artifara: Revista de lenguas y literaturas ibéricas y latinoamericanas*, 13, pp. 5-15.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2013b), “La novela ejemplar *La tía fingida*, atribuida a Cervantes”, *Artifara: Revista de lenguas y literaturas ibéricas y latinoamericanas*, 13 (bis), pp. 59-70.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2016), “El entremés de *La cárcel de Sevilla*, atribuido a Cervantes: edición anotada”, en “*Deste Artife*”: *Estudios dedicados a Aldo Ruffinatto en el IV centenario de las Novelas Ejemplares*, Guillermo Carrascón (ed. lit.), Daniela Capra (ed. lit.), Maria Consolata Pangallo (col.), Iole Scamuzzi (col.), pp. 129-144.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2016), Introducción a la edición modernizada de “*El entremés de los romances*, atribuido a Cervantes”, *Digilec: revista internacional de lenguas y culturas*, 3, pp. 92-106.
- RODRÍGUEZ LÓPEZ-VÁZQUEZ, Alfredo (2019), “*El entremés de los romances*, entre Cervantes y Góngora”, *Atalanta: Revista de Letras Barrocas*, 7 (2), pp. 221-239.
- RODRÍGUEZ MARÍN, Francisco (1905), “Discurso preliminar” en Miguel de Cervantes, *Rinconete y Cortadillo*, Sevilla, Tipografía de Francisco de P. Díaz, pp. 2-243.

BIBLIOGRAFÍA

- RODRÍGUEZ MARÍN, Francisco (1927-1928), “Apéndice IX. El *Entremés de los romances*”, en la *Nueva edición crítica de ‘El ingenioso hidalgo Don Quijote de la Mancha’*, Madrid, Tipografía de la Revista de Archivos, Bibliotecas y Museos, pp. 159-163.
- RODRÍGUEZ-MOÑINO, Antonio (ed.) (1957), “Introducción”, en *Las Fuentes del Romancero general*, tomo III, Madrid, Real Academia Española.
- RODRÍGUEZ-MOÑINO, Antonio (1962), “La carta de Cervantes al Cardenal Sandoval y Rojas”, *NRFH*, 16, pp. 81-89.
- RODRÍGUEZ, Juan Carlos (2003), *El escritor que compró su propio libro. Para leer el Quijote*, Barcelona, Debate.
- ROGEL VIDE, Carlos (2006), *Estudios completos de Propiedad Intelectual. Volumen segundo*, Madrid, Editorial Reus.
- ROJAS CASTRO, Antonio (2017), “Luis de Góngora y la fábula mitológica del Siglo de Oro: Clasificación de textos y análisis léxico con métodos informáticos”, *Studia Aurea*, 11, pp. 111-142.
- ROJAS CASTRO, Antonio (2018), “¿Cuántos Góngora podemos leer? Un análisis contrastivo de la poesía de Luis de Góngora”, *e-Spania*, pp.1-19.
- ROJAS VILLANDRANDO, Agustín (1995), *Viaje entretenido*, ed. Jean Pierre Ressonot, Madrid, Castalia.
- ROJO VEGA, Anastasio (2004), “Propuesta de nuevo autor para *La pícaro Justina*: fray Bartolomé Navarrete O.P. (1560-1640)”, *Dicenda. Cuadernos de Filología Hispánica*, 22, pp. 201-228.
- ROSA, Javier de la y Juan Luis SUÁREZ (2016), “The Life of *Lazarillo de Tormes* and of his Machine Learning Adversities. Non-Traditional Authorship Attribution Techniques in the Context of the *Lazarillo*”, *Lemir*. 20, pp. 373-438.
- ROSE, Phil y Geoffrey Stewart MORRISON (2009), “A response to the UK Position Statement on forensic speaker comparison”, *The International Journal of Speech, Language and the Law*, 16 (1), pp. 139-163.

BIBLIOGRAFÍA

- ROSENBLAT, Ángel (1978), *La lengua del «Quijote»*, Madrid, Gredos.
- RUANO DE LA HAZA, José María (1983), “An Early Rehash of Lope's Peribáñez”, *Bulletin of the Comediantes*, 35, pp. 5-29.
- RUANO DE LA HAZA, José María (1998), “Las dos versiones de *El mayor monstruo del mundo*, de Calderón”, *Criticón*, 72, pp. 35-47.
- RUANO DE LA HAZA, José María (2000), *La puesta en escena en los teatros comerciales del Siglo de Oro*, Madrid, Castalia.
- RUBIO MONTANER, Pilar (1990), “Sobre la necesaria integración de las poéticas de autor en la Teoría de la Literatura”, *Castilla: Estudios de literatura*, 15, pp. 183-197.
- RUDMAN, Joseph (1998), “The state of authorship attribution studies: Some problems and solutions,” *Computers and the Humanities*, vol. 31, pp. 351-365.
- RUFINATO, Aldo (1971), *Funzioni e variabili in una catena teatrale (Cervantes e Lope de Vega)*, Torino, G. Giappichelli.
- RUIZ GARCÍA, Elisa (2015), «Estudio paleográfico y diplomático de autógrafos cervantinos», en Darío Villanueva, en *Autógrafos de Miguel de Cervantes: edición conmemorativa del IV centenario de su muerte*, Madrid, Círculo Científico, 2015, pp. 23-115.
- RUIZ PÉREZ, Pedro (2003), “El concepto de autoría en el contexto editorial”, en Víctor Infantes, François Lopez y Jean-François Botrel (dirs.), *Historia de la edición y de la lectura en España, 1472-1914*, Madrid, Fundación Germán Sánchez Ruipérez, pp. 66-76.
- RUIZ PÉREZ, Pedro (2006), *La distinción cervantina: poética e historia*, Alcalá de Henares, Centro de Estudios Cervantinos.
- RUIZ PÉREZ, Pedro (2019a), “Subjetividad sentimental y sujeto autorial. Trayectoria y niveles en la lírica áurea”, *Theory Now. Journal of Literature, Critique, and Thought*, 2 (1), pp. 159-180.

BIBLIOGRAFÍA

- RUIZ PÉREZ, Pedro (2019b) (coord.), *Autor en construcción. Sujeto e institución literaria en la modernidad hispánica (siglos XVI-XIX)*, Universidad de Zaragoza, Prensas Universitarias de Zaragoza.
- RUIZ URBÓN, Cristina (2007), “Rojas Zorrilla autor de entremeses. *El doctor* como tema de larga herencia dramática”, en *Rojas Zorrilla en su IV centenario: Congreso Internacional (Toledo, 4-7 de octubre de 2007)*, pp. 753-768.
- RUIZ URBÓN, Cristina (2009), “De lo vivido a lo creado: Cervantes, espectador teatral”, en Begoña Regueiro Salgado y Ana M^a Rodríguez (eds.), *Lo real imaginado, soñado, creado: Realidad y literatura en las letras hispánicas*, Madrid, Universidad Complutense, pp. 94-106.
- RUIZ URBÓN, Cristina (2010), “El Entremés de los romances: una atribución cervantina largamente dubitada”, en Javier Blasco, Patricia Marín Cepeda y Cristina Ruiz Urbón (eds.), *Hos ergo versículos feci... Estudios de atribución y plagio*, Madrid, Iberoamericana, 2010, pp. 171-260.
- RUIZ URBÓN, Cristina (2014), “«Oyga mi vida y milagros»: controversias censorias en un polémico pasaje de Sólo en Dios la confianza, de Pedro Rosete Niño”, *Cincinnati Romance Review*, 37 (2014), pp. 93-131.
- RYBICKI, Jan y Maciej EDER (2011), “Deeper Delta across genres and languages: Do we really need the most frequent words?”, *Literary and Linguistic Computing*, 26 (3), pp. 315-321.
- SÁEZ, Adrián J. (2017), “Cuernos para todos: *El retablo de las maravillas* de Cervantes a Quiñones de Benavente”, *Anales cervantinos*, 49, pp. 153-167.
- SÁEZ, Adrián J. (2018), “Introducción”, en Miguel de Cervantes Saavedra, *La tía fingida*, Madrid, Cátedra, pp. 13-80.
- SÁEZ, Adrián J. (2019), “El bueno, el feo y el malo: los libros en Cervantes”, *Orillas*, 8, pp. 203-214.
- SÁEZ, Adrián J. (2020), “Introducción”, en Miguel de Cervantes Saavedra, *Entremeses*, Madrid, Cátedra, pp. 9-67.

BIBLIOGRAFÍA

- SÁNCHEZ GARCÍA, Raquel (2002), “La propiedad intelectual en la España contemporánea, 1847-1936”, *Hispania: Revista española de historia*, 62 (212), pp. 993-1020.
- SÁNCHEZ JIMÉNEZ, Antonio (2001), “Anonimia y censura en el teatro del siglo XVII: el caso de *El diablo predicador*”, *Hispanófila*, 131, pp. 9-19.
- SÁNCHEZ POL, Marta (2006), *Proposta d'un mètode d'estilística per a la verificació d'autoria: els límits de l'estil idiolectal*, tesis doctoral, Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra.
- SÁNCHEZ-ARJONA, José (1898), *Anales del teatro en Sevilla desde Lope de Rueda hasta fines del siglo XVII*, Sevilla, Imprenta de E. Rasco.
- SANDERSON, Conrad y Simon GUENTER (2006), “Short text authorship attribution via sequence kernels, Markov chains and author unmasking: An investigation”, en AA.VV., *Proceedings of International Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Morristown, NJ, Association for Computational Linguistics, pp. 482-491.
- SAUSSURE, Ferdinand de (1896-1952), *Curso de Lingüística general*, publicado por Charles Bally y Albert Sechehaye, con la colaboración de Albert Riedlinger traducción, prólogo y notas de Amado Alonso; edición crítica preparada por Tullio de Mauro, Madrid, Alianza.
- SBARBI, José María (1877), *El refranero general español*, Madrid, Imprenta de A. Gómez Fuentenebro, tomo VII.
- SCHEVILL, Rudolph y Adolfo BONILLA Y SAN MARTÍN (1922), “Introducción al teatro de Cervantes”, en Miguel de Cervantes, *Obras completas. Comedias y entremeses*. Madrid, Gráficas Reunidas, tomo VI.
- SCHEVILL, Rudolph y Adolfo BONILLA Y SAN MARTÍN (1928), “Notas”, en Miguel de Cervantes, *Obras completas. Don Qvixote [sic] de la Mancha*, Madrid, Gráficas reunidas, tomo I.
- SERRANO GÓMEZ, Eduardo (coord.) (2014), *Obras inéditas, anónimas, pseudónimas, póstumas y huérfanas*, Madrid, Reus.

BIBLIOGRAFÍA

- SEVILLA ARROYO, Florencio (1986), “Del *Quijote* al *Rufián dichoso*: capítulos de teoría dramática cervantina”, *Edad de Oro*, V, pp. 217-245.
- SEVILLA ARROYO, Florencio y Antonio REY HAZAS (eds.) (1987), *Teatro completo* de Miguel de Cervantes, Barcelona, Planeta.
- SHERGOLD, N. D. y J. E. VAREY (1985), *Genealogía, origen y noticias de los comediantes de España*, London, Tamesis books.
- SHUY, Roger W. (1993), *Language Crimes: the use and abuse of language evidence in the courtroom*, Cambridge, Blackwell.
- SHUY, Roger W. (2005), *Creating Language Crimes. How Law Enforcement Uses (and Misuses) Language*, Oxford, Oxford University Press.
- SIERRA DE CÓZAR, Ángel (1998), “El posible modelo de Cervantes en las comedias del P. Pedro Pablo de Acevedo”, *Cuadernos de Filología clásica: Estudios latinos*, 15, pp. 561-571.
- SIMÓN DÍAZ, José (1983), *El libro español antiguo: Análisis de su estructura*. Kassel, Reichenberger.
- SIMÓN DÍAZ, José: “El libro en Madrid durante el siglo de Oro”, en *Mundo del libro antiguo*, Madrid, Editorial complutense, 1996, pp. 43-54.
- SIMÓN PALMER, M^a del Carmen (1977), *Manuscritos dramáticos de los Siglos de Oro de la Biblioteca del Instituto del Teatro de Barcelona*, Madrid, CSIC.
- SITO ALBA, Manuel (1983), “La «Commedia dell’Arte», clave esencial en la gestación del *Quijote*”, *Arbor*, CXVI, 465, pp. 7-30.
- SMITH, Peter W. H. y W. ALDRIDGE (2011), “Improving authorship attribution: Optimizing Burrows’ delta method”, *Journal of Quantitative Linguistics*, 18 (1), pp. 63-88.
- SOLER BENITO, Cristina (2016), *Artes escénicas y Derechos de autor*, Madrid, Editorial Reus.

BIBLIOGRAFÍA

- SPASSOVA, Maria (2007), “Las perífrasis verbales en español como marcas distintivas en la atribución forense de autoría”, en *Actas del XXV Congreso Internacional de la Asociación Española de Lingüística Aplicada*, Murcia, Universidad de Murcia.
- SPASSOVA, Maria (2009), “El potencial discriminatorio de las secuencias de categorías gramaticales en la atribución forense de autoría de textos en español”, tesis doctoral, Institut Universitari de Lingüística Aplicada (IULA). Disponible en <https://www.tdx.cat/bitstream/handle/10803/7512/tmss.pdf;jsessionid=18DF4FF3A73A778F69A7F0C02CCEDE10?sequence=1> [Fecha de consulta: 08.12.2021].
- SPASSOVA, Maria Y M. Teresa TURELL (2007), “The use of Morpho-syntactically Annotated tag Sequences as forensic markers of authorship attribution”, Barcelona, Papers de l’IULA.
- STAGG, Geoffrey (2002), “Don Quijote and the «Entremés de los romances»: A Retrospective”, *Cervantes: Bulletin of the Cervantes Society of America*, XXII, 2, pp. 129-150.
- STAGG, Geoffrey (2003), “The Curious Case of the Suspect Epistle”, *Cervantes*, 23 (1), pp. 201-214.
- STAGG, Geoffrey y Daniel EISENBERG (2002), “Entremés de los romances”, *Bulletin of the Cervantes Society of America*, 22 (2), pp. 151-174.
- STAMATATOS, Efstathios (2009), “A Survey of Modern Authorship Attribution Methods”, *Journal of the American Society for Information Science and Technology*, 60, pp. 538-556.
- STAMATATOS, Efstathios, Nikos FAKOTAKIS y George KOKKINAKIS (2000), “Automatic text categorization in terms of genre and author”, *Computational Linguistics* 26 (4), pp. 471-495.
- STAMATATOS, Efstathios, Nikos Fakotakis y George Kokkinakis (2001), “Computer-Based Authorship Attribution Without Lexical Measures”, *Computers and the Humanities*, 35 (2), pp. 193-214.

BIBLIOGRAFÍA

- SVARTVIK, Jan (1968), *The Evans statements: a case for forensic linguistics*, Göteborg, University of Gothenburg Press.
- SVARVIK, Jan (1968), *The Evans Statements: A case for Forensic Linguistics*
- SYVERSON-STROK, Jill (1986), *Theatrical Aspects of the Novel: A Study of Don Quixote*, Valencia, Albatros Hispanofilia.
- TARLINSKAJA, Marina (2014), *Shakespeare and the versification of English Drama, 1561-1642*, Farnham, Ashgate.
- TOMASHEVSKY, Boris Viktorovich (1923), “Pjatistopnyj jamb Pushkina”, en *Ocherki po poetike Pushkina*, Berlin, Epokha, pp. 7-143.
- TURELL, M. Teresa (2003), “El temps aparent i el temps real en estudis de variació i canvi lingüístic”, *Noves SL.: Revista de Sociolingüística*, 3, pp. 1-11.
- TURELL, M. Teresa (2005a), “El plagio en la traducción literaria”, en M. Teresa Turell (ed.), *Lingüística forense. lengua y derecho. Conceptos, métodos y aplicaciones*, Barcelona, Institut Universitari de Lingüística Aplicada, pp. 275-298.
- Turell, M. Teresa (2005b), *Lingüística forense. lengua y derecho. Conceptos, métodos y aplicaciones*, Barcelona, Institut Universitari de Lingüística Aplicada.
- TURELL, M. Teresa (2007), “Plagio y traducción literaria”, *Vasos comunicantes*, 37, pp. 43-54.
- TURELL, M. Teresa (2010a), *Los retos de la lingüística forense en el siglo XXI. In Memoriam Enrique Alcaraz Varó*, Alicante, Universidad de Alicante / Diego Marín Librero-Editor.
- TURELL, M. Teresa (2010b), “The use of textual, grammatical and sociolinguistic evidence in forensic text comparison”, *The International Journal of Speech, Language and the Law*, 17, 2, pp. 211-250.
- TURELL, M. Teresa (2011), “La tasca del lingüista detectiu en casos de detecció de plagi i determinació d'autoria de textos escrits”, *Llengua, societat i comunicació*, 9, pp. 67-83.

BIBLIOGRAFÍA

- TURELL, M. Teresa y Nuria Gavaldà (2013), “Towards an Index of Idiolectal Similitude (or Distance) in Forensic Authorship Analysis”, *Journal of Law and Policy*, 21, 2, pp. 495-514.
- URZÁIZ, TORTAJADA, Héctor (1997), “Un entremés olvidado de Luis Vélez de Guevara: *Los atarantados*”, *Criticón*, 71, pp. 127-157.
- URZÁIZ, TORTAJADA, Héctor (2002), *Catálogo de autores teatrales del siglo*, Madrid, Fundación Universitaria Española, 2 vols.
- URZÁIZ, TORTAJADA, Héctor (2014), “Atribuciones polémicas y falsificaciones literarias: *Melisendra*, Lope y sus censores”, *Cincinnati Romance Review*, 37, pp. 132-153.
- VACCARI, Debora (2003), “Edición de una pieza inédita y de su plan en prosa: el *Entremés del paño*”, *Criticón*, 87-89, pp. 877-885.
- VAREY, John E. (1988), “Sobre un posible diccionario *biográfico* de actores españoles”, en *Varia Bibliographica. Homenaje a Simón Díaz*, Kassel, Reichenberger, pp. 641-644.
- VAREY, John. E. (1957), *Historia de los títeres en España*, Madrid, Revista de Occidente.
- VAREY, John. E. (1987), *Cosmovisión y escenografía: El teatro español en el Siglo de Oro*, Madrid, Castalia.
- VAREY, John. E. (1992), “Del *entrames* al *entremés*”, *Teatro y espectáculo*, pp. 65-79.
- VARO, Carlos (1968), *Génesis y evolución del Quijote*, Madrid, Alcalá ediciones.
- VEGA GARCÍA-LUENGOS, Germán (1991), “Teatro e imprenta en Sevilla durante el siglo XVIII: los entremeses sueltos”, *Archivo Hispalense*, 74 (226), pp.47-98.
- VEGA GARCÍA-LUENGOS, Germán (2002), “La transmisión del teatro en el siglo XVII”, en J. Huerta Calvo (ed.), *Historia del teatro español. I. De la Edad Media a los siglos de Oro*, Madrid, Iberoamericana, pp. 1289-1320.
- VEGA GARCÍA-LUENGOS, Germán (2021), “Las comedias de Lope de Vega: confirmaciones de autoría y nuevas atribuciones desde la estilometría”, *Talía. Revista de estudios teatrales*, 3, pp. 91-108.

BIBLIOGRAFÍA

- VEGA, Lope de (1997), *Comedias de Lope de Vega, Parte I*, dirigida por A. Blecua y G. Serés, Lleida, Millenium - Universitat Autònoma de Barcelona, vol. I***.
- VEGA, María José (dir.) (2004), *Poética y teatro. La teoría dramática del Renacimiento a la Posmodernidad*, Mirabel editorial.
- VÉLEZ-SAINZ, Julio (2000), “El *Recueil Fossard*, la compañía de los Gelosi y la génesis de *Don Quijote*”, *Cervantes: Bulletin of the Cervantes Society of America*, 20 (2), 2000, pp. 31-52.
- VICKERS, Brian (2002), *Counterfeiting Shakespeare: Evidence, Authorship, and “John Funerall Elegy”*, Cambridge, Cambridge UP.
- VIDAL DE VALENCIANO, Cayetano (1883), *El 'Entremés de refranes' ¿es de Cervantes? Ensayo de su traducción*, Barcelona, Librería de Bastinos / Madrid: Librería de Fernando Fé.
- VILLEGAS MORALES, Juan (2000), *Para la interpretación del teatro como construcción visual*, Irvine, Calif., USA, Ediciones de GESTOS.
- VITSE, Marc (1988), “Burla e ideología en los entremeses”, en Luciano García Lorenzo (ed.), *Los géneros menores en el teatro español del Siglo de oro*, Madrid, Ministerio de Cultura, pp. 163-176.
- WARDROPPER, Bruce (1973), “Las comedias”, en J. B. Avallé Arce y E. Riley (eds.), *Suma cervantina*, London, Tamesis, pp. 147-169.
- WILSON, E. M. (1961), “Calderón and the Stage-Censor in the seventeenth Century: A provisional Study”, *Symposium*, XV, pp. 165-184.
- WOOLLS, David (2003), “Better tools for the trade and how to use them”, *Forensic Linguistics. The International Journal of Speech, Language and Law*, 10 (1), pp. 102-112.
- WRISLEY, David Joseph (2016), “Modeling the Transmission of al-Mubashshir Ibn Fātik’s Mukhtār al-Ḥikam in Medieval Europe: Some Initial Data-Driven Explorations”, *Journal of Religion, Media and Digital Culture*, 5 (1), pp. 228-257.

BIBLIOGRAFÍA

- WOOLLS, David (2005), “La equivalencia y la diferencia en la determinación forense de la autoría de textos”, en M. Teresa Turell (ed.), *Lingüística forense, lengua y derecho. Conceptos, métodos y aplicaciones*, Barcelona, Institut Universitari de Lingüística Aplicada, pp. 249-266.
- WOOLLS, David y Malcolm COULTHARD (1998), “Tools for the trade”, *Forensic Linguistics. The International Journal of Speech, Language and Law*, 5 (1), pp. 33-57.
- WRIGHT, David (2013), “Stylistic variation within genre conventions in the Enron email corpus”, *The International Journal of Speech, Language and the Law*, 20 (1), pp. 45-75.
- YNDURÁIN, Domingo (1966), “*Rinconete y Cortadillo*. De entremés a novela”, *Boletín de la Real Academia Española (BRAE)*, 46 (177), pp. 321-332.
- ZAPATA, Juan Manuel (2011), “Muerte y resurrección del autor. Nuevas aproximaciones al estudio sociológico del autor”, *Lingüística y Literatura*, 60, pp. 35-58.
- ZHENG, Rong, Yi QIN, Zan HUANG y Hsinchun CHEN (2003), “Authorship Analysis in Cybercrime Investigation”, *ISI'03 Proceedings of the 1st NSF/NIJ Conference on Intelligence and Security Informatics*, pp. 59-73.
- ZIMIC, Stanislav (1981a), “La biografía satírica en *La guarda cuidadosa*”, *Segismundo*, pp. 95-149.
- ZIMIC, Stanislav (1981b), “Sobre dos entremeses cervantinos: *La elección de los alcaldes de Daganzo* y *El rufián viudo*”, *Anales cervantinos*, 19, pp. 119-160.
- ZIMIC, Stanislav (1982a), “La ejemplaridad de la burla en *El vizcaíno fingido*”, *Segismundo*, pp. 39-56.
- ZIMIC, Stanislav (1982b), “*El retablo de las maravillas*: Parábola de la mentira”, *Anales cervantinos*, 20, pp. 153-172.
- ZIMIC, Stanislav (1983), “*La cueva de Salamanca*: Parábola de la tontería”, *Anales cervantinos*, pp. 135-152.

BIBLIOGRAFÍA

ZIMIC, Stanislav (1984), “La ejemplaridad de los entremeses de Cervantes”, *BSH*, pp. 444-453.

ZIMIC, Stanislav (1992), *El teatro de Cervantes*, Madrid, Castalia.

ZIMIC, Stanislav (1996), *Las novelas ejemplares de Cervantes*, Madrid, Siglo XXI de España editores.

ZORRILLA, José (1880), *Recuerdos del tiempo viejo*, Barcelona, Imprenta de los sucesores de Ramírez y C^a.