

# Análisis cuantitativos aplicados a la poesía: la lírica indubitada sacra de Damián Cornejo y de León Marchante\*

**Zoraida Sánchez Mateos**

Universidad de Valladolid  
zoraida.sanchez@uva.es

Recepción: 09/04/2018, Aceptación: 04/06/2018, Publicación: 11/12/2018

## Resumen

Los análisis cuantitativos permiten delimitar y contrastar de manera objetiva las características estilísticas de corpus textuales. A través de este tipo de estudios, se intentarán definir rasgos morfosintácticos y léxicos que permitan diferenciar los poemas sacros indubitados de Damián Cornejo y de León Marchante. Estos datos serán comparados con los obtenidos en las poesías profanas indubitadas y en la prosa para comprobar si existen semejanzas significativas entre las producciones de un mismo autor. Los resultados establecerán una serie de parámetros que ayudarán a determinar la autoría de las poesías religiosas que se atribuyen a ambos.

## Palabras clave

poesía; sacra; Damián Cornejo; León Marchante; análisis cuantitativos; estilometría

## Abstract

*Quantitative analysis applied to poetry: the undoubted sacred lyric of Damian Cornejo and Leon Marchante*

The quantitative analysis allows to delimit and contrast in an objective way the stylistic characteristics of textual corpus. Through this type of studies, we have defined mor-

\* Esta investigación es el resultado de los trabajos realizados en el seno del proyecto de investigación «Ovidio versus Petrarca: nuevos textos de la poesía erótica española del Siglo de Oro (plataforma digital y edición)», Ref. FFI2015-68229-P.

phonsyntactic and lexical features which allow to differentiate the undoubted sacred poems of Damian Cornejo from Leon Marchante. These data have been compared with those obtained in undoubted profane poetry and in prose in order to check if there are similarities between the productions of the same author. The results have established a series of parameters that will help to determine the authorship of religious poems attributed to both of them.

### Keywords

poetry; sacred; Damian Cornejo; Leon Marchante; quantitative analysis; stylometry

Las innovaciones que se están realizando en el procesamiento de lenguaje natural y en el análisis masivo de textos<sup>1</sup> permiten conocer en profundidad el contenido y la estructura de grandes corpus literarios. La extracción automática de datos cuantificables en el campo de la lírica ha propiciado la aparición de novedosos estudios comparativos, los cuales abarcan desde la evolución de la poesía de un autor hasta el análisis de las obras poéticas creadas en diversas centurias.<sup>2</sup> Investigaciones como las llevadas a término por Florentin Smarandache (2007), Karol Opara (2015) o Sohrab Rezaei y Nasim Kashanian (2017) han delimitado las características (recurrentes y distintivas) sobre el uso que se hace de la gramática y del léxico en las producciones de varios grupos de poetas, ofreciendo así una visión precisa y completa de estas.

Aunque en la literatura hispana la aplicación de la estilometría en poesía todavía es una práctica poco extendida, en los últimos años se están haciendo im-

1. El procesamiento masivo de corpus textuales «*text mining*» abre el camino hacia el análisis distante (Moretti 2013) y el macroanálisis (Jockers 2013) de las obras literarias.

2. Entre los diversos estudios que han aplicado la estilometría a corpus poéticos que abarcan amplios periodos cronológicos, pueden destacarse por su actualidad los realizados por Liu Chao-Lin (2017), Borja Navarro (2016) Pablo Ruíz, Clara Martínez y Thierry Poibeau (2017). Estos dos últimos trabajos tienen como eje principal el análisis de sonetos castellanos.

portantes avances individuales y colectivos. Por un lado, destacan las iniciativas de Laura Hernández (2016), Antonio Rojas (2017) y Marie-Eglantine Lescasse (2018) para delimitar los rasgos estilísticos que singularizan las creaciones en verso de autores canónicos de los Siglos de Oro. Por otro, se están desarrollando dos proyectos de investigación relacionados con el análisis computacional de poemas: ADSO y POSTDATA. El primero, liderado por Borja Navarro, tiene como propósito determinar la semántica y la métrica del soneto áureo castellano; el segundo, dirigido por Elena Gonzalez-Blanco, intenta promover el análisis semántico de corpus poéticos con el objetivo de establecer vías de comunicación entre diferentes tradiciones literarias y culturales.

El empleo de métodos estadísticos de clasificación en las investigaciones sobre literatura también está ayudando a incorporar nuevas vías de estudio (como es el caso de la minería de opinión)<sup>3</sup> y a perfeccionar otras ya existentes como la resolución de problemas de autoría.<sup>4</sup> Este último campo resulta de especial interés para el género poético por los múltiples errores de atribución que se derivan de la transmisión manuscrita<sup>5</sup> e impresa. A pesar de que ya han transcurrido varias décadas desde los primeros trabajos que abordaron el tema, estos todavía no son muy frecuentes.<sup>6</sup> La artificiosidad del lenguaje poético, su tendencia a adaptarse a determinadas estructuras métricas y, sobre todo, la escasa extensión que suelen presentar las composiciones líricas dificultan la extracción automática de patrones lingüísticos distintivos.

3. Una compleja muestra del gran interés que supone para los estudios literarios la incorporación de análisis computacional de las opiniones la aportó Julia Silgen. En 2016, esta joven científica presentó en su blog el post «You must allow me to tell you how ardently I admire and love Natural Language Processing», donde expone cómo evolucionan los sentimientos (positivos y negativos) en la novela *Orgullo y Prejuicio*.

4. El empleo de análisis estadísticos en corpus de literatura que plantean dudas relacionadas con la autoría ha sido puesta en práctica por múltiples investigadores. Pueden citarse a modo de ejemplo los trabajos realizados por Hugh Craig (2009) y Javier Blasco (2016) y el proyecto digital *Computational Analysis, Author Attribution, and the Cheap Repository Tracts of the 18th Century*.

5. Los poetas que trasladaban sus versos a un códice, a veces, no ponían su nombre y añadían poesías de otros escritores sin indicarlo. El conjunto de composiciones resultantes podía ser copiado sin hacer mención a su creador o reproducidos en cadena con la indicación «del mismo». Si se añadían o suprimían cuadernillos, si se intercalaban poemas de distinta atribución en las hojas en blanco o si se guillotaban páginas, se perdía la relación con los verdaderos autores de las obras. Todo ello propiciaba la aparición de poemas anónimos y de falsas atribuciones, las cuales podían tener también su origen en el intento de copistas y de coleccionistas de determinar la autoría de una obra a partir de su bagaje cultural: «estos hermosos tercetos, marmóreos pero vibrantes, han de ser de Argensola sin vacilación; este soneto satírico ¡encaja tan bien con el tono de don Francisco de Quevedo! ¿Aparece en este romance Belardo? Sin duda es de Lope de Vega» (Rodríguez 1968: 40).

6. Entre los trabajos que se consideran pioneros en la aplicación de la estilometría a obras en verso para resolver problemas de autoría se halla el de Michael Smith (1984). También resulta de especial de interés, por el gran corpus poético que analiza, el llevado a término pocos años después por John Burrows y Harold Love (1999).

No obstante, artículos como el de Hugh Craig (2012) y el de Justin Kao y Dan Jurasfsky (2015) son una muestra de cómo los avances que se están produciendo en el campo de la estilometría aportan pruebas objetivas y cuantificables que permiten distinguir textos en verso, las cuales pueden ayudar a determinar la autoría de poesías que se consideran anónimas o que se atribuyen a varios creadores. Para ello, es imprescindible contar con un corpus de poemas indubitados de los escritores que se pretenden analizar. Cuanto mayor sea el número de palabras que poseen las composiciones que lo integran, mayor fiabilidad tendrán los datos obtenidos. Además, como señala Maciej Elder (2017), las obras seleccionadas deben manifestar huellas que singularicen con claridad el estilo de su autor, pues solo así podrán diferenciarse de las de otros poetas y manifestar características similares con las surgidas de la misma mano.

### Problemas de atribución en la obra poética de Damián Cornejo y de León Marchante

En los códigos de producción del Bajo Barroco no primaba la voluntad de plasmar el estilo personal del poeta, sino que se buscaba crear obras acordes a los gustos y a la demanda del público. Esto pudo provocar que las obras de escritores que cultivaban temas semejantes se confundieran de autoría, como sucede en el caso de Damián Cornejo y de Manuel de León Marchante. Su interés por cultivar poemas festivos, su tendencia al prosaísmo y, sobre todo, el hecho de que sus obras pudieron haberse transmitido en los mismos círculos y de manera simultánea (ambos estudiaron y trabajaron en la Universidad de Alcalá)<sup>7</sup> propiciaron que muchas de ellas se adjudicaran a ambos.

La casi total ausencia de poemas manuscritos que se atribuyen al dramaturgo manchego debido, principalmente, a su decisión de quemar estos poco antes de morir, habría dificultado aún más la posibilidad de identificar sus versos. Los poemas dubitados entre el cronista seráfico y el Maestro León ascienden a ochenta y seis: diez se hallan atribuidos al dramaturgo manchego en manuscritos que recopilan también la obra de Cornejo y setenta y seis fueron publicados<sup>8</sup> en sus *Obras poéticas póstumas* (1722 y 1733). La fama de poeta que poseía Marchante, por haber publicación en vida múltiples poemarios sacros,<sup>9</sup> podría haber

7. Existen múltiples puntos de confluencia entre las vidas de Marchante y de Cornejo: «nacidos en torno a 1630, estudiaron y vivieron bastante tiempo en Alcalá, uno era cura, el otro fraile, ambos defensores de la Inmaculada Concepción y tentados de las musas; no sabemos si fueron amigos o rivales, pero es imposible que no se conocieran» (Carreira 2008: 46).

8. Se imprimieron a nombre del Maestro León veintitrés poesías sacras y cincuenta y dos profanas. Estas composiciones dubitadas se distribuyen de la siguiente manera. En el primer volumen (1722), se hallan veinticuatro, en el segundo (1733), se encuentran catorce y, en el tercero (1733), treinta y nueve.

9. León Marchante fue el poeta de la segunda mitad del seiscientos que más publicó en vida. Desde 1658 hasta un par de años antes de su muerte (1680), sus versos sacros fueron llevados a la

provocado que las poesías que se asociaban a él y a un escritor ocasional (como era el caso de Cornejo) se terminaran imprimiendo con su nombre.<sup>10</sup>

Antonio Carreira (2008: 51) ha señalado algunas de las características que diferencian su producción poética. En fray Damián «el tono, muy desenfadado, no afecta unción como el de Marchante, sino que es demoledor [...] Hay, pues, en el mejor Cornejo un elemento esencial que no aparece en Marchante: la ironía. También la reflexión metapoética». Aunque el empleo de estudios cualitativos tradicionales podría ayudar a identificar la verdadera autoría de sus poemas dubitados, es necesario realizar una investigación más pormenorizada y objetiva de sus rasgos estilísticos. En este artículo, se analizará (a través de análisis cuantitativos) las poesías sacras indubitadas de fray Damián<sup>11</sup> y del Maestro León con el propósito de demostrar que existen diferencias significativas entre ambas<sup>12</sup> y que estas comparten patrones comunes con los versos profanos y la prosa<sup>13</sup> de sus autores.

## La estilometría aplicada a la lírica indubitada sacra de Damián Cornejo y León Marchante

### *Metodología aplicada y corpus analizado*

La escasa extensión que suelen poseer las composiciones líricas propicia la aparición de errores en los análisis estadísticos de clasificación que miden las palabras más frecuentes de un texto (Eder 2015). Con el propósito de evitar resultados equívocos, se ha optado por estudiar otros parámetros estilísticos relacionados con el uso de la morfosintaxis, del vocabulario distintivo y de la semántica, ya

---

impresa en diecinueve ocasiones. Además, el dramaturgo manchego imprimió las composiciones presentadas para la justa de 1657 (realizada en Alcalá de Henares), un poema de temática profana, una carta a un amigo suyo y la relación de un festejo taurino (Bègue 2010).

**10.** La popularidad que alcanzó Marchante como poeta sacro hizo que otros autores firmaran con su nombre composiciones que le eran ajenas. De tal suceso, se queja el dramaturgo manchego en su correspondencia: «ayer me remitió un poeta de Medellín (porque nació bonete) llorando, más que Jeremías sobre Jerusalén, en razón de que declaré que no es mío el villancico que está al talón del impreso; porque siendo fruto de su cosecha, la calumnia cruel ha divulgado que es mío» (León 1903: 516).

**11.** La lírica sacra en la producción poética de Cornejo es bastante escasa. El corpus de sus poemas asciende a unas doscientas composiciones, pero solo treinta y ocho son de temática religiosa. De estas, nueve se consideran indubitadas y veintinueve aparecen también atribuidas al Maestro León: veintitrés se imprimieron en las *Obras póstumas de Marchante* (1722-1733) y seis se atribuyen a ambos en manuscritos. La mayoría de la poesía divina de fray Damián fue editada por Klaus Pörtl (Cornejo 1978).

**12.** El reducido número de poemas sacros que se atribuyen exclusivamente a Cornejo y las múltiples poesías religiosas que disputa con León Marchante han hecho que investigadores como Antonio Carreira (2008: 53) pongan en cuestionamiento su atribución.

**13.** La prosa de fray Damián y Marchante se ha conformado a partir de algunas de las cartas amorosas que constituyen el extenso epistolario que León Marchante habría escrito, en un registro informal, a su prima monja (Foulché-del Bosc 1903) y de diversos capítulos, relacionados con la vida de Santa Rosa, del tercer tomo de la *Crónica seráfica* (Cornejo 1686).

que tales elementos se ven menos condicionados por las diferencias que existen en la longitud de las obras que se analizarán. Además, la extracción de múltiples características gramaticales y léxicas pueden complementarse y contrastarse para conformar un perfil lingüístico más completo y fiable de Damián Cornejo y de León Marchante.

En la presente investigación, se ha transcrito y normalizado un corpus de dieciséis mil palabras de lírica indubitada (seis mil de sacra y diez mil de profana) y treinta mil de la prosa de ambos autores.<sup>14</sup> Los textos seleccionados poseen más de cien palabras para garantizar que los programas de análisis cuantitativos utilizados (R, Treetagger, ContaWords y LIWC) sean capaces de extraer datos relevantes tanto en su estudio individual como en el de conjunto. En su codificación, se ha utilizado la inicial del apellido de Cornejo y de León («l» y «c»), las letras «si» y «pi» para indicar si se trata de poesías divinas o humanas indubitadas y el número asignados a cada una.<sup>15</sup> Las obras en prosa incluyen el mismo sistema de marcación e integran también las primeras letras de su título: C\_cron y L\_car.

### *Análisis morfosintáctico*

Gracias a la herramienta Treetagger, se ha podido obtener (con un grado de fiabilidad mayor al 90%) los porcentajes de algunas de las principales categorías gramaticales<sup>16</sup> que aparecen en la lírica indubitada y en las narraciones de fray Damián y del Maestro León. Estos han permitido generar varios análisis de multivariantes que han sido representados a través de biplots,<sup>17</sup> los cuales persiguen un doble propósito. Por un lado, mostrar las relaciones morfosintácticas que se manifiestan entre las producciones en verso y en prosa de un mismo autor y entre las poesías sacras de ambos. Por otro, probar la eficacia de este tipo de pruebas para clasificar obras de diversa extensión, casi todas ellas poseen una longitud bastante inferior a las dos mil palabras que aconseja Maciej Eder (2017) para distinguir con seguridad obras escritas por distintas manos. A continuación, se ofrecen dos gráficos (uno para cada poeta) con los resultados obtenidos al analizar la gramática de sus diversas creaciones.

14. El corpus textual de cada autor asciende a veintitres mil palabras (quince mil de prosa, cinco mil de poesía humana y tres mil de sacra), una cantidad suficiente para poder trazar con fiabilidad y precisión su perfil lingüístico.

15. Se han elaborado dos tablas (ver Anexo: Tabla 3 y Tabla 4) donde se especifican los títulos de los poemas analizados en esta investigación, se concreta su localización, se indica su extensión, el tipo de métrica y la codificación empleada para identificarlos.

16. Se han revisado (parcialmente) los resultados de los análisis morfosintácticos aportados por el programa Treetagger y el margen de error en ninguno de los textos seleccionados supera el diez por ciento del total.

17. La principal ventaja que aporta el biplot frente a otros tipos de pruebas estadísticas es que permite visualizar en un mismo gráfico el resultado de combinar múltiples variables. El número de estas siempre ha de ser igual o mayor que el de textos que se analicen.

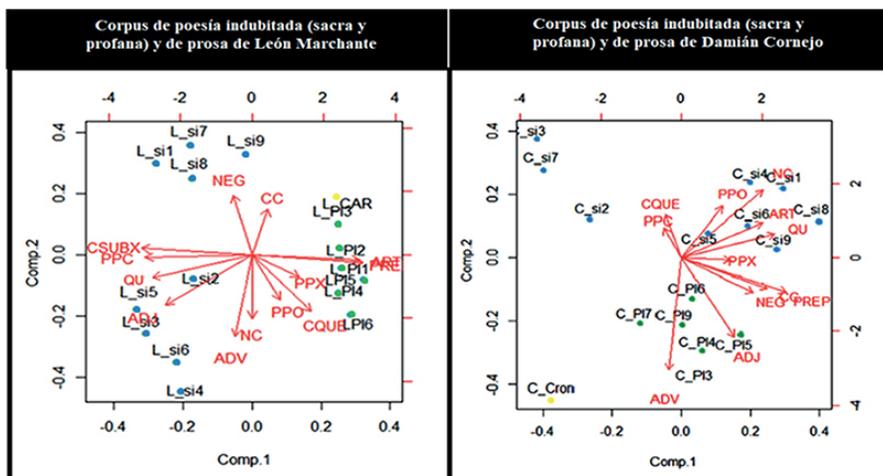


Figura 1. Análisis de multivariantes (biplot 1).

Se observan patrones comunes en las producciones de Cornejo y de Marchante. La temática de sus poemas parece influir de manera decisiva en el uso de la gramática; las composiciones humanas y las religiosas constituyen bloques independientes, aunque estas últimas se muestran más dispersas. Además, los versos que poseen más semejanzas con las narraciones (L\_car y C\_cron) son los de contenido profano,<sup>18</sup> posiblemente, por su estilo más prosaico y su mayor extensión.

En el biplot que recoge las obras de Maestro León, puede verse cómo sus poesías humanas se aglutinan en los cuadrantes de la derecha, que presentan un mayor número de artículos, preposiciones y pronombres. Respecto a su lírica divina, cabe señalar que conforma dos grupos. Las poesías situadas en la mitad superior (L\_si1, L\_si7, L\_si8 y L\_si9) tienen una longitud parecida y se caracterizan por incluir un número más elevado de marcas de negación, mientras que los que se hallan en la parte inferior (L\_si2, L\_si3, L\_si4, L\_si5 y L\_si6) albergan más cuantificadores, adverbios y adjetivos.

En el análisis de multivariantes que recopila las creaciones de fray Damián, se observa que gran parte de sus composiciones (divinas y humanas) se sitúan en la parte

18. Los resultados aportados en el estudio de multivariantes han sido corroborados con los obtenidos en el programa CopyCatch Gold. Las obras en prosa de Cornejo y de Marchante poseen en este un porcentaje de similitud más elevado con los poemas profanos (59% y 63% respectivamente) que con los sacros: 45% y 46%. Además, la semejanza entre los versos humanos y los divinos de fray Damián es algo más alta que en los del Maestro León: 48% y 46%.

derecha del gráfico. Tal dato pone de manifiesto que existe una mayor similitud entre estas que entre las del dramaturgo manchego. Los poemas profanos de Cornejo contienen más adverbios y adjetivos y los sacros (exceptuando el bloque constituido por C\_si2, C\_si3 y C\_si7) utilizan más nombres comunes, artículos y cuantificadores.

Para poder comprobar si se aprecian diferencias importantes entre los corpus de lírica sacra indubitada del cronista seráfico y del Maestro León, se han sometido a dos pruebas. En la primera, se han confrontado en un análisis de multivariantes y, en la segunda, se ha añadido al biplot resultante la poesía religiosa de Sor Juana Inés de la Cruz<sup>19</sup> con el objetivo de poder determinar si su distancia aumenta o disminuye al incluir esta. Los resultados de ambos estudios se muestran en la figura 2.

En el gráfico de la izquierda, se conforman tres bloques independientes. Por un lado, se hallan los poemas divinos del dramaturgo manchego, menos L\_si9, los cuales destacan por utilizar con más frecuencia artículos, adverbios, adjetivos, cuantificadores, interrogativos y clíticos. Por otro, se encuentran la mayoría de las poesías sacras del obispo de Orense (exceptuando C\_si2, C\_si3 y C\_si7), que presentan un mayor uso de demostrativos, preposiciones, posesivos y pronombres personales. Por último, en los cuadrantes de la derecha, aparecen cuatro composiciones: una de marchante (L\_si9) y las tres últimas citadas de fray Damián.

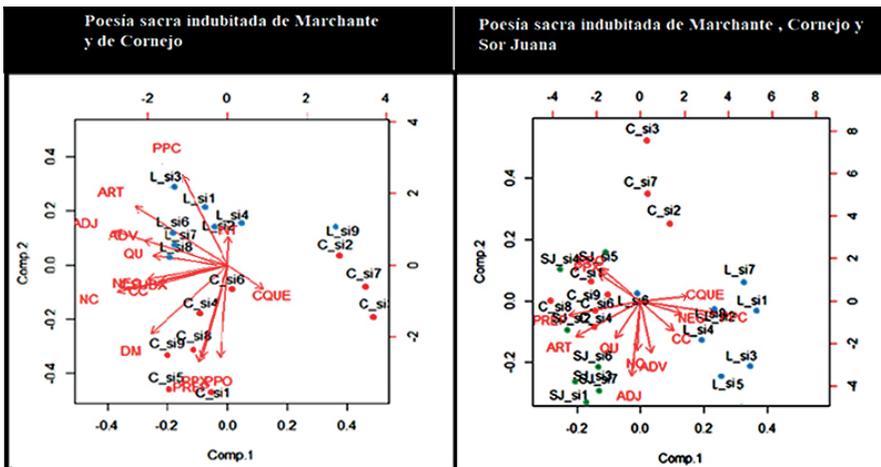


Figura 2.  
Análisis de multivariantes (biplot 2).

19. La lírica sacra indubitada de la escritora novohispana ha sido elegida porque es contemporánea a la de Cornejo y Marchante y porque en sus versos se puede apreciar la influencia de este último (Tenorio 2002). Los datos sobre los poemas seleccionados de Sor Juana pueden consultarse en la tercera tabla del «Anexo».

La incorporación del corpus de la escritora novohispana, en el biplot de la derecha, mantiene separados todos los textos del Maestro León y los del cronista seráfico. No obstante, C\_si2, C\_si3 y C\_si7 continúan alejados del resto de los atribuidos al poeta palentino.<sup>20</sup> Cabe señalar, además, que la morfosintaxis de las poesías sacras de Sor Juana se asemeja bastante a la del Cornejo e, incluso, algunos de sus poemas (SJ\_si2, SJ\_si4 y SJ\_si5) llegan a entremezclarse, pues poseen una frecuencia parecida de preposiciones, artículos, pronombres y clíticos. A pesar de las similitudes que comparten sus producciones divinas, estas pueden distinguirse con claridad si no se añaden en el análisis de multivariantes las poesías de Marchante (ver Anexo: Figura 7).

Tras comprobar que existen diferencias relevantes entre los versos religiosos del cronista seráfico y los del Maestro León, se ha procedido a identificar qué elementos gramaticales son los más recurrentes en estos. Con el propósito de poder delimitar con mayor exactitud los datos y para que estos puedan ser cotejados en futuras investigaciones con los de otros poemas (dubitadas e indubitadas) de ambos, se han extraído las medias de los porcentajes obtenidos y se han representado en dos gráficos.

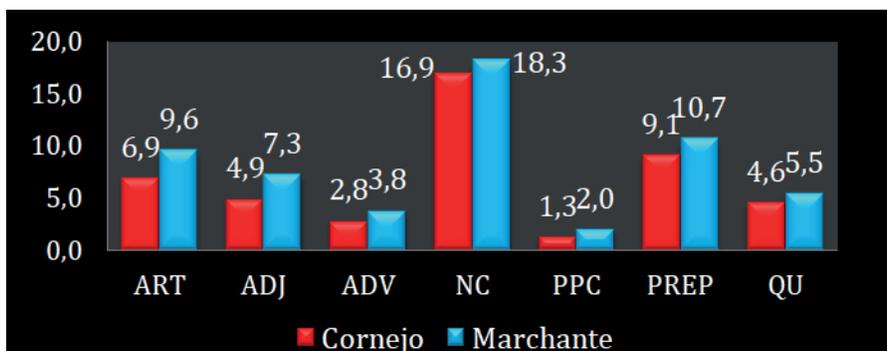


Figura 3.  
Análisis morfológico (I).

20. La división de la lírica religiosa indubitada de Cornejo en dos grupos podría estar relacionada con las siguientes causas. Por un lado, los versos de C\_si3 están constituidos por títulos de comedias y, por tanto, el uso que se hace en estos de la morfosintaxis no manifiesta el estilo personal de su autor. Por otro, su mayor extensión (como también sucede con C\_si2 y C\_si7) propiciaría que aparecieran diferencias en el empleo de las categorías gramaticales entre este y las poesías sacras de menor longitud de fray Damián.

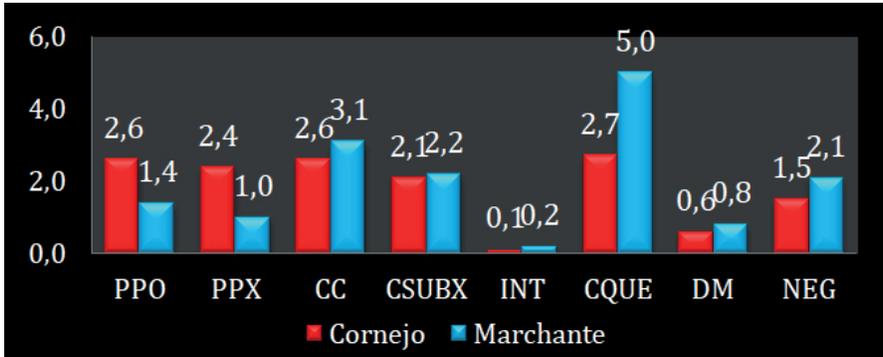


Figura 4.  
Análisis morfosintáctico (II).

Todas las categorías gramaticales seleccionadas son más abundantes en la lírica religiosa del Maestro León.<sup>21</sup> No obstante, las medias son bastante similares, excepto en el uso de artículos y adjetivos. Tal equidad<sup>22</sup> podría ser justificada si se probara que ambos autores empleaban ciertas pautas de escritura a la hora de componer poesía sacra. Las principales diferencias, por tanto, deberían hallarse en elementos morfosintácticos más difíciles de controlar. En la figura 4 se recogen los porcentajes de algunos de ellos.

Se observa una distancia significativa en la proporción del uso de posesivos y de los pronombres personales que hace fray Damián y en el empleo de la conjunción «que» y de las marcas de negación por parte de Marchante. Estos datos coinciden con los aportados en sus poesías humanas indubitadas y en sus prosas. Con el propósito de seguir estableciendo diferencias en el uso que hacen ambos de la gramática, se han analizado las combinaciones más recurrentes de nombres, verbos y adjetivos mediante nueve bigramas: N+N, N+V, N+A, V+N, V+V, V+A, A+N, A+V y A+A. Los resultados obtenidos se ofrecen en la figura 5.

21. Se han comparado tales datos con los resultados de la lírica indubitada profana de fray Damián y el Maestro León y existen contrastes importantes. El dramaturgo manchego en sus versos humanos destaca en el uso de sustantivos comunes, preposiciones y artículos, mientras que el cronista franciscano utiliza más adjetivos, adverbios y cuantificadores. Lo mismo sucede, exceptuando en el empleo de los cuantificadores, en la *Crónica seráfica* y en el epistolario de Marchante.  
22. Si se extraen del corpus de Cornejo los poemas C\_si2, C\_si3 y C\_si7, la varianza disminuye de manera considerable y se obtienen unas medias aún más similares a las de Marchante: ART (8,3), ADJ (6,4), ADV (3,5), NC (21,3), PPC (1,5) y QU (5,6).

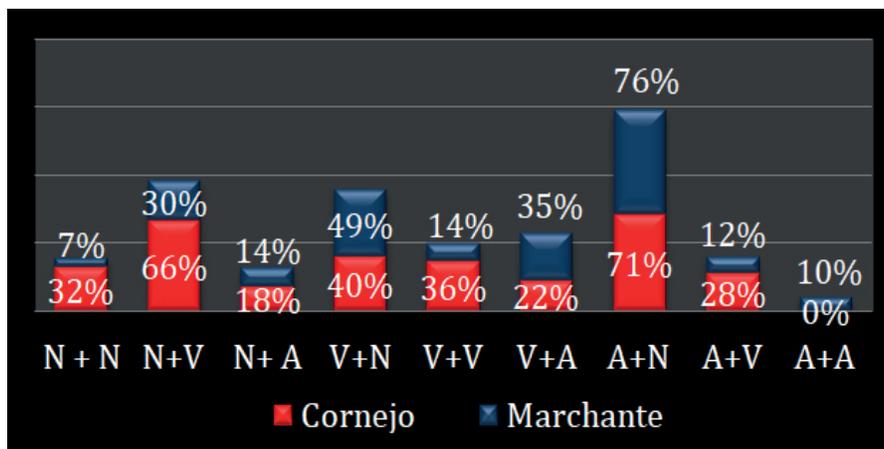


Figura 5.  
Frecuencia de bigramas.

El bigrama más recurrente en sus poesías indubitadas sacras es la combinación de un adjetivo más un nombre (A+N) y el menos utilizado es el de dos adjetivos seguidos (A+A). Cornejo emplea con mucha más frecuencia la estructura de un sustantivo más otra categoría (nombre, verbo o adjetivo) y Marchante explota más la de verbos con adjetivos y nombres. Tales resultados coinciden en su mayoría<sup>23</sup> con los obtenidos en sus versos humanos. No sucede lo mismo con las narraciones de ambos. El único bigrama que ofrece datos similares a los recientemente comentados es el de «N+A», pues este aparece con mayor frecuencia en la *Crónica Seráfica* que en el epistolario del Maestro León.

#### *Análisis léxico-semántico*

Se ha realizado el análisis del léxico y de la semántica de la lírica sacra indubitada de fray Damián y Marchante para completar los datos estilísticos aportados en el estudio morfosintáctico y comprobar si existen diferencias significativas en el vocabulario empleado por ambos. A partir del programa LIWC se ha indagado en el tipo de emociones (positivas y negativas) que predominan en sus poemas, en su subjetividad, en el contenido social de estos y en su riqueza léxica. Los resultados se han contrastado con su poesía profana indubitada y su prosa con el objetivo de establecer relaciones entre las producciones de cada autor. En la tabla 1 se ofrecen los porcentajes obtenidos.

23. Cornejo en sus poemas indubitados profanos hace un menor uso que Marchante de los bigramas «N+V» y «V+V», mientras que el dramaturgo manchego emplea menos las combinaciones de «A+N» y «A+V».

**Tabla 1.** Análisis léxico-semántico

	Self-Reference	Social Words	Positive emotions	Negative emotions	Ratio Type-Token*
<b>Damián Cornejo</b>					
Poesía sacra	0.39	0.41	0.17	0.58	38.1
Poesía profana	1.21	0.93	0.21	0.56	39.3
Prosa religiosa	0.13	0.44	0.46	0.70	37.5
<b>León Marchante</b>					
Poesía sacra	0.31	0.74	0.48	0.83	34.9
Poesía profana	0.54	0.99	0.24	0.66	37.9
Prosa profana	1.57	0.94	0.28	0.33	34.7

\* Como la extensión de los corpus analizados influye en la obtención del porcentaje de riqueza léxica, se han equiparado la longitud (2890 palabras) de todos los textos seleccionados.

En primer lugar, puede señalarse que la lírica de Cornejo, especialmente la profana, da más relevancia al uso de la primera persona que la de Marchante. No sucede lo mismo con la prosa, ya que la correspondencia íntima del Maestro de León posee una presencia elevada del yo. En segundo lugar, cabe destacar que las composiciones del dramaturgo manchego manifiestan una mayor voluntad de conectar emocional y socialmente con el receptor, sobre todo, en sus versos sacros.<sup>24</sup> En tercer lugar, respecto a las emociones es importante mencionar que en las creaciones de ambos autores, exceptuando la crónica de fray Damián, predominan los sentimientos positivos. Estos, al igual que los negativos, son más abundantes en las poesías de Marchante.

Por último, si se observa la riqueza léxica, las obras del obispo de Orense, principalmente su poesía humana, albergan una mayor diversidad de vocabulario. No obstante, los porcentajes no se distancian demasiado de los del Maestro León, pues recibieron una formación muy parecida y sus creaciones no buscan la erudición. Tras determinar la existencia de diferencias significativas en la lírica sacra de Cornejo y Marchante y concretar algunas de las semejanzas que esta presenta con sus poemas profanos y su prosa, se ha procedido a recopilar las palabras que se reiteran en sus obras y que no comparten los corpus de ambos. Con el objetivo de aportar una muestra equitativa de estas, se ha restringido el número de ejemplos que aparecen en la tabla 2 a diez.

**24.** La capacidad de León Marchante para ganarse el beneplácito del público hizo que compusiera versos sacros para las Capillas Reales (1661-1679) y que se imprimieran múltiples pliegos sueltos de sus poesías religiosas.

Tabla 2. Palabras distintivas entre Cornejo y León

Poesías sacras (léxico común)	Poesía sacra y profana	Poesía sacra y prosa	Poesía (sacra y profana) y prosa	
León Marchante	<i>Barro</i>	<i>Barro</i>	Aldea	Condición
	Ratos	Oscuro	Aliño	Coral
		Polvo	Chanza	León (animal)
			Deidad	
			<i>Diablo</i>	
			Gallego	
			Madera	
			Parto	
			Plomo	
			Seglar	
Damián Cornejo	Clavo	Borrar	Desdicha	Afán
	<i>Llagas</i>	Calma	Juez	Conforme
	<i>Pueblo</i>	Cristales	<i>Llagas</i>	<i>Dorada</i>
	Pureza	<i>Dorada</i>	Martirio	Inmenso
	Tormentos	Madeja	Omnipotencia	Estruendo
	Vientre	Manteles	Perpetuo	Hacienda
		Sentidos	Práctica	Mesa
		Toledana	Propósito	Peña
			<i>Pueblo</i>	
			Zozobra	

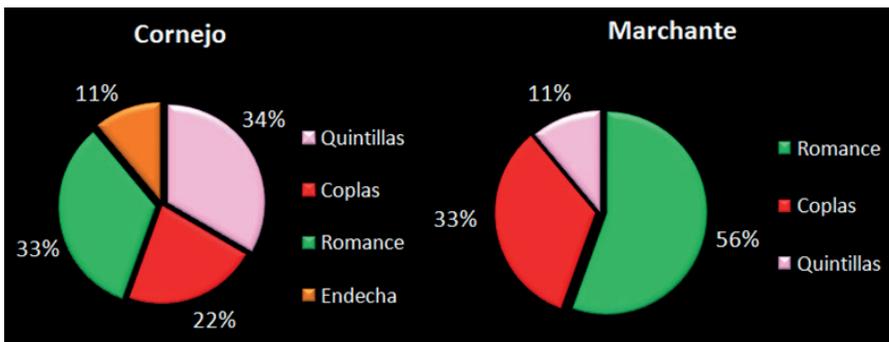
Las creaciones de fray Damián manifiestan un mayor número de vocabulario distintivo. Llama la atención la gran cantidad de términos que se hallan tanto en su poesía (profana y sacra) como en su prosa y también los que aparecen en sus versos humanos y divinos. Además, las palabras que comparten los poemas sacros del obispo de Orense son más del doble que las del dramaturgo manchego. En el corpus de este último, solo su lírica religiosa y sus cartas presentan múltiples palabras comunes. Resulta relevante señalar, sobre todo para

las futuras investigaciones que se realizarán sobre las composiciones dubitadas entre ambos, la preferencia de Cornejo por usar «pueblo» frente a León que opta por «aldea».

A partir del vocabulario extraído del corpus de ambos poetas, se pueden establecer una serie de relaciones semánticas. Las obras del escritor palentino parecen mostrar una tendencia por describir detalles de la fisonomía de los personajes (madeja, vientre, llagas...) y de su entorno:<sup>25</sup> mesa, peña, manteles, cristales y clavo. Además, en sus producciones religiosas se pretende captar los infortunios que estos sufren: tormentos, martirio, desdicha y zozobra. Respecto a las producciones del dramaturgo manchego, es importante señalar su especial interés hacia la descripción de materiales (barro, plomo, madera, polvo y coral) y su afán por identificar entidades o profesiones:<sup>26</sup> seglar, deidad y león.

### *Análisis métrico*

Con el propósito de completar la caracterización de los poemas religiosos indubitados de Marchante y de Cornejo, se ha procedido a delimitar las formas métricas que se utilizan en ellos. A pesar de que el número de textos seleccionados es bajo (nueve por autor), conocer los porcentajes de estas puede aportar datos relevantes sobre sus preferencias en el uso de los metros. Los resultados obtenidos se han representado a través de dos gráficos (figura 6).



**Figura 6.**  
Análisis métrico.

**25.** Las palabras distintivas que aparecen en los poemas profanos indubitados de Cornejo (ver Anexo: Tabla 4) y en los fragmentos transcritos de la *Crónica Seráfica* muestran también un especial interés por retratar el físico de los personajes (flaqueza, mejilla, barba o flaca) y por mostrar con exactitud los colores «ceniciento» o «carmín» o los elementos del entorno: cierzo, andrajo o almohada.

**26.** En los poemas humanos indubitados de Marchante y en sus cartas, se incluyen palabras no compartidas con los corpus de fray Damián que hacen referencia a profesiones o individuos: alcalde, mesonero, carretero, galeno, comisario, pícaro, sabio, preso o chulo.

Aunque en las composiciones de ambos escritores predominan las quintillas, los romances y las coplas, se observan diferencias significativas en sus porcentajes. La proporción de quintillas en la poesía sacra indubitada de Cornejo<sup>27</sup> triplica la de Marchante, siendo este el metro más utilizado por el cronista seráfico. El dramaturgo manchego, por el contrario, realiza un mayor empleo de los romances (56%) y de las coplas (33%) y no usa endechas. Se han comparado estos datos con los extraídos en las poesías religiosas dubitadas entre ambos y se ha comprobado que los metros más frecuentes en ellas son también los romances (55%) y las quintillas (33%).<sup>28</sup>

## Conclusiones

Los avances que se están llevando a término en el procesamiento de lenguaje natural permiten desarrollar estudios de corpus textuales más complejos y fiables. Estos ayudan a distinguir, mediante datos objetivos, incluso creaciones breves de escritores con un estilo, una formación y un contenido similar. Los análisis cuantitativos aplicados a la morfosintaxis y al léxico de las producciones literarias de fray Damián y del Maestro León han permitido singularizar sus poemas sacros indubitados y establecer elementos comunes entre estos y los versos profanos y las narraciones de cada autor.

Gracias al uso de biplots, se han podido visualizar las diferencias morfosintácticas que existen entre las poesías divinas seleccionadas de Cornejo y las de Marchante. Tales rasgos han sido corroborados a través de la obtención de las medias de las principales categorías gramaticales que aparecen en el conjunto de estas, las cuales coinciden (en algunos casos) con las de las poesías profanas. En el uso de los bigramas, también se han observado similitudes importantes entre los versos religiosos y los humanos de un mismo escritor. El cronista seráfico utiliza con mucha más frecuencia la estructura de un sustantivo más otra categoría (nombre, verbo o adjetivo) y el dramaturgo manchego explota más la combinación de verbos con adjetivos y nombres.

El análisis cuantitativo del léxico ha ofrecido más datos relevantes sobre sus obras. Por un lado, ha puesto en evidencia que los poemas de fray Damián, en especial los profanos, contienen más referencias a la primera persona que los de

**27.** En la poesía profana indubitada de Cornejo los metros más empleados son los romances (31%), los sonetos (27%) y las décimas (17%). Además, en las composiciones de temática humana se utiliza una mayor diversidad de formas métricas (ovillejos, silva o septeto), las cuales no adquieren un papel significativo en el total (8%).

**28.** Los romances, las quintillas, los villancicos y las seguidillas son las formas métricas más recurrentes en la lírica religiosa cómico-festiva del Bajo Barroco español. Exceptuando las quintillas, el resto de estas están «históricamente ligadas en sus orígenes a la producción lírica popular y tradicional, de difusión básicamente oral y mayoritariamente escrita para ser cantada» (López 2011: 77).

Marchante. Por otro, ha permitido mostrar que las poesías de este último, sobre todo las de temática sacra, poseen una mayor voluntad de conectar emocional y socialmente con el receptor y, por ello, intentan retratar más sentimientos, tanto positivos como negativos. Por último, ha puesto de manifiesto que las creaciones del obispo de Orense, principalmente su poesía humana, albergan una mayor diversidad de vocabulario.

La extracción de las palabras distintivas que aparecen en los poemas divinos estudiados ha ayudado, además, a confirmar los vínculos que existen entre las composiciones en prosa y en verso de cada autor, siendo mucho más elevado el número de estas en los de Cornejo. La semántica de los términos recopilados parece mostrar que las poesías sacras del poeta palentino, al igual que sucede en las profanas, tienden a describir con detalles la fisonomía de los personajes y el entorno y las del dramaturgo manchego manifiestan un interés por identificar materiales, entidades o profesiones.

Los diversos análisis gramaticales y léxicos empleados en este artículo pueden ayudar, en futuras investigaciones, a determinar la verdadera autoría de los múltiples poemas dubitados (sacros y profanos) entre León Marchante y fray Damián. Además, estos también pueden contribuir a seguir indagando en cómo aplicar herramientas estadísticas al estudio de corpus textuales breves para extraer información objetiva y cuantificable sobre su contenido o sus características estilísticas.

## Bibliografía

- BÈGUE, Alain, «Relación de la poesía española publicada entre 1648 y 1750», *La luz de la razón: literatura y cultura del siglo XVIII: a la memoria de Ernest Lluch*, Aurora Egido, Zaragoza, IFC, 2010.
- BURROWS, John y LOVE, Harold, «Attribution Tests and the Editing of Seventeenth-Century Poetry», *The Yearbook of English Studies*, 29, Modern Humanities Research Association, Cambridge, 1999, 151-175.
- BLASCO PASCUAL, Javier, «Avellaneda desde la estilometría», *Cervantes, los viajes y los días*, Pedro Ruiz, Sial Ediciones, 2016.
- CARREIRA, Antonio, «La obra poética de Damián Cornejo: cuatro manuscritos más y uno menos», *Criticón*, 103-104 (2008), 39-54.
- CORNEJO, Damián, *Crónica seráfica y vida del glorioso patriarca San Francisco y de sus primeros discípulos*, 3, Madrid, Imprenta de Juan García Infanzón, 1686.
- , *Das lyrische Werk des Damian Cornejo (1629-1707). Erster Teil*, Klaus Pörtl, München, Wilhelm Fink, 1978.
- CRAIG, Hugh, «George Chapman, John Davies of Hereford, William Shakespeare, and «A Lover's Complaint»», *Shakespeare Quarterly*, 63, 2 (2012), 147-174.
- CRAIG, Hugh y KINNEY, Arthur, *Shakespeare, Computers, and the Mystery of Authorship*, Nueva York, Cambridge University Press, 2009.
- CRUZ, Juana, *Villancicos: Lírica coral*, México, Fondo de Cultura Económica, 2006.
- , *Obras completas de Sor Juana Inés de la Cruz*, 1. Lírica personal, Antonio Alatorre, México, Fondo de Cultura Económica, 2012.
- EDER, Maciej, «Does size matter? Authorship attribution, small samples, big problema», *Literary and Linguistic Computing*, 30, 2 (2015), 167-182.
- , «Short Samples in Authorship Attribution: A New Approach», *Digital Humanities* (2017, Montreal), 221-223. <<https://dh2017.adho.org/abstracts/341/341.pdf>>.
- JOCKERS, Matthew, *Macroanalysis, Digital Methods and Literary History*, Illinois, University of Illinois Press, 2013.
- HERNÁNDEZ LORENZO, Laura, «The Poetic Word of Fernando de Herrera. An Approach through Corpus and Computational Linguistics», *EPiC Series in Language and Linguistics*, 1 (2016), 170-180.
- KAO, Justin y JURASFSKY, Dan, «A computational analysis of poetic style: Imagism and its influence on modern professional and amateur poetry», *LILT*, 12, 3 (2015), 1-31. <<http://csli-lilt.stanford.edu/ojs/index.php/LiLT/article/view/56/45>>.
- LEÓN MARCHANTE, Manuel, *Letras de los villancicos de Navidad que se han de cantar en la Santa Iglesia de Toledo*, Toledo, Imprenta de Francisco Calvo, 1662.
- , *Obras poéticas póstumas que a diversos asuntos escribió el maestro León Marchante*, 1, Madrid, Real Capilla de su Majestad, 1722.

- , *Obras poéticas póstumas que a diversos asuntos escribió el maestro León Marchante*, 2, Madrid, Real Capilla de su Majestad, 1733.
- , «Coplas de “Trescientas cosas más”», *Revue Hispanique*, 10 (1903), 234-235.
- LESCASSE, Marie-Eglantine, «L'emploi du concept de «propiedad» dans la polémique gongorine», *e-Spania*, 29 (2018). <<https://journals.openedition.org/e-spania/27486>>.
- LIU, Chao-Lin, «Flexible Computing Services for Comparisons and Analyses of Classical Chinese Poetry», *Digital Humanities* (2017, Montreal), 507-510. <<https://dh2017.adho.org/abstracts/612/612.pdf>>.
- LÓPEZ GUIL, Itziar, *Poesía religiosa cómico-festiva del bajo barroco español: estudio y antología*, Bern, Peter Lang, 2011.
- MORETTI, Franco, *Distant Reading*, Verso, London, 2013.
- NAVARRO COROLADO, Borja, «Hacia un análisis distante del endecasílabo áureo: patrones métricos, frecuencias y evolución histórica», *Rhythmica. Revista española de métrica comparada*, 14 (2016), 89-118. <<http://revistas.uned.es/index.php/rhythmica/article/view/18459>>.
- OPARA, Karol, «Grammatical rhymes in Polish poetry: A quantitative analysis», *Digital Scholarship in the Humanities*, 30, 4 (2015), 589-598.
- REZAEI, Sohrab y KASHANIAN, Nasim, «A stylometric analysis of iranian poets», *Theory and Practice in Language Studies*, 7, 1 (2017), 55-64.
- RODRÍGUEZ-MOÑOINO, Antonio, *Construcción crítica y realidad histórica en la poesía española de los siglos XVI y XVII*, Madrid, Castalia, 1968.
- ROJAS CASTRO, Antonio, «Luis de Góngora y la fábula mitológica del Siglo de Oro: clasificación de textos y análisis léxico con métodos informáticos», *Studia Aurea*, 11 (2017), 111-142. <<https://doi.org/10.5565/rev/studiaaurea.260>>
- RUIZ FABO, Pablo, MARTÍNEZ CANTÓN, Clara y POIBEAU, Thierry, «Distant Rhythm: Automatic Enjambment Detection on Four Centuries of Spanish Sonnets», *Digital Humanities* (2017, Montreal), 572-576. <<https://dh2017.adho.org/abstracts/485/485.pdf>>.
- SMARANDACHE, Florentin, «Linguistic-Mathematical Statistics in Recent Romanian Poetry», *Collected Papers*, 1, Florentin Smarandache, Ann Arbor, InfoLearnQuest, 2007.
- SMITH, Michael, «The Authorship of «A Lover's Complaint»: An Application of Statistical Stylemetry to Poetry», *Computers and the Humanities*, 18, 1 (1984), 23-37.
- TENORIO, Martha, «Sor Juana y León Marchante», *Nueva Revista de Filología Hispánica*, 50, 2 (2002), 543-561.



## Anexo

**Tabla 3.** Corpus de poesía sacra indubitada

Poesía indubitada sacra				
Primer verso	Localización	Extensión	Métrica	Nombre
<b>Damián Cornejo</b>				
En tu alabanza mi musa	BNE: Mss/ 2245 (f. 185r)	164	Romance	C_si1
De santa resurrección	BNE: Mss/ 2245 (f. 186r)	378	Romance	C_si2
En títulos, Virgen pura	BNE: Mss/ 2245 (f. 198r)	294	Quintillas	C_si3
Dulce Jesús, Dios mío	BNE: Mss/ 2245 (f. 190r)	166	Endecha	C_si4
Hoy con gusto a cantar llego	BNE: Mss/ 2245 (f. 200r)	166	Quintillas	C_si5
Oigan, que el bravo del cielo	BNE: Mss/ 2245 (f. 191v)	189	Romance	C_si6
Divina Madre de Gracia,	BNE: Mss/ 2245 (f. 188v)	382	Romance	C_si7
Haced que la Gracia encuentre	BNE: Mss/ 2245 (f. 196r)	116	Quintillas	C_si8
A su esposo, el alma, escribe	BNE: Mss/ 2245 (f. 193v)	191	Coplas	C_si9
<b>León Marchante</b>				
En la noche más helada	León (1722: 17)	244	Romance	L_si1
Atención, que hoy de José	León (1722: 46)	372	Romance	L_si2
En la noche más oscura	León (1662: s.p.)	175	Coplas	L_si3
¡Oh qué bien que navega	León (1662: s.p.)	298	Coplas	L_si4
Hoy un Francisco hace fiesta	León (1722: 38)	243	Romance	L_si5
Pendiente a morir, de un leño,	León (1722: 26)	194	Romance	L_si6
La vida quiero contar	León (1722: 43)	225	Quintillas	L_si7
Vaya de jácara , y sea	León (1722: 9)	242	Romance	L_si8
Maravillas del Señor	León (1722: 12)	239	Coplas	L_si9
<b>Sor Juana Inés de la Cruz</b>				
Hoy la maestra divina	Cruz (2003: s.p.)	249	Coplas	SJ_si1
La soberana doctora	Cruz (2003: s.p.)	240	Romance	SJ_si2
Hoy es del divino amor	Cruz (2003: s.p.)	200	Redondillas	SJ_si3
Cuán grande, José, seréis,	Cruz (2012: s.p.)	203	Redondillas	SJ_si4
Por celebrar del Infante	Cruz (2006: s.p.)	399	Coplas	SJ_si5
¡Allá va, fuera, que sale	Cruz (2006: s.p.)	301	Romance	SJ_si6
Antes que todas las cosas	Cruz (2012: s.p.)	184	Coplas	SJ_si7

