



Problemas de autoría en la Comedia Nueva

José Luis Madrigal
Graduate Center (CUNY)

RESUMEN:

Los problemas de atribución en el teatro áureo español son innumerables. Mediante el programa JGAAP ideado por Patrick Juola se hace un análisis estadístico de la producción dramática de Lope de Vega, Tirso de Molina, Alarcón, Guillén de Castro, Claramonte y Mira de Amescua. El método empleado se vale de la métrica de similitud de coseno aplicada a n-gramas de caracteres y de palabras. Con un grado de eficacia notable y resultados a veces sorprendentes, se determina la fiabilidad autorial en 268 comedias de Lope y decenas de comedias atribuidas a otros dramaturgos contemporáneos.

PALABRAS CLAVE: JGAAP, atribución textual, Comedia Nueva, Lope de Vega, Tirso, Alarcón, Claramonte

ABSTRACT:

Authorship in Spanish Golden Age drama is mired in innumerable disputes. Through the JGAAP program devised by Patrick Juola, a statistical analysis has been carried out with hundreds of plays attributed to Lope de Vega, Tirso de Molina, Alarcón, Guillén de Castro, Claramonte and Mira de Amescua. By applying the cosine similarity measure on character and word n-grams, we have been able to test hundreds of plays with a fair degree of success and surprising results.

KEY WORDS: JGAAP, authorship attribution, N-gram-based author profile, Spanish Golden Age Drama, Lope de Vega, Tirso, Alarcón, Claramonte.

La producción teatral en la España del Siglo de Oro es ingente y muchas veces de difícil adscripción a un determinado autor. El manuscrito original constituía la base de todo, pero en seguida, en cuanto salía de las manos del poeta, se volvía un producto maleable, intercambiable, anónimo casi. El poeta dramático lo sabía mejor que nadie. Cuando vendía su comedia a una compañía teatral, perdía de inmediato todo control sobre el producto. La compañía, una vez en posesión del original, lo utilizaba a su antojo. Quitaba y añadía de aquí o de allá y, si venía al caso, no dudaba en refundir la comedia en su totalidad. Lo prioritario en todo momento era crear un espectáculo que diera gusto al vulgo. A la vez, los hurtos entre las compañías teatrales estaban a la orden del día. Unas veces alguien dentro de la compañía copiaba del original y se lo pasaba a la competencia; otras,

Fecha de recepción: 28/08/2021

Fecha de aceptación: 30/09/2021

la competencia contrataba a un memorión para que memorizase la comedia, que luego se transcribía con infinidad de errores, de morcillas y de partes inventadas. El trapicheo era constante. Algunos dramaturgos lo llevaban mejor que otros, aunque parece que la mayoría aceptaba tal orden de cosas sin mayores alharacas. No en balde habían vendido los originales a muy buen precio por lo general. Más duro les resultaba cuando los impresores imprimían estas mismas comedias sin su consentimiento, pues ni sacaban provecho con ello ni, dado el pésimo estado con que muchas llegaban a la imprenta, tampoco la publicación redundaba en su prestigio, caso de que se las publicaran a su nombre. A este respecto, Lope de Vega tuvo más de un pleito y, harto de perderlos todos, se decidió a ser él mismo quien publicara sus comedias. Sus palabras describen muy bien la situación:

Viendo imprimir cada día mis comedias de suerte que era imposible llamarlas mías y que en los pleitos de esta defensa siempre me condenaban los que tenían más solicitud y dicha para seguirlos, me he resuelto a imprimirlas por mis originales, que aunque es verdad que no las escribí con este ánimo ni para que los oídos del teatro se trasladaran a la censura de los aposentos, ya lo tengo por mejor que ver la crueldad con que despedazan mi opinión algunos intereses¹

Tal declaración va en el prólogo a la Parte IX publicada en 1617 y es de suponer que, a partir de esa fecha, las demás Partes que se publican en vida del Fénix pueden calificarse de obras auténticas. En todo caso, Lope no disponía siempre de los originales y, en más de una ocasión, especialmente en las comedias escritas quince o veinte años antes, se atenía a las versiones que le llegaban de tal o cual compañía, ya muy viciadas, como él mismo nos dice².

La misma situación —y quizá agravada— se pudo extender a otros dramaturgos, de Tirso de Molina a Ruiz de Alarcón, de Guillén de Castro a Mira de Amescua, de Moreto a Calderón. Este último mostraba su resignación con palabras que resumen muy bien la precariedad del poeta dramático en el siglo XVII:

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 de tercero, ¿quién quiere Vuestra Merced que se meta en advertir (los hurtos) el día que no los advierte la conciencia...?³

Los problemas de atribución son, como se ve en estas pocas pinceladas, innumerables en el teatro del Siglo de Oro. Unas veces la comedia nos llega anónima; otras, la autoría es dudosa; otras no sabemos el grado de participación. Frente a ello, el estudioso suele fa-

1.– Parte IX, (Madrid, 1617). En otros prólogos arrojará más luz sobre el *modus operandi* de autores de comedias e impresores. Así, bajo la máscara de Gaspar de Porres, recuerda que se imprimen «sus comedias tan bárbaras como las han hallado, después de muchos años que salieron de sus manos, donde apenas hay cosa concertada, y los que padece de otros que por sus particulares intereses imprimen o representan las que no son suyas con su nombre, me han obligado, por el amor y amistad que ha muchos años que le tengo, a dar luz a estas doce, que yo tuve originales...». DIXON (45-63).

2.– Lope, o más bien su alter ego, Sebastián de Medrano, así lo refiere: «no de todas (las comedias) se hallan los originales...»; y de ahí que haya «tantos librillos de Romances, y otros versos con su nombre, así divinos como humanos, que no le ha pasado en el pensamiento escribirlos, fuera de lo que algunos ciegos, Gitanos y mulatos van pregonando por las calles, levantándole mil testimonios a sus ojos, que en otras ciudades es cosa de lástima el poco respeto con que algunos hombres dan a un sacerdote docto y bien nacido por autor de sus desatinos, ya por acreditarlos, ya por ganar de comer con ellos.» Parte XVIII, 1623, f. IV r y v. (Moner: 123).

3.– Recogido en COUDERIC (126).

vorecer, primeramente, las pruebas externas. Un autógrafo con la firma del autor, el testimonio encontrado en una carta o la documentación relativa a la compraventa de una determinada comedia han solido ser pruebas inequívocas para atribuírsela a alguien. Desgraciadamente, no siempre se manejan documentos así. Lo normal es tener que dilucidar la atribución a través del análisis interno del texto. En el pasado los análisis se centraban casi siempre en el tipo de métrica empleada. Todo dramaturgo tiene preferencias. Unos están más a gusto con las redondillas y otros con las quintillas o los romances. A principios del siglo XX S. Griswold Morley clasificó las comedias de Tirso a tenor del mayor o menor empleo de las redondillas que, según sus cálculos, solía ser la composición poética favorita del mercedario⁴. Décadas después semejante criterio fue el empleado por el mismo Morley y su colega Courtney Bruerton no solo para cuestiones de atribución, sino para establecer la cronología de las comedias de Lope. Es un estudio seminal que todavía sigue vigente⁵.

En realidad, las pruebas internas se basan en un principio básico de la conducta humana, y es que todo individuo se comporta por la mayor parte de manera rutinaria. Lo que hacemos una vez, solemos repetirlo otras veces en cualquier actividad, sea consciente o inconscientemente. El uso del lenguaje no resulta muy distinto. El repertorio verbal de un individuo es recurrente. Tal fenómeno lo podemos demostrar mediante el acopio de paralelismos o, de manera quizá más eficaz, gracias al cálculo estadístico de un determinado rasgo lingüístico diferenciador. A veces pueden bastarnos los primeros versos de una comedia. En *El arenal de Sevilla* de Lope de Vega hay una dama llamada Laura que dice:

No tiene, a mi parecer,
todo el mundo vista igual;
tanta galera y navío
mucho al Betis engrandece.

Y su amiga le responde:

Otra Sevilla parece
que está fundada en el río. (ACTO I, vv. 5-10)

Un rápido cotejo con un corpus de solamente diez comedias de Lope nos reporta el siguiente paralelismo procedente de *El amante agradecido*:

¿Qué Salamanca ni corte
como aquel famoso río?
Ver la galera, el navío
del mar del sur o del norte.
Ver aquella variedad
que es imposible decilla,
porque el río de Sevilla
tiene otra tanta ciudad. (*El amante agradecido*, ACTO I, v. 22)

En los dos pasajes se divisa una *galera* y un *navío* en el río Guadalquivir. El paralelismo <río-galera-navío-Sevilla> sitúa claramente ambas comedias en la misma órbita. Unos versos después la amiga de Laura dice:

4.- MORLEY (1914).

5.- MORLEY, BRUERTON (1968).

Mejor será que lleguemos
 hasta la Torre del Oro,
 y todo ese gran tesoro
 que va a las Indias veremos

Estamos en Sevilla y no puede extrañar que se nombre la *Torre del Oro* junto con *ir a las Indias*, pero resulta bastante inusual que *El amante agradecido* (1602), una vez más, presente una constelación verbal muy semejante:

Vienen de Sanlúcar,
 rompiendo el agua,
 a **la Torre del Oro**,
 barcos de plata.
 ¿Dónde te has criado
 la niña bella,
 que, sin ir a **las Indias**,
 toda eres perla?» (Acto III, vv. 2834-2841)

El escritor tiende a formar agrupaciones verbales parecidas cuando describe una misma situación. Pondremos un ejemplo más. Esta vez el pasaje procede de *El cuerdo loco* (1602):

Atento más al rüido...
 cajas y trompetas sienta,
 como en galeras o naves
 retumba el mar desde lejos....
 Al fin, acercado más,
 declaradamente veo
 un ejército formado
 dando banderas al viento,
 como parece escuadrón
 de abejas pasado enero,
 marchar con las varias flores
 de los romeros y brezos. (Acto III, vv. 2164-2178)

En *El juez en su causa* (1610) se describe un despliegue naval con muy parecidas palabras:

Se vio la **mar** coronada / de **naves**, urcas y barcos.... De las entenas pendientes
 / tantos estandartes varios / que de lejos parecían/ **un ejército formado**. / Las
cajas y las **trompetas** / clavan ecos al **mar** cano, / que de bullir con la espuma /
 encanean los peñascos. (Acto III, vv. 2171-2182)

Y todavía en otra comedia, *El alcaide de Madrid* (1599), leemos este otro paralelismo:

como el **escuadrón de abejas**
 que cuando las **flores** comen
 con algún espanto vuelan
 y van dejando las **flores**,
 así siembran por el campo
 de los bagajes y cofres
 tocas, bandas y almalafas
 y cendales de colores (Acto II, vv. 2120-2127)

Los ejemplos se podrían multiplicar. En cuanto cotejamos textos encuadrados dentro de un mismo género tarde o temprano aparecen correspondencias; y si los textos los escribió un mismo autor inevitablemente las correspondencias aumentan en cantidad y, sobre todo, en calidad. El fenómeno era perceptible ya desde la Antigüedad, como lo atestigua Aulio Gelio en uno de los capítulos de sus *Noches áticas* dedicado al problema de atribución en las comedias de Plauto. Allí se concluye que, a la hora de determinar qué comedias escribió en realidad el comediógrafo, debemos confiar no tanto en los catálogos de los eruditos, como en la particular manera que tenía el propio Plauto de utilizar su lengua (*ipsi Plauto moribusque ingenii atque linguae eius*)⁶. Tal había sido, al menos, la vara de medir (*iudicii norma*) de Marco Varrón, el cual, además de confirmar la autoría de 21 comedias, había atribuido a Plauto otras cuatro por «la índole del estilo y el carácter de los chistes» (*adductus filo atque facetia sermonis*)⁷.

El análisis estilístico y temático, acompañado de analogías verbales, ha sido tradicionalmente el método seguido en la atribución de textos, a falta de pruebas externas. El estudio del *modus scribendi* presenta, sin embargo, dos obstáculos importantes. El primero está en la posible imitación y hasta el plagio que se haya podido llevar a cabo. Dos obras pueden ser de distinto autor y compartir unos mismos *estilemas* o frases idénticas. En el teatro áureo, sin ir más lejos, era algo muy común. Cualquier cotejo que hagamos con comedias de diferentes dramaturgos nos deparará de inmediato un acervo de frases hechas compartidas entre sí. El segundo obstáculo, muchas veces insalvable, es la llamada paradoja del montón. ¿En qué momento podemos decir que una obra es de uno y no de otro? ¿En dónde ponemos la línea identitaria entre textos? ¿Cuántos paralelismos son necesarios para ello? ¿Cinco, diez, veinte? Todos repetimos frases. No es fácil reconocer las voces de los ecos.

Cierto que actualmente, gracias al ingente corpus de textos digitalizados, se puede discriminar en un instante una secuencia entre cientos de miles. Si algo ha quedado demostrado en estos últimos años es el extraordinario grado de singularidad en el repertorio verbal de un hablante. Las correspondencias vistas arriba con esos pasajes de Lope de Vega no suelen ser producto del azar. Puede haber imitación o plagio, desde luego, pero raramente dos paralelismos con secuencias idénticas ocurren casualmente entre textos no relacionados entre sí.

Pensemos en *La tía fingida*, por ejemplo. El manuscrito apareció en un cartapacio que contenía, entre otros textos anónimos, dos novelas de Miguel de Cervantes: *El celoso extremeño* y *Rinconete y Cortadillo*. Isidoro Bosarte, su descubridor, no dudó en adscribir la procaz novelita al manco de Lepanto. Ciertamente todo apuntaba a ello, desde los temas y la caracterización de personajes hasta el empleo de la lengua, tan semejante a la prosa cervantina. Bosarte nunca publicó el estudio donde pensaba ofrecer «un gran número de frases y expresiones tomadas de las demás obras de Cervantes» y murió poco después (LUCÍA: 238). En las décadas siguientes eruditos de la talla de Martín Fernández de Navarrete o Bartolomé Gallardo no creyeron necesario demostrar nada, de tan obvia que les parecía la paternidad, hasta que un francés, Raymond Foulché-Delbosc, haciendo uso de

6.- *Noctes atticae*, III, 1.

7.- *Noctes atticae*, III, 3.

la duda cartesiana, puso la atribución cervantina en el disparadero. Según él, los supuestos paralelismos se encontraban, a poco que se hiciera un esfuerzo, en cualquier otra obra contemporánea de Cervantes y no podía descartarse la mano de un imitador. Además, la calidad de *La tía fingida* estaba muy lejos de cualquier novela ejemplar. De poco sirvió que otros aportaran una catarata de coincidencias en años posteriores. Las coincidencias se veían como usos de época, mientras que cualquier supuesta diferencia se magnificaba, desde tal o cual palabra a la cruda procacidad del asunto tratado. A lo largo de los años la novelita ha permanecido anónima o, a lo sumo, con el nombre de Cervantes entre paréntesis, pese a que hoy, con los medios que tenemos a nuestro alcance, resulta poco menos que imposible negar su paternidad. Un simple cotejo con el corpus cervantino hecho con el más rudimentario de los programas descubre un paralelismo o una secuencia singular casi en cada renglón (MADRIGAL). Ya no vale decir que con un poco de esfuerzo esas correspondencias se encontrarían en otras obras. No es así. El repertorio de un hablante es recurrente y extraordinariamente singular, mucho más de lo creído hasta hace muy poco. Esta singularidad no está en tal o cual frase aislada, pues aisladamente las palabras y las frases son comunes a muchos, sino en el conjunto de combinaciones que cada individuo hace con las palabras de su repertorio, algunas de las cuales terminan por ser idiosincráticas.

Últimamente existen programas que nos facilitan enormemente la labor a este respecto. Uno, al alcance de cualquiera, es JGAAP, un programa de *software* desarrollado por Patrick Juola de la Universidad de Dushesne⁸. Todos los métodos y técnicas empleados hasta ahora en los estudios de estilometría pueden aplicarse con facilidad en este programa (JUOLA). El criterio que yo sigo es simple. Dado que la recurrencia verbal es el rasgo más característico del idiolecto de un hablante, selecciono un corpus de textos ad hoc y calculo luego el grado de similitud con n-gramas de cuatro caracteres (C4), n-gramas de dos palabras (2W) y n-gramas de tres palabras (3W)⁹. El programa ofrece múltiples medidas de similitud, pero la métrica que yo empleo es la similitud de coseno. Debajo pongo los resultados hechos con el cotejo de *La tía fingida* y un conjunto de novelas cortas del periodo, entre las que se encuentran cuatro novelas ejemplares de Cervantes.

	C4	W2	W3
LA TIA FINGIDA	1. CERVANTES Coloquio de los perros	1. CERVANTES Coloquio de los perros	1. CERVANTES El celoso extremeño
	2. CERVANTES El celoso extremeño	2. CERVANTES El celoso extremeño	2. CERVANTES Coloquio de los perros
	3. CERVANTES La Ilustre fregona	3. CERVANTES La Ilustre fregona	3. CERVANTES Rinconete
	4. SALAS BARBADILLO La hija de la Celestina	4. SALAS BARBADILLO El sagaz Estacio	4. CERVANTES La Ilustre fregona
	5. CASTILLO SOLORZANO Bachiller Trapaza	5. SALAS BARBADILLO La hija de la Celestina	5. ZAYAS EL castigo de la miseria
	6. TIRSO Los tres maridos burlados	6. LOPE Las fortunas de Diana	6. ZAYAS Aventurarse perdiendo
	7. CERVANTES Rinconete	7. TIRSO Los tres maridos burlados	7. CASTILLO SOLORZANO Bachiller Trapaza
	8. LOPE La desdicha por la honra	8. CASTILLO SOLORZANO Bachiller Trapaza	8. ZAYAS La esclava de su amante
	9. ALEMAN Cuentos	9. LOPE La prudente venganza	9. LOPE Las fortunas de Diana
	10. LOPE Guzmán el Bravo	10. CERVANTES Rinconete	10. CASTILLO SOLORZANO Lisardo enamorado

8.- JGAAP (Java Graphical Authorship Attribution Program) se puede bajar gratis en <http://www.jgaap.com>.

9.- Los n-gramas de caracteres en combinación con los n-gramas de 2 y 3 palabras empiezan a imponerse como el rasgo lingüístico más eficaz para la clasificación de textos por autor (KESELJ, STAMATATOS, ANTONIA, SARI).

La proximidad de las cuatro novelas cervantinas con respecto a *La tía fingida* es casi total. Solamente *Rinconete* y *Cortadillo* se aleja algo entre las obras cotejadas, pero nótese que en el apartado donde se analizan los n-gramas de tres palabras (W3) todas las novelas ejemplares, incluida la de *Rinconete*, ocupan las primeras cuatro posiciones. Si hacemos la misma operación con *La desdicha por la honra* de Lope el resultado es muy parecido:

	C4	W2	W3
LA DESDICHA POR LA HONRA	1. LOPE Guzmán el Bravo	1. LOPE La prudente venganza	1. LOPE Guzmán el Bravo
	2. LOPE La prudente venganza	2. LOPE Las fortunas de Diana	2. LOPE La prudente venganza
	3. LOPE Las fortunas de Diana	3. SALAS BARBADILLO El sagaz Estacio	3. LOPE Las fortunas de Diana
	4. SALAS BARBADILLO La hija de la Celestina	4. CERVANTES Coloquio de los perros	4. SALAS BARBADILLO El sagaz Estacio
	5. CERVANTES Coloquio de los perros	5. LOPE Guzmán el Bravo	5. CERVANTES Coloquio de los perros
	6. CASTILLO SOLORZANO Bachiller Trapaza	6. SALAS BARBADILLO La hija de la Celestina	6. CERVANTES El celoso extremeño
	7. SALAS BARBADILLO El sagaz Estacio	7. CERVANTES La Ilustre fregona	7. ZAYAS EL castigo de la miseria
	8. ZAYAS Aventurarse perdiendo	8. CASTILLO SOLORZANO Bachiller Trapaza	8. ZAYAS Aventurarse perdiendo
	9. CERVANTES La Ilustre fregona	9. SALAS BARBADILLO La peregrinación sabia	9. ZAYAS La esclava de su amante
	10. ZAYAS La esclava de su amante	10. CERVANTES El celoso extremeño	10. CASTILLO SOLORZANO Bachiller Trapaza

Las cuatro novelas a Marcia Leonarda se escribieron en un trecho de tiempo corto (1621-1624) y están las cuatro escritas en formato muy semejante. Veamos ahora qué ocurre con un texto de solamente 2.000 palabras extraído de *El peregrino en su patria* (1604), novela publicada dos décadas antes.

	C4	W2	W3
PEREGRINO EN SU PATRIA	1. LOPE Las fortunas de Diana	1. LOPE Las fortunas de Diana	1. LOPE Las fortunas de Diana
	2. CASTILLO SOLORZANO Bachiller Trapaza	2. LOPE La prudente venganza	2. LOPE Guzmán el Bravo
	3. ZAYAS Aventurarse perdiendo	3. CASTILLO SOLORZANO Bachiller Trapaza	3. CERVANTES El celoso extremeño
	4. LOPE La prudente venganza	4. CERVANTES El celoso extremeño	4. ZAYAS Aventurarse perdiendo
	5. LOPE La desdicha por la honra	5. CERVANTES Coloquio de los perros	5. ZAYAS La esclava de su amante
	6. SALAS BARBADILLO La hija de la Celestina	6. LOPE La desdicha por la honra	6. LOPE La prudente venganza

Las fortunas de Diana (1621) aparece en primera posición en cada uno de los tres parámetros, y allí donde analizamos secuencias de dos y tres palabras otra novela de Lope está en segunda posición y una tercera en sexto lugar. Volvamos a Cervantes y hagamos esta vez la operación con tres textos: *La gitanilla* y extractos de *La Galatea* (2.000 palabras) y *Persiles* (11.000 palabras).

	<i>La Galatea</i>	<i>La gitanilla</i>	<i>Persiles</i>
C4	1. LOPE Las fortunas de Diana 2. SALAS BARBADILLO La hija de la Celestina 3. CERVANTES Coloquio de los perros 4. ZAYAS Aventurarse perdiendo 5. CASTILLO SOLORZANO Bachiller Trapaza	1. CERVANTES Coloquio de los perros 2. CERVANTES La Ilustre fregona 3. SALAS BARBADILLO La hija de la Celestina 4. CERVANTES Rinconete 5. CERVANTES El celoso extremeño	1. CERVANTES Coloquio de los perros 2. CERVANTES La Ilustre fregona 3. SALAS BARBADILLO La hija de la Celestina 4. LOPE Las fortunas de Diana 5. CERVANTES El celoso extremeño

W2	1. CERVANTES Coloquio de los perros 2. LOPE Las fortunas de Diana 3. SALAS BARBADILLO La hija de la Celestina 4. CERVANTES El celoso extremeño 5. LOPE La prudente venganza	1. CERVANTES Coloquio de los perros 2. CERVANTES La Ilustre fregona 3. CERVANTES El celoso extremeño 4. SALAS BARBADILLO La hija de la Celestina 5. LOPE Las fortunas de Diana	1. CERVANTES Coloquio de los perros 2. CERVANTES El celoso extremeño 3. CERVANTES La Ilustre fregona 4. LOPE Las fortunas de Diana 5. SALAS BARBADILLO La hija de la Celestina
W3	1. CERVANTES Coloquio de los perros 2. CERVANTES La Ilustre fregona 3. ALEMAN Cuentos 4. CERVANTES El celoso extremeño 5. LOPE Las fortunas de Diana	1. CERVANTES Coloquio de los perros 2. CERVANTES El celoso extremeño 3. CERVANTES La Ilustre fregona 4. CERVANTES Rinconete 5. LOPE Las fortunas de Diana	1. CERVANTES El celoso extremeño 2. CERVANTES Coloquio de los perros 3. CERVANTES La Ilustre fregona 4. LOPE Las fortunas de Diana 5. ZAYAS Aventurarse perdiendo

La gitanilla, igual que habíamos visto con *La tía fingida*, ofrece un grado de cercanía extraordinario con las cuatro novelas ejemplares. Los resultados bajan algo en el *Persiles*, pero no tanto como esperaríamos teniendo en cuenta la distinta temática. Incluso *La Galatea*, escrita mucho antes, presenta un resultado muy favorable, pese a que el texto no tiene más de 2.000 palabras. No me extenderé en más ejemplos. Parece claro que el corpus de un mismo autor se caracteriza por compartir un mayor número de secuencias verbales, lo cual se acentúa cuando los textos pertenecen a un mismo género y tocan una temática semejante.

¿Podríamos aplicar este sencillo método para las comedias del siglo XVII? El barullo allí es grande, pero tal como vamos a ver, el programa JPAAG facilita enormemente la clasificación textual en lo relativo a la autoría, sin que sea, naturalmente, la panacea universal. Empezaré con el corpus de comedias de Lope. Afortunadamente contamos con hasta 268 comedias digitalizadas procedentes de la base de datos de ARTELOPE¹⁰. Todas ellas han sido catalogadas y nos ofrecen un magnífico surtido de información, desde datos bibliográficos hasta datación, cómputo de versos y temática. ARTELOPE, a partir de lo establecido por la crítica, divide el grado de fiabilidad autorial en cuatro apartados: fiable, probable, dudosa y no de Lope. El experimento que voy a llevar a cabo consiste en lo siguiente. Elijo primeramente un corpus de 24 comedias pertenecientes a seis autores canónicos con cuatro comedias cada uno:

ALARCÓN	<i>La verdad sospechosa</i> <i>La cueva de Salamanca</i> <i>El semejante a sí mismo</i> <i>Las paredes oyen</i>
CLARAMONTE	<i>De este agua no beberé</i> <i>Púsoseme el sol</i> <i>De lo vivo a lo pintado</i> <i>El valiente negro</i>
GUILLÉN DE CASTRO	<i>El amor constante</i> <i>La humildad soberbia</i> <i>El nacimiento de Montesinos</i> <i>Las mocedades del Cid</i>

10.- <<https://artelope.uv.es>>. La gran mayoría de textos empleados en este trabajo proceden de ARTELOPE, Biblioteca Virtual Miguel de Cervantes y The Association for Hispanic Classical Theater (AHCT).

LOPE	<i>El castigo sin venganza</i> <i>El perro del hortelano</i> <i>Fuente Ovejuna</i> <i>El caballero de Olmedo</i>
MIRA DE AMESCUA	<i>La Fénix de Salamanca</i> <i>El ejemplo mayor</i> <i>El esclavo del demonio</i> <i>Callar en buena ocasión</i>
TIRSO	<i>El vergonzoso en palacio</i> <i>Caballero de Gracia</i> <i>Marta la piadosa</i> <i>Don Gil de las calzas verdes</i>

En todos los casos analizaré estadísticamente los n-gramas de cuatro caracteres (4C), de dos palabras (2W) y de tres palabras (3W) mediante la métrica de coseno en JPAAG y luego daré una puntuación de 1 a 24 en referencia al puesto que cada comedia consiga en cada cotejo. La comedia que aparezca en primer lugar conseguirá 24 puntos, la que esté en segundo lugar 23, etc. Lo ilustraré con un ejemplo antes de entrar con Lope. Elijamos para ello *Palabras y plumas*, comedia que es indiscutiblemente de Tirso. Si nos vamos al programa de JPAAG y lo contrastamos con las 24 comedias del corpus elegido obtenemos estos resultados:

	C4	W2	W3
Palabras y plumas	1. TIRSO Caballero de Gracia	1. TIRSO El vergonzoso en palacio	1. TIRSO El vergonzoso en palacio
	2. TIRSO Marta la piadosa	2. TIRSO Caballero de Gracia	2. MIRA DE AMESCUA La Fénix de Sal.
	3. TIRSO El vergonzoso en palacio	3. MIRA DE AMESCUA La Fénix de Sa	3. TIRSO Caballero de Gracia
	4. ALARCÓN La verdad sospechosa	4. TIRSO Marta la piadosa	4. TIRSO Marta la piadosa
	5. CLARAMONTE Deste agua no beberé	5. LOPE El castigo sin venganza	5. CLARAMONTE Deste agua no bebe.
	6. ALARCÓN Las paredes oyen	6. LOPE El perro del hortelano	6. LOPE El perro del hortelano
	7. LOPE El caballero de Olmedo	7. CLARAM. Deste agua no beberé	7. LOPE El castigo sin venganza
	8. LOPE El castigo sin venganza	8. CASTRO El nacimiento de Montes.	8. GUILLÉN DE CASTRO El nacimiento de M.
	9. CLARAMONTE Púsoseme el sol	9. CLARAMONTE Púsoseme el sol	9. MIRA DE AMESCUA El ejemplo mayor
	10. LOPE El perro del hortelano (...)	10. ALARCÓN La verdad sospechosa (...)	10. LOPE Fuente Ovejuna (...)
	23. TIRSO Don Gil de las calzas verdes	24. TIRSO Don Gil de las calzas verd.	17. TIRSO Don Gil de las calzas verdes

La puntuación en C4 es de 71 pts. (78%), en W2 de 69 pts. (76%) y en W3 de 76 pts. El porcentaje de cercanía con respecto a las cuatro comedias de Tirso es superior al resto de los autores del corpus, aunque queda lastrado por *Don Gil de las calzas verdes*, muy al final de la tabla en las tres categorías. Si excluimos esta última comedia y contabilizamos solo las tres primeras, el grado de cercanía es de más del 95%:

PALABRAS Y PLUMAS	C4	W2	W3
	69 (100%)	68 (98%)	69 (100%)

No encuentro explicación a la lejanía de *Don Gil de las calzas verdes* respecto a *Palabras y plumas*. Su filiación tirsiana está fuera de toda duda. También lo está en el programa de JGAAP: basta sustituirla en el corpus por *La primera parte de la Santa Juana* y realizar el mismo procedimiento con ella para comprobar que las comedias de Tirso se apiñan en los primeros puestos en las tres categorías:

	C4	W2	W3
Don Gil de las calzas verdes	1. TIRSO El vergonzoso en palacio 2. TIRSO Marta la piadosa 3. LOPE El caballero de Olmedo 4. TIRSO Caballero de Gracia 5. CLARAMONTE Deste agua no beberé 6. ALARCÓN La verdad sospechosa 7. TIRSO La santa Juana	1. TIRSO El vergonzoso en palacio 2. TIRSO La santa Juana I 3. TIRSO Caballero de Gracia 4. TIRSO Marta la piadosa 5. MIRA DE AMESCUA La Fénix de S.	1. TIRSO El vergonzoso en palacio 2. MIRA DE AMESCUA La Fénix de Sal. 3. TIRSO La santa Juana I 4. TIRSO Marta la piadosa 5. TIRSO Caballero de Gracia

Don Gil de las calzas verdes	C4	W2	W3
	86 (95%)	90 (100%)	87 (96%)

Daré otro ejemplo, esta vez con *La Segunda parte de las mocedades del Cid* de Guillén de Castro. El procedimiento es el mismo y el resultado, tal como se observa en las tablas de abajo, vuelve a demostrar que textos de un mismo autor suelen atraerse entre sí como un imán:

	C4	W2	W3
Mocedades del Cid II	1. CASTRO Las mocedades del Cid I 2. CASTRO La humildad soberbia 3. CASTRO El nacimiento de Montesinos 4. CLARAMONTE Deste agua no beberé 5. CASTRO El amor constante	1. CASTRO Las mocedades del Cid I 2. CASTRO El nacimiento de Montesinos 3. CLARAMONTE Deste agua no beberé 4. CASTRO El amor constante 5. CASTRO La humildad soberbia	1. CASTRO Las mocedades del Cid I 2. CASTRO El nacimiento de Montesinos 3. CASTRO El amor constante 4. CASTRO La humildad soberbia 5. CLARAMONTE Deste agua no beberé

Mocedades del Cid II	C4	W2	W3	%
	98	97	100	98%

Pasemos ahora a Lope. Debajo pongo el cuadro completo con las 268 comedias y los cálculos llevados a cabo. La información procede de ARTELOPE. En la primera columna se indica el grado de fiabilidad autorial de acuerdo con lo dicho por la crítica hasta ahora. Las cuatro alternativas son *fiable*, *probable*, *dudosa* y *no de Lope*. Nótese que a veces existen discrepancias entre el consenso general y nuestros cálculos. La segunda columna es la fecha en que se supone que se debió escribir la comedia. La tercera columna muestra el título. Las tres siguientes ofrecen el porcentaje de similitud en el cálculo de n-gramas (C4, W2, W3). La última columna incluye el resultado final tras hacer la media con los resultados de las tres categorías. La puntuación máxima es 90 puntos (100%) y la mínima 10 (11%). Las cuatro comedias de Lope, tal como ya vimos arriba, son *Fuente Ovejuna* (1612-1614), *El perro del hortelano* (1613-1615), *El caballero de Olmedo* (1620-25), y *El castigo sin venganza* (1631). Hay algo de arbitrario en ello, salvo que las cuatro han salido de la pluma del Fénix y cubren los últimos veinte años de su carrera literaria. La clasificación de esta lista se organiza de mayor a menor a partir del máximo grado de proximidad con respecto a estas cuatro comedias de Lope.

93%

AUTORÍA	FECHA	TÍTULO	C4	W2	W3	%
FIABLE	1620-1622	LA MAYOR VICTORIA	100	93	86	93%

92%

FIABLE	1606-1609	DEL MAL LO MENOS	97	88	92	92%
FIABLE	1615?-1620	LA VENGADORA DE LAS MUJERES	98	92	86	92%
FIABLE	1620-1625	GUARDAR Y GUARDARSE	98	91	88	92%
FIABLE	1623	LA CORONA DE HUNGRÍA	100	90	88	92%

91%

FIABLE	1620-1628	LOS TELLOS DE MENESES	96	88	91	91%
--------	-----------	-----------------------	----	----	----	-----

90%

FIABLE	1608?	LO FINGIDO VERDADERO	98	86	87	90%
FIABLE	1615-1616	EL VALOR DE LAS MUJERES	100	77	94	90%
FIABLE	1615-1621	NADIE SE CONOCE	97	91	82	90%
FIABLE	1620-1623	EL MEJOR ALCALDE, EL REY	95	84	91	90%
FIABLE	1624	EL MARQUÉS DE LAS NAVAS	97	91	83	90%

89%

FIABLE	1601	LOS AMANTES SIN AMOR	92	92	84	89%
FIABLE	1617-1621	DIOS HACE REYES	100	77	91	89%
FIABLE	1626	EL PIADOSO ARAGONÉS	97	88	84	89%

88%

FIABLE	1604-1610	LA HERMOSURA ABORRECIDA	96	85	83	88%
FIABLE	1613-1618	EL HIJO SIN PADRE	97	85	84	88%
FIABLE	1624-1625	POR LA PUENTE, JUANA	93	84	88	88%
FIABLE	1625	AY, VERDADES, QUE EN AMOR...	92	90	84	88%

87%

FIABLE	1612-1613	LOS ENEMIGOS EN CASA	96	77	90	87%
FIABLE	1613-1614	DE CUANDO ACÁ NOS VINO	93	81	87	87%
FIABLE	antes 1618?	LA MOZA DE CÁNTARO	94	86	83	87%
FIABLE	1620-1622	EL LABRADOR VENTUROSO	100	86	75	87%
PROBABLE	1624-1628	PORFIAR HASTA MORIR	98	87	76	87%
FIABLE	1633	EL DESPRECIO AGRADECIDO	94	85	82	87%

86%

FIABLE	1599-1603	LOS TRES DIAMANTES	96	78	84	86%
FIABLE	1599-1603	EL LACAYO FINGIDO	95	83	80	86%
FIABLE	1605-1608	PERIBÁÑEZ	90	86	83	86%
FIABLE	1612?- 1615?	LAS FLORES DE DON JUAN	91	84	83	86%
PROBABLE	1622-1623	LA AMISTAD Y OBLIGACIÓN	92	81	85	86%
FIABLE	1627?	MÁS PUEDEN CELOS QUE AMOR	90	83	86	86%
FIABLE	1634	LAS BIZARRÍAS DE BELISA	93	82	85	86%

85%

AUTORÍA	FECHA	TÍTULO	C4	W2	W3	%
FIABLE	1598?	EL GALÁN CASTRUCHO	90	85	82	85%
FIABLE	1601-1605	EL HALCÓN DE FEDERICO	94	81	82	85%
FIABLE	1602	EL AMANTE AGRADECIDO	86	86	83	85%
FIABLE	1606-1612	CASTELVINES Y MONTESES	92	80	84	85%
FIABLE	1611	LA DISCORDIA DE LOS CASADOS	98	84	75	85%
FIABLE	1612-1615	EL VALIENTE CÉSPEDES	91	81	84	85%
FIABLE	1613	LA DAMA BOBA	94	90	72	85%
FIABLE	1613-1614	CON SU PAN SE LO COMA	87	85	84	85%
FIABLE	1616?	EL PRÍNCIPE PERFECTO II	95	91	70	85%
FIABLE	1617	LO QUE PASA EN UNA TARDE	97	86	73	85%
FIABLE	1617-1621	EL MARIDO MÁS FIRME	95	85	76	85%
PROBABLE	1620-1625	AMAR COMO SE HA DE AMAR	93	82	80	85%
FIABLE	1620-1625	LA BELLA AURORA	95	78	84	85%
FIABLE	1620-1630	LOS TRABAJOS DE JACOB	93	78	84	85%
FIABLE	1623	EL PODER EN EL DISCRETO	90	86	80	85%
FIABLE	1627	DEL MONTE SALE	96	83	77	85%

84%

AUTORÍA	FECHA	TÍTULO	C4	W2	W3	%
FIABLE	1588-1598	EL GALÁN ESCARMENTADO	87	85	81	84%
FIABLE	1598-1601	EL CASTIGO DEL DISCRETO	95	78	80	84%
FIABLE	1599-1608	EL PIADOSO VENECIANO	91	80	82	84%
FIABLE	1604?-1612?	EL ALCALDE MAYOR	93	84	75	84%
FIABLE	1606-1612	EL ACERO DE MADRID	92	76	84	84%
FIABLE	1608-1612	LA CORTESÍA DE ESPAÑA	88	82	82	84%
FIABLE	1610-1612	LA FIRMEZA EN LA DESDICHA	100	74	78	84%
FIABLE	1610-1615	LA MALCASADA	95	82	76	84%

FIABLE	1612-1615	AMOR SECRETO, HASTA CELOS	91	83	80	84%
FIABLE	1613?	LA NECEDAD DEL DISCRETO	94	82	76	84%
FIABLE	1620	QUIEN TODO LO QUIERE	91	83	80	84%
FIABLE	1621	AMOR, PLEITO Y DESAFÍO	94	86	72	84%
PROBABLE	1623-1626	LAS BURLAS Y VERAS	90	82	82	84%
FIABLE	1625-1635	EL AMOR ENAMORADO	94	85	74	84%
FIABLE	1630?	LA BOBA PARA LOS OTROS	95	74	84	84%
FIABLE	1630?	NO SON TODO RUISEÑORES	94	81	78	84%

83%

FIABLE	1588-1595	EL HIJO VENTUROSO	95	80	76	83%
FIABLE	1590-1595	LOS LOCOS DE VALENCIA	92	78	81	83%
FIABLE	1600-1606	EL ASALTO DE MASTRIQUE	91	84	76	83%
FIABLE	1608-1611	LOS HIDALGOS DE ALDEA	97	81	73	83%
FIABLE	1610	LA BUENA GUARDA	97	82	70	83%
FIABLE	1610	LA HERMOSA ESTER	94	76	81	83%
FIABLE	1610-1612	LA LOCURA POR LA HONRA	95	83	73	83%
FIABLE	1611-1612	EL ANIMAL DE HUNGRÍA	96	80	73	83%
FIABLE	1615	EL GALÁN DE LA MEMBRILLA	96	74	80	83%
FIABLE	1615	EL MAYOR IMPOSIBLE	93	82	74	83%
FIABLE	1615?-1616?	GUERRAS DE AMOR Y DE HONOR	85	83	81	83%
FIABLE	1622	EL VELLOCINO DE ORO	83	84	84	83%
FIABLE	1623-1626	LA CARBONERA	97	83	71	83%
FIABLE	1624-1635	AMAR, SERVIR Y ESPERAR	95	78	77	83%
FIABLE	1631	LA NOCHE DE SAN JUAN	94	83	80	85%

82%

FIABLE	1598-1600	LAS BATUECAS DEL DUQUE DE ALBA	93	80	74	82%
FIABLE	1604?-1606?	EL MAYORAZGO DE LA DUQ. AMALFI	90	76	81	82%
FIABLE	1605	LA NOCHE TOLEDANA	84	80	82	82%
FIABLE	1608	LA BATALLA DEL HONOR	100	73	75	82%
FIABLE	1610-1612	LA NIÑA DE PLATA	91	81	74	82%
FIABLE	1610-1615	EL LABERINTO DE CRETA	85	77	85	82%
FIABLE	1611	EL VILLANO EN SU RINCÓN	94	76	78	82%
FIABLE	1612-1618	EL ABANILLO	90	80	76	82%
FIABLE	1613	LA BURGALESA DE LERMA	93	80	74	82%
FIABLE	1614-1619	LAS CUENTAS DEL GRAN CAPITÁN	92	83	72	82%
FIABLE	1616	AL PASAR DEL ARROYO	91	82	75	82%

FIABLE	1620-1622	DE AMAR SIN SABER A QUIÉN	92	82	74	82%
FIABLE	1624	LO QUE HA DE SER	98	84	65	82%

81%

FIABLE	1585-1595?	LA INGRATITUD VENGADA	91	77	75	81%
FIABLE	1599	EL ARGEL FINGIDO	90	78	76	81%
FIABLE	1604-1606	GUANCHES DE TENERIFE	88	80	77	81%
FIABLE	1606-1608	EL CUERDO EN SU CASA	94	77	73	81%
FIABLE	1612-1615	EL HOMBRE POR SU PALABRA	96	73	74	81%
FIABLE	1615	LA ARCADIA	93	78	73	81%
PROBABLE	1625-1630	LA ESCLAVA DE SU GALÁN	88	82	75	81%

80%

FIABLE	1587-1595	LAS BURLAS DE AMOR	94	76	72	80%
FIABLE	1588-1595	EL MESÓN DE LA CORTE	92	78	72	80%
FIABLE	1590-1595	LOS AMORES DE ALBANIO Y ISMENIA	91	70	81	80%
FIABLE	1597-1603	LOS BANDOS DE SENA	90	77	74	80%
FIABLE	1599	LA FE ROMPIDA	97	72	71	80%
FIABLE	1601-1603	LA GALLARDA TOLEDANA	90	82	70	80%
FIABLE	1604-1606	EL HOMBRE DE BIEN	88	78	75	80%
FIABLE	1606-1612	EL AMIGO HASTA LA MUERTE	96	74	72	80%
FIABLE	1606-1615	LA HISTORIA DE TOBÍAS	82	77	81	80%
FIABLE	1611?	BARLAÁN Y JOSAFAT	84	78	78	80%
FIABLE	1614	EL PRÍNCIPE PERFECTO I	91	81	70	80%
PROBABLE	1615-1622	EL ROBO DE DINA	92	75	74	80%
NO ES DE LOPE	1620	EL ALCALDE DE ZALAMEA	91	84	65	80%
FIABLE	1620-1622	EL HIJO DE LOS LEONES	93	75	73	80%
FIABLE	1624-1625	EL PREMIO DEL BIEN HABLAR	86	76	78	80%
DUDOSA	1630-1635	ACERTAR ERRANDO	93	72	75	80%

79%

AUTORÍA	FECHA	TÍTULO	C4	W2	W3	%
DUDOSA	Antes 1593?	LAS BURLAS Y ENREDOS DE BENITO	87	76	74	79%
FIABLE	1598-1600	LA HERMOSA ALFREDA	92	71	75	79%
FIABLE	1600	LOS EMBUSTES DE CELAURO	84	77	78	79%
FIABLE	1604	LA DESDICHADA ESTEFANÍA	87	70	80	79%
FIABLE	1604-1608	LA INOCENTE LAURA	93	77	68	79%

FIABLE	1610-1612	LAS FAMOSAS ASTURIANAS	86	66	86	79%
FIABLE	1612	EL BASTARDO MUDARRA	81	80	77	79%

78%

FIABLE	1589-1590	EL GRAO DE VALENCIA	93	70	71	78%
FIABLE	1593	EL FAVOR AGRADECIDO	94	70	70	78%
FIABLE	1593-1598	EL ENEMIGO ENGAÑADO	90	77	67	78%
FIABLE	1595-1599	LA VIUDA VALENCIANA	84	76	76	78%
FIABLE	1595-1603	LA FUERZA LASTIMOSA	90	71	73	78%
FIABLE	1599	EL ARAUCO DOMADO	87	71	77	78%
FIABLE	1602	EL CABALLERO DE ILLESCAS	91	71	72	78%
FIABLE	1603	EL ARENAL DE SEVILLA	87	75	72	78%
FIABLE	1604	CARLOS V EN FRANCIA	90	73	73	78%
FIABLE	1604-1606	EL ANZUELO DE FENISA	84	76	74	78%
FIABLE	1609	LA OCTAVA MARAVILLA	82	70	84	78%
FIABLE	1611	DON LOPE DE CARDONA	85	72	78	78%

77%

FIABLE	1593-1598	EL CABALLERO DEL MILAGRO	84	74	73	77%
FIABLE	1594	EL MAESTRO DE DANZAR	83	73	75	77%
FIABLE	1599-1608	EL GENOVÉS LIBERAL	88	76	68	77%
FIABLE	1599-1603	LOS ESCLAVOS LIBRES	88	73	70	77%
FIABLE	1599	EL BLASÓN DE LOS CHAVES	84	81	66	77%
FIABLE	1604-1608	LOS PORCELES DE MURCIA	91	76	65	77%
FIABLE	1604-1607	LAS MUDANZAS DE FORTUNA	80	77	74	77%
FIABLE	1608	LOS MELINDRES DE BELISA	81	75	77	77%
FIABLE	1608?-1610?	EL HAMETE DE TOLEDO	84	78	70	77%
FIABLE	1610?-1612?	LAS PACES DE LOS REYES	87	78	68	77%
FIABLE	1610?-1615?	LA DONCELLA TEODOR	87	75	71	77%
PROBABLE	1612-1614	LA ALDEHUELA	96	73	62	77%
FIABLE	1625	LA NIÑEZ DEL PADRE ROJAS	92	66	73	77%
FIABLE	1625-1630	VALOR, FORT. Y LEALT. T DE MENESES	84	81	68	77%

76%

FIABLE	1585-1588	LAS FERIAS DE MADRID	81	76	73	76%
FIABLE	1598-1600	EL HONRADO HERMANO	86	72	71	76%
FIABLE	1597-1603	ADONIS Y VENUS	82	76	70	76%
DUDOSA	1599-1603	EL ESCLAVO FINGIDO	86	76	66	76%
DUDOSA	1599	LOS CAUTIVOS DE ARGEL	93	63	74	76%

FIABLE	1600	LA CONTIENDA DE JUAN DE PAREDES	83	67	78	76%
FIABLE	1604-1606	LOS PRADOS DE LEÓN	86	71	73	76%
FIABLE	1606?	EL GRAN DUQUE DE MOSCOVIA	96	61	73	76%
FIABLE	1610-1612?	EL CONDE FERNÁN GONZÁLEZ	82	75	71	76%
FIABLE	1610-1612	LOS PONCES DE BARCELONA	86	75	67	76%
FIABLE	1610-1614	LA VILLANA DE GETAFE	88	75	67	76%
NO ES DE LOPE	/	SANTA CASILDA	90	70	70	76%

75%

FIABLE	1588-1595	EL GANSO DE ORO	87	76	62	75%
FIABLE	1590	EL PERSEGUIDO	91	68	66	75%
FIABLE	1595-1597?	JORGE TOLEDANO	95	67	63	75%
FIABLE	1596	EL REMEDIO EN LA DESDICHA	88	73	66	75%
FIABLE	1596-1602	LA ESCOLÁSTICA CELOSA	81	71	73	75%
FIABLE	1596-1603	EL ESCLAVO DE ROMA	83	72	71	75%
FIABLE	1597-1606	LOS ESPAÑOLES EN FLANDES	86	76	65	75%
PROBABLE	1597-1608	EL LOCO POR FUERZA	82	70	73	75%
FIABLE	1598-1603	EL NUEV. MUN. DESCU. POR C. COLÓN	84	68	75	75%
FIABLE	1602	EL PRÍNCIPE DESPEÑADO	92	67	68	75%
FIABLE	1612-1614	EL TRIUNFO DE LA HUMILDAD	86	68	71	75%
FIABLE	1625-1630	CONTRA VALOR NO HAY DESDICHA	91	68	66	75%
						75%

74%

FIABLE	1588-1595	EL VERDADERO AMANTE	86	78	60	74%
FIABLE	1599-1603	LA RESISTENCIA HONRADA	85	64	74	74%
FIABLE	1599-1603	EL VAQUERO DE MORAÑA	87	66	70	74%
FIABLE	1604-1608	DON JUAN DE CASTRO I	83	70	70	74%
FIABLE	1606	EL BOBO DEL COLEGIO	80	70	73	74%
FIABLE	1608-1612	EL JUEZ EN SU CAUSA	86	63	75	74%
FIABLE	1615?-1616?	EL PREMIO DE LA HERMOSURA	82	67	73	74%
FIABLE	1625	EL BRASIL RESTITUIDO	77	70	76	74%

73%

FIABLE	1585-1595	EL MOLINO	80	77	64	73%
FIABLE	1588-1595	LOS EMBUSTES DE FABIA	88	73	60	73%
FIABLE	1590	EL PRÍNCIPE INOCENTE	92	67	60	73%
FIABLE	1591-1595	EL DÓMINE LUCAS	78	75	68	73%
FIABLE	1594	BELARDO EL FURIOSO	76	78	65	73%
FIABLE	1595-1603	LA SANTA LIGA	73	63	83	73%
FIABLE	1598-1603	EL INGRATO ARREPENTIDO	81	71	68	73%
FIABLE	1602	EL CUERDO LOCO	91	55	74	73%
FIABLE	1603	EL NIÑO INOCENTE DE LA GUARD.	81	68	71	73%
FIABLE	1604-1606	SAN ISIDRO LABRAD. DE MADRID	78	67	75	73%
FIABLE	1610-1612	LA MADRE DE LA MEJOR	80	65	75	73%
FIABLE	1611-1612	JUAN DE DIOS Y ANTÓN MARTÍN	77	71	71	73%
DUDOSA	1612?- 1615?	LA FIANZA SATISFECHA	86	63	71	73%
FIABLE	1615	EL DESCONFIADO	80	68	72	73%
FIABLE	1620-1625	LA FILISARDA	82	73	66	73%
FIABLE	1629?	LA SELVA SIN AMOR	75	81	63	73%
DUDOSA	/	LA INTENCIÓN CASTIGADA	92	61	68	73%

72%

FIABLE	1595-1598	LA FRANCESILLA	82	70	64	72%
DUDOSA	1597?- 1603?	EL REY POR SU SEMEJANZA	87	62	67	72%
FIABLE	1597-1603	EL DESPOSORIO ENCUBIERTO	85	72	61	72%
FIABLE	1598-1600	ROMA ABRASADA	85	63	68	72%
FIABLE	1599?- 1600?	EL POSTRER GODO DE ESPAÑA	80	67	71	72%
FIABLE	1600	LOS BENAVIDES	80	66	70	72%
FIABLE	1603	PEDRO CARBONERO	77	74	67	72%
DUDOSA	1609-1615	LA MEJOR ENAMORADA, LA MAGDAL.	91	67	60	72%

71%

FIABLE	Antes 1596	LOS CELOS DE RODAMONTE	91	61	63	71%
DUDOSA	1597-1603	EL PLEITO POR LA HONRA	85	67	63	71%
FIABLE	1599	ANGÉLICA EN EL CATAY	75	67	71	71%
FIABLE	1608	DON JUAN DE CASTRO II	76	74	65	71%

70%

DUDOSA	1590-1605	EL PREMIO RIGUROSO	82	67	63	70%
FIABLE	1594	EL LEAL CRIADO	75	70	65	70%
FIABLE	1597	EL CASAMIENTO EN LA MUERTE	78	63	70	70%
FIABLE	1597	EL AMOR DESATINADO	83	64	64	70%
FIABLE	1599	EL ALCAIDE DE MADRID	86	57	68	70%
PROBABLE	1596?-1603	EN LOS INDICIOS, LA CULPA	78	67	65	70%
FIABLE	1599	EL AMIGO POR FUERZA	77	60	72	70%
FIABLE	1593-1603	EL BAUTIS. DEL PRÍNCIPE DE MARR.	83	72	57	70%
FIABLE	1604-1608	LAS GRANDEZAS DE ALEJANDRO	83	63	64	70%

69%

FIABLE	1588-1595	EL PRÍNCIPE MELANCÓLICO	81	66	62	69%
FIABLE	1588?-1595?	LA INFANTA DESESPERADA	83	53	72	69%
FIABLE	Antes 1596	LAS JUSTAS DE TEBAS	80	61	68	69%
FIABLE	1596-1603	EL SOL PARADO	76	65	68	69%
DUDOSA	1593-1603	LA COMPETENCIA EN LOS NOBLES	85	60	64	69%

68%

FIABLE	1594	LAURA PERSEGUIDA	86	62	58	68%
DUDOSA	1596	LOS DONAIRES DE MATICO	86	62	57	68%
FIABLE	1596	EL MARQUÉS DE MANTUA	72	68	66	68%
FIABLE	Antes 1598	LOS COMENDADORES DE CORBOBA	81	61	63	68%
FIABLE	1598	LA IMPERIAL DE OTÓN	82	56	66	68%
DUDOSA	1604	LA DEVOCIÓN DEL ROSARIO	76	62	67	68%
FIABLE	1605-1606	EL HIDALGO BENCERRAJE	86	62	57	68%

67%

DUDOSA	1596-1603	EL TOLEDANO VENGADO	80	66	57	67%
FIABLE	1598-1603	LOS PLEITOS DE INGLATERRA	90	50	63	67%
DUDOSA	Antes 1605	SANTO ANGELO	78	58	65	67%

66%

FIABLE	1599-1603	EL NEGRO DEL MEJOR AMO	73	61	64	66%
FIABLE	1618	LA LIMPIEZA NO MANCHADA	73	62	65	66%

DUDOSA	/	JULIÁN ROMERO	75	64	61	66%
FIABLE	1629	LA VIDA DE SAN PEDRO NOLASCO	77	63	58	66%

65%

NO ES DE LOPE	/	LOS MÁRTIRES DE JAPÓN	80	60	56	65%
NO ES DE LOPE	/	EL TRUHÁN DEL CIELO	82	60	55	65%

64%

AUTORÍA	FECHA	TÍTULO	C4	W2	W3	%
FIABLE	1595?- 1596?	LA BELLA MALMARIDADA	68	65	61	64%
FIABLE	1598-1600	LA CAMPANA DE ARAGÓN	78	57	57	64%
FIABLE	1599-1603	LA DIVINA VENCEDORA	74	57	62	64%
DUDOSA	1620-1635	EL INGRATO	76	60	56	64%
FIABLE	1625	EL NACIMIENTO DE CRISTO	60	66	68	64%
NO ES DE LOPE	/	EL REY POR TRUEQUE	81	55	57	64%

63%

FIABLE	1588-1595	URSÓN Y VALENTÍN	72	62	55	63%
DUDOSA	1597-1603	LAS DOS BANDOLERAS	77	62	51	63%

62%

PROBABLE	1593-1603	LA DIFUNTA PLEITEADA	77	54	56	62%
FIABLE	1596-1603	EL TESTIMONIO VENGADO	72	56	58	62%
DUDOSA	1599-1603	LA PRÓSP. FORT. DE D BERN. CABRERA	76	49	61	62%
DUDOSA	1616	LA ILUSTRE FREGONA	67	55	64	62%

61%

DUDOSA	/	LAS DONCELLAS DE SIMANCAS	65	63	57	61%
DUDOSA	1599-1603	LOS GUZMANES DEL TORAL	83	50	52	61%

60%

DUDOSA	1613?- 1615?	EL ANTICRISTO	78	49	55	60%
NO ES DE LOPE	/	ANTONIO ROCA	65	55	60	60%
NO ES DE LOPE	/	EL VASO DE ELECCIÓN SAN PABLO	71	50	61	60%

58%

DUDOSA	/	AUDIENCIAS DEL REY DON PEDRO	58	57	61	58%
NO ES DE LOPE	/	DAVID PERSEGUIDO	68	56	51	58%

57%

DUDOSA	-	LA CORONA DERRIBADA	72	48	53	57%
--------	---	---------------------	----	----	----	-----

56%

DUDOSA	1596-1600	EL HIJO POR ENGAÑO	62	48	60	56%
--------	-----------	--------------------	----	----	----	-----

55%

DUDOSA	/	LOS TERCEROS DE SAN FRANCISCO	61	56	48	55%
--------	---	-------------------------------	----	----	----	-----

54%

NO ES DE LOPE	-	LA ADVERSA FORT. DE D. B DE CABRERA	55	56	53	54%
NO ES DE LOPE	-	LA CIUDAD SIN DIOS	60	53	51	54%

52%

NO ES DE LOPE	/	LA MAYOR CORONA	56	49	51	52%
---------------	---	-----------------	----	----	----	-----

51%

FIABLE	1579-1583?	LOS HECHOS DE GARCILASO	46	43	64	51%
DUDOSA	/	DONDE NO ESTÁ SU DUEÑO	47	51	55	51%

48%

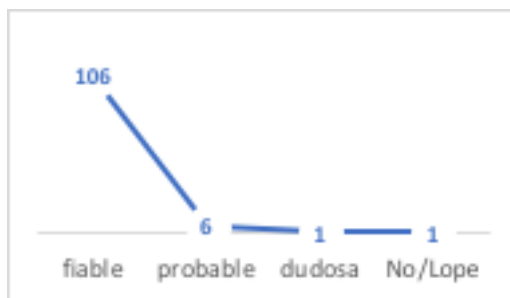
FIABLE	Antes 1607	EL SANTO NEGRO ROSAMBUCO	51	47	48	48%
NO ES DE LOPE		LA ESTRELLA DE SEVILLA	46	46	52	48%

45%

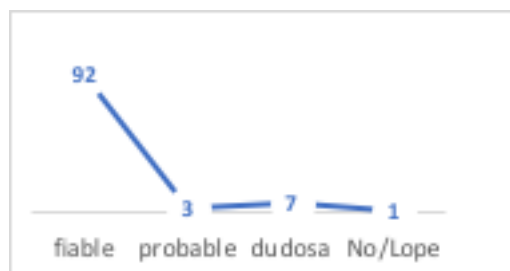
PROBABLE	1590-1604	LA MADRE TERESA DE JESÚS	58	37	40	45%
----------	-----------	--------------------------	----	----	----	-----

En el cómputo total de las 268 comedias analizadas encontramos los siguientes datos pertinentes.

- 1) La media de proximidad de las 268 comedias estudiadas con respecto a las cuatro comedias de Lope es de 76%, siendo el máximo de proximidad 93% (*La mayor victoria*) y el mínimo 45% (*La madre Teresa de Jesús*).
- 2) De las ciento catorce (114) comedias que promedian entre 93% y 80% de proximidad con respecto a las de Lope, ciento seis (106) están consideradas de autoría fiable, seis (6) de autoría probable, una (1) de autoría dudosa y otra (1) como no de Lope, según la crítica especializada.



- 3) De las ciento tres (103) comedias que promedian entre 79% y 70% de proximidad con respecto a las de Lope, noventa y dos (92) aparecen como comedias de autoría fiable, tres (3) de autoría probable, siete (7) como de autoría dudosa y una sola (1) como no de Lope, siempre según la crítica especializada.



- 4) Del 69% al 45% de proximidad con respecto a las de Lope, las comedias consideradas de autoría fiable son veintiuna (21), de autoría probable dos (2), de autoría dudosa dieciocho (18) y etiquetadas como «no de Lope» diez (10).



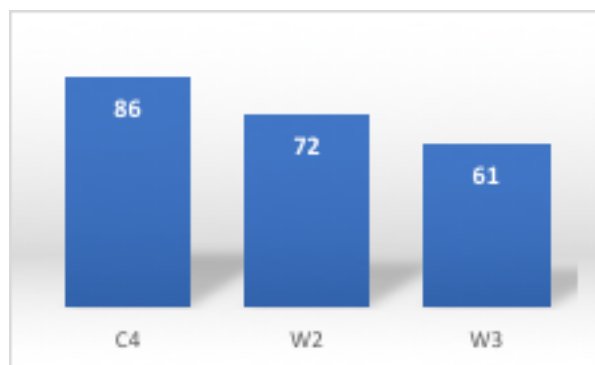
- 5) Debe notarse que hasta 69 comedias con un promedio de proximidad inferior al 80% y consideradas de autoría fiable están fechadas antes de 1599. El caso más evidente es la primera comedia escrita supuestamente por Lope en su adolescencia, *Los hechos de Garcilaso*, que apenas llega al 51% de proximidad.

Con estos datos puede concluirse que:

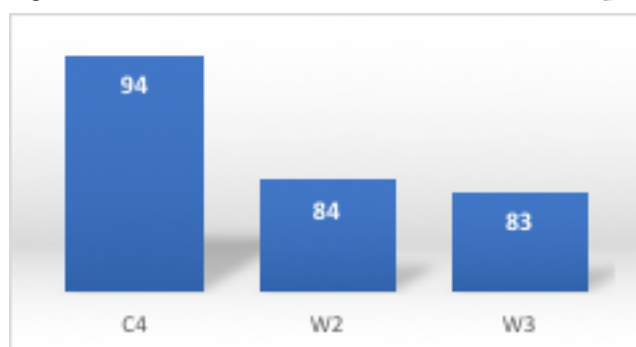
- Todas las comedias analizadas con un promedio de más del 80% de proximidad respecto a las de Lope deben considerarse escritas por Lope o ser vistas como una refundición muy cercana al original lopesco, como sería el caso de *Acertar errando* o *El alcalde de Zalamea*.
- Las comedias fechadas después de 1600 que son vistas por la crítica como dudosas o no de Lope y con un promedio inferior al 70% deben excluirse del canon lopesco.
- Las comedias que la crítica ha considerado de autoría fiable hasta la fecha y que tienen un promedio inferior al 70% deben revisarse a fondo, especialmente si están fechadas entre 1600 y 1635, pues es muy probable que se trate de obras ajenas a Lope o refundiciones.

Deberíamos también aclarar lo siguiente:

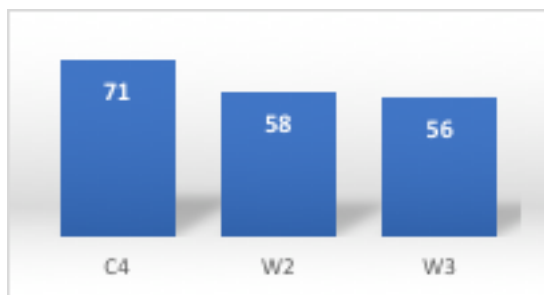
Si hacemos el promedio con las 268 comedias computadas, C4 supera a W2 en un 14% y a W3 en un 25%:



El cálculo se hace con todas las comedias, incluidas las que se consideran no de Lope o de autoría dudosa. Ahora bien, si nos ceñimos a aquellas que alcanzaron un promedio de al menos 85%, los n-gramas de 4 caracteres (C4) rondan el 95% de proximidad.



En cuanto a las comedias por debajo del 70%, C4 vuelve a obtener un promedio algo mayor que W2 y W3, probablemente debido a que los n-gramas de caracteres están menos condicionados por el léxico o por el asunto.



El cóctel de C4, W2 y W3 con una métrica de similitud de coseno es, creo, una fórmula muy eficaz en la atribución de textos, siempre y cuando tengamos presente que los n-gramas de palabras tienden a agrupar textos en virtud del léxico empleado. Así las cosas, las comedias de autoría probable que, según el análisis llevado a cabo, parecen estar claramente en la órbita de Lope son:

- Porfiar hasta morir* (87%)
- La Amistad y obligación* (86%)
- Amar como se ha de amar* (85%)
- Las burlas y veras* (84%)
- La esclava de su galán* (81%)
- El robo de Dina* (80%)

Podrían añadirse *La Aldehuela* (77%) y *El loco por la fuerza* (75%), aunque ya con reparos. En cuanto a *La difunta pleiteada* (62%) y *La madre Teresa de Jesús* (45%), las otras dos comedias vistas como probables, están ya muy alejadas de la órbita de Lope y deben etiquetarse, más bien, como comedias «no de Lope».

Casi ninguna comedia considerada de autoría dudosa alcanza el 70% de proximidad y la mayoría están por debajo. Las únicas excepciones son *Acertar errando* (80%) y *Las burlas y enredos de Benito* (79%). Esta última parece anterior a 1593, según Morley y Bruerton, lo cual podría apoyar la autoría de Lope o, cuando menos, sembrar dudas respecto a esa supuesta dudosa autoría. Las demás que pongo debajo sí creo que deben considerarse de otra mano o refundiciones hechas a partir de comedias de Lope:

- La competencia en los nobles* (69%)
- La devoción del rosario* (68%)
- Santo Angelo* (67%)
- El toledano vengado* (67%)
- Julián Romero* (66%)
- El ingrato* (64%)
- Las dos bandoleras* (63%)
- La próspera fortuna de don Bernardo de Cabrera* (62%)
- La ilustre fregona* (62%)
- Las doncellas de Simancas* (61%)
- Los Guzmanes del Toral* (61%)
- El anticristo* (60%)
- Las audiencias del rey don Pedro* (58%)

La corona derribada (57%)
El hijo por engaño (56%)
Los terceros de San Francisco (55%)
Donde no está su dueño está su duelo (51%)

Todas las comedias etiquetadas por ARTELOPE como «no de Lope» están en mi análisis por debajo del 70%, con dos solas excepciones: *El alcalde de Zalamea* y *Santa Casilda*. Debajo pongo la lista:

El alcalde de Zalamea (80%)
Santa Casilda (76%)
Los mártires de Japón (65%)
El truhán del cielo (65%)
El rey por trueque (64%)
Antonio Roca (60%)
El vaso de elección (60%)
David perseguido (58%)
La adversa Fortuna de don Bernardo de Cabrera (55%)
La ciudad sin Dios (54%)
La mayor corona (52%)
La estrella de Sevilla (48%)

Me detendré muy someramente en alguna de estas comedias consideradas como «no de Lope». La más famosa es, sin duda, *La estrella de Sevilla*. El Fénix, con una proximidad del 48%, debe quedar definitivamente descartado. No mejor parados salen los otros cinco dramaturgos empleados en el cotejo, tal como se observa en la tabla:

90	C4	W2	W3	
MIRA	56	70	71	65%
TIRSO	58	65	62	61%
CLARAMONTE	57	57	60	58%
CASTRO	40	51	56	49%
ALARCÓN	53	26	27	35%

Bien es verdad que Claramonte obtiene un resultado óptimo si contabilizamos solo las dos primeras comedias clasificadas entre las 24 posibles¹¹.

47	C4	W2	W3	
CLARAMONTE	100	100	95	98%
LOPE	80	53	72	68%

11.- En este experimento la comedia que aparece en la primera casilla obtiene 24 puntos, en la segunda 23, en la tercera, 22, y así sucesivamente hasta la última. La máxima puntuación que un autor puede lograr con sus dos comedias en cada columna es 47 puntos (100%) y la peor 3 puntos (2%).

Debo resaltar que las comedias *Deste agua no beberé* y *Púsoseme el sol, saliome la luna*¹² de Claramonte aparecen con frecuencia cerca de las comedias de Lope, mientras que las otras dos (*El valiente negro* y *De lo vivo a lo pintado*) se sitúan una y otra vez al final de la lista. Es difícil interpretar este fenómeno. Puede que las dos primeras sean refundiciones hechas a partir de otras comedias, y de ahí la confusión que crea en el programa, pero no barajo ningún dato para fundamentar tal conjetura.

La ciudad sin Dios, etiquetada en ARTELOPE, como «no de Lope», aparece también con números muy semejantes a *La estrella de Sevilla*:

47	C4	W2	W3	
CLARAMONTE	100	100	95	98%
LOPE	76	36	57	56%

Un 98% de proximidad no debería hacernos dudar, pero este alto porcentaje se diluye en cuanto incluimos en el cómputo *El negro más valiente* y *De lo vivo a lo pintado*, mientras que desaparece del todo si estas dos últimas comedias se analizan de manera independiente.

La mayor corona creo que es la única comedia de las aquí analizadas en donde la autoría de Claramonte (o su participación en ella) parece probable, ya que, a diferencia de otros casos, sus cuatro comedias ofrecen un grado de proximidad notable en los tres parámetros empleados:

90	C4	W2	W3	
CLARAMONTE	86	78	91	85%
LOPE	56	49	51	52%

Andrés de Claramonte ha sido de siempre un problema en la atribución textual de la Comedia Nueva. Hartzenbusch, Menéndez Pelayo, Cotarelo y otros mostraban a menudo irritación hacia el dramaturgo murciano por lo que consideraban inaceptables intrusiones en las obras de los grandes autores de la época. Puede que tal actitud nos resulte actualmente un poco anacrónica respecto al *modus operandi* en el teatro de la época. Al fin y al cabo, eran los «autores» de las compañías y no los poetas los verdaderos dueños del producto. Lope, hartado con los abusos, será el primero en publicar sus comedias con el fin de controlar su obra, pero es muy posible que más de una comedia publicada por él esté ya muy deturpada o sea una mera refundición, especialmente las que estrenó quince o veinte años antes de la publicación. El mismo Lope parece concederle en algún prólogo. Solamente las obras de su última época deben resultarnos verdaderamente fiables.

Desde luego las cuestiones de atribución en el teatro del Siglo de Oro resultan una labor hartada complicada. Se requiere rigor y mucha cautela, aunque a veces la cosa parece clara. Así pasa con la atribución de *Donde no está su dueño está su duelo*. En efecto, no es de Lope y, tal como ha visto la crítica¹³, todo apunta a que la comedia haya sido concebida por Guillén de Castro. El análisis que llevo a cabo así lo atestigua:

12.– La atribución de *Púsoseme el sol, saliome la luna* no está nada clara. El primer manuscrito conocido tiene tachado el nombre de Lope y luego, en su lugar, se ha escrito el de Claramonte (BNE, ms. 14955). Otra parte “extravagante” (P XXVI) está a nombre de Lope. (Lo recojo de JAURALDE POU).

13.– JULIÁ MARTÍNEZ (1926), ARJONA (1965).

90	C4	W2	W3	
CASTRO	78	88	90	85%
LOPE	47	51	55	51%

Según se va viendo, toda comedia que promedia un resultado por encima del 80% de proximidad con respecto al grupo de comedias de uno de los autores seleccionados tiene muchas papeletas para pertenecer al corpus de ese autor en cuestión.

Otro procedimiento paralelo al utilizado hasta ahora sería calcular exclusivamente las comedias que ocupan las cuatro primeras plazas en cada una de las tres categorías y puntuar de 4 a 1 con un máximo de 30 puntos (100%). Lo ilustraré con dos ejemplos. En el primer caso elijo *La mayor victoria*, que es la comedia que ha sacado el promedio más alto (93%) en el análisis anterior:

	C4	W2	W3
La mayor victoria	1. LOPE El castigo sin venganza 2. LOPE El caballero de Olmedo 3. LOPE El perro del hortelano 4. LOPE Fuente Ovejuna	1. TIRSO El vergonzoso en palacio 2. LOPE El perro del hortelano 3. LOPE El castigo sin venganza 4. LOPE El caballero de Olmedo	1. LOPE El perro del hortelano 2. MIRA La Fénix de Salamanca 3. TIRSO El vergonzoso en palacio 4. LOPE Fuente Ovejuna

Lope obtiene 21 puntos de 30 (70%), Tirso 6/30 (20%) y Mira 3/30 (10%). En el gráfico de debajo vemos mejor cómo se reparte el pastel:



Analicemos ahora la comedia que obtuvo el promedio más bajo en el cotejo anterior, *La madre Teresa de Jesús*, con un 45%.

	C4	W2	W3
La madre Teresa	1. TIRSO El vergonzoso en palacio 2. CLARAMONTE Deste agua no 3. ALARCÓN La verdad sospechosa 4. TIRSO Caballero de Gracia	1. TIRSO El vergonzoso en palacio 2. MIRA La Fénix de Salamanca 3. TIRSO Don Gil de las calzas verdes 4. CASTRO El nacimiento de Montes.	1. MIRA La Fénix de Salamanca 2. TIRSO El vergonzoso en palacio 3. CLARAMONTE Deste agua no beberé 4. CASTRO El nacimiento de Montesinos

Lope no logra un solo punto y es Tirso quien alcanza el promedio más alto 14/30 (»47%).



Si calculamos el promedio con las cuatro comedias, Tirso llega hasta el 78%; y si con las tres primeras, tenemos un respetable 84% de proximidad. No digo que el mercedario sea el autor de *La madre Teresa de Jesús*, pero se debe estudiar esa posibilidad una vez descartada la autoría lopesca. En Apéndice van todos los resultados con este segundo procedimiento en relación con las 268 comedias de Lope. Aquí yo solo analizaré aquellas del Fénix que estaban por debajo del 70% en el primer análisis y suman ahora, con este nuevo procedimiento, al menos 11 puntos de 30 (37%) en relación con alguno de los otros cinco autores del corpus cotejado. En total, he identificado 28 comedias con estas características, tal como se ve en la tabla de abajo:

	ALARCÓN	CASTRO	CLARAMONTE	LOPE	MIRA	TIRSO
ANGÉLICA EN EL CATAY			56	10	17	17
EL ANTICRISTO			43	6	20	23
EL ARAUCO DOMADO			43	33	10	13
BARLAÁN Y JOSAFAT			47	16	14	23
BARLAÁN Y JOSAFAT REFUNDICIÓN			60	13	16	10
LA BELLA MALMARIDADA	3	3	10	27	43	13
EL BRASIL RESTITUIDO			60	10	13	17
LOS CELOS DE RODAMONTE			43	20	20	16
EL CABALLERO DE ILLESCAS	3			33	23	40
LA CIUDAD SIN DIOS			64		27	10
DAVID PERSEGUIDO	6	10	43		10	30

	ALARCÓN	CASTRO	CLARAMONTE	LOPE	MIRA	TIRSO
EL DESCONFIADO	10		13	10	23	40
LAS DONCELLAS DE SIMANCAS		37	20		20	23.3
DONDE NO ESTÁ SU DUEÑO		60	13.3		16.6	10
LAS DOS BANDOLERAS			40	10	23.3	26.6
LAS GRANDEZAS DE ALEJANDRO		7	43	13	16	23
JULIÁN ROMERO	10		26	10	40	20

LA LIMPIEZA NO MANCHADA	3	3	43	10	30	13
EL NACIMIENTO DE CRISTO			70	17	13	
EL NUEVO MUNDO... CRISTÓBAL COL	3		47	13	20	17
EL PRÍNCIPE MELANCÓLICO		40		17	20	23
LA PRÓSPERA FORTUNA			37	3	33	27
EL REY POR TRUEQUE	3	53		6	23	13
SANTO ANGELO	13			6	33	47
EL SANTO NEGRO ROSAMBUCO			53		27	20
EL SOL PARADO	3		40	13	23	13
LOS TERCEROS DE SAN FRANCISCO	3		13	6	23	53
EL VASO DE ELECCIÓN SAN PABLO			47	6	23	20

Claramonte obtiene los mejores números en 19 comedias. Los otros son, por este orden, Guillén de Castro (4), Tirso (3) y Mira de Amescua (2). De Claramonte ya comprobamos que la proximidad casi siempre se da con *Deste agua no beberé* y *Púsoseme el sol* y raramente con sus otras dos comedias (*De lo vivo a lo pintado* y *El valiente negro*), pero veamos más de cerca algunos casos. Empezaré por *El nacimiento de Cristo* impresa por primera vez en 1641 dentro de la Parte XXIV de la colección de comedias de Lope de Vega. En el gráfico de debajo se aprecia nítidamente la desproporción entre los tres autores que comparten las primeras plazas en las tres categorías analizadas:



La crítica cataloga esta comedia de autoría probable, pero en los cálculos anteriores Lope no llegaba al 64% y aquí no pasa del 17%. En cuanto a Claramonte, volvemos a toparnos con que dos de sus comedias se sitúan al principio y las otras dos muy al final. Es muy difícil así sacar conclusiones, salvo barajar dos hipótesis: 1) *Deste agua* y *Púsoseme el sol* están en buena medida escritas por otra mano de la que escribió *El valiente negro* y *De lo vivo a lo pintado*; o 2) *Deste agua*, *Púsoseme el sol* y *El Nacimiento de Cristo*, entre otras, son refundiciones hechas por Claramonte o por alguien cercano a él. Debajo se puede apreciar la gran disparidad:

	C4	W2	W3
NACIMIENTO DE CRISTO	1. CLARAMONTE Pú- soseme el sol	1. CLARAMONTE Púsoseme el sol	1. CLARAMONTE Deste agua no beberé
	2. CLARAMONTE Deste agua no beberé	2. CLARAMONTE Deste agua no beberé	2. CLARAMONTE Púsoseme el sol
	3. LOPE Fuente Ovejuna	3. MIRA DE AMESCUA El ejemplo mayor	3. MIRA DE AMESCUA La Fénix de Sal.
	4. LOPE El caballero de Ol- medo	4. LOPE Fuente Ovejuna	4. LOPE Fuente Ovejuna

	22. CLARAMONTE De lo vivo a lo	22. CLARAMONTE De lo vivo a lo	19. CLARAMONTE De lo vivo a lo
23. CLARAMONTE El va- liente negro	23. CLARAMONTE El vali- ente negro	23. CLARAMONTE El vali- ente negro	

El nacimiento de Cristo, según ya dije, está impresa en una de las Partes de la colección de comedias de Lope y, por ello, su autoría no puede descartarse sin más. Es cierto que los números que manejamos ponen a Claramonte (o dos de sus comedias, más bien) en la misma órbita de *El nacimiento*, pero el tema religioso-alegórico es muy afín a Lope. En los primeros versos del primer Acto un personaje abstracto, La Soberbia (que no es otro que Lucifer), dice:

Vivo en mi opinión tan firme,
que a un primero movimiento
no tengo de arrepentirme;
porque al arrepentimiento
no puede Dios persuadirme.
Verdad es que no lo intenta;
pero, cuando lo intentara,
fuera mi obediencia exenta,
y le dijera en su cara
que era arrepentirme afrenta. (vv. 6-15).

En *Juan de Dios y Antón Martín* tenemos algo muy parecido en boca del demonio:

¿Yo confesor? No puedo confesarme,
porque si nunca **arrepentirme** pude
¿cómo diré que es culpa haber querido
alzarme con el cielo...? (Acto III, vv. 3103-3106)

Es posible que tal situación se pueda dar en otras obras de corte religioso, pero es que poco después entra la Envidia en acción y es así descrita:

Sale ENVIDIA con un corazón en las manos, ceñida la cabeza de culebras

Pues bien, en *El peregrino en su patria* tenemos casi lo mismo:

Con éste venía la **Envidia**, casi en el hábito que la pinta Ovidio, crinada **la cabeza de culebras**.

Y todavía en *La dragrantea* leemos:

Sale la fiera abominable Aletto
Por mil volcanes de diversas quiebras

Del Erebo espantoso, a un triste efeto
Crinada **la cabeza de culebras**.

Entremedias recojo en los primeros versos otras secuencias compartidas con el corpus de Lope:

<i>El nacimiento de Cristo</i>	LOPE
le dijera en su cara	...osó decirle en su cara (<i>La contienda de Diego G. Paredes</i>)
a dos Reyes que ha criado, el cetro del mundo ha dado	Hizo este capitán tales hazañas En años treinta y tres... Mató la muerte, reparó la vida, Encadenó al pecado, y al demonio Quitó el cetro del mundo , y con mil triunfos, Con mil palmas angélicas y lauros, Subió a la diestra de su Eterno Padre. (<i>El viaje del alma</i>)
Y anduvo tan liberal, que todo cuanto ha criado comen con licencia igual	Pues, Otavio, si codicias Ser tan liberal conmigo, Yo me contento con dos (<i>Quien ama no haga fieros</i>) Julio, hago testigo al cielo, a cuanto ha criado , a ti, a mi honra, a ese poco entendimiento mío... (<i>La Dorotea</i>)

Por lo que se ve en estos pocos ejemplos, Lope no se puede desvincular tan fácilmente de *El nacimiento de Jesucristo*. Hagamos un nuevo experimento con el programa JPAAG, pero esta vez poniendo en el corpus comparativo las comedias *La mayor victoria* y *Juan de Dios* y *Antón Martín*, en lugar de *El castigo sin venganza* y *El caballero de Olmedo*.

	C4	W2	W3
NACIMIENTO DE CRISTO	1. CLARAMONTE Púsoseme el sol	1. LOPE <i>La mayor victoria</i>	1. LOPE <i>Juan de Dios</i> y <i>Antón Martín</i>
	2. CLARAMONTE Deste agua no beberé	2. CLARAMONTE Púsoseme el sol	2. CLARAMONTE Deste agua no beberé
	3. LOPE <i>Fuente Ovejuna</i>	3. LOPE <i>Juan de Dios</i> y <i>Antón Martín</i>	3. LOPE <i>La mayor victoria</i>
	4. TIRSO <i>Marta la piadosa</i>	4. CLARAMONTE Deste agua no beberé	4. CLARAMONTE Púsoseme el sol

21. CLARAMONTE De lo vivo a lo pint	22. CLARAMONTE De lo vivo a lo pint.	19. CLARAMONTE De lo vivo a lo	
22. CLARAMONTE El valiente negro	23. CLARAMONTE El valiente negro	23. CLARAMONTE El valiente negro	
24. LOPE <i>Juan de Dios</i> y <i>Antón Martín</i>			

Ahora Lope (47%) y Claramonte (50%) promedian casi lo mismo, pero es perentorio señalar que en la categoría C4 (n-gramas de 4 caracteres), la comedia *Juan de Dios* queda en último lugar, mientras que en la categoría W3 (n-gramas de tres palabras) pasa a ocupar el primer puesto entre las 24 comedias cotejadas. ¿Cómo se explica algo así? En principio, el asunto religioso que ambos textos comparten puede influir. Al fin y al cabo,

la métrica de similitud de coseno está especialmente pensada para clasificar los textos en virtud de su cercanía temática. Elijamos entonces solo comedias religiosas del Fénix y veamos lo que pasa:

	C4	W2	W3
NACIMIENTO DE CRISTO	1. CLARAMONTE Púsoseme el sol 2. CLARAMONTE Deste agua no beberé 3. LOPE La hermosa Ester 4. TIRSO Marta la piadosa 8. LOPE San Isidro 17. LOPE Lo fingido verdadero 21. CLARAMONTE De lo vivo a lo pinta 22. CLARAMONTE El valiente negro 24. LOPE Juan de Dios y Antón Martín	1. LOPE La hermosa Ester 2. LOPE Lo fingido verdadero 3. CLARAMONTE Púsoseme el sol 4. LOPE San Isidro 5. LOPE Juan de Dios y Antón Martín	1. LOPE Juan de Dios y Antón Martín 2. LOPE La hermosa Ester 3. LOPE Lo fingido verdadero 4. LOPE San Isidro 5. CLARAMONTE Deste agua no beberé

El resultado de esta última operación aclara varias cosas. En primer lugar, que Claramonte se mantiene firme en los n-gramas de 4 caracteres (C4). Segundo, que los n-gramas de dos y tres palabras (W2 y W3) priman mucho más el aspecto léxico, y de ahí que las cuatro comedias religiosas de Lope tengan un mejor rendimiento. Tercero, que la métrica de similitud de coseno es de gran ayuda para clasificar y discriminar textos, pero no es en absoluto la fórmula mágica que vaya a resolvernos todos los problemas de atribución. Un solo ejemplo debe alertarnos al respecto. Elijo para ello una comedia hagiográfica, *La segunda parte de la Santa Juana* de Tirso de Molina, uno de los pocos autógrafos conocidos del mercedario.

	C4	W2	W3
LA SANTA JUANA II	1. TIRSO Caballero de Gracia 2. CLARAMONTE Púsoseme el sol 3. CLARAMONTE Deste agua no beberé 4. TIRSO Marta la piadosa	1. TIRSO El vergonzoso en palacio 2. TIRSO Caballero de Gracia 3. TIRSO Marta la piadosa 4. CLARAMONTE Deste agua no beberé	1. LOPE Juan de Dios y Antón Martín 2. LOPE La hermosa Ester 3. LOPE Lo fingido verdadero 4. LOPE San Isidro

En C4 y W2 los resultados son muy favorables a Tirso e incluso a Claramonte, mientras que en W3 nos encontramos con que las cuatro comedias religiosas de Lope copan las primeras plazas. No parece haber duda, pues. Los n-gramas de tres palabras (W3) dentro de la métrica de similitud de coseno clasifican los textos por la proximidad léxica y son, por ello, menos fiables en cuestiones de atribución, al menos de manera aislada.

Pasemos a *El Brasil restituido*, comedia de Lope que viene avalada, por si fuera poco, con un autógrafo parcial, aunque no se menciona en *El peregrino* ni está en ninguna de las Partes. En el primer examen que hicimos las cuatro comedias del Fénix promediaban un 74% de proximidad:

	C4	W2	W3	
EL BRASIL RESTITUIDO	77	70	76	74%

En el segundo análisis Claramonte vuelve a llevarse la parte del león con un 60%.

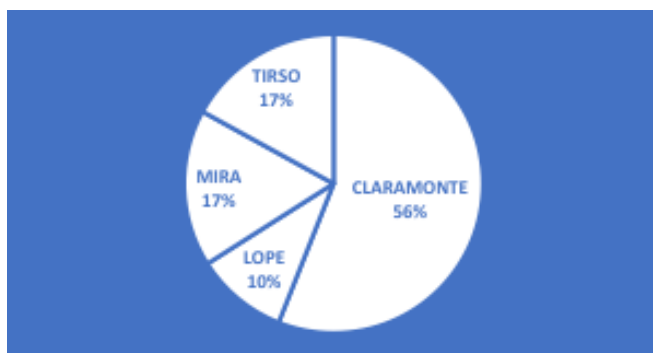


Lo mismo que dijimos con el *Nacimiento de Cristo* podemos repetirlo aquí. Claramonte supera a todos una vez más, pero solamente con dos comedias, pues en las otras dos queda muy al final. Y si espigamos paralelismos entre el corpus de Lope y *El Brasil restituido*, bastan 20 versos para persuadirnos de la autoría lopesca. Aquí pongo unos cuantos:

EL BRASIL RESTITUIDO (I, vv. 1-5)	LOPE
¿Quién duda que lo son, siendo de amor, de un alma, en tu rigor, de todo favor desnuda?	¿ Quién duda que le soys más inclinado? (<i>La dama boba</i>) Quién duda que me suceda lo que temo y adivino (<i>El molino</i>)
EL BRASIL RESTITUIDO (I, vv. 6-10)	LOPE
Luego que un hombre se muda, todo le parece injusto, que como os obliga el gusto a las locuras que hacéis, después que no le tenéis, ¿qué habéis de tener por justo?	Celia, a quien no tiene gusto , el alegría entristece; sólo le alegra el disgusto, el mal justo le parece, el bien le parece injusto . (<i>El secretario de sí mismo</i>) Pienso, mi bien, que te pago, puesto que siempre te debo, pero oblígame de nuevo a las locuras que hago (<i>La ocasión perdida</i>)

EL BRASIL RESTITUIDO (I, vv. 10-15)	LOPE
Justo fuera que yo hubiera mirado más por mi honor, pero siendo ciego amor, o no viera o no lo fuera...	<p>Cel. ¿Pues no ha de haber diferencia? Sol. Justo fuera que la hubiera (<i>Los tres diamantes</i>)</p> <p>No harán, que estaré escondido; aunque mucho mejor fuera que yo ser tú me fingiera, y hablara a Paula atrevido (<i>El ausente en el lugar</i>, III, 3)</p>

Parecida situación nos encontramos con la comedia *Angélica en el Catay*. Está incluida en la Parte VIII de la colección de comedias de Lope de Vega y aparentemente bajo la supervisión de su autor, con lo cual se hace difícil pensar que esté escrita por otro. Cabe la posibilidad, eso sí, de la colaboración (aunque no se conocen muchos casos hasta bastante más adelante) o de alteraciones en el original. Morley y Bruerton sugieren la fecha de 1599 para su redacción. En general, las comedias de Lope fechadas antes de 1600 obtienen peores resultados en los análisis llevados a cabo. En el primero que hicimos con *Angélica en el Catay*, Lope promediaba el 71% y en el segundo Claramonte ocupa más del 50%:



Tales números podrían hacernos cuestionar la autoría de *Angélica en el Catay*, pero una barrida rápida por los primeros versos del primer acto nos ofrecen varios paralelismos con el corpus de Lope:

ANGÉLICA EN CATAY	LOPE
¿Tú me amenazas a mí?	Pues, ¿tú me amenazas ya? (<i>El rústico del cielo</i>)
Reinaldos: ¿Hablas como primo? Roldán: No. Reinaldos: ¿Y como enemigo?	Ah, mi Fabio, hablas como amigo en todo (<i>El halcón de Federico</i>)
Pues qué, no soy hombre en quien cabrá tu amenaza bien	No soy hombre yo que mato sin culpa a nadie (<i>Fuente Ovejuna</i>) Elisa: ¿Quién es? Feliciano: Un hombre en quien hoy hablaste (<i>El ausente en el lugar</i>)

¿Que aún no vale un hombre honrado, por lo menos caballero, para ser amenazado, trayendo otro tanto acero como agora traes al lado?	¿No valdrá el hombre que digo para ser testigo, pues se confederó contigo? ¿Tan malo es que aún no vale por testigo? (Comedia de los amigos enojados)
---	---

El teatro áureo está todo él repleto de frases hechas y versos compartidos, pero tantas correspondencias en menos de 20 versos suelen ser indicio de una misma mano. La mayoría, además, son secuencias singulares, como es el caso de la colocación léxica <no valer hombre... para ser...>, inserta dentro de un n-grama de cuatro palabras (*que aún no vale*), sin equivalente en ningún documento de CORDE.

Las otras comedias atribuidas a Lope con alta representación de Claramonte¹⁴ siguen el guion previsto: *Deste agua y Púsoseme* se reparten las primeras plazas, mientras *De lo vivo a lo pintado* y *El negro más valiente* aparecen una y otra vez al final. Veremos este mismo fenómeno cuando analicemos *El condenado por desconfiado* y *El burlador de Sevilla*. En cuanto a las otras comedias que se alejan de Lope, *Donde no está su dueño está su duelo* es seguramente de Guillén de Castro, tal como vimos. Podría ser también de Castro *El rey por trueque* a tenor de los números que obtiene. En el primer análisis sus cuatro comedias promedian 87% y en el segundo ocupa más de la mitad 16/30 (53%). Los números son, pues, favorables, pero al hacer un rápido peinado por los primeros cien versos apenas se encuentra una correspondencia relevante con su corpus de comedias; y, aún peor, varias colocaciones léxicas extraídas de esos primeros versos no presentan un solo equivalente en las comedias del valenciano. Puede ser casualidad y no cuento con el corpus total, ciertamente, pero por experiencia sé que una frase del tipo «sin duda alguna» escrita dos o tres veces en un texto debería estar también en otros escritos similares del mismo autor. Debajo pongo esta frase y otras que gravitan en torno a la palabra «cosa».

	EL REY POR TRUEQUE
(SABER) QUÉ COSA *	¿ Sabes qué cosa es pasar amor y sufrir desdenes? ... mas como su mal no es burla, no sabe qué cosa sea
LA PRIMER COSA *	Prometió Jepté a Dios sacrificarle la primer cosa que en su casa viese la primer cosa que el encuentro vía, la palabra cumplió que dado había.
COSA NO ME *	Cosa no me satisfizo, si no fue la bota pura. Conde, cosa no me digas, ni me alegues más razón.

14.- *El santo negro Rosambuco, Barlaán, El nuevo mundo descubierto por Cristóbal Colón, El vaso de elección, El anticristo, El Arauco domado, Los celos de Rodamonte, David perseguido, Las grandezas de Alejandro y La limpieza manchada.*

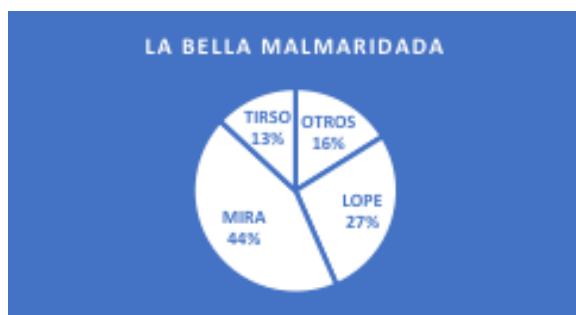
	EL REY POR TRUEQUE
SIN DUDA ALGUNA *	Y entiendo, sin duda alguna , que la Fortuna, atrevida, por no ponerte en caída, quiso mostrarse importuna Que si vos os acordáis de darme el mal que queréis, sin duda alguna creeréis

La ausencia de todas estas frases no ayuda a su candidatura, como no ayuda que la palabra «continuo», con hasta diez casos en *El rey por trueque*, ocurra una sola vez en una de sus comedias: *Los mal casados de Valencia*. Se necesitaría quizá un análisis más exhaustivo, aunque con los pocos datos que manejo creo que Castro debe descartarse. ¿Y Lope? Pues no lo sé. La crítica lo ha excluido del canon hace tiempo y los análisis que hemos hecho nosotros tampoco le son favorables. Sin embargo, sorprende que algún topónimo (del *Antártico a Calixto*) o personajes históricos como *Scévola* y *Porsena* o mitológicos como *Palas* tengan su correspondencia en la obra del Fénix. O que nos topemos con un paralelismo tan exclusivo como el que pongo debajo:

<i>El rey por trueque</i>	<i>Los hechos de Garcilaso</i>
No es bastante tu muerte, dulce esposo , hermoso espejo donde yo me vía, a buscar otro dueño poderoso, a quien entregue la libertad mía; que, si atada con lazo riguroso, te entregué el alma desde el primer día, y se fue con la tuya en mortal sueño, ¿qué alma tendré que dar a otro algún dueño?	Dulce esposa de mi vida, a quien el alma entregué , sabe el cielo si quisiera hacer aquesta jornada, si en la ciudad de Granada otro mi segundo hubiera. Yo no lo puedo escusar (<i>Los hechos de Garcilaso</i>)

Conceptos amorosos como los arriba expuestos son del acervo común, pero es mucho más difícil reunir en una misma constelación verbal <dulce espos*... a quien entregué / el alma>. Nótese, por otro lado, que *Los hechos de Garcilaso* es una de las primeras comedias de Lope, compuesta probablemente antes de 1580, siendo todavía adolescente. En nuestro análisis toda comedia con datación tan temprana presenta números muy bajos, sin que necesariamente deba cuestionarse la autoría lopesca. ¿Podría ser este el caso de *El rey por trueque*? No tengo respuesta. Un análisis de la métrica podría aportar nuevos datos, labor que dejo a otros, si se sienten con ganas. Por el momento digamos que *El rey por trueque* es de autoría dudosa.

La bella malmaridada está considerada por la crítica de autoría fiable, entre otras cosas porque poseemos un manuscrito apógrafo. Los análisis con JPAAG, sin embargo, presentan pobres resultados. La fecha de composición, según Morley y Bruerton, es de 1595. Con el segundo procedimiento Mira de Amescua alcanza un respetable 44%.



¿Se puede inferir algo de ello? Pues probablemente que la fecha de composición condiciona el grado de proximidad con las comedias que se cotejan, ya que la huella del Fénix en esta comedia se siente desde los primeros versos, como lo atestiguan estos pocos paralelismos:

<i>La bella malmaridada</i>	LOPE
... en todo Madrid la llama, por excelencia, la bella	En todo Madrid se dice que Lucrecia ha sido dama de don Juan... (<i>La malcasada</i>)
Vuelve la tuya cansada de lo [que] sufriendo está, que hay mil que la llaman ya la bella malmaridada.	Hay mil que la miran bien y impidesle su remedio (<i>El desposorio encubierto</i>)
Da gracias, Leonardo, al cielo, que fue Lisbella la que es...	Da gracias, Mengo, a los cielos, que te hicieron sin amor. (<i>Fuente Ovejuna</i>)
Vuelve, Teodoro ¿do vas, pues, siendo un alma los dos ? Pues ¿así te piensas ir y dejar muerto a un amigo?	... Somos un alma los dos. Carlos: Eso no, que si eso fuera, supieras que te engañaba... (<i>El perseguido</i>)
Cantan y dan dulce guerra, llevando el cielo el compás a los tonos de Juan Blas, que es un ángel en la tierra.	Arroyuelos cristalinos, ruido sonoro y manso, que parece que corréis tonos de Juan Blas, cantando (<i>El acero de Madrid</i>)

Permítaseme hacer ahora un brevísimo peinado sobre dos comedias cuyos resultados no han sido muy favorables en los análisis que hemos llevado a cabo y que, sin embargo, presentan concordancias muy significativas con el corpus de comedias del Fénix. La primera es *Los donaires de Matico*. Se imprimió en la Parte 1 (1604), se cita en las dos listas de *El peregrino en su patria* y, según el consenso general, es de autoría fiable. Morley y Bruerton la fechan hacia 1596. Me fijaré solo en las colocaciones léxicas que se forman en torno a la palabra *Dios*.

LOS DONAIRES DE MATICO

DIOS	
santo Dios	V
si pluguiese a Dios	
gracias a Dios	V
que Dios haya	V
ay, Dios	V
Dios no quiera / no quiera Dios	V
vive Dios	V
plegue a Dios	V
maldígate Dios	V

De las nueve frases hasta ocho concurren en el corpus de Lope. Además, una de ellas, «maldígate Dios», es bastante rara. En CORDE solamente existen 14 casos, cinco en obras del Fénix. Y si rastreamos documentos que compartan <maldígate Dios / que Dios haya>, solamente hay tres: la colección de romances editada por Agustín Durán en 1851, *Los donaires de Matico* y *El acero de Madrid*:

Consulta: *maldígate Dios y que Dios haya*, en todos los medios, en CORDE
 Resultado: 6 casos en 3 documentos.

OBTENCIÓN DE EJEMPLOS Documentos:

Clasificación: Selección
 Agrupación: Marcas:

Cómo citar el CORPUS Documentos.
 Pantalla: 1 de 1.

	Casos	Año	Autor	Obra	País	Tema	Publicación
<input type="checkbox"/>	2	1600 - 1604	Anónimo	Romances, en Romancero general [Romancero general]	ESPAÑA	21.Lirica romanceril	Agustín Durán, Rivedeneira (Madrid), 1851
<input type="checkbox"/>	2	a 1596	Vega Carpio, Lope de	Los donaires de Matico	ESPAÑA	23.comedia	Marco Presotto, Editorial Milenio (Lleida), 1997
<input type="checkbox"/>	2	c 1608 - 1612	Vega Carpio, Lope de	El acero de Madrid	ESPAÑA	23.comedia	Aline Bergounioux, Jean Lemartinel y Gilbert Zenana, Éditions Klincksieck (Paris), 1971

La otra comedia es *La devoción del rosario*. No está presente en la colección de Partes de Lope de Vega ni tampoco aparece en las listas de *El peregrino en su patria*. ARTELOPE etiqueta esta comedia como de autoría dudosa y los números que nosotros barajamos así parecen confirmarlo. Sin embargo, en el rastreo de colocaciones léxicas en torno a *Dios* nos encontramos con que un romance de Lope inserto en sus *Rimas sacras* está estrechamente relacionado con los primeros versos del primer Acto de *La devoción del rosario*. El romance desarrolla la conocida historia de San Agustín y el niño junto al mar, de proce-

dencia medieval, presente en la *Legenda aurea* y también en el *Flos sanctorum* de Villegas¹⁵. Debajo pongo algunas coincidencias verbales que apuntan a una relación causal:

LA DEVOCIÓN DEL ROSARIO	LOPE, <i>Rimas sacras</i>
¡Dios uno y Personas tres, que entender quiso Agustín, y en el ejemplo del mar, que el niño encerrar quería en tan pequeño lugar, vio que ninguno podía tan gran piélagos aplacar!	En las riberas del mar se paseaba Agustino... Lo que presume entender , ningún mortal lo ha entendido: cómo es Dios uno en esencia, siendo en las personas trino.

Otras coincidencias (*Dios sin principio / divino Dios / Dios permita / Dios le ruego*) ponen la comedia en la órbita del Fénix, si acaso como refundición o en colaboración con otros.

Quiero terminar esta sección dedicada a Lope con *El alcalde de Zalamea*. Hartzenbusch ya puso en duda su autoría y los análisis métricos de Morley y Bruerton terminaron por sacarla del canon loopesco. El consenso entre la crítica ha sido casi unánime desde entonces. Nuestros análisis con JPAAG, sin embargo, resultan mucho más favorables al Fénix. En el primero *El alcalde de Zalamea* promediaba un 80% de proximidad con respecto a las cuatro comedias de Lope empleadas en el análisis y en el segundo casi un 40%. Un cotejo apresurado entre el primer acto y el corpus loopesco depara varias concordancias muy significativas. Fijémonos en los cuatro primeros versos con que se inicia la comedia:

Ginesillo: Yo las vi.
Alcalde: ¿A entrambas?
Ginesillo: No hay que dudar;
digo que hablar las oí.

En *El hijo de Reduán*, escrita posiblemente antes de 1595, leemos esto:

... escojo en fin a las dos
entre otras muchas que vi.
Celora: ¿A entrambas?
Gomel: **Digo que sí**
por no ofenderos a vos. (Acto I, v. 767-768)

La constelación verbal <vi ¿A entrambas? / digo que> reproducida en otro texto suele ser sello de autoría en la mayoría de los casos. Poco después, don Diego, uno de los capitanes, explica así el repudio de la gente del pueblo hacia ellos:

Quejas acumula
siempre el vulgo sin consejo
que como es aborrecido
un soldado, siempre ha sido

15.- «Aquí se dice que compuso los libros de la Trinidad y le acaeció la revelación del niño que le vio estar cavando con sus manos junto a la marina una pequeña hoyita. Preguntóle San Agustín qué era su intento. Dijo que pasar allí el mar. Sonriose de esto el santo doctor y díjole que era cosa dificultosa aquello. El niño le respondió: «¿A ti se te hace esta cosa dificultosa? Pues yo te digo que no es más el negocio en que tú entiendes de la Trinidad, pretendiendo comprender y apear tan alto misterio con tu flaco entendimiento». Desapareció el niño por donde entendió San Agustín que era enviado de Dios para advertirle de su atrevimiento...» VILLEGAS (600).

o con envidia mirado,
o sin razón murmurado

En *El peregrino en su patria* las quejas del vulgo van dirigidas contra los gobernantes:

... una cierta malignidad **quejosa** tiene **siempre el vulgo** contra los que gobiernan

Otras secuencias bastante inusuales en CORDE están también compartidas con el corpus de Lope. Debajo pongo unas cuantas:

con igual porfía / ¿Qué queréis hermano(s)? / Amplius lava me / salmos penitenciales / Pesárame que / ¿qué es lo que quieres? / se... cuece el pan / ¿Hay desvergüenza...?

Algunas de estas frases puede que pertenezcan a la guardarropía teatral, pero tal cúmulo de coincidencias entre textos suele evidenciar una misma mano. En principio, pues, *El alcalde de Zalamea* es en su mayor parte de Lope, pero otros deberán continuar con la investigación y confirmar o desechar tal aserto.

Las Comedias de Tirso, Mira de Amescua, Guillén de Castro y Ruiz de Alarcón

Me he preocupado también de revisar algunas de las comedias de Tirso, Mira y Alarcón. Esta vez, sin embargo, computo solamente las tres primeras comedias de las cuatro posibles, tal como hicimos arriba con *Palabras y plumas*¹⁶. La máxima puntuación en este caso es 69 (100%) y la menor 6 (8%).

TIRSO

	C4	W2	W3	%
PALABRAS Y PLUMAS	100	100	98	99
LA ELECCIÓN POR LA VIRTUD	100	100	94	98
LA VILLANA DE VALLECAS	95	100	98	97
EL AMOR MÉDICO	98	100	95	97
EL HONROSO ATREVIMIENTO	98	100	92	96
LA CELOSA DE SÍ MISMA	97	98	95	96
LA PRUDENCIA EN LA MUJER	92	100	98	96
EL CASTIGO DEL PENSEQUE	95	98	94	95
EL MELANCÓLICO	95	95	97	95
AMAR POR ARTE MAYOR	86	100	97	94
EL AMOR Y EL AMISTAD	97	98	89	94
LA SANTA JUANA II	88	100	94	94
LA VENGANZA DE TAMAR	86	100	97	94

16.- Sustituyo en este caso *Don Gil de las calzas verdes* por *La primera parte de la Santa Juana* en el corpus de 24 comedias.

POR EL SÓTANO Y EL TORNO	92	95	95	94
LA SANTA JUANA III	91	98	92	93
EL PRETENDIENTE AL REVÉS	89	95	97	93
ESTO SÍ QUE ES NEGOCIAR	95	95	89	93
LA FINGIDA ARCADIA	88	98	94	93
AMOR Y CELOS HACEN DISCRETOS	89	94	95	92
DEL ENEMIGO EL PRIMER CONSEJO	92	95	91	92
ANTONA GARCÍA	91	97	86	91
DOÑA BEATRIZ DE SILVA	91	94	88	91
LA VIDA DE HERODES	82	97	94	91
VENTURA TE DÉ DIOS	95	92	88	91
AMAR POR SEÑAS	92	95	85	90
DESDE TOLEDO A MADRID	82	97	92	90
LA SANTA JUANA I	92	91	88	90
QUIEN CALLA OTORGA	92	91	89	90
TODO ES DAR EN UNA COSA	86	98	86	90
AMAZONAS DE LAS INDIAS	76	98	95	89
AVERIGÜELO VARGAS	92	92	84	89
EL AQUILES	85	97	85	89
BELLACO SOIS, GÓMEZ	82	94	91	89
ESCARMIENTOS PARA EL CUERDO	81	97	91	89
AMAR POR RAZÓN DE ESTADO	84	94	88	88
EL MAYOR DESENGAÑO	84	91	89	88
LA GALLEGA MARI HERNÁNDEZ	84	98	84	88
LA MUJER QUE MANDA EN CASA	84	95	86	88
EL ARBOL DEL MEJOR FRUTO	79	94	88	87
LA DAMA DEL OLIVAR	85	97	78	86
LA VILLANA DE LA SAGRA	89	92	79	86
QUIEN NO CAE NO SE LEVANTA	76	91	91	86
SANTO Y SASTRE	82	92	81	85
TANTO ES LO DEMÁS COMO LO DE	78	98	81	85
CELOS CON CELOS SE CURAN	78	92	82	84
LOS HERMANOS PARECIDOS	81	88	81	83
COMO HAN DE SER LOS AMIGOS	85	84	79	82
LA PEÑA DE FRANCIA	82	85	81	82
HABLADME EN ENTRANDO	73	84	75	77
EL BURLADOR DE SEVILLA	81	79	66	75

EL CONDENADO POR DESCONFIADO	60	85	82	75
TAN LARGO ME LO FIAIS	73	82	71	75
CAUTELA CONTRA CAUTELA	69	85	65	73
LA ROMERA DE SANTIAGO	69	79	73	73
LOS AMANTES DE TERUEL	68	79	62	69
LA NINFA DEL CIELO	56	75	69	66
LOS ALCALDES	69	73	55	65

Entre las 57 comedias computadas 48 están por encima del 80% de proximidad y más de la mitad rondan un 90% o más. Hay solamente nueve que están por debajo, algunas claramente no de Tirso, como *Los amantes de Teruel* y *Los alcaldes*, y otras de autoría dudosa, caso de *La ninfa del cielo*, *La Romera de Santiago* y *Habládmelo en entrando*. La comedia *Cautela contra cautela*, atribuida tradicionalmente a Tirso, es con seguridad de Mira de Amescua. En nuestro análisis alcanza casi el 90%:

	C4	W2	W3	%
MIRA	79	92	98	89%
TIRSO	69	85	65	73%
LOPE	91	59	50	66%

Nos quedan tres de las comedias más controvertidas: *El burlador de Sevilla*, *Tan largo me lo fiais* y *El condenado por desconfiado*. Las tres ofrecen un grado de proximidad con respecto a las cuatro comedias de Tirso sensiblemente inferior a la media. El promedio mejora algo si en lugar de *Don Gil de las calzas verdes* empleamos *La Santa Juana I* en el corpus, pero todavía lejos de la zona de confort en cuestiones de autoría. Analicemos más de cerca los resultados de *El burlador*.

	C4	W2	W3
EL BURLADOR DE SEVILLA	1. CLARAMONTE Deste agua no beberé	1. CLARAMONTE Deste agua no beberé	1. CLARAMONTE Deste agua no beberé
	2. CLARAMONTE Púsoseme el sol	2. CLARAMONTE Púsoseme el sol	2. MIRA DE AMESCUA La Fénix de Salam.
	3. TIRSO La santa Juana I	3. TIRSO La santa Juana I	3. CLARAMONTE Púsoseme el sol
	4. TIRSO El vergonzoso en palacio	4. TIRSO El vergonzoso en palacio	4. MIRA DE AMESCUA El esclavo del d.
	5. LOPE El caballero de Olmedo	5. MIRA DE AMESCUA La Fénix de Salaman.	5. MIRA DE AMESCUA El ejemplo mayor
	6. LOPE Fuente Ovejuna	6. MIRA DE AMESCUA El ejemplo mayor	6. TIRSO El vergonzoso en palacio
	7. ALARCÓN La verdad sospechosa	7. ALARCÓN La verdad sospechosa	7. TIRSO La santa Juana I
	8. TIRSO Marta la piadosa	8. TIRSO Marta la piadosa	8. LOPE El castigo sin venganza
	9. MIRA DE AMESCUA El esclavo del de.	9. MIRA DE AMESCUA El esclavo del demon.	9. GUILLÉN DE CASTRO El nacimiento de M.
	10. TIRSO Caballero de Gracia (...)	10. LOPE El castigo sin venganza (...)	10. TIRSO Caballero de Gracia (...)
	23. CLARAMONTE De lo vivo a lo pintado	23. CLARAMONTE De lo vivo a lo pintado	19. CLARAMONTE De lo vivo a lo pintado
	24. CLARAMONTE El valiente negro	24. CLARAMONTE El valiente negro	23. CLARAMONTE El valiente negro

Andrés de Claramonte, propuesto desde hace tiempo como autor de la comedia¹⁷, es quien más se aproxima, pero como tantas otras veces lo hace solo con *Dește agua y Púsoseme el sol*, mientras que sus otras dos comedias ocupan las últimas plazas en las tres categorías. Una cosa así es muy poco habitual. A veces pasa, como veíamos con *Don Gil de las calzas verdes*, que una comedia de autoría fiable se quede rezagada, pero resulta in-sólito lo ocurrido aquí. La posible explicación, ya adelantada antes cuando tratábamos de *La estrella de Sevilla*, es que *El burlador*, *Dește agua* y *Púsoseme* no sean sino refundiciones hechas por Claramonte a partir de comedias de Tirso o de Lope. Desde luego la relación del murciano con *El burlador* parece innegable. Otra cosa es que la obra sea originalmente de su cosecha. Tomemos el paralelismo más obvio entre *El burlador* y *Dește agua no beberé*:

<i>El burlador de Sevilla</i>	Claramonte, <i>Dește agua no beberé</i>
Celos, reloj (de) cuidado, que a todas las horas dais tormentos con que matáis, aunque dais desconcertado; celos, del vivir desprecios, con que ignorancias hacéis, pues todo lo que tenéis de ricos, tenéis de necios:	Celos, reloj de cuidados, que a todas las horas dais tormentos con que matáis aunque estéis desconcertados, Gutierre Alfonso Solís muchos años me sirvió, y la palabra me dio; ¿cómo no se la pedís?

Los cuatro primeros versos de la redondilla son casi iguales en las dos comedias. Sabido es que todo autor a veces se plagia a sí mismo y a veces toma prestado de otros. La cuestión aquí es saber quién escribe primero estos versos y quién los copia. La crítica decimonónica, poco favorable a Claramonte, lo tenía muy claro¹⁸, pero desde hace ya tiempo hay varios estudiosos que piensan que el autor murciano es el responsable de ambos pasajes. Yo no estoy tan seguro. Otra redondilla en *El pretendiente al revés* desarrolla una idea muy parecida con correspondencias verbales idénticas a las que tenemos arriba:

<i>El burlador</i>	<i>El pretendiente al revés</i>
Celos, reloj (de) cuidado, que a todas las horas dais tormentos con que matáis, aunque dais desconcertado ; celos, del vivir desprecios, con que ignorancias hacéis, pues todo lo que tenéis de ricos, tenéis de necios	Es reloj la voluntad: desconcertada una rueda, no hay quien concertalla pueda, si no es con dificultad. La rueda han desconcertado los celos que amor labró, y pues no tengo orden yo, nada ha de andar ordenado.

17.- LEAVITT (1931), RODRÍGUEZ LÓPEZ-VÁZQUEZ (1987).

18.- «(Claramonte) is perhaps the most maligned dramatist of the period. Even his contemporaries used the name Claramonte as a synonym for literary thievery: Alonso de Salas Barbadillo, for example, calls Ruiz de Alarcón a «Segundo Claramonte» to emphasize his borrowings. Modern critics have followed suit. Luis Fernández Guerra y Orbe, Emilio Cotarelo y Mori, and Marcelino Menéndez Pelayo are largely responsible for this continuing view of Claramonte as an author who «se dedicó a la piratería literaria»... Blanca de los Ríos also speaks against his «piraterías»... and calls him an «usurpador y profanador». DE ARMAS (58).

Y en *La gallega Mari-Hernández* tenemos este otro paralelismo:

<i>El burlador</i>	<i>La gallega Mari-Hernández</i>
... celos, del vivir desprecios, con que ignorancias hacéis, pues todo lo que tenéis de ricos, tenéis de necios:	«Teme segundos desprecios; que aunque ausente de la sierra, su memoria os hará guerra, los celos pecan de necios.

Incluso si escogemos la secuencia <hacéis dist/10 tenéis dist/15 necio*>, vemos que, al menos en CORDE, no existe ningún otro caso, aparte de *El burlador* y *Don Gil de las calzas verdes*.

Consulta: hacéis dist/10 tenéis dist/15 necio*, en todos los medios, en CORDE
Resultado: 2 casos en 2 documentos.

OBTENCIÓN DE EJEMPLOS

Recuperar Concordancias. Normal. Clasificación: Agrupación: Marcas:

Cómo citar el CORPUS Concordancias. Pantalla: 1 de 1. Ver párrafos

nos referir los dos. DON JUAN O hacéis de mí poco caso, o tenéis poco valor; pero a v ** 1615
sprecios, con que ignorancias hacéis, pues todo lo que tenéis de ricos, tenéis de n ** c 1613 - 1617

AÑO AUTOR
1615 Tirso de Molina
c 1613 - 1617 Tirso de Molina

Daré otro caso. En la Jornada II de *El burlador* nos encontramos con estos versos en boca del duque Octavio:

Hablé al Rey, viome y honrome,
César con el César fui,
pues vi, peleé y vencí,
y hace que esposa tome
de su mano, y se prefiere
a desenojar al Rey
en la fulminada ley.

Unos versos muy semejantes, aunque en un contexto algo distinto, leemos también en *La venganza de Tamar*:

En muestras de este deseo,
siempre que a la guerra fui,
o partí, llegué, vi y vencí;
y agora llego, entro y veo
amores abominables,
ofensas de Dios, del Rey,
de tu sangre, de tu ley.

El programa JPAAP ha dejado clara la proximidad verbal de *Deste agua y Púsoseme* en relación con *El burlador*, pero estos pocos paralelismos vistos arriba dejan entrever un problema de fondo en el texto que nos ha llegado. Una explicación podría ser que Tirso escribió el original, hoy perdido, y que luego pasó a las manos de Claramonte, que la refundió de arriba abajo, como hizo seguramente con *Tan largo me lo fiais*. Otra conjetura, quizá no tan

descabellada, puede ser esta: *Deste agua no beberé* y *Púsoseme el sol* son refundiciones a partir de comedias de Tirso o incluso de Lope y, de ahí, que en tantos cotejos estas dos comedias supuestamente de Claramonte aparezcan de vez en cuando tan próximas al corpus de los dos grandes dramaturgos. En todo caso, no podemos soslayar los fríos datos de la estadística, los cuales nos dicen que al menos dos comedias atribuidas a Andrés de Claramonte están estrechamente asociadas al *Burlador* o a su otra versión, como lo muestra la tabla de abajo:

	C4	W2	W3
TAN LARGO ME LO FIAIS	1. CLARAMONTE <i>Deste agua no beberé</i>	1. CLARAMONTE <i>Púsoseme el sol</i>	1. CLARAMONTE <i>Deste agua no beberé</i>
	2. CLARAMONTE <i>Púsoseme el sol</i>	2. CLARAMONTE <i>Deste agua no beberé</i>	2. CLARAMONTE <i>Púsoseme el sol</i>
	3. LOPE <i>Fuente Ovejuna</i>	3. TIRSO <i>La santa Juana I</i>	3. MIRA DE AMESCUA <i>La Fénix de Salamanca</i>
	4. LOPE <i>El caballero de Olmedo</i>	4. TIRSO <i>El vergonzoso en palacio</i>	4. TIRSO <i>El vergonzoso en palacio</i>
	5. TIRSO <i>Marta la piadosa</i>	5. MIRA DE AMESCUA <i>El ejemplo mayor</i>	5. MIRA DE AMESCUA <i>El esclavo del demonio</i>
	6. TIRSO <i>Caballero de Gracia</i>	6. MIRA DE AMESCUA <i>La Fénix de Salam.</i>	6. TIRSO <i>La santa Juana I</i>
	7. ALARCÓN <i>La verdad sospechosa</i>	7. ALARCÓN <i>La verdad sospechosa</i>	7. MIRA DE AMESCUA <i>El ejemplo mayor</i>
	8. ALARCÓN <i>Las paredes oyen</i>	8. TIRSO <i>Marta la piadosa</i>	8. GUILLÉN DE CASTRO <i>El nacimiento de Mont</i>
	9. LOPE <i>El castigo sin venganza</i>	9. TIRSO <i>Caballero de Gracia</i>	9. LOPE <i>El perro del hortelano</i>
	10. MIRA DE AMESCUA <i>El esclavo</i>	10. MIRA DE AMESCUA <i>El esclavo</i>	10. ALARCÓN <i>La verdad sospechosa</i>
	13. TIRSO <i>El vergonzoso en palacio</i>	(...)	11. TIRSO <i>Marta la piadosa</i>
	16. TIRSO <i>La santa Juana I</i>	23. CLARAMONTE <i>De lo vivo a lo pintado</i>	17. CLARAMONTE <i>De lo vivo a lo pintado</i>
	23. CLARAMONTE <i>De lo vivo a lo pintado</i>	24. CLARAMONTE <i>El valiente negro</i>	23. CLARAMONTE <i>El valiente negro</i>
	24. CLARAMONTE <i>El valiente negro</i>		

El condenado por desconfiado es otra de las comedias en disputa desde hace tiempo. La crítica parece haber descartado la autoría del mercedario y los cálculos hechos por mí no hacen sino confirmarlo. Debajo pongo los resultados:

	C4	W2	W3
EL CONDENADO POR DESCONFIADO	1. CLARAMONTE <i>Púsoseme el sol</i>	1. TIRSO <i>La santa Juana I</i>	1. MIRA DE AMESCUA <i>La Fénix de Sal</i>
	2. CLARAMONTE <i>Deste agua no beberé</i>	2. TIRSO <i>El vergonzoso en palacio</i>	2. CLARAMONTE <i>Deste agua no beberé</i>
	3. MIRA DE AMESCUA <i>El esclavo del de.</i>	3. CLARAMONTE <i>Púsoseme el sol</i>	3. TIRSO <i>El vergonzoso en palacio</i>
	4. LOPE <i>El caballero de Olmedo</i>	4. CLARAMONTE <i>Deste agua no beberé</i>	4. MIRA DE AMESCUA <i>El esclavo del demo.</i>
	5. ALARCÓN <i>La verdad sospechosa</i>	5. MIRA DE AMESCUA <i>La Fénix de Salam.</i>	5. CLARAMONTE <i>Púsoseme el sol</i>
	6. GUILLÉN DE CASTRO <i>El amor constante.</i>	6. MIRA DE AMESCUA <i>El esclavo del demo.</i>	6. GUILLÉN DE CASTRO <i>El amor constante</i>
	7. LOPE <i>Fuente Ovejuna</i>	7. MIRA DE AMESCUA <i>El ejemplo mayor</i>	7. TIRSO <i>La santa Juana I</i>
	8. TIRSO <i>Marta la piadosa</i>	8. TIRSO <i>Caballero de Gracia</i>	8. TIRSO <i>Marta la piadosa</i>
	9. LOPE <i>El castigo sin venganza</i>	9. TIRSO <i>Marta la piadosa</i>	9. TIRSO <i>Caballero de Gracia</i>
	10. LOPE <i>El perro del hortelano</i>	10. ALARCÓN <i>La verdad sospechosa</i>	10. GUILLÉN DE CASTRO <i>El nacimiento de</i>

El problema que veíamos antes con *El burlador* (y con tantas otras comedias vistas aquí) nos los volvemos a encontrar ahora. Claramonte ocupa las primeras posiciones, pero solamente con *Deste agua no beberé* y *Púsoseme el sol*; sus otras dos comedias descansan en el fondo de la lista. Un cálculo con las tres primeras comedias depara este resultado:

69	C4	W2	W3	%
TIRSO	60	92	82	78%
MIRA	62	82	84	76%
CLARAMONTE	71	65	68	68%

Ninguno supera el 80%, pero naturalmente Claramonte mejora exponencialmente si calculamos solo las dos primeras comedias:

47	C4	W2	W3	%
CLARAMONTE	100	91	91	94%
MIRA	70	82	95	82%
TIRSO	63	100	82	81%

Con estos datos es muy difícil pronunciarse. Quien escribe *Deste agua* y *Púsoseme* está muy próximo a *El condenado*, pero falta por saberse si en verdad Claramonte es su autor; y de serlo, si son obras enteramente suyas o meras refundiciones. El «caso Claramonte» es, como se ve, un embrollo grande. No resulta nada fácil dilucidar qué escribió y qué tomó prestado, cuáles son sus comedias originales y cuáles falsas atribuciones. Alfredo Rodríguez López-Vázquez (RLV) ha hecho una meritoria labor en la rehabilitación del autor murciano y en lo esencial sus propuestas en todo lo relativo a las atribuciones no parecen ir descaminadas. El debate sobre las más famosas sigue ahí abierto, pero a RLV le debemos la recuperación de textos olvidados y, sobre todo, el haber traído a la palestra cuestiones de atribución en comedias de gran solera, como las aquí estudiadas, o como *El rey don Pedro en Madrid* o *el Infanzón de Illescas*.

Esta última comedia ha llegado hasta nosotros en una plétora de manuscritos, impresos y sueltas del siglo XVII en donde se le atribuye su autoría indistintamente a Lope, a Tirso, a Calderón y al mismo Andrés de Claramonte. ¿Cuál de los cuatro tiene más visos de serlo? Menéndez Pelayo pensó que su autor, dada la calidad literaria del texto, no podía ser otro que Lope, tras desechar las otras tres, incluida la de autor murciano, con quien se despachaba a gusto¹⁹. Nuestros análisis, sin embargo, concuerdan con la autoría de Claramonte y, a diferencia de los demás casos, esta vez sus cuatro comedias copan los primeros lugares:

19.– Las palabras que el gran polígrafo reserva al murciano son demoledoras: «En Andrés de Claramonte no hay que pensar como autor original. Era un dramaturgo vulgar y adocenado, que, siendo comediante de oficio y viéndose obligado a abastecer la escena con novedades propias o ajenas, se dedicó a la piratería literaria con el candor con que ésta se practicaba en aquel tiempo, y del cual daban ejemplo grandes poetas... N(o) se le pueden negar ciertas cualidades, inferiores sin duda, pero muy recomendables: conocimiento de la escena y cierto brío y desgarro popular, que principalmente lucen en su comedia soldadesca de *El valiente negro en Flandes*. Lo intolerable de Claramonte, y lo que prueba la penuria de su educación literaria, es el estilo. Por raro caso en su tiempo, Claramonte escribe mal, no ya por culteranismo o conceptismo, como muchos otros, sino por incorrección gramatical grosera, que hace enmarañados y oscuros sus conceptos. Este desaseo y torpeza de expresión es, por decirlo así, la marca de fábrica de su teatro, y sirve de indicio casi infalible para deslindar lo que realmente le pertenece en las obras que llevan su nombre. Así sucede en *El Rey Don Pedro en Madrid*... (328-29).

	C4	W2	W3
EL INFANZÓN DE ILLESCAS	1. CLARAMONTE De lo vivo a lo pintado	1. CLARAMONTE Deste agua no beberé	1. CLARAMONTE Deste agua no beberé
	2. CLARAMONTE El valiente negro	2. CLARAMONTE Púsoseme el sol	2. MIRA DE AMESCUA La Fénix de Salamanca
	3. CLARAMONTE Púsoseme el sol	3. CLARAMONTE De lo vivo a lo pintado	3. MIRA DE AMESCUA El ejemplo mayor
	4. CLARAMONTE Deste agua no beberé	4. CLARAMONTE El valiente negro	4. CLARAMONTE Púsoseme el sol
	5. CASTRO La humildad soberbia	5. MIRA DE AMESCUA El ejemplo mayor	5. TIRSO El vergonzoso en palacio
	6. MIRA El ejemplo mayor	6. MIRA La Fénix de Salamanca	6. CLARAMONTE El valiente negro
	7. LOPE Fuente Ovejuna	7. CASTRO La humildad soberbia	7. CLARAMONTE De lo vivo a lo pintado
	8. MIRA El esclavo del demonio	8. TIRSO El vergonzoso en palacio	8. ALARCÓN La verdad sospechosa
	9. ALARCÓN La verdad sospechosa	9. ALARCÓN La verdad sospechosa	9. GUILLÉN DE CASTRO La humildad soberbia
	10. CASTRO El amor constante	10. CASTRO El amor constante	10. MIRA DE AMESCUA La adúltera virtuosa

El grado de proximidad es nada menos que de un 97% y en dos categorías (C4 y W2) obtenemos un pleno:

90	C4	W2	W3	%
CLARAMONTE	100%	100%	91%	97%

El resultado no deja margen a la duda en cuanto a la autoría de Claramonte en esta obra y, a la vez, nos hace ver que *Deste agua* y *Púsoseme* no son comedias tan ajenas a quien escribió las otras tres, incluida *El infanzón de Illescas*. Pero entonces, ¿por qué tal disparidad en tantos y tantos casos? Decido incluir en el corpus *El infanzón* para ver qué pasa cuando se contrasta con *El Burlador*, *La estrella de Sevilla* y *El condenado*. Los resultados no pueden ser más desalentadores para la causa de Claramonte: en todos los casos *El infanzón* ocupa siempre el último lugar en las tres categorías.

	C4	W2	W3
BURLADOR	1. CLARAMONTE Deste agua	1. CLARAMONTE Deste agua	1. CLARAMONTE Deste agua
	2. CLARAMONTE Púsoseme el sol	2. CLARAMONTE Púsoseme el sol	3. CLARAMONTE Púsoseme el sol
	22. CLARAMONTE De lo vivo a lo pint.	22. CLARAMONTE De lo vivo a lo	20. CLARAMONTE De lo vivo a lo
	24. CLARAMONTE El valiente negro	23. CLARAMONTE El valiente negro	24. CLARAMONTE El valiente
	25. El infanzón de Illescas	25. El infanzón de Illescas	25. El infanzón de Illescas
ESTRELLA	1. CLARAMONTE Deste agua no beberé	1. CLARAMONTE Deste agua no beberé	1. CLARAMONTE Deste agua no beberé
	2. CLARAMONTE Púsoseme el sol	2. CLARAMONTE Púsoseme el sol	5. CLARAMONTE Púsoseme el sol
	22. CLARAMONTE De lo vivo a lo pint.	22. CLARAMONTE De lo vivo a lo	19. CLARAMONTE De lo vivo a lo
	23. CLARAMONTE El valiente negro	23. CLARAMONTE El valiente negro	24. CLARAMONTE El valiente
	25. El infanzón de Illescas	25. El infanzón de Illescas	25. El infanzón de Illescas

CONDENADO	1. CLARAMONTE Deste agua no beberé	1. TIRSO El vergonzoso	1. MIRA La Fénix de Salamanca
	2. CLARAMONTE Púsoseme el sol	2. CLARAMONTE Púsoseme el sol	2. CLARAMONTE Deste agua
		3. CLARAMONTE Deste agua	5. CLARAMONTE Púsoseme
	22. CLARAMONTE De lo vivo a lo pint.	22. CLARAMONTE De lo vivo a lo	22. CLARAMONTE De lo vivo a lo
	23. CLARAMONTE El valiente negro	23. CLARAMONTE El valiente negro	24. CLARAMONTE El valiente
25. El infanzón de Illescas	25. El infanzón de Illescas	25. El infanzón de Illescas	

Habrà que ahondar mäs, naturalmente. Por ahora lo único claro, con los datos de nuestro análisis, es que Lope no es el autor de *La estrella de Sevilla* ni de *El rey don Pedro en Madrid*, mientras que Claramonte podría serlo de la primera y, con casi toda certeza, de la segunda. En cuanto a Tirso, resulta dudoso que haya escrito el texto actual de *El Burlador* y aún más el de *El Condenado*, sin que nos quede clara la atribución de Claramonte tampoco.

El doctor Antonio Mira de Amescua tuvo buena fama en su época (COTARELO: 5), pero abandonó el ejercicio de las letras en 1632 sin haber imprimido personalmente ninguna de sus comedias²⁰. El resultado es el previsible: al menos la mitad de las que nos han llegado bajo su nombre o bien no son suyas o son refundiciones o están escritas en colaboración con otros dramaturgos. Los análisis que hemos realizado así parecen confirmarlo. Fijémonos, primeramente, en tres comedias autógrafas: *La casa del tahúr*, *La tercera de sí misma* y *El mártir de Madrid*. De la primera se conserva un manuscrito firmado a 20 de diciembre de 1616. Su autoría es, pues, de lo más fiable. Al hacer los cálculos encontramos un 95% de proximidad en el primer análisis y casi un 80% en el segundo:

69	C4	W2	W3	%
LA CASA DEL TAHÚR	89	97	100	95



Algo bastante parecido ocurre en la segunda comedia autógrafa que analizamos:

	C4	W2	W3	%
LA TERCERA DE SÍ MISMA	81	89	100	90

20.- Las que se publican en vida están siempre incluidas en colecciones impresas con otros dramaturgos (VALLADARES: 7-15).

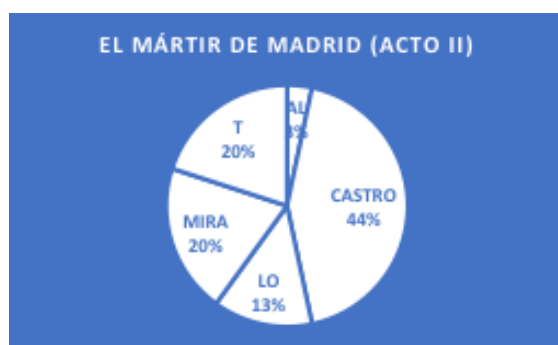


La comedia *El mártir de Madrid* nos llega ya en un manuscrito parcialmente autógrafo y aquí los números bajan sustancialmente:

	C4	W2	W3	%
EL MÁRTIR DE MADRID	57	84	97	79

Hace algunos años una profesora italiana estudió detenidamente este manuscrito y llegó a la conclusión de que mientras las grafías del primer acto y del tercero parecían de Mira o de su amanuense, la del segundo acto era casi con toda seguridad de Belmonte Bermúdez (ALVITI: 70-72), de lo cual la investigadora infería que estábamos ante una comedia escrita en colaboración. En el corpus que manejo cuento solamente con *El diablo predicador*, pero en los dos análisis con JGAAP la comedia de Belmonte aparece muy alejada del Segundo Acto, como lo están también, hasta cierto punto, las cuatro comedias de Mira. De los seis autores quien más se aproxima a ese Segundo Acto es Guillén de Castro:

	C4	W3	W2	%
CASTRO	92	85	79	85
MIRA	63	84	94	80
TIRSO	56	66	58	60



¿Debemos descartar la colaboración de Belmonte y pensar, más bien, en Guillén de Castro? Pues no lo sé, pero al hacer un rápido cotejo en CORDE, *La hispánica* de Belmonte ofrece paralelismos significativos con *El mártir de Madrid*. Veamos primero tres colocaciones léxicas en torno a la palabra *corazón*:

CORAZÓN

a quien el corazón robusto humilla	<i>El mártir de Madrid</i>
corazón robusto y fuerte	<i>La hispánica</i>
Por eso rendirle pienso el humilde corazón	<i>El mártir de Madrid</i>
mi humilde corazón , decir podría	<i>La hispánica</i>
¿cuánto más valor tendrán alma y corazón rendidos...?	<i>El mártir de Madrid</i>
temeroso cierra, con alma y corazón sencillo y puro	<i>La hispánica</i>

O esta otra con la palabra *alma*, sin ningún equivalente en CORDE:

Pedro, al fin, con alma fiera entró en casa de mi esposa	<i>El mártir de Madrid</i>
con mano acometió y con alma fiera ...	<i>La hispánica</i>

Al poco de empezar el Segundo Acto nos topamos con un paralelismo exclusivo:

mas de la suerte que envía luz hermosa, ausente el día	<i>El mártir de Madrid</i>
temerosa y fría de ver tan nueva luz ausente el día	<i>La hispánica</i>

Y, poco después, se describe la violenta acometida de un guerrero en parecidos términos a otra descrita en el poema épico de Belmonte:

<i>El mártir de Madrid Acto II</i>	<i>La hispánica</i>
¿No has visto, trepando cerros , manchar de espuma las flores espín coronado a hierros , derribando cazadores y desbaratando perros , que con el rabioso diente, mirando a l[a] escuadra enfrente, con el fuego en que se abrasa, tronchando venablos pasa más veloz que rayo ardiente ?	Como el lobo feroz, hecha la presa, que estima en poco, penetrando cerros , que resonando el valle y selva espesa, silben pastores ni que ladren perros el monstruo así las ondas atraviesa, sirviendo a un tiempo de pesados hierros , pues fueron en su claustro solo y mudo alas al Geta y grillos a Bermudo... Rompió, en efecto, la coyunda el toro; el preso la cadena; el mar, la nave... el venablo el espín , la jaula el ave... que jamás contra el muro lanzas juegue, para el suelo andaluz que rayo ardiente al sol que esparce lo verán presente

El mártir de Madrid es una comedia que mezcla de manera bastante deslavazada el enredo amoroso con el tema morisco. Al final del Primer Acto los protagonistas caen presos a manos de corsarios, que se los llevan a Argel. Allí, a poco de iniciarse el Segundo Acto, entra en escena Celaura, la hermana del rey, que se enamorará de inmediato de uno de los cautivos cristianos. Pues bien, otra Celaura mora aparece en el poema épico de Belmonte. ¿Casualidad? Evidentemente, no. Belmonte está a buen seguro involucrado en la composición del Segundo Acto. No podemos precisar cuál es el grado de su colaboración, pues la mano de Mira puede que esté presente en otras partes del Segundo Acto. Dos ejemplos:

<i>El mártir de Madrid</i>	MIRA DE AMESCUA
Deme Dios dolor eterno por descanso y por regalo, pues que soy árbol tan malo...	Deme Dios dolor tan fuerte que no le acabe la muerte para que viva y lo sienta... (<i>El ejemplo mayor de la desdicha</i>)
Mi amor es niño y gigante, y con tirana violencia me persüade a seguiros.	¿Yo he de ser ejecutor de esta tirana violencia ? (<i>El conde Alarcos</i>)

Se necesitará un análisis exhaustivo, pero por lo visto aquí la hipótesis de la colaboración entre Mira y Belmonte tiene buenos fundamentos.

En todo caso, los análisis llevados a cabo con más de 40 comedias de Mira o atribuidas a él vuelven a confirmar que, por lo general, toda comedia considerada de autoría fiable suele alcanzar un muy alto porcentaje de proximidad (entre el 95% y 90% si contabilizamos las tres primeras comedias), mientras que el resto, especialmente las publicadas en sueltas o muerto ya él, no pasan del 80%. Debajo pongo la primera lista hecha tras el análisis realizado con JGAAP:

	C4	W2	W3	%
LA CASA DEL TAHÚR	89	97	100	95
EXAMINARSE DE REY	91	92	98	93
NO HAY DICHA NI DESDICHA HASTA LA MUERTE	88	92	100	93
EL PRIMER CONDE DE FLANDES	85	91	100	92
EL CONDE ALARCOS	85	92	97	91
LA RUEDA DE LA FORTUNA	84	89	98	90
LA TERCERA DE SÍ MISMA	81	89	100	90
LAS DESGRACIAS DEL REY ALFONSO	81	89	100	90
CAUTELA CONTRA CAUTELA	79	92	98	89
LAS LISES DE FRANCIA	81	89	97	89
LOS CARBONEROS DE FRANCIA	84	92	91	89
EL PALACIO CONFUSO	85	91	88	88
LA PRÓSPERA FORT. ALVARO DE LUNA	76	92	98	88
EL ARPA DE DAVID	79	88	94	87
GALÁN, VALIENTE Y DISCRETO	73	92	97	87
NO HAY REINAR COMO VIVIR	78	94	91	87
CUATRO MILAGROS DE AMOR	65	97	97	86
HERO Y LEANDRO	73	92	95	86
LA PRÓSPERA FORT. BERN. CABRERA	73	92	91	85
EL RICO AVARIENTO	63	89	100	84
EL MÁRTIR DE MADRID	57	84	97	79
EL ANIMAL PROFETA	56	86	92	78

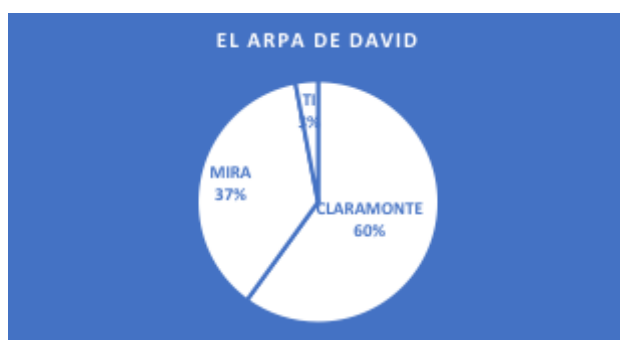
EL AMPARO DE LOS HOMBRES	76	88	68	77
EL HOMBRE DE MAYOR FAMA	59	84	89	77
LA JUDÍA DE TOLEDO	55	82	89	75
LA MONJA DE PORTUGAL	66	81	78	75
NARDO ANTONIO, BANDOLERO	52	81	89	74
EL CABALLERO SIN NOMBRE	56	79	86	73
EL CLAVO DE Jael	56	81	84	73
NO HAY BURLAS CON LAS MUJERES	60	71	86	72
EL MÁS FELIZ CAUTIVERIO	50	75	88	71
EL NEGRO DEL MEJOR AMO	57	79	78	71
LA CONFUSIÓN DE HUNGRÍA	50	81	84	71
LA HIJA DE CARLOS V	55	78	79	70
LA MESONERA DEL CIELO	53	75	84	70
LO QUE NO ES CASARSE A GUSTO	46	72	91	69
LA ADÚLTERA VIRTUOSA	40	82	84	68
AMOR, INGENIO Y MUJER	50	75	76	67
LOS PRODIGIOS DE LA VARA	49	79	75	67
OBLIGAR CONTRA SU SANGRE	47	69	81	65
LO QUE PUEDE UNA SOSPECHA	47	66	81	64

A la vista de los resultados la marca de separación entre comedias de autoría fiable y las que deben considerarse dudosas o escritas en colaboración estaría alrededor del 85%. Por lo mismo, las que no llegan al 75% deben descartarse o revisarse a fondo. En el segundo análisis hecho a partir de las cuatro comedias mejor clasificadas de cada categoría, observamos unos resultados parejos respecto al análisis anterior, con alguna que otra discrepancia. Al menos 17 comedias de Mira están por encima del 40% con respecto a los demás autores cotejados:

	AL	CA	CL	LO	MI	TI
LA CASA DEL TAHÚR			10	7	77	6
EXAMINARSE DE REY	7		13		70	10
EL CONDE ALARCOS		4	13		67	16
LA RUEDA DE LA FORTUNA			33		67	
NO HAY DICHA NI DESDICHA HASTA LA MUERTE		6	17		67	10
LAS DESGRACIAS DEL REY ALFONSO			30		57	13
LAS LISES DE FRANCIA		3	33		57	7
LOS CARBONEROS DE FRANCIA			30		53	17
EL RICO AVARIENTO		4	23	6	50	17
HERO Y LEANDRO		3	37	7	50	3
LA PRÓSPERA FORTUNA DE DON ALVARO DE LUNA		17	17	3	50	13
LA TERCERA DE SÍ MISMA	13		7	23	50	7

NO HAY REINAR COMO VIVIR	7	3	10	3	50	27
CUATRO MILAGROS DE AMOR	20	6	17	4	47	6
EL PALACIO CONFUSO			27		47	26
EL AMPARO DE LOS HOMBRES			10	17	43	30
LA PRÓSPERA FORT. DON BERNARDO DE CABRERA			20		43	37
GALÁN, VALIENTE Y DISCRETO	3		10	17	40	30

De esta segunda lista se descuelga *El arpa de David*, comedia que, según Cotarelo, «ofrece muy pocas garantías de ser el texto genuino de Mira» (620). Claramonte sería, en cambio, quien más se aproxima:



No es en la única. Como tantas otras veces, *Deste agua* y *Púsoseme el sol* de Claramonte obtienen un porcentaje de proximidad por encima del 40% en siete de las comedias aquí analizadas:

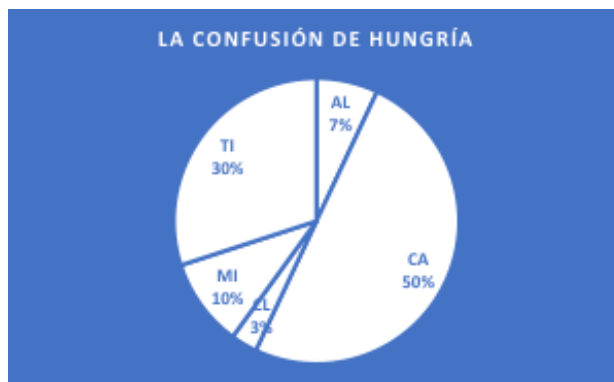
	AL	CA	CL	LO	MI	TI
EL CLAVO DE JAEL			70	10	10	10
EL ARPA DE DAVID			60		37	3
LA HIJA DE CARLOS V			57		10	33
EL NEGRO DEL MEJOR AMO	4		53		20	23
LA MONJA DE PORTUGAL	10	7	53		17	13
EL ANIMAL PROFETA	3		47	13	17	20
LA ADÚLTERA VIRTUOSA			40	17	13	30

No me atrevería a asegurar que quien escribió *Deste agua* o *Púsoseme el sol* es el autor de alguna de estas siete comedias, pero en cambio sí me resulta fácil afirmar que Mira —o, si se quiere, el autor de *El esclavo del demonio*— no ha escrito ninguna de ellas, salvo quizá *El arpa de David*. No mejor parado sale el dramaturgo granadino en el resto de comedias. En tres al menos los mejores números se los lleva Tirso:

	AL	CA	CL	LO	MI	TI
EL MÁS FELIZ CAUTIVERIO			23	10	20	47

LO QUE NO ES CASARSE A GUSTO	13		23	10	10	44
EL CABALLERO SIN NOMBRE		17	33		10	40

Guillén de Castro obtiene un muy buen resultado en *La confusión de Hungría*:



Y si hacemos el cálculo con las tres primeras comedias, Castro promedia más de un 85%, claramente en la zona de autoría probable:

	C4	W2	W3	%
LA CONFUSIÓN DE HUNGRÍA	89	82	88	86

Debería estudiarse todo esto con más detenimiento. Yo solamente pongo aquí las bases para una futura investigación. Desde luego las comedias de Guillén de Castro analizadas aquí, como se puede ver en la tabla que pongo debajo, obtienen números altísimos, la que menos cuatro puntos por encima de *La confusión de Hungría*:

	C4	W2	W3	%
LAS MOCEDADES DEL CID II	100	98	100	99
EL CURIOSO IMPERTINENTE	100	91	97	96
EL CONDE ALARCOS	100	89	98	95
EL RENEGADO ARREPENTIDO	98	91	81	90

Quiero terminar con Ruiz de Alarcón, en parte para subrayar la efectividad del método, aunque encontremos en el camino alguna sorpresa que luego comentaré. Debajo están los resultados de las 15 comedias cotejadas:

69	C4	W2	W3	%
EXAMEN DE MARIDOS	100	97	97	98
LA INDUSTRIA Y LA SUERTE	100	97	95	97
TODO ES VENTURA	100	95	94	96
LA PRUEBA DE LAS PROMESAS	100	95	88	94
MUDARSE PARA MEJORARSE	100	94	89	94
LA CRUELDAD POR EL HONOR	100	91	85	92

LA AMISTAD CASTIGADA	100	88	86	91
EL DUEÑO DE LAS ESTRELLAS	100	89	82	90
LA MANGANILLA DE MELILLA	100	82	89	90
GANAR AMIGOS	100	88	81	89
EL TEJEDOR DE SEGOVIA	98	88	76	87
LOS FAVORES DEL MUNDO	79	59	52	63
EL ANTICRISTO	78	55	46	59
EL DESDICHADO EN FINGIR	76	60	35	57
LOS PECHOS PRIVILEGIADOS	79	53	33	55

La disparidad entre las once primeras comedias con una media del 92% y las cuatro últimas con menos del 60% llama poderosamente la atención. A través de lo visto hasta ahora deberíamos descartarlas *ipso facto* del canon del novohispano, pero las cuatro fueron publicadas en vida de Alarcón, seguramente bajo su supervisión²¹. En relación con *El anticristo* tenemos hasta una carta de don Luis de Góngora en donde refiere el escándalo que se formó el día de su estreno cuando alguien entre el público tiró una bomba fétida. Lope y Mira de Amescua parece que fueron a la cárcel por el incidente, aunque todo hace pensar que el responsable del desaguisado fue Juan Pablo Rizo (ORTIZ: 148; PEÑA: 66). ¿Podemos dudar entonces de su autoría? En principio no, pero los datos estadísticos son ciertamente poco favorables. Es sabido que alguna vez Alarcón solicitó la colaboración de poetas, ocasionándole más de un quebradero de cabeza, como cuando encargó toda una serie de composiciones a distintos autores con ocasión de la visita del príncipe de Gales en 1623, lo cual provocó la rechifla de sus colegas y una cascada de letrillas hirientes por parte de casi todos, de Lope a Quevedo, y hasta Mira de Amescua, pese a ser uno de los que habían colaborado con Alarcón. Una décima de Salas Barbadillo lo compara con Claramonte a la hora de plagiar o aprovecharse de obras ajenas:

El segundo Claramonte
 Por llenar más presto el vaso
 No fue al monte del Parnaso
 Por agua, sino a Belmonte...

Este Belmonte no es otro que Luis Belmonte Bermúdez (1587-1650), dramaturgo que colaboró con otros ingenios en la confección de varias comedias, tal como vimos antes cuando estudiamos *El mártir de Madrid*. Más difícil es determinar si Belmonte, en efecto, participó directamente en la gestación de alguna de las comedias que hoy pasan por obras del novohispano, pero los versos de Barbadillo sí que parecen sugerir que Alarcón, lo mismo que Claramonte, se valía de otras «fuentes» para «llenar más presto el vaso» de la comedia que tenía entre manos. ¿Pasó así con esas cuatro comedias y de ahí las discrepancias con el resto de obras computadas? Podría ser, pero yo no me siento ni con conocimiento ni con más fuerzas para profundizar en el asunto. Otros deberán retomar esta labor, si es que sienten la curiosidad.

21.— *Los favores del mundo* y *El desdichado en fingir* se imprimieron en 1628 en Madrid (Juan González); *El anticristo* y *Los pechos privilegiados* en 1634 en Barcelona (Sebastián de Cormellas).

Recapitulemos

Todo estudio de atribución basado en datos internos analiza el grado de similitud entre textos. A mayor similitud, mayor porcentaje de fiabilidad. La cuestión es determinar: 1) qué rasgos lingüísticos se deben analizar, 2) cómo se cuantifican estos rasgos; y 3) cuál es la línea divisoria o el porcentaje que separa a un autor de los demás. Últimamente se ha visto que uno de los discriminadores más eficaces en la clasificación por autor es el n-grama de caracteres/palabras. Así, en esta investigación me he preocupado de calcular las secuencias verbales de un texto bajo tres categorías: n-gramas de cuatro caracteres (C4), n-gramas de dos palabras (W2) y n-gramas de tres palabras (W3). El programa utilizado ha sido JPAAG y el método, entre los muchos disponibles, ha consistido en aplicar la métrica de similitud de coseno. Tras probar la eficacia del método con textos como *La tía fingida* o las novelas a Marcia Leonarda, he analizado 268 comedias de Lope o atribuidas a él, además de una buena cantidad de comedias de Tirso, Alarcón y Mira de Amescua. Todos los cotejos se han hecho con 24 comedias de seis autores canónicos, cada uno de ellos con cuatro comedias. Por lo general, los cálculos con comedias consideradas de autoría fiable han ofrecido un alto porcentaje de proximidad en las tres categorías estudiadas y, por el contrario, las de autoría dudosa o no de Lope no han pasado normalmente del 70%. Pero no siempre ha sido así. Algunas consideradas de autoría fiable por la crítica han presentado un bajo porcentaje. La explicación parece estar en las fechas de composición, ya que casi todas las que obtuvieron un porcentaje por debajo del 80% están escritas antes de 1600 y otras mucho antes. En estos casos la mejor opción sigue siendo el tradicional acopio de paralelismos. Así lo he hecho con *Los donaires de Matico*, por ejemplo. Otras veces, con comedias consideradas de autoría probable, como *El nacimiento de Cristo*, o dudosa, como *La devoción del rosario*, también se ha visto que comparten paralelismos muy significativos con el corpus de Lope, pese al pobre rendimiento en los análisis. En cuanto a las comedias etiquetadas como no de Lope, todas han conseguido números bajos en nuestro análisis. La única excepción es *El alcalde de Zalamea*, aunque en este caso la atribución debe ser revisada, pues junto al buen resultado de nuestro análisis, se suma un rimerero de paralelismos raros o exclusivos. La comedia *La estrella de Sevilla* sí que debe descartarse definitivamente del canon lopesco, sin que Andrés de Claramonte sea necesariamente su autor, pese a tener dos comedias atribuidas a él en los primeros lugares. En efecto, *De este agua no beberé* y *Púsoseme el sol, saliome la luna* presentan en muchas de las comedias cotejadas un grado de proximidad extraordinario, mientras que sorprendentemente sus otras dos comedias (*De lo vivo a lo pintado* y *El negro más valiente*) aparecen continuamente al final la lista. No hay clara explicación. Una conjetura es que *De este agua* y *Púsoseme* surgen del círculo de Lope de Vega, Tirso de Molina y quizá Mira de Amescua, lo cual explicaría por qué aparecen también muy próximas a *El burlador de Sevilla* (o su otra versión, *Tan largo me lo fiáis*) y *El condenado por desconfiado*. Independientemente de la participación o no del dramaturgo murciano, parece claro, a tenor de los resultados, que *El condenado por desconfiado* no es de Tirso y que el texto que nos ha llegado de *El burlador*, si en verdad lo escribió el mercedario, dista mucho del original. Finalmente, casi todas las comedias de Tirso, Mira y Alarcón de autoría fiable alcanzan muy buenos resultados, mientras que las cuestionadas o dudosas bajan el porcentaje. Hay alguna excepción.

El caso más llamativo sería *El anticristo* de Alarcón, autoría refrendada por una carta de Góngora, pero con números bajísimos en nuestro análisis.

Bibliografía

- ALVITI, Roberta. (2006). *I manoscritti autografi delle commedie del Siglo de Oro scritte in collaborazione. Catalogo e estudio*. Florencia: Alinea.
- ANTONIA, A. y Hugh ELIOT. (2014). «Language chunking, data sparseness, and the value of a long marker list: explorations with word n-grams and authorial attribution». *Literary and Linguistic Computing*, Volume 29, n. 2 (147–163).
- ARELLANO, Ignacio [coordinador]. (2004). *Paraninfos, segundones y epígonos de la comedia del Siglo de Oro*. Navarra: Anthropos.
- ARJONA, J. H. (1965). «Two Plays Attributed to Lope de Vega and Guillen de Castro». *Hispanic Review* (v. 33).
- COTARELO Y MORI, Emilio. (1930). *Mira de Amescua y su teatro*. Madrid: Boletín de la Real Academia de la lengua. XVII.
- COUDERC, Christophe (2009). «El autor ante la edición de sus obras». Incluido en MONER (2009)
- CASTRO, GUILLÉN DE (1926). *Obras de don Guillén de Castro y Bellvís*. Madrid: Real Academia Española, Tipografía de la Revista de Archivos, Bibliotecas y Museos.
- DE ARMAS, Frederick A. (1994). «Balthasar's Doom: Letters that Heal/Kill in Claramonte's El secreto en la mujer». Incluido en GANELIN (2009).
- DIXON, Victor (1996). «La intervención de Lope en la publicación de sus comedias». *Anuario Lope de Vega*, núm. 2.
- GANELIN, Charles et al (1994). *The Golden Age Comedia: Text, Theory, and Performance*. Purdue University Press.
- JAURALDE POU, Pablo (2012). *Diccionario filológico de literatura española (Siglo XVII): Volumen I*. Madrid: Castalia.
- JULIÁ MARTÍNEZ, Eduardo. (1926). Incluido en CASTRO (1926: 40-91).
- JUOLA, Patrick (2009). «JGAAP: A System for Comparative Evaluation of Authorship Attribution». *Chicago Colloq. Digit. Hum. Comp. Sci.* 1(1).
- KESELJ, Vlado et al. (2003). «N-gram-based author profiles for authorship attribution». *Proceedings of the Conference Pacific Association for Computational Linguistics*. Stroudsburg: ACL Press. (255-264).
- LEAVITI, Sturgis E. (1931), *The Estrella de Sevilla and Claramonte*. Boston: Harvard University Press.
- LUCÍA MEGÍAS, José Manuel (2018). «El código Porras (casi) recuperado (la copia del Cigarral del Carmen de *La tía fingida*)» *Anales cervantinos* v. 50.
- MADRIGAL, José Luis (2003). «De cómo y por qué *La tía fingida* es de Cervantes». *Artifara: Revista de lenguas y literaturas ibéricas y latinoamericanas*, N.º. 2.
- MENÉNDEZ PELAYO, Marcelino (1949). *Estudios sobre el teatro de Lope de Vega— IV: Crónicas y leyendas*. Madrid: Consejo Superior de Investigaciones Científicas.
- MONER, Michel et al (2009). *Paratextos en la literatura española (siglos XV-XVIII)*. Madrid: Casa Velázquez.

- MORLEY, Sylvanus Griswold (1914). «El uso de las combinaciones métricas en las comedias de Tirso de Molina». *Bulletin Hispanique* XVI. N.º. 2: 177-208
- MORLEY, Sylvanus Griswold y C. BRUERTON (1968). *Cronología de las comedias de Lope de Vega*. Madrid: Gredos.
- PEÑA, Margarita. (2004). «Juan Ruiz de Alarcón: biografía y comedias». Incluido en ARELLANO (2004)
- RODRIGUEZ LÓPEZ-VÁZQUEZ, Alfredo (1987). *La autoría de «El Burlador de Sevilla»: Andrés de Claramonte*. Kassel: Reichenberger.
- SARI, Yunita *et al.* (2017). «Continuous N-gram Representations for Authorship Attribution». *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. V. 2. Stroudsburg: ACL Press. (267-273).
- STAMATATOS, Efstathios. (2009). «A Survey of Modern Authorship Attribution Methods». *Journal of the American Society for Information Science and Technology*, 60(3):538–556.
- (2013). «On the Robustness of Authorship Attribution Based on Character n-gram Features». *Journal of Law and Policy*. 21(2):421–439.
- VALLADARES REGUERO, Aurelio. (2004). *Bibliografía de Antonio Mira de Amescua*. Kassel: Reichenberger
- VEGA, LOPE DE (1614). *Doze comedias de Lope de Vega Carpio... sacadas de sus originales*. Cuarta parte. Barcelona: Sebastián de Cormellas.
- (1617). *Doze comedias de Lope de Vega, sacadas de sus originales por él mismo*. Novena parte. Madrid: Viuda de Alonso Martín de Balboa.
- VILLEGAS, Alonso de (1796). *Flos sanctorum. Historia general de la vida y hechos de Jesu-Chisto*. Barcelona: Imprenta de Isidro Aguasvivas.

Apéndice

En este análisis contabilizo únicamente las primeras cuatro comedias que aparecen al principio de la lista, con una puntuación de 4 a 1 en cada categoría (C4, W2, W3), siendo 30 pts. la puntuación más alta (100%) y 1 pt. la más baja (3%). Baste un ejemplo. En el coitejo con *La estrella de Sevilla* el programa JGAAP nos da el siguiente resultado al calcular el grado de similitud de los n-gramas de 4 caracteres:

```

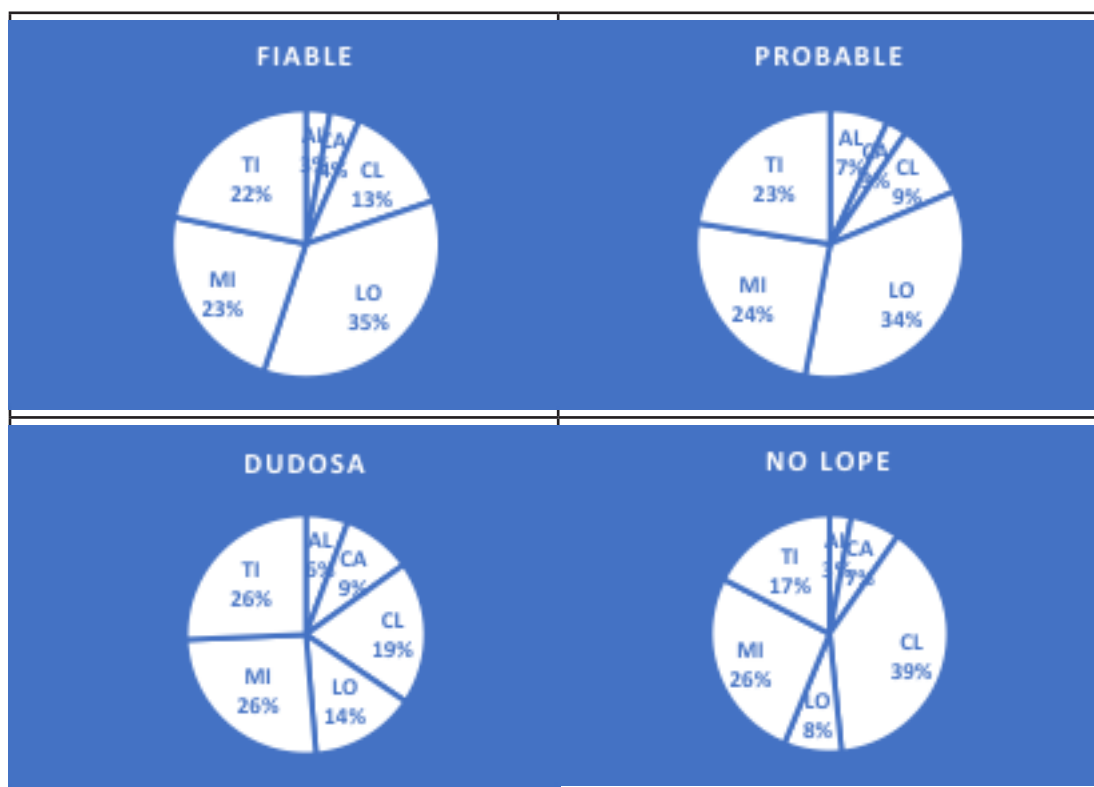
La estrella de Sevilla.docx L:\Users\jose_luis\Dropbox\JGAAP\HITS\ALARCÓN\MI
Canonicizes:
  Normalize ASCII
  Normalize Whitespace
  Strip Numbers
  Strip Punctuation
EventDrivers:
  Character NGrams n: 4
Analysis:
  Absolute Centroid Driver with metric Cosine Distance
1. CLARAMONTE Deste agua no beberé 0.06367704376633554
2. CLARAMONTE Púsoseme el sol 0.07782523951307707
3. TIRSO Caballero de Gracia 0.07903163220136178
4. ALARCÓN La verdad sospechosa 0.07943482589458295
5. LOPE Fuente Ovejuna 0.09089196157177113
6. TIRSO Marta la piadosa 0.08143773276944735
7. LOPE El caballero de Olmedo 0.09391092951674516
8. MIRA DE AMESCUA La Fénix de Salamanca 0.08477566502881573
9. MIRA DE AMESCUA El esclavo del demonio 0.08529914194685106
10. GUILLÉN DE CASTRO El amor constante 0.08666036573401392
11. MIRA DE AMESCUA El ejemplo mayor 0.08785538964443174
12. ALARCÓN Las paredes oyen 0.089566670399995
13. LOPE Castigo sin venganza 0.08965795341457816
14. TIRSO El vergonzoso en palacio 0.09000371776806615
  
```

Pues bien, en nuestro experimento las dos comedias de Claramonte suman 7 puntos (4 + 3), la de Tirso 2 y la de Alarcón 1. El mismo procedimiento se aplica a las otras dos categorías (W2, W3). El resultado final es:

	Pts.	%
CLARAMONTE	18	60%
MIRA	6	20%
TIRSO	5	17%
ALARCÓN	1	3%

	ALarcón	CAstro	CLaramonte	LOpe	MIRA	TIrso
LA ESTRELLA DE SEVILLA	3		60		20	17

Debajo está la lista completa con las comedias que se han analizado. He dividido la lista en cuatro apartados. El primero incluye las comedias consideradas por la crítica de autoría fiable (F); el segundo, las de autoría probable (P); el tercero, de autoría dudosa (D) y, por último, las consideradas como no de Lope (N). En un rápido cálculo de porcentajes se puede comprobar que Lope promedia más de un tercio de proximidad con respecto al resto de autores en las comedias de autoría fiable o probable, mientras que en las dudosas o en las etiquetadas como «no de Lope», el Fénix se ve superado por Tirso, por Mira de Amescua y, especialmente, por Claramonte, con un porcentaje aún más alto del que tiene Lope en las comedias de autoría fiable, según se ve en el gráfico:



AUT		AL	CA	CL	LO	MI	TI
F	EL ABANILLO	7		3	40	27	23
F	EL ACERO DE MADRID	7			47	30	16
F	ADONIS Y VENUS	7	13	30	13	7	30
F	AL PASAR DEL ARROYO	13		7	43	27	10
F	EL ALCAIDE DE MADRID	7	10	20	20	23	20
F	EL ALCALDE MAYOR	7		3	40	20	30
F	EL AMANTE AGRADECIDO	3			47	33	17
F	LOS AMANTES SIN AMOR	3		7	50	30	10
F	DE AMAR SIN SABER A QUIÉN	7			43	20	30
F	AMAR, SERVIR Y ESPERAR	3			54	23	20
F	EL AMIGO HASTA LA MUERTE	3	3	10	44	23	17
F	EL AMIGO POR FUERZA			3	27	50	20
F	LA AMISTAD Y OBLIGACIÓN			3	44	23	30
F	EL AMOR DESATINADO	3	7	17	20	30	23
F	EL AMOR ENAMORADO	4	10	20	40	13	13
F	AMOR SECRETO HASTA CELOS				54	16	30
F	AMOR, PLEITO Y DESAFÍO	10			40	27	23

F	LOS AMORES DE ALBANIO Y ISMENIA		7	3	30	27	33
F	ANGÉLICA EN EL CATAY			57	10	17	16
F	EL ANIMAL DE HUNGRÍA			3	44	20	33
F	EL ANZUELO DE FENISA	3		10	44	27	16
F	EL ARAUCO DOMADO			44	33	10	13
F	LA ARCADIA		4	10	53	23	10
F	EL ARENAL DE SEVILLA	6	4	4	40	23	23
F	EL ARGEL FINGIDO		3	10	40	27	20
F	EL ASALTO DE MASTRIQUE	3		17	37	26	17
F	AY VERDADES QUE EN AMOR	6			54	20	20
F	LOS BANDOS DE SENA	7			50	23	20
F	BARLAÁN Y JOSAFAT			46	17	14	23
F	EL BASTARDO MUDARRA			27	43	17	13
F	LA BATALLA DEL HONOR			4	60	23	13
F	LAS BATUECAS DEL DUQUE DE AL			37	26	17	20
F	EL BAUTISMO DEL PRÍNCIPE MAR	4		13	30	30	23
F	BELARDO EL FURIOSO		6	10	27	30	27
F	LA BELLA AURORA			20	44	23	13
F	LA BELLA MALMARIDADA	3	3	10	27	44	13
F	LOS BENAVIDES		20	17	17	23	23
F	LAS BIZARRÍAS DE BELISA	7		10	44	16	23
F	EL BLASÓN DE LOS CHAVES	3		20	30	33	14
F	LA BOBA PARA LOS OTROS	3	3		54	20	20
F	EL BOBO DEL COLEGIO	3	3		40	37	17
F	EL BRASIL RESTITUIDO			60	10	13	17
F	LA BUENA GUARDA			14	40	23	23
F	LA BURGALESA DE LERMA			7	40	30	23
F	LAS BURLAS DE AMOR	7	10		37	23	23
F	EL CABALLERO DE ILLESCAS	4			33	23	40
F	EL CABALLERO DEL MILAGRO	7	3	3	40	27	20
F	LA CAMPANA DE ARAGÓN	3	6	27	10	27	27
F	LA CARBONERA			17	47	20	16
F	CARLOS V EN FRANCIA		4	30	20	23	23
F	EL CASAMIENTO EN LA MUERTE			13	27	20	40
F	CASTELVINES Y MONTESES		7	3	57	20	13
F	EL CASTIGO DEL DISCRETO	4		10	50	23	13
F	LOS CELOS DE RODAMONTE			44	20	20	16

F	LOS COMENDADORES DE CORD	7		30	17	23	23
F	CON SU PAN SE LO COMA			10	44	23	23
F	EL CONDE FERNÁN GONZALEZ		6	20	17	30	27
F	LA CONTIENDA DE JUAN DE PARE			27	23	27	23
F	CONTRA VALOR NO HAY DESDIC		4	20	37	16	23
F	LA CORONA DE HUNGRÍA				63	10	27
F	LA CORTESÍA DE ESPAÑA	3		7	40	27	23
F	LAS CUENTAS DEL GRAN CAPITÁN	7		13	27	23	30
F	EL CUERDO EN SU CASA	3		4	53	27	13
F	EL CUERDO LOCO		7	33	17	20	23
F	LA DAMA BOBA	7			53	23	17
F	DE CUANDO ACÁ NOS VINO	7			43	27	23
F	DEL MAL LO MENOS	4			53	23	20
F	DEL MONTE SALE	3			57	17	23
F	EL DESCONFIADO	11		13	13	23	40
F	LA DESDICHADA ESTEFANÍA	4		23	30	23	20
F	EL DESPOSORIO ENCUBIERTO			10	37	30	23
F	EL DESPRECIO AGRADECIDO	4	3		50	23	20
F	DIOS HACE REYES			4	60	13	23
F	LA DISCORDIA DE LOS CASADOS			10	43	30	17
F	DIVINA VENCEDORA		10	17	23	30	20
F	EL DÓMINE LUCAS	14		6	27	23	30
F	DON JUAN DE CASTRO I	7		23	24	23	23
F	DON JUAN DE CASTRO II			23	24	30	23
F	DON LOPE DE CARDONA			34	33		33
F	LOS DONAIRES DE MATICO		33		17	27	23
F	LA DONCELLA TEODOR			23	30	33	14
F	LOS EMBUSTES DE CELAURO		3		37	37	23
F	EL ENEMIGO ENGAÑADO	6	4		43	20	27
F	LOS ENEMIGOS EN CASA	4			53	23	20
F	EL ESCLAVO DE ROMA		7	16	27	37	13
F	LOS ESCLAVOS LIBRES	6	7	10	27	27	23
F	LA ESCOLÁSTICA CELOSA		3	4	30	33	30
F	LOS ESPAÑOLES EN FLANDES			16	27	20	37
F	EL FAVOR AGRADECIDO	4	7		23	43	23
F	LAS FAMOSAS ASTURIANAS			34	23	13	30
F	LA FE ROMPIDA			4	50	23	23

F	LO FINGIDO VERDADERO	6			57	17	20
F	LA FIRMEZA EN LA DESDICHA		3		50	17	30
F	LAS FLORES DE DON JUAN	7			40	23	30
F	LA FRANCESILLA		7	7	33	33	20
F	LA FUERZA LASTIMOSA		10	17	30	20	23
F	EL GALÁN CASTRUCHO	7			47	23	23
F	EL GALÁN DE LA MEMBRILLA			14	43	20	23
F	EL GALÁN ESCARMENTADO	10			53	27	10
F	LA GLLARDA TOLEDANA	7	3		43	20	27
F	EL GANSO DE ORO		7	3	33	20	37
F	EL GENOVÉS LIBERAL	4		10	23	33	30
F	EL GRAN DUQUE DE MOSCOVIA	6	7		33	27	27
F	LAS GRANDEZAS DE ALEJANDRO		4	43	13	17	23
F	EL GRAO DE VALENCIA	7		13	33	27	20
F	GUANCHES DE TENERIFE			47	23	17	13
F	GUARDAR Y GUARDARSE	3			63	17	17
F	GUERRAS DE AMOR Y DE HONOR			23	37	17	23
F	EL HALCÓN DE FEDERICO	3			57	20	20
F	EL HAMETE DE TOLEDO	10		10	40	23	17
F	LOS HECHOS DE GARCILASO	7	33	33	4		23
F	LA HERMOSA ALFREDA		13	7	37	23	20
F	LA HERMOSA ESTER			17	53	10	20
F	LA HERMOSURA ABORRECIDA				40	13	47
F	EL HIDALGO BENCERRAJE	3	7	17	20	30	23
F	LOS HIDALGOS DE ALDEA	6	10		50	17	17
F	HIJO DE LOS LEONES	10		4	30	36	20
F	EL HIJO SIN PADRE		6		57	20	17
F	EL HIJO VENTUROSO	7			47	23	23
F	LA HISTORIA DE TOBÍAS			30	30	27	13
F	EL HOMBRE DE BIEN	7			40	33	20
F	EL HOMBRE POR SU PALABRA			20	37	20	23
F	EL HONRADO HERMANO		10	7	17	26	40
F	LA IMPERIAL DE OTÓN		7	27	13	23	30
F	LA INFANTA DESESPERADA	6	6	6	20	36	26
F	LA INGRATITUD VENGADA	3	7		40	40	10
F	EL INGRATO ARREPENTIDO	4	16	4	33	23	20
F	LA INOCENTE LAURA	4	7		40	23	26
F	JORGE TOLEDANO	10	10		27	30	23

F	JUAN DE DIOS Y ANTÓN SAN M			17	17	33	33
F	EL JUEZ EN SU CAUSA	7	4	23	26	23	17
F	LAS JUSTAS DE TEBAS	6	23	17	14	17	23
F	EL LABERINTO DE CRETA	3		33	33	17	14
F	EL LABRADOR VENTUROSO				56	27	17
F	EL LACAYO FINGIDO	6			43	24	27
F	LAURA PERSEGUIDA	4	20	7	23	33	13
F	EL LEAL CRIADO		21		23	23	33
F	LA LIMPIEZA NO MANCHADA	4		43	10	30	13
F	LO QUE HA DE SER		3	3	50	24	20
F	LA MADRE DE LA MEJOR			53	23	14	10
F	EL MAESTRO DE DANZAR		7		37	23	33
F	LA MALCASADA			3	47	30	20
F	EL MARIDO MÁS FIRME			23	40	17	20
F	EL MARQUÉS DE LAS NAVAS			10	50	27	13
F	EL MARQUÉS DE MANTUA			33	17	20	30
F	MÁS PUEDEN CELOS QUE AMOR		3		40	27	30
F	EL MAYOR IMPOSIBLE			10	47	20	23
F	LA MAYOR VICTORIA	3		3	64	10	20
F	EL MAYORAZGO DE LA DUQUESA DE AMALFI		7		33	23	37
F	EL MEJOR ALCALDE, EL REY			7	53	23	17
F	LOS MELINDRES DE BELISA		6	3	54	24	13
F	EL MESÓN DE LA CORTE	4		3	40	23	30
F	EL MOLINO		27		17	23	33
F	LA MOZA DE CÁNTARO		4		53	20	23
F	LAS MUDANZAS DE FORTUNA ... DE DON BELTRÁN	10	4	7	33	23	23
F	EL NACIMIENTO DE CRISTO			70	17	13	
F	NADIE SE CONOCE	4	4	6	50	23	13
F	LA NECEDAD DEL DISCRETO	7			53	23	17
F	EL NEGRO DEL MEJOR AMO			57	10	10	23
F	LA NIÑA DE PLATA	6	4		43	20	27
F	LA NIÑEZ DEL PADRE ROJAS	10		6	30	27	27
F	EL NIÑO INOCENTE DE LA GUAR.			33	30	20	17
F	NO SON TODO RUISEÑORES	7		10	43	27	13
F	LA NOCHE DE SAN JUAN	6			47	27	20
F	LA NOCHE TOLEDANA	6		4	37	37	16
F	EL NUEVO MUNDO DESC. CRISTÓBAL C.	3		47	13	20	17

F	LA OCTAVA MARAVILLA			14	33	23	30
F	LAS PACES DE LOS REYES	4		27	23	33	13
F	PEDRO CARBONERO		4	23	30	16	27
F	EL PERSEGUIDO		10	10	34	23	23
F	EL PIADOSO ARAGONÉS			3	57	13	27
F	EL PIADOSO VENECIANO	3			57	23	17
F	LOS PLEITOS DE INGLATERRA			23	27	23	27
F	EL PODER EN EL DISCRETO				67	13	20
F	LOS PONCES DE BARCELONA	7		10	27	33	23
F	POR LA PUENTE, JUANA				47	20	33
F	LOS PORCELES DE MURCIA	4		17	36	23	20
F	PORFIAR HASTA MORIR	3		10	57	20	10
F	EL POSTRER GODO DE ESPAÑA			23	24	23	30
F	LOS PRADOS DE LEÓN	6	6	6	37	24	21
F	EL PREMIO DE LA HERMOSURA			53	20	17	10
F	EL PRÍNCIPE DESPEÑADO	7	6	27	17	23	20
F	EL PREMIO DEL BIEN HABLAR	4		10	33	23	30
F	EL PRÍNCIPE INOCENTE	7	7	7	30	26	23
F	EL PRÍNCIPE MELANCÓLICO		40		17	20	23
F	EL REMEDIO EN LA DESDICHA			20	30	23	27
F	LA RESISTENCIA HONRADA		10	17	30	23	20
F	ROMA ABRASADA	10		10	27	23	30
F	SAN ISIDRO LABRADOR DE MADRID			33	37	17	13
F	LA SANTA LIGA			27	20	23	30
F	EL SANTO NEGRO ROSAMBUCO			53		27	20
F	LA SELVA SIN AMOR	6		27	17	23	27
F	EL SOL PARADO	3		47	14	23	13
F	LOS TELLOS DE MENESES			3	60	20	17
F	EL TESTIMONIO VENGADO		14	33	10	23	20
F	EL TOLEDANO VENGADO	20	10		10	30	30
F	LOS TRABAJOS DE JACOB			17	50	13	20
F	LOS TRES DIAMANTES	3		3	57	17	20
F	EL TRIUNFO DE LA HUMILDAD		23	30	10	14	23
F	URSÓN Y VALENTÍN		37	7	13	20	23
F	EL VALIENTE CÉSPEDES			10	43	30	17
F	EL VALOR DE LAS MUJERES		3		67	17	13
F	VALOR...DE LOS TELLOS DE MENESES	7		23	23	17	30
F	EL VAQUERO DE MORAÑA		10	24	20	23	23

F	EL VELLOCINO DE ORO			37	40	3	20
F	LA VENGADORA DE LAS MUJERES	3			63	17	17
F	EL VERDADERO AMANTE	3	30	10	20	20	17
F	LA VIDA DE SAN PEDRO NOLASCO	7	30	10	23	30	
F	LA VILLANA DE GETAFE	10		10	30	27	23
F	EL VILLANO EN SU RINCÓN	7	7		40	23	23

AUT		AL	CA	CL	LO	MI	TI
P	LA ALDEHUELA	13		7	27	30	23
P	AMAR COMO SE HA DE AMAR	3			57	13	27
P	LAS BURLAS Y VERAS	3			53	27	17
P	LA DIFUNTA PLEITEADA	10		17	13	37	23
P	EN LOS INDICIOS, LA CULPA	23	3	4	20	33	17
P	LA ESCLAVA DE SU GALÁN				53	20	27
P	EL LOCO POR FUERZA	7		14	33	23	23
P	LA MADRE TERESA DE JESÚS	4	20	7	23	23	23
P	EL ROBO DE DINA			33	30	14	23

AUT		AL	CA	CL	LO	MI	TI
D	ACERTAR ERRANDO	7			40	30	23
D	EL ANTICRISTO		4	43	10	20	23
D	AUDIENCIAS DEL REY DON PEDR	13		30		30	27
D	LAS BURLAS Y ENREDOS DE BENIT	7	7	6	30	30	20
D	LOS CAUTIVOS DE ARGEL			27	33	23	17
D	LA COMPETENCIA EN LOS NOBL	17	10	23	3	33	14
D	LA CORONA DERRIBADA	7	23	10	13	33	14
D	LA DEVOCIÓN DEL ROSARIO		4	33	4	23	36
D	LAS DONCELLAS DE SIMANCAS		37	20		20	23
D	DONDE NO ESTÁ SU DUEÑO		60	13		17	10
D	LAS DOS BANDOLERAS			40	10	23	27
D	EL ESCLAVO FINGIDO	6	7		40	27	20
D	LA FIANZA SATISFECHA	7	30	7	23	6	27
D	LOS GUZMANES DEL TORAL	7	3	30	4	23	33
D	EL HIJO POR ENGAÑO	4	23	23	10	23	17
D	LA ILUSTRE FREGONA	24	13	10		20	33
D	EL INGRATO		3	33	17	30	17
D	LA INTENCIÓN CASTIGADA			27	33	20	20

D	JULIÁN ROMERO	4		20	10	50	16
D	LA MEJOR ENAMORADA, LA MAGDALENA			13	27	23	37
D	EL PLEITO POR LA HONRA...	10	3	20	17	27	23
D	EL PREMIO RIGUROSO	13	3		17	30	37
D	LA PRÓXPERA FORTUNA DE...		4	33	4	33	26
D	EL REY POR SU SEMEJANZA	7	13	10	14	23	33
D	SANTO ANGELO	13			7	33	47
D	LOS TERCEROS DE SAN FRANCISCO	4		14	6	23	53

AUT		AL	CA	CL	LO	MI	TI
N	LA ADVERSA FORTUNA DE DON BERNARDO DE C			27		53	20
N	EL ALCALDE DE ZALAMEA	7		16	37	20	20
N	ANTONIO ROCA O LA MUERTE		20	23	4	23	30
N	BARLAÁN Y JOSAFAT REFUNDICIÓN			60	13	17	10
N	LA CIUDAD SIN DIOS			64		26	10
N	DAVID PERSEGUIDO	6	10	44		10	30
N	LA ESTRELLA DE SEVILLA	3		60		20	17
N	LOS MÁRTIRES DE JAPÓN			53		40	7
N	LA MAYOR CORONA			77	13		10
N	EL REY POR TRUEQUE	4	53		7	23	13
N	SANTA CASILDA	3		40	20	17	20
N	EL TRUHÁN DEL CIELO	13		23	10	37	17
N	EL VASO DE ELECCIÓN SAN PABLO			47	6	27	20