

RITMO E INTELIGENCIA ARTIFICIAL: NUEVAS PERSPECTIVAS SOBRE
TEATRO LOPESCO DESDE LAS HUMANIDADES DIGITALES*

SIMON KROLL (Universidad de Viena)

FERNANDO SANZ-LÁZARO (Universidad de Viena)

CITA RECOMENDADA: Simon Kroll y Fernando Sanz-Lázaro, «Ritmo e inteligencia artificial: nuevas perspectivas sobre teatro lopesco desde las Humanidades Digitales», *Anuario Lope de Vega. Texto, literatura, cultura*, XXIX (2023), pp. 351-375.

DOI: <<https://doi.org/10.5565/rev/anuariolopedevega.491>>

Fecha de recepción: 14 de septiembre de 2022 / Fecha de aceptación: 4 de noviembre de 2022

RESUMEN

Se analiza con un nuevo método de humanidades digitales diferentes aspectos de la obra teatral de Lope de Vega. De esta manera se presentan las nuevas herramientas escritas para analizar desde una perspectiva distante aspectos métricos de los textos. En concreto el ritmo del octosílabo resulta una marca importante que puede señalar autorías y en algunos casos también el género literario, al aplicarse distintos métodos de análisis estadístico a la diferente frecuencia de los patrones rítmicos del octosílabo. Finalmente se observa que también a nivel de los personajes de una obra (en este caso *El castigo sin venganza*) pueden darse diferencias en el uso de los diferentes octosílabos. La lectura distante del metro produce por tanto resultados prometedores tanto para el análisis de autorías, como para los debates de género literario y análisis de personajes concretos.

PALABRAS CLAVE: Lectura distante; inteligencia artificial; métrica; Lope de Vega; regresión logística.

ABSTRACT

This article analyses different aspects of Lope de Vega's plays using new digital methods. In this way, new written tools are presented to analyse metrical aspects of the texts from a distant perspective. In particular, the rhythm of the octosyllable is an important mark that can indicate authorship and in some cases also the literary genre, applying different methods of statistical analysis to the different frequency of the rhythmic patterns of the octosyllable. Finally, the authors show that also at the level of the characters in a play (in this case *El castigo sin venganza*) there can be differences in the use of the different octosyllables. The distant reading of metre therefore produces promising results for the analysis of authorship, debates on literary genre and the analysis of specific characters.

KEYWORDS: Distant reading; artificial intelligence; metrics; Lope de Vega; logistic regression.

* Este artículo forma parte del proyecto «Sound and Meaning in Spanish Golden Age Literature» (FWF, Austrian Science Fund, P32563-G).

CONTEMOS PALABRAS Y MÁS

Contar las palabras de un autor, calcular así la distancia estilística que separa a un poeta de otro, una época de otra o dos géneros entre sí es fascinante y abre ante nosotros una serie de nuevas e interesantes perspectivas para observar la literatura.

Desde los trabajos de Álvaro Cuéllar y Germán Vega García-Luengos (2017-2022 y en prensa, también Vega García-Luengos 2021), la estilometría cuantitativa se ha establecido como un método de investigación importante en los estudios del teatro del Siglo de Oro. Más allá del uso instrumental del paquete Stylo desarrollado por Eder, Rybicki y Kestemont [2016], habría que mencionar que Cuéllar y Vega aplicaron varios métodos matemáticos digitales a sus resultados, de manera que conceptos como *inteligencia artificial*, *machine learning* y máquinas de vectores de soporte (SVM) han dejado de ser términos totalmente ajenos a nuestra disciplina. Ya era hora.

Todos estos estudios estilométricos —llamémosles *clásicos*— comparten una característica, pues con independencia de los métodos que emplean, coinciden en el criterio textual que analizan: la frecuencia de las palabras. Además, claro está, los trabajos aludidos no son recuentos a ciegas, sino que se interpretan a partir de una revisión detallada del contexto filológico de las obras, teniendo en cuenta los impresos, manuscritos, fechas de representación y trabajos previos de datación, entre estos últimos, también los basados en la métrica, como viene a ser el ya canónico trabajo de Morley y Bruerton [1968].

De esta manera, no obstante los fructíferos resultados de investigadores como Germán Vega y Álvaro Cuéllar, creemos que sería interesante mover el foco del análisis y evaluar no solo la frecuencia de las palabras, sino la de otros elementos estilísticos. El proyecto de investigación *Sound and Meaning in Spanish Golden Age Literature* ha producido en los dos últimos años una colección de programas que automatizan el análisis digital de la estructura métrica del texto dramático (Sanz-Lázaro 2021a, Sanz-Lázaro 2021b, Sanz-Lázaro 2022). El resultado permite ir un paso más allá de los trabajos pioneros de Morley y Bruerton, en tanto que este constituye un detallado corpus que recoge tanto la caracterización métrica del verso como su

contexto dramático. De esta forma, además de la estructura rítmica de cada verso, se indica la secuencia de sus núcleos vocálicos y la rima, pero también metadatos como el número de verso en la obra, la jornada y el número de parlamento al que pertenece el verso, el personaje que lo pronuncia o la tipología y sexo de este, así como información extratextual como el autor de la obra, título¹ y fecha de creación. Este corpus ya constituye en sí mismo una utilísima herramienta para encontrar instancias de elementos dados, pero, además, al estar estructurado en un formato tabular que cualquier paquete estadístico puede utilizar, da la posibilidad de hacer análisis de «lectura distante»² procesando cantidades descomunales de versos. Actualmente contamos con 285 obras completamente analizadas en este archivo, 144 de Calderón y 58 de Lope. Hay que hacer la prevención de que no todas las pruebas cuantitativas permiten usar el corpus completo, ya que determinados procedimientos obligan a seleccionar tan solo un subconjunto para evitar la potencial inconsistencia de los resultados que podría derivar de la desproporción de obras entre autores.

Veamos, pues, cómo se aplica esto en la práctica, y qué mejor que por la vía del ejemplo. Empezaremos por comparar si existen diferencias entre las frecuencias de los diversos ritmos. En nuestro caso, limitamos el análisis a los octosílabos en consideración a trabajos anteriores en los que observamos que la frecuencia de determinados patrones octosilábicos es una señal de autoría similar a la producida por las palabras más frecuentes (Kroll y Sanz-Lázaro 2022). Esto es, el análisis del uso que hace cada autor de determinados elementos —palabras más frecuentes en la estilometría léxica, patrones rítmicos en nuestras investigaciones previas— produce un patrón reconocible que permite identificar la mano que lo escribió con un alto

1. Nos referimos al título como uno de los elementos del paratexto, no embargante otras posibles menciones metatextuales autorreferenciales a este en los propios parlamentos.

2. Entiéndase «lectura distante» no como una descripción del volumen concreto del corpus sino en el sentido de *formalismo cuantitativo* que le dio Moretti [2004] al acuñar el término, como una aproximación a la interpretación en sentido opuesto a la de la lectura atenta o *close reading*: «you reduce the text to a few elements, and abstract them, and construct a new, artificial object. A model. And at this point you start working at a ‘secondary’ level, removed from the text: a map, after all, is always a look from afar—or is useless, like Borges’s map of the empire. Distant reading, I have called this work elsewhere; where distance is however not an obstacle, but a *specific form of knowledge*: fewer elements, hence a sharper sense of their overall interconnection. Shapes, relations, structures. Patterns» (Moretti 2004:94). Dicho de otra manera, asumimos que lo que media conceptualmente entre la lectura distante y la lectura atenta no es el tamaño del corpus sino la escala del análisis y las observaciones a partir de las que este se lleva a cabo. En nuestro caso, esto se traduce en la búsqueda de los patrones macroscópicos formados por elementos mínimos tales como el ritmo de cada verso.

grado de precisión. Probaremos este procedimiento de nuevo con otro conjunto de datos para verificar la validez de esta aproximación. En el artículo referido sobre *El divino Jasón* (Kroll y Sanz-Lázaro 2022) solo tuvimos en cuenta autos sacramentales. Ampliaremos por lo tanto el enfoque para probar los siguientes puntos:

1. ¿Se mantiene la señal de autoría al estudiar todas las obras de Lope y Calderón, independientemente del género?
2. ¿Qué influencia tiene el género literario sobre la señal?
3. ¿Cómo se comportan obras con problemas de atribución cuando las comparamos con obras de Lope y Calderón de atribución segura?
4. ¿Hay diferencias dentro de las obras, personajes que usan con mayor o menor frecuencia un determinado ritmo?

FRECUENCIA DE RITMOS: ¿SEÑAL DE AUTORÍA?

Como primera toma de contacto, escudriñaremos los valores absolutos de nuestro corpus para obtener una impresión preliminar de lo que vamos a encontrarnos. Tomamos para ello todas las obras de Calderón y Lope de las que dispone nuestra base de datos.³ Si comparamos los octosílabos más frecuentes de los dos autores, observamos que el ritmo con sílabas 3 y 7 tónicas es con creces el ritmo más frecuente en Calderón, mientras que Lope parece preferir acentos en 2, 4 y 7 (Tabla 1).

Lope		Calderón	
Ritmo	Instancias	Ritmo	Instancias
-+ + + + + -	6.398	-+ + + + + -	23.175
-+ + + + + -	5.257	-+ + + + + -	20.930
-+ + + + + -	4.968	-+ + + + + -	20.761
-+ + + + + -	4.949	-+ + + + + -	19.392
+ + + + + + -	4.554	+ + + + + + -	14.291

Tabla 1. Ritmos más frecuentes de Lope y Calderón.

3. Un listado de los títulos estudiados para este artículo puede verse al final del texto.

Para entrar más en detalle, aplicaremos procedimientos de cálculo estadístico para verificar si las diferencias en la distribución de estos ritmos entre los corpus de Lope y de Calderón son significativas. La distribución de tres de estos cinco ritmos en las obras de cada autor resultó presentar una diferencia relevante con una confianza del 95 %.⁴ Los patrones $-+ -+ -+ -$ y $+ - -+ -+ -$, por el contrario, no presentaban diferencias significativas y, por lo tanto, resultan menos susceptibles de ser empleados para discriminar por autoría. En realidad, la prueba estadística sirve para poco más que para confirmar matemáticamente lo que ya resulta evidente en la Figura 1,⁵ donde se observa una divergencia notable entre las distribuciones de los ritmos 4-7 y, sobre todo, 3-7 y 2-5-7 en las obras de Lope y de Calderón.

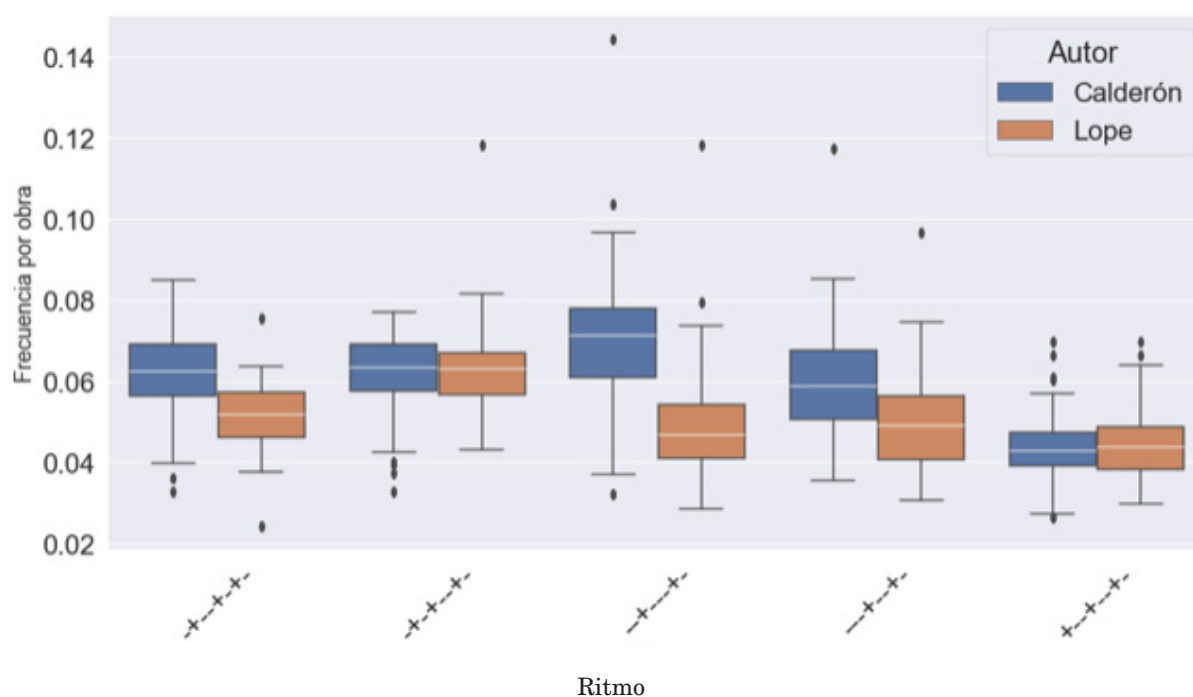


Figura 1. Distribución de los ritmos más frecuentes por autor.

4. Aplicamos una prueba paramétrica robusta de análisis de la varianza con $\alpha = 0,05$.

5. El gráfico muestra un diagrama de cajas, que permite observar la distribución de la frecuencia de un ritmo (en el eje de ordenadas) —esto es, la proporción de este (en el eje de abscisas) respecto al total de versos octosílabos en cada texto— en las obras. Así, la línea que divide la caja en dos corresponde a la mediana y los *bigotes* inferior y superior corresponden a las obras con las frecuencias mínima y máxima del ritmo correspondiente. Entre el mínimo y la caja se ubica el primer cuartil, el segundo es la parte inferior de la caja hasta la mediana, la parte superior de la caja es el tercer cuartil y, desde ahí hasta el máximo es el cuarto cuartil. Los puntos representan valores atípicos.

Esto quiere decir que las distribuciones de al menos tres de los octosílabos más frecuentes apuntan a una diferencia estilística claramente marcada entre ambos autores. Podemos reducir la parte calderoniana a una sola obra para comparar obras singulares con todas las de Lope. Si comparamos de esta manera *La dama duende* con el corpus de Lope, comprobamos que este caso escogido ejemplifica el fenómeno observado en el gráfico anterior. En dos de los cinco octosílabos más frecuentes *La dama duende* queda muy claramente fuera de los cuartiles centrales lopescos, como se muestra en Figura 2.

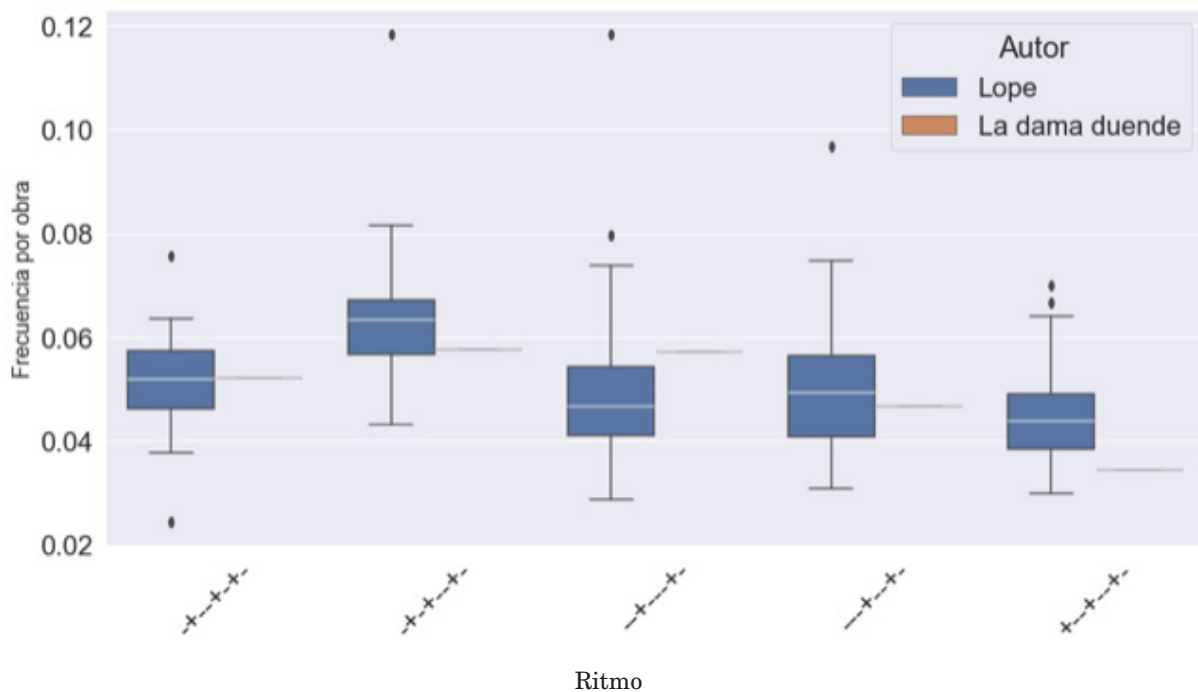


Figura 2. Ritmos de *La dama duende* frente a la distribución calderoniana.

El caso es parecido si comparamos a Lope con otros dramaturgos. Para evaluarlo, tomamos al azar tres obras de diversos autores de nuestro corpus: *Los empeños de una casa* (Sor Juana), *El lindo don Diego* y *Antíoco y Seleuco* (Moreto) para verificar si asimismo se cumplía con estos la distinción en las preferencias rítmicas observadas entre Lope y Calderón. En efecto, parece que existe una diferencia entre los usos rítmicos, que se hace notar de forma especial en los ritmos 3-7, que Lope tiende a usar de una manera notablemente exigua con respecto a otros autores. Para ese ritmo, la obra de Sor Juana (fig. 3) queda totalmente fuera de la norma

lopesca, y fuera de los cuartiles centrales para los otros dos ritmos. Dicho de otra manera, la prolijidad en el empleo del ritmo 3-7 haría improbable que la pieza pro-

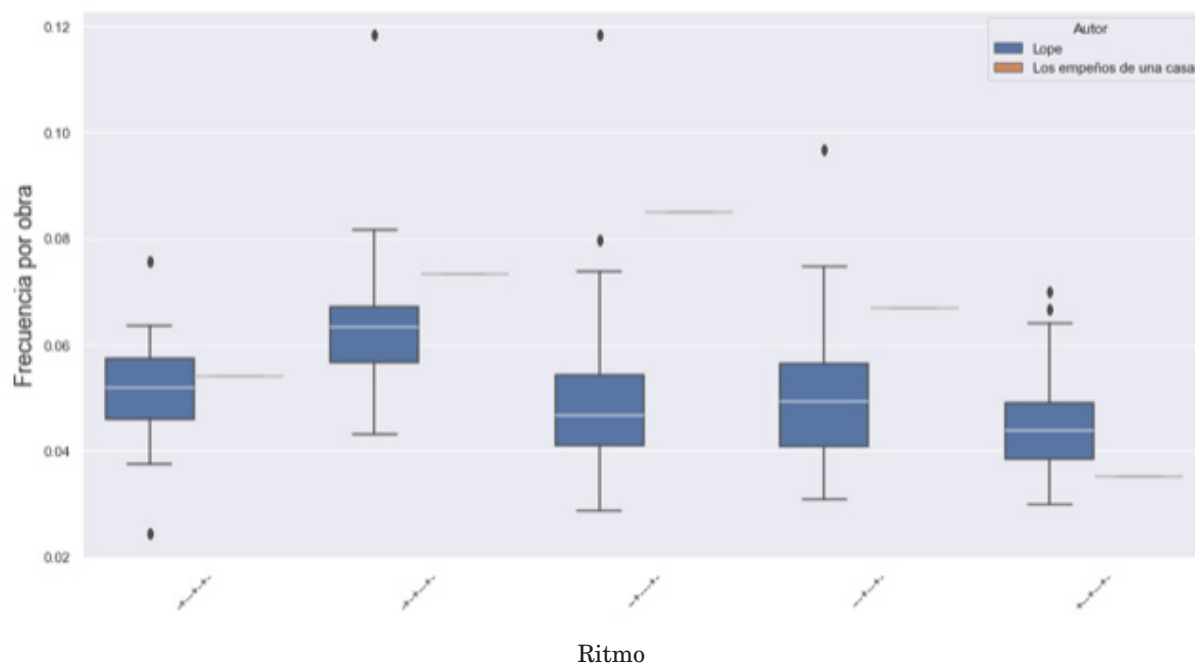


Figura 3. Ritmos de *Los empeños de una casa* frente a la distribución lopesca.

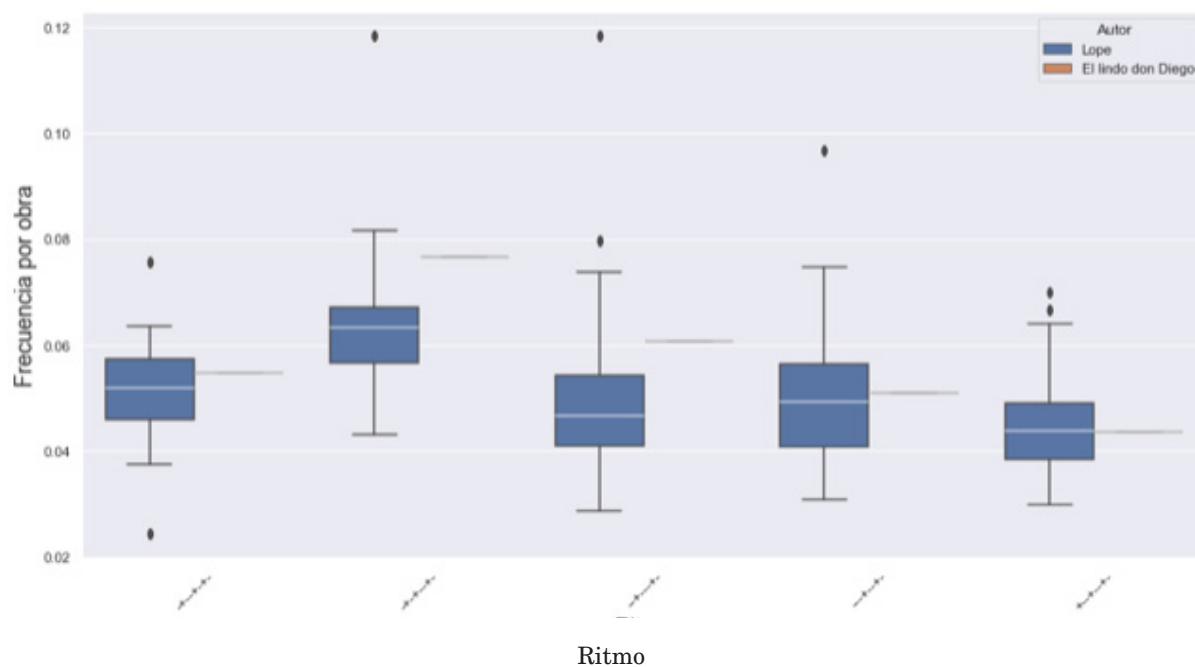


Figura 4. Ritmos de *El lindo don Diego* frente a la distribución lopesca.

viniera de la pluma de Lope. Nótese que hemos dicho improbable y no imposible; esta observación ha de tomarse con cierta precaución, pues alude a una tendencia general que, en casos concretos, puede no cumplirse, como vemos en los dos puntos del gráfico que representan dos obras de Lope con valores atípicos.

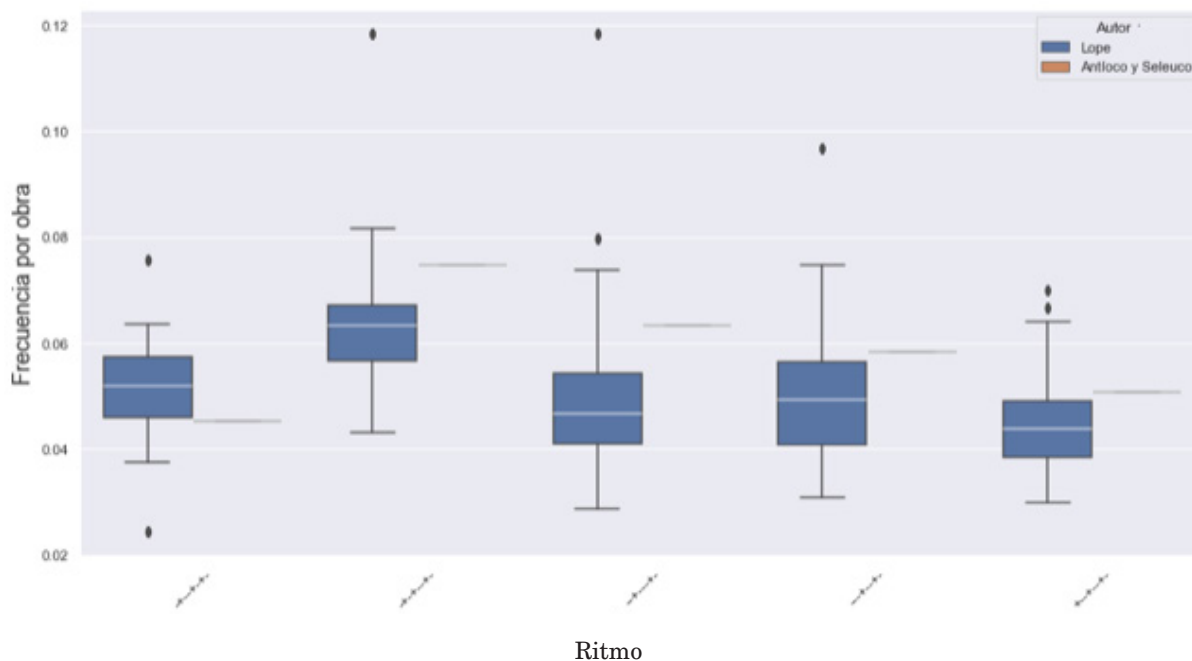


Figura 5. Ritmos de *Antíoco* y *Seleuco* frente a la distribución lopesca.

Hasta aquí hemos visto ejemplos con los patrones rítmicos de interés reducidos a los cinco más comunes de Lope. Esta simplificación responde a un propósito ilustrativo puesto que es la única manera de garantizar la legibilidad de los gráficos conservando unas dimensiones moderadas. No obstante, es posible ampliar el foco para observar todos los octosílabos que aparezcan un número suficiente de veces⁶ y seguir comparando de forma análoga. Si de nuevo cotejamos todo Calde-

6. Esta restricción responde al motivo ya mencionado. Representar todos los ritmos posibles requiere reducir los elementos del gráfico a una escala que imposibilita su correcta lectura o ampliar el gráfico a unas dimensiones que lo hacen poco práctico. Así pues, hay que encontrar un compromiso entre reflejar tanta información como sea posible en el gráfico y que este siga siendo legible. Hay además ritmos tan poco frecuentes que son irrelevantes en análisis estadísticos y pueden excluirse sin que afecte a cualitativamente al resultado de los cálculos. Aunque precisamente esa excepcionalidad hace de estos versos *raros* elementos dignos de estudio casuístico, por lo que creemos que tal vez merecerían dedicarles un trabajo aparte con una metodología más tradicional.

rón con todas las obras que tenemos de Lope, se evidencian las diferencias rítmicas. Entre estas, algunas sugieren a simple vista una marca autorial más fuerte, por lo que cabe esperar que ofrezcan buenos resultados. Otras, por el contrario, no presentan distribuciones con una distinción clara, por lo que probablemente su contribución a discriminar entre autores sería menos evidente. Para afinar incluso más, podríamos tomar no un ritmo sino un conjunto de ritmos con marca autorial clara, de manera que, en caso de que un ritmo presentara valores atípicos para un autor, esto se compensara mediante el cotejo del resto de ritmos. Dicho de otra manera, en lugar de contrastar si un ritmo determinado se ajusta a los usos del autor, comparamos si se adecua el patrón que forman una serie de ritmos en su conjunto.

Llegados a este punto empezamos a entrar de nuevo en el terreno donde el volumen de los datos complica su evaluación. Habíamos sintetizado los versos en gráficos, que resultan manejables, por lo menos cuando tomamos solo cinco ritmos; la situación es bien distinta a medida que empezamos a manejar más, como en la Figura 6. En este caso, no recurriremos a representaciones visuales para salvar el obstáculo, sino a modelos matemáticos capaces de evaluar todas las variables.

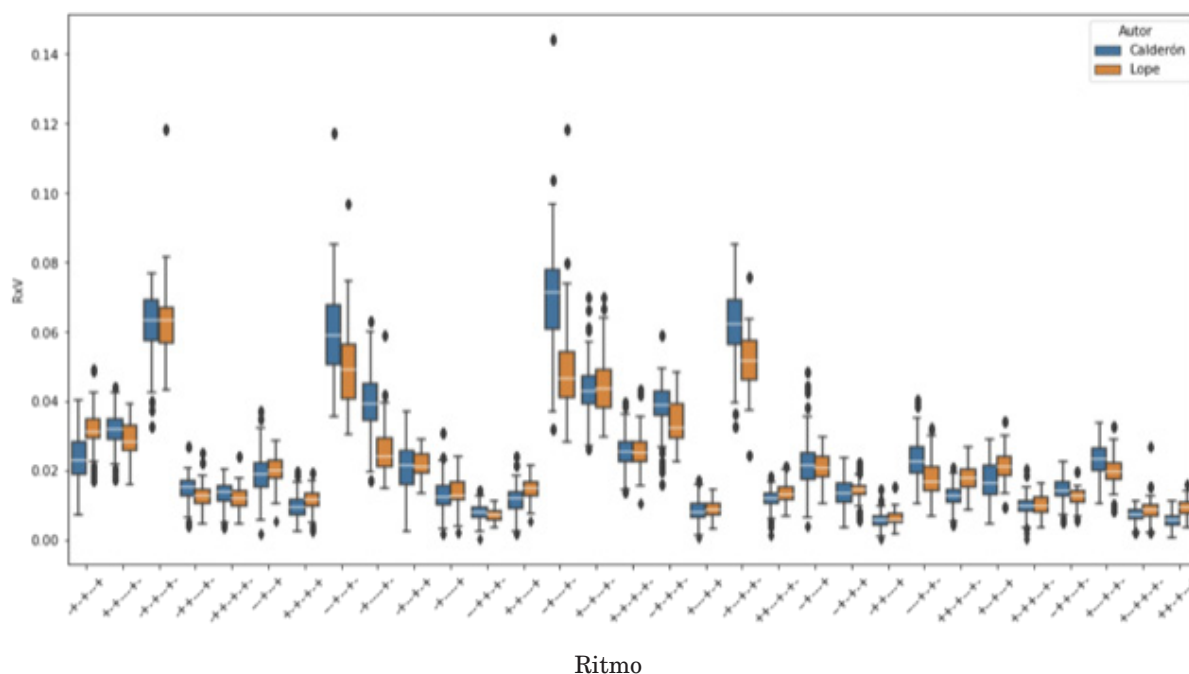


Figura 6. Distribuciones de las proporciones de ritmos por obra.

En concreto, es posible probar la validez del criterio de las diferentes frecuencias de los octosílabos también mediante *machine learning*. Limitamos para eso el corpus de entrada a 37 comedias calderonianas y otras 23 lopescas. Comenzamos a comprobar si el programa es capaz de identificar obras de autoría segura como de su autor tomándolas de una en una. Probamos varios métodos, y la regresión logística dio los mejores resultados, pues hizo solo una falsa atribución en las comedias de Calderón (*El astrólogo fingido*) y dos en las de Lope (*El perro del hortelano* y *El cardenal de Belén*) comparando solo diez ritmos (++-+---+, +-+---+, -+---+--, +-+---+, ++++---+, +-+---+, +-+---+, +-+---+, +-+---+, +-+---+, +-+---+, +-+---+). Dicho de otra manera, el procedimiento adjudica la autoría a Calderón y Lope con márgenes de error del 2,7 % y del 4,3 %. No debe sorprendernos la diferencia, pues Calderón parece dejar una huella autorial en los ritmos más marcada de lo habitual (Kroll y Sanz-Lázaro 2022).

Pero, al igual que los colegas de la estilometría léxica no se quedaron en la comparación de dramaturgos, nosotros también elevaremos la mirada más allá de las disputas de autoría. ¿Podemos, por ejemplo, encontrar diferencias si comparamos géneros literarios o tipos de personaje?

GÉNERO LITERARIO

El teatro del Siglo de Oro mezcla elementos trágicos y cómicos, lo que cristaliza en la obra de Lope como en pocas y, de esta forma, da lugar a una extraordinaria variedad de géneros y subgéneros. A pesar de esta enorme diversidad, podemos simplificar quizá esta riqueza en obras religiosas, obras serias y obras cómicas. Compararemos, pues, a continuación autos sacramentales, comedias cómicas y obras trágicas de Lope, siguiendo para ello la clasificación genérica que proponen los editores de ARTELOPE.⁷ Veamos, para empezar, un gráfico que visualiza la frecuencia de los octosílabos por obra, comparando dramas, tragedias y comedias.

7. Con la excepción de *El castigo sin venganza*. Ya que Lope la tildaba en dos ocasiones como tragedia, no nos cabe duda de que le importaba esta distinción. En ARTELOPE la clasifican como un drama imaginario: <<https://artelope.uv.es/basededatos/browserecord.php?action=browse&-recid=71#caracterizaciones>>, consulta del 2 de septiembre de 2022. Ver al respecto también el artículo de Florence d'Artois [2012].

En tanto que *El castigo sin venganza* es la única obra del género trágico en nuestro corpus, solo se representa en el gráfico con una rayita horizontal entre los dramas y las comedias en la Figura 7. Esta aún en su unidimensionalidad todos

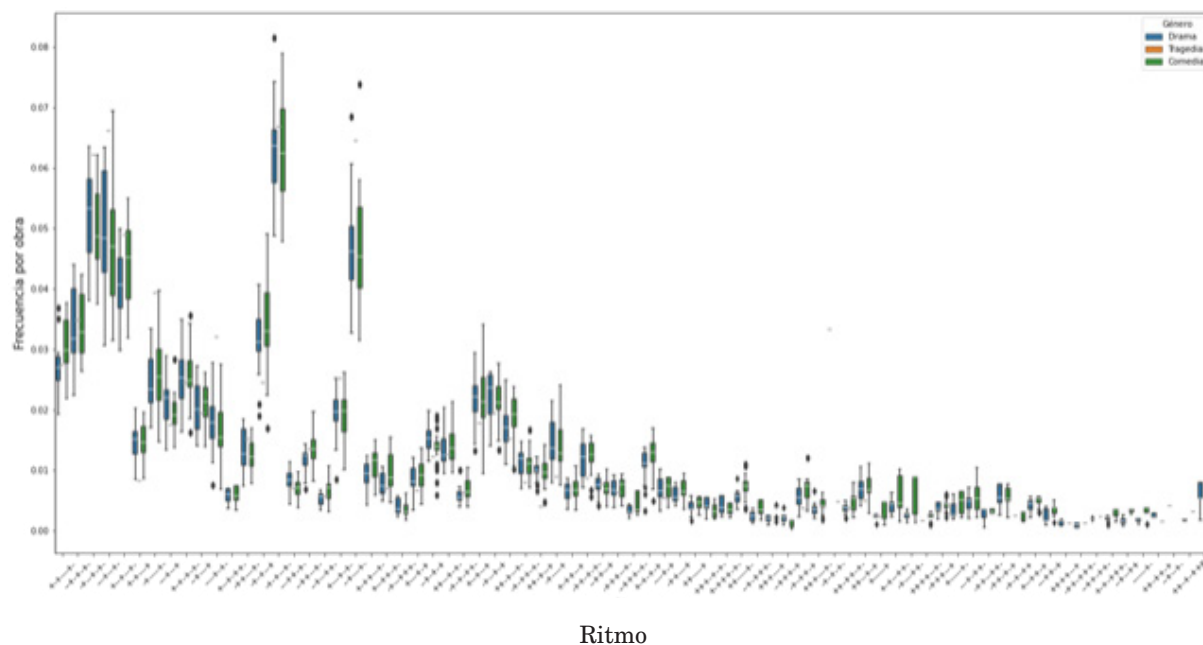


Figura 7. Distribuciones de las proporciones de ritmo por obra.

los percentiles de una distribución monomuestral, al contrario que otros géneros, de los que nuestro corpus anda mejor provisto. Llama la atención la ausencia de diferencias notables entre los dramas y las comedias. *El castigo sin venganza*, sin embargo, parece destacar. En casi todos los ritmos se aleja de forma perceptible de la mediana del uso del ritmo en comedias y dramas, incluso de los cuartiles centrales.

El gráfico de la Figura 8 excluye los dramas, de manera que permite apreciar mejor la distancia entre texto trágico y cómico. Como explica Alejandro García-Reidy en su nota al prólogo de *El castigo sin venganza*, Lope usaba poco esta asignación genérica para sus obras. Él mismo parece que sabía que la intensidad de las emociones y el final funesto (tomo prestadas estas formulaciones de García-Reidy 2009: 75) distinguen esta obra de muchas otras de su vasto caudal de comedias. De momento, lo formulamos como una hipótesis de trabajo, pero hay indicios de cierto peso que apuntan a que esta intensidad afectiva también se vertía en una estructu-

ra rítmica diferente, tanto de las obras cómicas como de las serias que, a pesar de su seriedad, no alcanzan el grado de intensidad dramático de *El castigo sin venganza*. Esta auténtica tragedia española destaca en muchos sentidos, y también lo hace en su uso diferente del octosílabo.

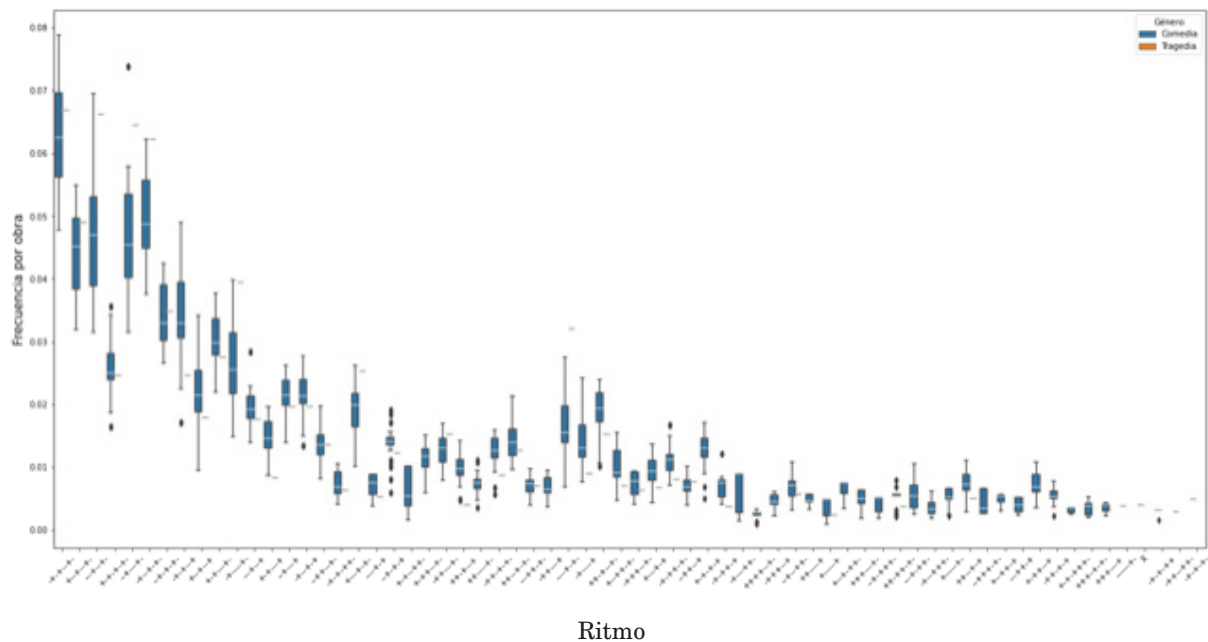


Figura 8. Distribuciones de las proporciones de ritmo por obra en comedias y tragedia. La tragedia se muestra como una línea al concentrarse todos los cuartiles en una sola obra.

Quedaría por ver el género del auto sacramental. Para una mejor legibilidad comparamos en la Figura 9 los autos solo con obras cómicas. A la luz de lo que muestra el gráfico, cabe concluir que hay una diferencia entre los dos géneros en varios octosílabos, especialmente en los más frecuentes. Esto quiere decir que los diferentes octosílabos se usan con frecuencias distintas según sea la obra un auto o una comedia. O, dicho de otra forma, el autor adecua su estilo rítmico al género de la obra.

En resumidas cuentas, las observaciones sugieren que hay géneros —como el del auto sacramental— y obras —como *El castigo sin venganza*— singulares que se caracterizan por su uso peculiar de determinados octosílabos. Cabe plantearse si esto responde no solo a una cuestión estructural sino a elementos temáticos, por lo que las relaciones concretas entre afectos y ciertos octosílabos serán objeto de futuros trabajos. Consideremos por ahora qué diferencias rítmicas pueden darse dentro de una obra.

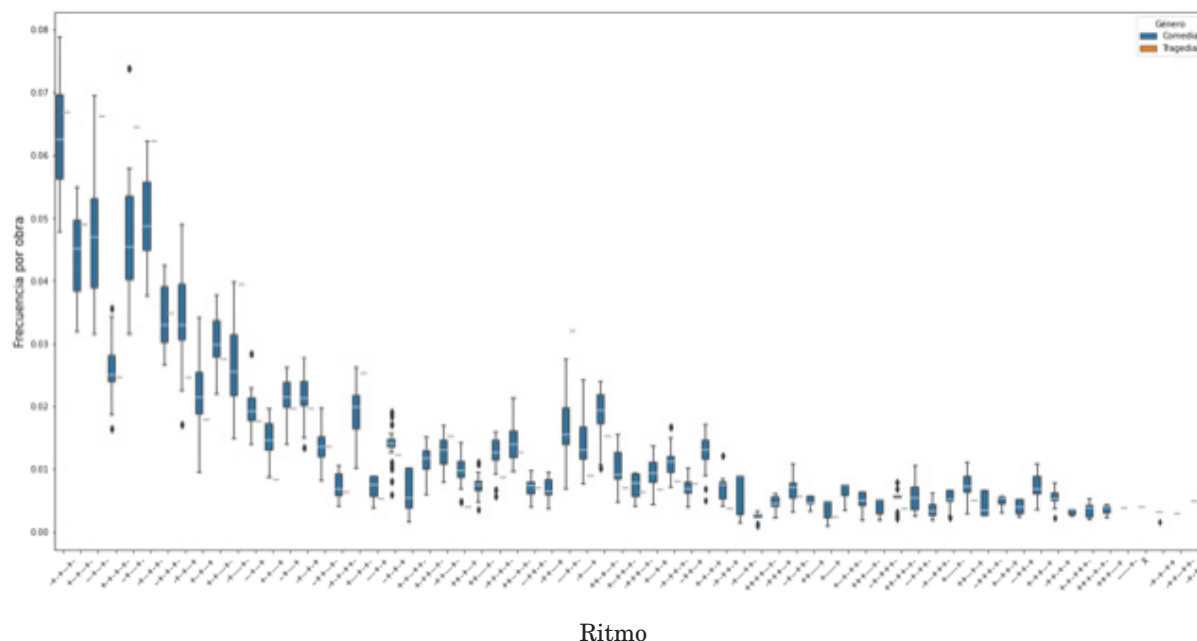


Figura 9. Distribuciones del porcentaje de ritmos por obra en autos y comedias.

PERSONAJES

De la misma manera que comparamos los diferentes géneros literarios del teatro del Siglo de Oro, podemos plantear si se aprecian diferencias entre personajes masculinos y femeninos, sea en una obra concreta, o de forma general a lo largo de la producción de un autor. Proponemos analizar de entrada una sola obra como ejemplo para no sacrificar la claridad por la cantidad. Elegimos para este propósito *El castigo sin venganza*, ya que esta obra cuenta con dos papeles protagonistas de rango parecido pero sexos opuestos. Se trata de Federico y Casandra: él es hijo ilegítimo, pero heredero de un potentado; ella es mujer, pero esposa de dicho potentado, el duque de Ferrara, y por tanto madrastra de Federico.

Como se observa en la Figura 10, hay diferencias rítmicas entre los discursos de Federico y de Casandra. Es cierto que no hay una disparidad notable en la mayoría de los ritmos de los octosílabos producidos por los personajes de ambos sexos, no obstante, encontramos ritmos que utiliza uno u otro personaje en exclusiva. También se observa que Casandra pronuncia más del doble de veces que Federico versos con ritmo $-+++---+-$, mientras que este usa $--+--+$ alrededor del doble de veces que ella.

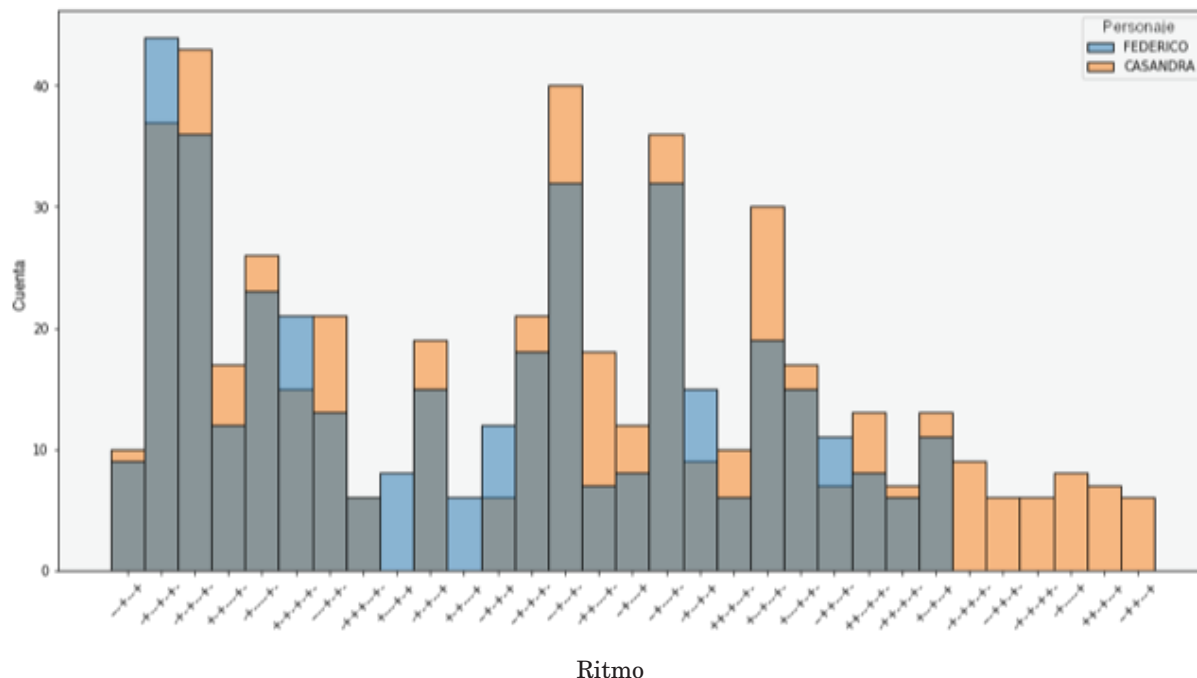


Figura 10. Recuento de ritmos por personaje. Las zonas grises corresponden al solapamiento entre ambos personajes.

Ahora tratemos de contextualizar estos resultados en la obra de Lope. Para ello, repetiremos el procedimiento, pero, esta vez, tomaremos todas las obras que tenemos de Lope y compararemos todos los versos pronunciados por mujeres frente a todos los versos declamados por varones.

Apenas un vistazo basta para deducir que hay un dominio masculino en los versos recitados, lo que no resulta extraño: sin pretender profundizar en este artículo en el tema, resulta evidente que los personajes masculinos hablan mucho más que los femeninos. También es cierto que hay considerablemente más personajes masculinos que femeninos. En cualquier caso, siempre hay que tener en cuenta que los resultados son provisionales, ya que todavía se basan en una selección de obras pequeña. A pesar de ello, el gráfico (ver fig. 11) resulta en gran medida representativo, pues reúne prácticamente todos los géneros teatrales de la época y también obras con un protagonismo femenino destacado. Sea como fuere, la diferencia numérica es aplastante. En el corpus que estamos analizando hay aproximadamente el doble de personajes masculinos que femeninos. Comparando las Figuras 10 y 11, parece confirmarse la mayor tendencia de algunos patrones rítmicos en los dos sexos. El 1-3-7, reservado a Federico

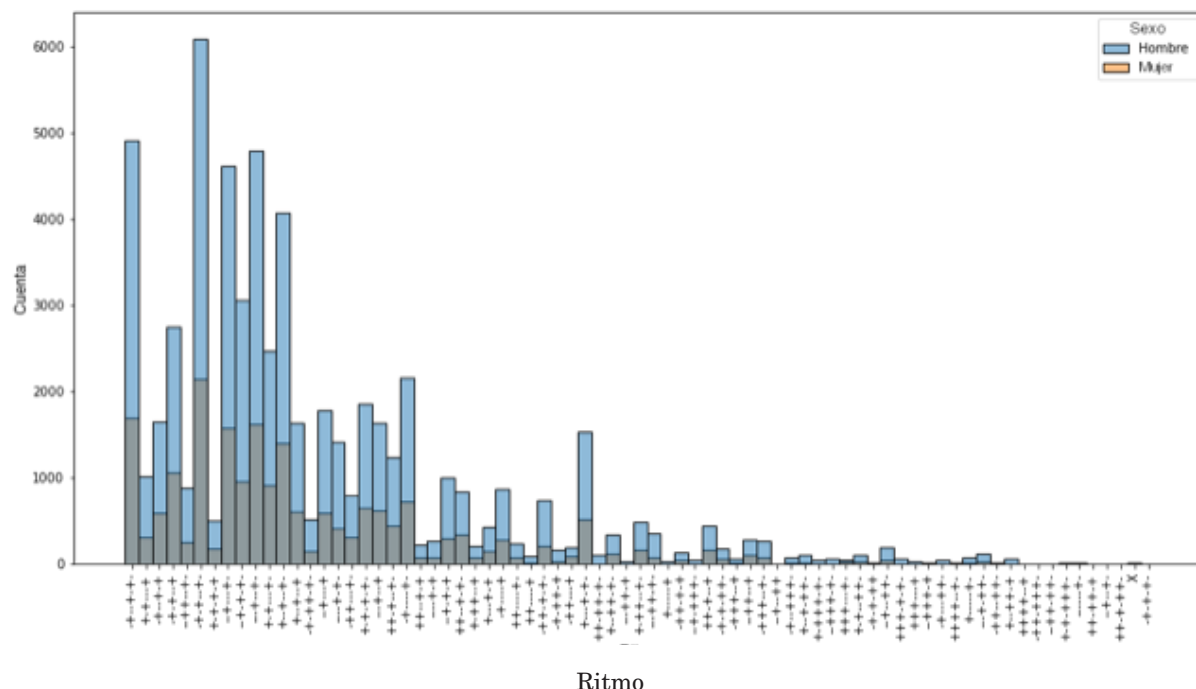


Figura 11. Recuento de ritmos en todo el corpus de Lope por sexo.

en *El castigo sin venganza*, también resulta un verso de mayor frecuencia masculina en toda la obra lopesca.

Visto así, llama la atención la casi equivalencia en el número de octosílabos entre Casandra y Federico en *El castigo sin venganza*. Casandra es una mujer muy dominante en la segunda parte de la obra e incluso llega a reprocharle a Federico su presunta debilidad. Si miramos el momento culminante de la relación entre ambos —Federico decide dejar a Casandra y ella le echa en cara su cobardía— es llamativo que Federico no se sirva ni una vez del patrón 2-4-7, mientras que Casandra lo aplique precisamente en el momento de recriminarle su pusilanimidad:

¡Oh, *cobarde*, mal nacido!
 Las lágrimas y los ruegos,
 hasta hacernos volver locas
 robando las honras nuestras
 —que de las traiciones vuestras
 cuerdas se libraron pocas—,
 ¿agora son *cobardías*?
 (vv. 2730-2736, las cursivas son nuestras)

No obstante, si miramos la preferencia de todos los personajes por este patrón, observaremos que Federico y Casandra emplean este ritmo de manera parecida (fig. 12).

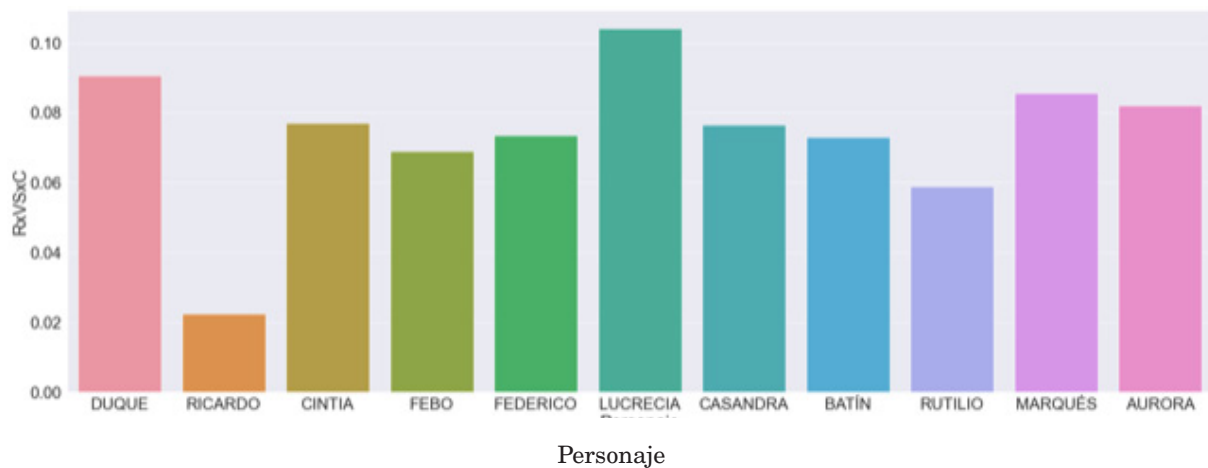


Figura 12. Preferencia por el patrón -+--+--+ en los octosílabos.

CONCLUSIÓN

Hemos revisado en este artículo la posibilidad de ampliar los estudios estilométricos basados en el léxico mediante la introducción de métodos de análisis métrico y aplicarlo al estudio del ritmo. Se ha visto que hay diferencias autoriales palpables en la frecuencia de los diferentes patrones del octosílabo. Estas diferencias pueden aprovecharse para aplicar procedimientos de inteligencia artificial: entrenar a la máquina con los datos rítmicos para clasificar las obras. La regresión logística ofrece resultados interesantes, pero obliga a preparar bien el corpus, es decir, trabajar con un corpus bien balanceado y con obras de un solo género.

Asimismo, hemos visto que, al menos en Lope, existen diferencias en el ritmo que atañen a diferentes niveles de su producción dramática. En el conjunto de su obra —la parte de la que disponemos—, hemos observado diferencias entre géneros literarios, pero también lo hemos hecho dentro de una obra como elemento individual cuando hemos comparado personajes. Este último punto, sin embargo, no parece presentar todavía resultados muy concluyentes. Quedan pues por emprender en el futuro trabajos sobre el ritmo y el género de los personajes.

Podemos extraer dos conclusiones de todo ello. En primer lugar, la evaluación estadística de elementos rítmicos abre la puerta a la *lectura distante* de las obras. Este tipo de análisis, entre otros muchos usos, tiene potencial para constituirse como método complementario a los estudios estilométricos léxicos *clásicos*. En segundo lugar, más allá de los cálculos estadísticos y los grandes corpus, el procesamiento digital del metro proporciona una potente herramienta para extraer y presentar información sobre aspectos muy concretos de textos en verso de cierto tamaño, como hemos visto en las observaciones a *El castigo sin venganza*. Esta información puede ser co-tejada a simple vista con otros elementos de la obra e interpretada de acuerdo a su contexto filológico, algo que, de otra manera, requeriría un tedioso proceso manual de escansión y clasificación antes de poder evaluar los datos así obtenidos.

BIBLIOGRAFÍA

- CUÉLLAR, Álvaro, y Germán VEGA GARCÍA-LUENGOS. *ETSO. Estilometría aplicada al Teatro del Siglo de Oro*, 2017-2022, en línea, <<http://etso.es/>>. Consulta del 12 de noviembre de 2022.
- CUÉLLAR, Álvaro, y Germán VEGA GARCÍA-LUENGOS. *Examen de autorías en el teatro del Siglo de Oro. Propuestas desde la estilometría*, Reichenberger, Kassel (en prensa).
- D'ARTOIS, Florence, «La tragedia lopesca en los años 1620-1630: ¿de un paradigma a otro?», *Mélanges de la Casa de Velázquez. Nouvelle série*, 42 1 (2012), pp. 95-119.
- EDER, Maciej, Jan RYBICKI y Mike KESTEMONT, «Stylometry with R: a package for computational text analysis», *R Journal*, VIII 1 (2016), pp. 107-121.
- GARCÍA-REIDY, Alejandro, «Prólogo», en Lope de Vega, *El castigo sin venganza*, ed. A. García-Reidy, Crítica, Barcelona, 2009, pp. 7-76.
- KROLL, Simon, y Fernando SANZ-LÁZARO, «Romances teatrales entre Mira de Amescua, Calderón y Lope: ritmo, asonancia y cuestiones de autoría», *Revista de Humanidades Digitales*, 7 (2022), pp. 1-18.
- MORETTI, Franco, «Graphs, maps, trees. Abstract Models for Literary History (2)», *New Left Review*, 26 (2004), pp. 79-103.
- MORLEY, Sylvanus Griswold, y Courtney BRUERTON. *Cronología de las comedias de Lope de Vega: con un examen de las atribuciones dudosas, basado todo ello en un estudio de su versificación estrófica*, Gredos, Madrid, 1968.
- OLEZA, Joan, dir., *ARTELOPE. Base de datos y argumentos del teatro de Lope de Vega*. 2011-2022, en línea, <<https://artelope.uv.es>>. Consulta del 12 de noviembre de 2022.
- SANZ-LÁZARO, Fernando, *Fonemas* (version 1.0.0), GitHub, 2021a, en línea, <<https://github.com/fsanzl/fonemas>>. Consulta del 12 de noviembre de 2022.
- SANZ-LÁZARO, Fernando, *Silabeador* (version 1.0.2), GitHub, 2021b, en línea, <<https://github.com/fsanzl/silabeador>>. Consulta del 12 de noviembre de 2022.
- SANZ-LÁZARO, Fernando, *LibEscansión* (version 1.0.0), GitHub, 2022, en línea, <<https://github.com/fsanzl/libescansion>>. Consulta del 12 de noviembre de 2022.
- VEGA CARPIO, Lope de, *El castigo sin venganza*, ed. A. García-Reidy, Crítica, Barcelona, 2009.
- VEGA GARCÍA-LUENGOS, Germán, «Las comedias de Lope de Vega: confirmaciones de autoría y nuevas atribuciones desde la estilometría (I)», *Talía. Revista de estudios teatrales*, 3 (2021), pp. 91-108.

APÉNDICE

Lista de obras de Lope

Adonis y Venus

Auto del pan y el palo

Boda entre dos maridos

El amigo por fuerza

El anzuelo de Fenisa

El asalto de Mastrique por el príncipe de Parma

El cardenal de Belén

El casamiento en la muerte

El castigo sin venganza

El cerco de Santa Fe

El cuerdo en su casa

El galán Castrucho

El genovés liberal

El heredero del Cielo

El hijo pródigo

El laberinto de Creta

El lacayo fingido

El letrado (entremés)

El mayorazgo dudoso

El mayordomo de la Duquesa de Amalfi

El mejor alcalde, el rey

El misacantano

El niño pastor

El nombre de Jesús

El nuevo mundo descubierto por Cristóbal Colón

El perro del hortelano

El primero Benavides

El santo negro Rosambuco, de la ciudad de Palermo

El testimonio vengado

El tirano castigado

El viaje del alma
El villano, en su rincón
Fuenteovejuna
La amistad pagada
La batalla del honor
La bella malmaridada o la cortesana
La dama boba
La escolástica celosa
La fe rompida
La maya
La noche toledana
La ocasión perdida
La privanza del hombre
La resistencia honrada y Condesa Matilde
La viuda valenciana
Las bizarrías de Belisa
Las bodas entre el Alma y el Amor Divino
Las ferias de Madrid
Las hazañas del segundo David
Las mudanzas de fortuna y sucesos de don Beltrán de Aragón
Laura perseguida
Los acreedores del hombre
Los comendadores de Córdoba
Los donaires de Matico
Los locos de Valencia
Los torneos de Aragón
Peribáñez y el Comendador de Ocaña
Santiago el verde

Lista de obras de Calderón

A Dios por razón de Estado
A María el corazón

A secreto agravio, secreta venganza
Afectos de odio y amor
Agradecer y no amar
Amigo, amante y leal
Amor, honor y poder
Andrómeda y Perseo
Apolo y Climene
Argenis y Poliarco
Basta callar
Bien vengas, mal si vienes solo
Cada uno para sí
Casa con dos puertas mala es de guardar
Celos aun del aire matan
Cómo se comunican dos estrellas contrarias
¿Cuál es mayor perfección, hermosura o discreción?
Dar tiempo al tiempo
De una causa, dos efectos
Dicha y desdicha del nombre
Eco y Narciso
El Faetonte
El José de las mujeres
El Maestrazgo del Tusón
El agua mansa
El alcalde de Zalamea
El arca de Dios cautiva
El astrólogo fingido
El año sancto de Roma
El año santo en Madrid
El conde Lucanor
El cordero de Isaías
El cubo de la Almudena
El diablo mudo
El divino Orfeo
El divino Orfeo (segunda versión)

El divino cazador
El día mayor de los días
El encanto sin encanto
El galán fantasma
El gran mercado del mundo
El gran príncipe de Fez
El gran teatro del mundo
El hombre pobre todo es trazas
El indulto general
El laberinto del mundo
El lirio y el azucena
El maestro de danzar
El mayor encanto, amor
El mayor monstruo del mundo
El mágico prodigioso
El médico de su honra
El nuevo hospicio de pobres
El nuevo palacio del Retiro
El orden de Melquisedec
El pastor Fido
El pintor de su deshonra
El pleito matrimonial
El primer blasón del Austria
El primer refugio del hombre y probática piscina
El príncipe constante
El purgatorio de san Patricio
El sacro Pernaso
El santo rey don Fernando (segunda parte)
El santo rey don Fernando (primera parte)
El secreto a voces
El segundo Scipión
El segundo blasón del Austria
El sitio de Bredá
El socorro general

El valle de la Zarzuela
El veneno y la triaca
El verdadero Dios Pan
El viático cordero
El árbol de mejor fruto
Fortunas de Andrómeda y Perseo
Hado y divisa de Leonido y Marfisa
Judas Macabeo
La aurora en Copacabana
La banda y la flor
La cena del rey Baltasar
La cura y la enfermedad
La dama duende
La dama y el galán Aquiles (El monstruo de los jardines)
La desdicha de la voz
La devoción de la Cruz
La devoción de la misa
La divina Filotea
La exaltación de la cruz
La gran Cenobia
La hidalga del valle
La hija del aire (parte primera)
La hija del aire (parte segunda)
La humildad coronada
La iglesia sitiada
La inmunidad del sagrado
La lepra de Constantino
La nave del mercader
La piel de Gedeón
La primer flor del Carmelo
La protesta de la fe
La puente de Mantible
La redención de cautivos
La selva confusa

La semilla y la cizaña
La serpiente de metal
La siembra del Señor (Los obreros del señor)
La torre de Babilonia
La universal redención
La vacante general
La vida es sueño
La vida es sueño (primera versión)
La vida es sueño (segunda versión)
La viña del Señor
Lances de amor y fortuna
Las espigas de Ruth
Las tres justicias en una
Las órdenes militares
Llamados y escogidos.
Lo que va del hombre a Dios
Loa para la comedia de Los tres mayores prodigios
Loa sacramental de El primer refugio del hombre o la probática piscina
Los alimentos del hombre
Los cabellos de Absalón
Los empeños de un acaso
Los encantos de la culpa
Los misterios de la misa
Los tres mayores prodigios
Los tres mayores prodigios Loa
Mañana será otro día
Mística y real Babilonia
No hay burlas con el amor
No hay cosa como callar
No hay instante sin milagro
No hay más Fortuna que Dios
Peor está que estaba
Primero y segundo Isaac
Psiquis y Cupido (Madrid)

Psiquis y Cupido (Toledo)

¿Quién hallará mujer fuerte?

Saber del mal y el bien

Sueños hay que verdad son

Tu prójimo como a ti (primera versión)

Un castigo en tres venganzas