

#23

GRAFOS DE ESCENAS Y ESTUDIOS LITERARIOS DIGITALES: UNA PROPUESTA COMPUTACIONAL CRÍTICA

Concepción-María Jiménez-Fernández

Universidad Internacional de La Rioja (UNIR)

José Calvo Tello

Göttingen State and University Library (Alemania)

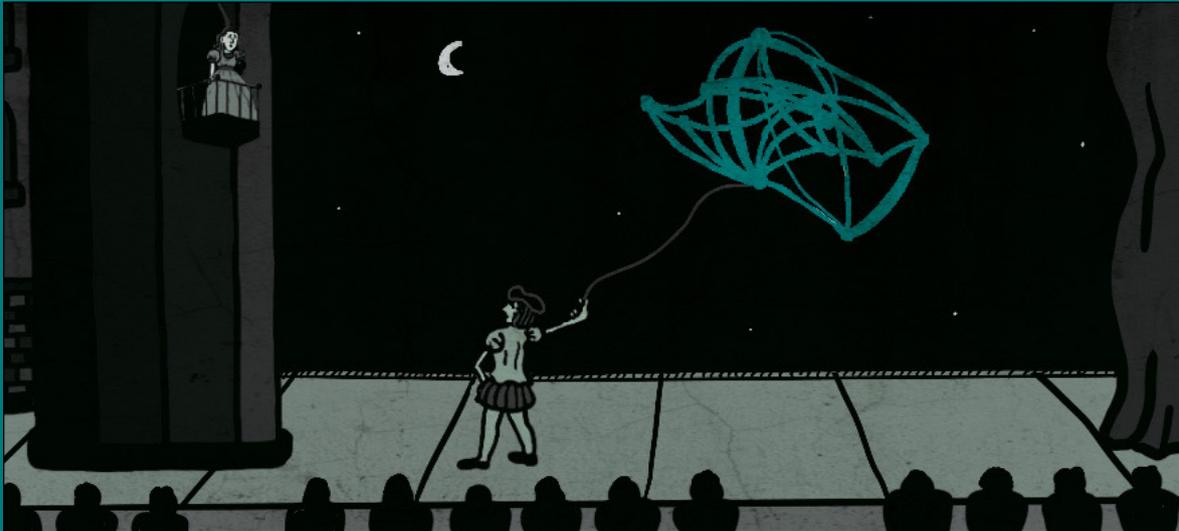
Ilustración || Andrés Müller

Artículo || Recibido: 13/02/2020 | Apto Comité Científico: 04/06/2020 | Publicado: 07/2020

DOI: [10.1344/452f.2020.23.4](https://doi.org/10.1344/452f.2020.23.4)

concepcionm.jimenez@unir.net / jose.calvo@uni-wuerzburg.de

Licencia || Reconocimiento-No comercial-Sin obras derivadas 3.0 License



Resumen || Desde que en 2005 Moretti los popularizase, los grafos se han utilizado para representar la interacción de los personajes en obras de teatro y novelas. Sin embargo, en los últimos años han aparecido algunos trabajos evaluando estos grafos, cuestionando su validez. En este trabajo proponemos una modificación de estos grafos, obteniendo una representación de las obras dramáticas que prioriza la estructura en escenas. A diferencia de otros trabajos de Humanidades Digitales, realizamos una evaluación crítica mediante hipótesis que aúna variables cuantitativas y cualitativas.

Palabras clave || Grafos | Obras de teatro | Evaluación estadística | Edad de Plata

Abstract || Since their popularization by Moretti in 2005, graphs have been used to represent the interaction of characters in plays and novels. However, in recent years some works have evaluated these graphs, questioning their validity. In this research, we propose a modification of these graphs, obtaining a representation of the dramatic works that prioritizes the structure in scenes. Unlike other DH works, we carry out a critical evaluation by means of hypotheses that combine quantitative and qualitative variables.

Keywords || Graphs | Theatre plays | Statistical evaluation | Silver age

Resum || Des que en 2005 Moretti els popularitzés, els grafs s'han utilitzat per a representar la interacció dels personatges en obres de teatre i novel·les. No obstant això, en els últims anys han aparegut alguns treballs avaluant aquests grafs, qüestionant la seva validesa. En aquest treball proposem una modificació d'aquests grafs, obtenint una representació de les obres dramàtiques que prioritza l'estructura en escenes. A diferència d'altres treballs d'Humanitats Digitals, realitzem una avaluació crítica mitjançant hipòtesis que conjumina variables quantitatives i qualitatives.

Paraules clau || Grafs | Obres de teatre | Avaluació estadística | Edat de Plata

0. Críticas a las Humanidades Digitales

Las respuestas de muchos sectores de la sociedad a la tecnología tienden a situarse en dos extremos: la fobia y la filia. Dependiendo de quién realice el discurso y de qué rincón de la sociedad proceda, es probable que la tecnología aparezca como la salvación o como la destrucción de la civilización. Tendencias similares se observan en los ámbitos académicos, más concretamente en las Humanidades. Hace una década, la aplicación de metodologías informáticas a objetos culturales era algo absolutamente marginal, llevado a cabo principalmente por lingüistas de corpus o computacionales, informáticos interesados en cultura o aficionados al margen del mundo académico. Ese ámbito ha pasado a recibir el nombre de *Humanidades Digitales* (HD) y se ha establecido como una disciplina atractiva que se ha desarrollado académicamente mediante revistas, congresos regulares, asociaciones (ADHO, HDH, Red de México, Umanística, DHd, Humanum, etcétera), proyectos de infraestructura (DARIAH, CLARIN, actualmente en proceso de fusión como CLARIAH), etcétera.

En el marco internacional, se observa una constante profesionalización del área, con la aparición de diferentes carreras y programas de doctorado. Al margen de las nuevas posibilidades que la informática aporta a la investigación, creemos que este proceso ha sido acelerado por la percepción de que participar en esta disciplina digital podría traer ventajas profesionales. De esta manera, las universidades creen poder atraer más estudiantes si ofrecen cursos en cuyos títulos aparecen la palabra *digital*, los grupos de investigación aumentarían las probabilidades de obtener fondos, y en general lo digital resulta interesante para amasar méritos académicos en la carrera por una posición más estable o mejorar la empleabilidad de humanistas en empresas.

Este proceso ha despertado reacciones críticas por parte de algunos humanistas, dentro y fuera de las HD. Estas críticas tienden a utilizar principalmente tres tipos de argumentos. El primer grupo, más nutrido, contiene trabajos que cuestionan el marco epistemológico de las HD y su capacidad de llegar a conclusiones aceptables (Priani, 2015; Fiorimonte, 2018; Sanz, 2018).

En segundo lugar, se encuentra un grupo de trabajos críticos que utilizan argumentos axiomáticos señalando por ejemplo una supuesta posición política de la disciplina entre lo conservador y lo neoliberal (Allington *et al.*, 2016; Puentes, 2018).

A pesar de que ambos tipos de críticas son frecuentes en congresos de HD, ambos disienten en la metodología para realizar su crítica. Es decir, aunque las HD se asientan en buena medida en medios

computacionales para realizar su análisis, los dos tipos de crítica vistos hasta ahora no aplican metodologías informáticas para analizar la disciplina. De esta manera, se establece un diálogo crítico en el que ambas posiciones no comparten ni sus conclusiones (la aceptabilidad de las HD) ni su metodología.

Sin embargo, hay una tercera manera de criticar las HD que se basa principalmente en aspectos metodológicos. Este tipo de crítica comparte los medios informáticos utilizados en el área para examinar la disciplina o trabajos específicos de ella. Es decir, esta posición sí acepta los medios computacionales como herramienta de investigación.

Recientemente, Da (2019) ha publicado un artículo que ha tenido una notable repercusión en círculos de investigación. En él, la autora realiza una feroz crítica contra los estudios literarios digitales, diseccionando los trabajos de algunos de los investigadores más importantes de Estados Unidos y Canadá (Underwood, Jockers, Piper o Algee-Hewit). Su principal argumento es que muchos estudios de HD aplicados a literatura no aplican test estadísticos. Además, señala que muchos de los trabajos llegan a resultados negativos (en inglés *no results*), es decir, resultados que desmienten la hipótesis planteada o que no llegan a resultados concluyentes. En particular, Da critica la utilización de frecuencias de palabras (*tokens*), así como la utilización de técnicas de aprendizaje automático no supervisado (como *principal component analysis* o *topic modelling*). Este artículo ha recibido una respuesta excepcional y hasta cierto punto ofensiva por parte de los autores criticados, pero también por otros (Herrmann *et al.*, 2020; Jannidis, 2020; o Piper, 2020).

No compartimos con Da la crítica categórica de que la literatura no puede ser objeto de análisis de métodos computacionales. Ella misma acepta que trabajos de HD aplicados a otras áreas pueden ser de interés. La autora no explica cuál es la particularidad que hace que para la investigación en docenas de áreas puedan aplicarse ordenadores, pero esta opción quede excluida para textos literarios. Sin embargo, sí compartimos su crítica de que muchos trabajos de HD no sostienen sus conclusiones con test estadísticos. En nuestra opinión, muchos trabajos del área se han centrado en probar herramientas que llevaban algunos años o décadas funcionando en la informática, aplicarlo a conjuntos de datos humanísticos, explorar de manera muy selectiva ciertos resultados, mostrar visualizaciones espectacularmente misteriosas y concluir reflexionando sobre su futuro potencial. Hasta cierto punto, el objetivo de estos trabajos es convencer a otros humanistas del interés y potencial de estos métodos.

Un aspecto notablemente débil sobre el trabajo de Da es su crítica a los artículos de resultados negativos. En algunas disciplinas, especialmente en medicina, este tipo de artículos resultan la norma. Es importante saber qué nuevas medicinas o tratamientos tienen un efecto positivo, pero también es necesario saber lo que no funciona para poder descartarlo. En las HD varios trabajos de estilometría han mostrado resultados negativos, como los realizados por De la Rosa y Suárez (2016) sobre el *Lazarillo* y de Reißler-Pipka (2018) sobre el *Quijote* de Avellaneda. Ambos trabajos no llegan a concluir de manera contundente sobre quién habría sido el autor de ambas obras. Sin embargo, el hecho de que un análisis no llegue a conclusiones sólidas no debe ser un argumento en contra de su publicación. Los resultados negativos son señales para el resto de la comunidad de que ese es un callejón sin salida, al menos con los actuales datos, métodos e ideas. Si el sistema académico impidiese la publicación de resultados negativos, De la Rosa no habría publicado su trabajo y cada departamento de Hispanística estaría probando análisis similares sobre el *Lazarillo* para llegar a los mismos resultados negativos.

Otro aspecto débil del artículo de Da es el sesgo en la selección de los estudios criticados. La autora analiza exclusivamente investigadores masculinos de Estados Unidos y Canadá que trabajan en corpus de miles de textos. Numerosos trabajos provenientes de otras áreas como la estilometría han aportado luz sobre preguntas relevantes (Burrows, 2002; Juola, 2013; Kestemont *et al.*, 2016), también para el español (Cerezo Soler y Calvo Tello, 2019; Hernández, 2018). Algunos centros de HD en Europa han publicado análisis literarios computacionales con fuertes metodologías estadísticas (Hettinger *et al.*, 2016; Evert *et al.*, 2017; Koolen, 2018; o van Cranenburgh *et al.*, 2019).

Sin embargo, creemos que este artículo puede funcionar como piedra de toque y producir una reacción constructiva en las HD. Eso es lo que proponemos con este escrito. Creemos que la estructura de los artículos no debe terminar en la propuesta y exploración de nuevas metodologías, sino que su evaluación necesita ser un aspecto central. Cuando los investigadores proponen una nueva metodología o modelización de los datos, deben en primer lugar convencer al resto de la comunidad de que su propuesta no produce exclusivamente ruido. Esta evaluación no debería limitarse a uno o dos casos, sino aplicarse a la totalidad de los datos disponibles. Además, los resultados de la propuesta deben ser contrastados con anotaciones cualitativas, en el mejor de los casos realizadas por humanistas que sean especialistas en el área. De esta manera, sugerimos que los investigadores propongan hipótesis operacionalizables que deben evaluarse estadísticamente contra variables cualitativas en

el conjunto total de los datos. Solo cuando la evaluación aporta resultados positivos, el investigador puede pasar a la interpretación.

En concreto, en este artículo proponemos una nueva formalización de obras de teatro en forma de grafo. En los últimos años numerosos trabajos han realizado similares transformaciones tanto en teatro como en novelas, colocando los personajes como centro del análisis, como describiremos en las siguientes secciones. La formalización que proponemos modifica su centro de interés y coloca la estructura textual de escenas en principal lugar. Esta se complementa con una evaluación en la que asumimos, como uno de los posibles resultados, que en nuestra propuesta no hayamos condensado de la manera que esperábamos la información sobre la obra de teatro. Este rechazo de las hipótesis de trabajo se comprueba mediante test estadísticos que comparan la anotación cualitativa de la obra. De esta manera, integramos en nuestro análisis un punto de vista crítico que combina test estadísticos realizados sobre el corpus completo, así como información cualitativa sobre la estructura de la obra y sus personajes.

2. Grafos y Humanidades Digitales

Dos aspectos que se han desarrollado en las HD son la innovación en visualización y la formalización de la interacción entre grupos de datos. Hasta el momento, y como advierten (Medrano, Alonso y Figuerola, 2011) estos esfuerzos se han centrado en una descripción estática de los datos (tablas, gráficos de tarta, histogramas, etc.). Sin embargo, estas técnicas no representan correctamente la interacción dinámica de información (es decir, agrupaciones de datos y sus relaciones entre sí) como en el caso de la distribución de personajes en escenas de una obra.

La llamada teoría de grafos posibilita agrupar datos definiendo relaciones entre nodos. Como es sabido, gracias a los grafos «el usuario se mueve por la información que se muestra a través de enlaces» lo que posibilita su experiencia lectora haciéndola más rica «gracias a la formación de un grafo por el que se pueda *navegar* siguiendo los enlaces entre sus nodos» (Palacios, García y Granollers, 2015). Cabría destacar que esta área de investigación, proveniente de las matemáticas, tiene sus primeros orígenes en el siglo XVIII (Murray y Tillett, 2011). Así, un grafo sería un conjunto de nodos (normalmente representados como puntos) y aristas (relaciones entre esos nodos) que representan objetos relacionados entre sí (Zweig, 2016). Esta estructura puede representarse gráficamente lo que posibilita un novedoso tipo de interacción entre el investigador y los datos (de cantidad muy variable) para describirlos y extraer conclusiones (Medrano, Alonso y Figuerola, 2011). Así los grafos se

pueden utilizar para representar las relaciones entre los diferentes actores (término acuñado por sociólogos) o elementos. Por su lado, la biología ha explorado procesos celulares como el metabolismo a través de esas *redes* o estructuras; la química ha utilizado esta metodología para la representación molecular y sus reacciones; la economía la usa para estudios de mercado en relación con efectos de la globalización; o la sociología para mostrar las interacciones entre personas (Powell *et al.*, 2011).

Los grafos, por tanto, hacen posible la modelización de datos complejos, aportando diferentes usos (Zweig, 2016): para los lectores aporta novedosos tipos de visualizaciones y para investigadores facilita la descripción numérica de esos datos. A mediados del siglo XX la teoría de los grafos comenzó a ser adoptada por investigadores en psicología y en ciencias sociales cuyos experimentos dieron lugar al análisis de redes (de las que luego hablaremos), que han sido asumidas por las HD en las últimas décadas (Rochat y Kaplan, 2014). En esta área los grafos pueden representar los datos de la estructura textual recogiendo aspectos hasta ahora difícilmente analizables (Keim, 2002).

Uno de los primeros usos de la representación de grafos para información humanística fue la representación léxica, como en *WordNet* (Fellbaum, 1998), donde los nodos representan palabras que son sinónimas entre ellas (*synsets*) conectados con otros (comparten canto) cuando son entre ellos hiperónimos, merónimos, etcétera. Un uso similar se puede visualizar actualmente en el *Nuevo diccionario histórico del español* (Pascual y Campos Souto, 2013), e incluso en el prólogo del *Diccionario* de María Moliner puede observarse un grafo que representa una taxonomía semántica (Moliner, 1966).

Sin embargo, probablemente el campo de mayor utilización de las redes sociales dentro de las humanidades es el popularizado por Moretti (2005) como manera de realizar *distant reading*: trabajar con representaciones abstractas y condensadas a partir de una gran cantidad de obras literarias. Realizando una red por obra teatral, sus personajes están representados como nodos que están conectados según la coaparición en escenas. En realidad, este uso proviene del matemático rumano Marcus, quien publicó en 1973 su *Poética matemática*. Diez años después de la propuesta de Moretti, la aplicación de estos grafos se popularizó y numerosos proyectos han aplicado grafos a una cantidad variable (desde decenas hasta miles, dependiendo del proyecto) de obras teatrales o novelas (Trilcke, Fischer y Kampkaspar, 2015; Jannidis *et al.*, 2016; Trilcke *et al.*, 2016; Glorieux, 2016; Algee-Hewitt, 2017; Isasi, 2017; Martínez Carro, 2019).

Continuando los pasos planteados por Marcus y Moretti, y seguidos por otros investigadores, podemos crear una matriz binaria relacionando personajes (filas) y escenas (columnas). Cada celda representa si el personaje (cuyo identificador se encuentra en la primera columna) habla en la escena (cuyo número aparece en la primera fila). Si lo hace, la celda contendrá un uno, si no, un cero. A partir de este punto utilizaremos obras de nuestro corpus BETTE, en concreto de *Luces de bohemia*:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
#max	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0
#lati	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1
#pisa	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1
#dori	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0
#rube	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
#coll	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0

Ilustración 1: Matriz binaria de personajes y escenas de *Luces de Bohemia*

De esta matriz binaria podemos extraer relaciones entre personajes sobre si aparecen juntos (coaparecen) en cada escena. Por ejemplo, Max y Latino coaparecen en nueve escenas, mientras que Max y Dorio solo coaparecen en una escena (la cuarta). De esta manera la relación de Max y Latino será de nueve, mientras que la relación de Max y Dorio será de uno:

Source	Target	Weight	Type
#max	#lati	9	Undirected
#max	#guar	3	Undirected
#lati	#clar	3	Undirected
#lati	#pere	3	Undirected
#dori	#pere	3	Undirected
#lati	#dori	3	Undirected
#lati	#pisa	3	Undirected

Ilustración 2: Coaparición de personajes en escenas de *Luces de bohemia*

Esta última tabla ya está formalizada como aristas de un grafo, en el que cada personaje es un nodo origen (*source*) u objetivo (*target*), cada relación o arista tiene una cantidad o peso (*weight*). Debido a que la relación se ha definido como coaparición, no tiene dirección (*undirected*).

Las aristas y la información de los nodos pueden ahora visualizarse como grafo:

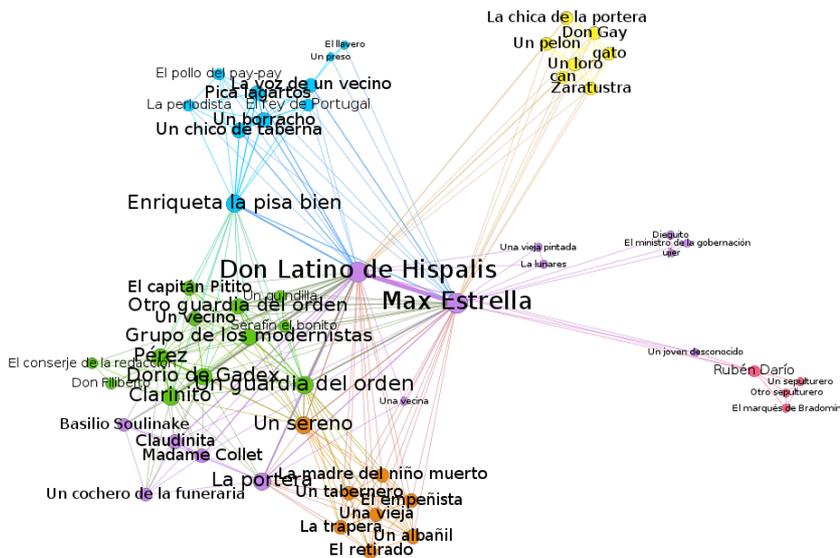


Ilustración 3: Grafo de relaciones entre personajes de *Luces de bohemia*, de Valle-Inclán, según coaparición en escenas

Podemos observar que Latino y Max aparecen en el centro del grafo, posición que tendemos a interpretar con su importancia. Los grafos no solo invitan a una lectura digital original, sino que también podemos utilizarlos como un modelo matemático que representa aspectos literarios. Por ejemplo, en este caso el diámetro del grafo es de tres, lo que quiere decir que entre dos personajes diferentes hay como máximo otros dos personajes. Así, para ir desde el nodo del Marqués de Bradomín al de Zaratustra tendremos que pasar por otros dos nodos: el de Rubén Darío y el de Max.

A pesar de la espectacularidad de estas visualizaciones y la promesa de que esto podría ser una manera de realizar *distant reading*, en los últimos años, ciertos trabajos han cuestionado la validez de estas representaciones de las obras. En concreto se ha aplicado aprendizaje automático supervisado (concretamente clasificación) para predecir la categoría de protagonista (Jannidis *et al.*, 2015; Krautter *et al.*, 2018), o correlaciones para predecir la importancia percibida de cada personaje (Santa María *et al.*, 2018, Santa María *et al.*, 2020). Los resultados de estos trabajos ponen claramente en duda que este tipo de grafos capturen información relevante en cuanto a la importancia de personajes o la estructura social de las obras.

De esta manera, se observa que esta metodología ha pasado por tres fases históricas:

1. Propuesta en 2005.
2. Exploración entre 2015 y 2017
3. Evaluación y crítica a partir de 2018

3. Propuesta de grafos de escenas

Los trabajos citados en la anterior sección han centrado su atención en los personajes, aunque en realidad para la obtención de los grafos se utiliza otro aspecto: las escenas o capítulos. Esto puede observarse en la segunda ilustración, que contiene personajes como filas y escenas como columnas. Sin embargo, hasta ahora no se ha experimentado con la posibilidad de realizar grafos en los que las escenas sean el centro del análisis. En esta sección proponemos representar las escenas como nodos del grafo. Dos escenas estarán conectadas si comparten personajes. Es decir, la idea sería hacer pivotar 90 grados la matriz de escenas y personajes (Ilustración 1).

Por ejemplo, la escena 14 de *Luces de bohemia* (que tiene lugar en el cementerio) y la escena 9 (en la que Max y Latino encuentran a Darío en el café) comparten un único personaje: Rubén Darío. Esas dos escenas estarán por lo tanto conectadas con un peso de 1. Si calculamos todas las posibles relaciones entre escenas, el resultado será una tabla de aristas entre escenas:

Source	Target	Weight	Type
3	15	5	Undirected
4	5	5	Undirected
4	13	4	Undirected
4	7	4	Undirected
7	13	4	Undirected
4	11	4	Undirected
11	12	3	Undirected

Ilustración 4: Aristas entre escenas de *Luces de bohemia*

Esta tabla puede ser cargada como tabla de aristas junto con la información de las escenas como nodos. Esto produce un grafo como el siguiente:

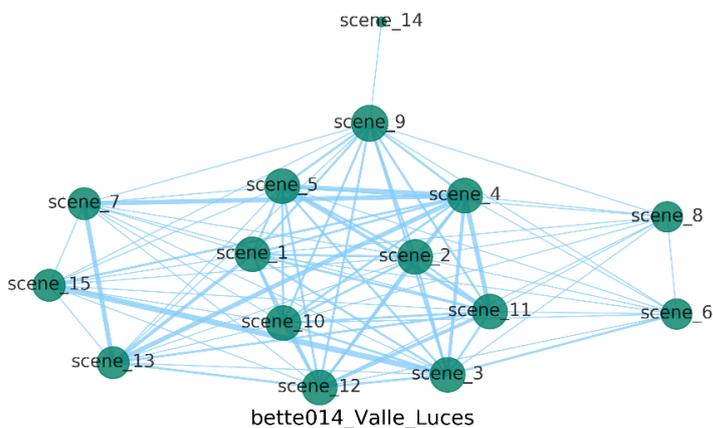


Ilustración 5: Grafo de las escenas de *Luces de bohemia* según la cantidad de personajes compartidos

En este grafo de escenas, la escena novena (la escena del café en la que Max, Latino y Rubén Darío charlan) posee el mayor grado¹ y es la única escena que está conectada con todas las demás escenas (en todas ellas aparecen o Latino, o Max o Rubén Darío). En el grafo también se observa el carácter excepcional de la escena decimocuarta (la del cementerio): como ya habíamos señalado, solo Rubén Darío conecta esta escena con el resto de las escenas. De esta manera, este grafo puede resultar una nueva manera de representar las obras de teatro subrayando su atención en las escenas, y por lo tanto una nueva manera de realizar *distant reading* sobre la estructura de las obras.

En el caso de que las escenas estén estructuradas en diferentes actos, los nodos pueden representar esto mediante su color. La mayoría de obras del corpus así lo están, por ejemplo, *Juan José de Dicenta*. En la siguiente imagen, las escenas en el primer acto tienen colores verdes oscuros, las escenas del segundo acto son nodos de color verde claro, mientras que las escenas del tercer y último acto son nodos amarillos.

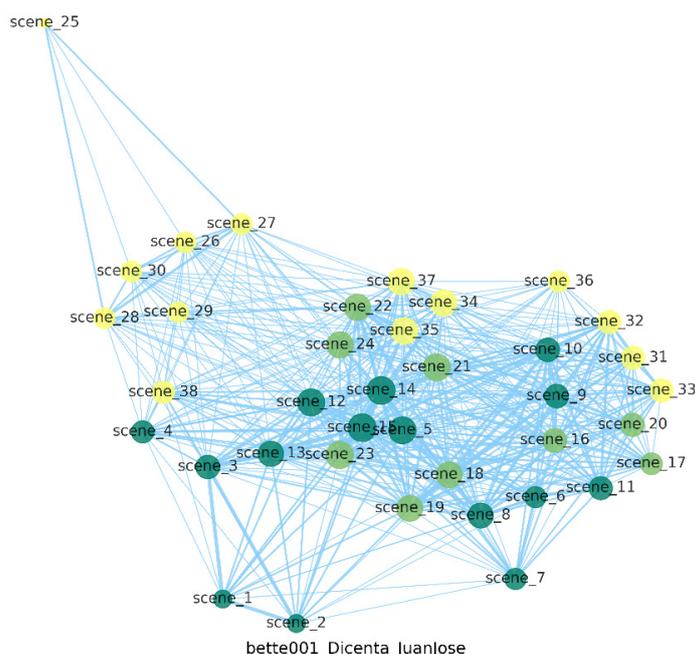


Ilustración 6: Grafo de las escenas de Juan José de Dicenta según la cantidad de personajes compartidos

El grafo representaría la estructura en escenas utilizando los personajes como información básica, pero a la vez mostrando la estructura de actos. Las primeras escenas de la obra se encuentran en la sección inferior del grafo, muy cercanas entre sí hasta la quinta escena. Por ejemplo, la primera escena y segunda escena se encuentran muy cerca entre sí, al igual que la tercera y cuarta. Esto quiere decir que estas escenas comparten la mayoría de los personajes entre sí.

NOTAS

1 | El grado es el número de aristas de cada nodo (Zweig, 2016: 63).

En general la estructura de la obra queda dividida en tres grupos de escenas que se visualizan de manera vertical en el grafo: el grupo que se encuentra en el centro del grafo (con las primeras escenas en la sección inferior, hasta la escena 37 en la parte superior). A su izquierda se encuentra un grupo de escenas (de la 25 a la 29 y 38, todas del tercer acto). El tercer grupo se encuentra a la derecha del grafo, con escenas de los tres actos (escena séptima en la parte inferior, 36 en la parte superior). Es decir, las escenas de la obra estarían estructuradas en tres típicas de personajes.

Sin embargo, el grafo no solo representaría las escenas según sus personajes en su verticalidad. En la horizontalidad se observa el desarrollo de la obra: las escenas del primer acto se encuentran principalmente en la parte inferior del grafo (color verde oscuro), las escenas del segundo acto se encuentran principalmente en el medio, mientras que las escenas del último acto se posicionan en la parte superior (color amarillo). De esta manera el grafo capturaría la secuencia y estructura en actos.

Como se puede observar, al rotar la matriz de escenas y personajes, se obtendría un grafo que representaría de manera visual diferentes aspectos de las obras. Esto podría ser utilizado con el objetivo de realizar lectura distante sobre una gran cantidad de textos. Esa es la hipótesis básica de este trabajo.

4. Evaluación de hipótesis sobre grafos de escenas

En la sección anterior hemos explorado los grafos de escenas, de manera meramente tentativa. Aunque algunas características de los grafos pueden ser observables para ciertas obras, desconocemos si esto ocurre en el corpus completo. Para sostener interpretaciones sobre esta formalización de las obras de teatro, en las siguientes subsecciones, realizaremos diferentes hipótesis y evaluaremos los resultados mediante test estadísticos en la totalidad del corpus. De esta manera analizamos de manera crítica nuestra propia propuesta, aunando datos cuantitativos y cualitativos. A diferencia de una exploración selectiva sobre ciertos aspectos de las visualizaciones, como hemos realizado en la anterior sección y como suele ser habitual en trabajos de HD, en esta sección uno de los posibles resultados es que ninguna de nuestras hipótesis sea observable en el conjunto del corpus. Es decir, es posible que terminemos la evaluación sin poder constatar que los grafos de escenas representan alguna característica interesante de las obras. De esta manera estamos aunando en nuestro trabajo una propuesta metodológica (como Moretti hizo con redes de personajes de obras de teatro en 2005) con su evaluación empírica y crítica (como los

trabajos de Jannidis *et al.*, 2015; Krautter *et al.*, 2018; Santa María *et al.*, 2018; Santa María *et al.*, 2020). Es decir, este artículo representa lo que en los grafos de personajes costó más de una década de diferentes investigaciones.

Para la evaluación utilizaremos los datos cuantitativos de los grafos de escenas. En algunos casos estos serán extraídos directamente de los valores de las aristas; en otros, de valores de centralidad del grafo². Como en otras situaciones, habría diferentes valores que podríamos extraer del grafo para evaluar las hipótesis. En general, hemos preferido extraer las variables cuyos cálculos requieren menos operaciones respecto a los datos originales. Estas variables cuantitativas son contrastadas con información cualitativa de cada una de las escenas. Es decir, las variables cuantitativas deben en primer lugar demostrar que representan ciertas características cualitativas de las obras anteriormente anotadas. Solo tras esta evaluación, los datos pueden ser interpretados.

4.1. Hipótesis de secuencia de escenas

La primera hipótesis es que el grafo de escenas representa la secuencia de escenas entre sí. Esto se ha señalado en la ilustración sexta, en la que ya hemos comentado que muchas escenas consecutivas se encuentran cercanas entre ellas. Si esto fuese observable en cada una de las obras del corpus, querría decir que escenas consecutivas tienden a compartir una gran cantidad de personajes entre sí. Consecuentemente, las obras irían modificando su cantidad de personajes en pequeñas dosis.

Para analizar esta hipótesis, la variable cuantitativa que extraemos de cada grafo de cada obra es el valor de peso medio en relación al resto de escenas, es decir, la cantidad de personajes que comparten dos escenas. La variable cualitativa utilizada en este caso es la distancia en la secuencia de la obra: la distancia entre la primera y la segunda escena es de uno; entre la primera y la tercera, la distancia es de dos. De esta manera, por cada obra habría dos vectores: uno cuantitativo (peso medio), uno cualitativo (distancia en secuencias). Si los datos sostienen nuestra hipótesis, ambos vectores deben tener una correlación negativa. Es decir, cuanto menor sea la distancia de las escenas en la obra, más cerca deberían estar sus nodos en el grafo. Para medir esta correlación, utilizamos el test estadístico Spearman para datos ordinales. Este test aporta un valor entre -1 y 1, representando el 0 dos vectores sin correlación. De esta manera, esperamos que las diferentes obras tiendan a tener valores negativos.

NOTAS

2 | «Centrality indices quantify the importance of a node in a given network» (Zweig, 2016: 243).

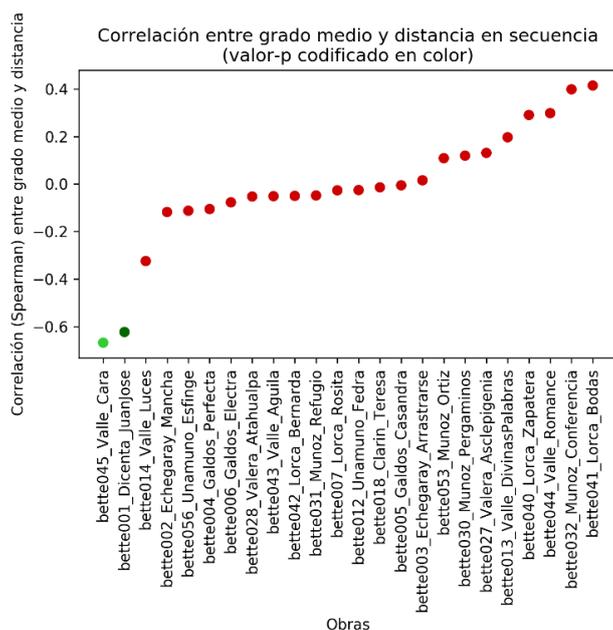


Ilustración 7: Correlación entre peso medio y distancia de las escenas en la obra

Como se puede observar, solo dos obras obtienen claros valores negativos: *Cara de Plata* de Valle-Inclán (valor p por debajo de 0.01, representado como verde claro) y *Juan José* de Dicenta (valor p por debajo de 0.001, representado como verde oscuro). Es decir, los resultados de correlación para esta obra de Valle-Inclán serían esperables en una de cada mil obras, mientras que la de Dicenta lo sería en una de cada diez mil obras o más. Es decir, las probabilidades de que este efecto sea aleatorio en estas obras son muy bajas. En otras palabras, en estas dos obras sí ocurre que los personajes van cambiando poco a poco entre escena y escena. Sin embargo, el resto de las obras obtienen valores p por encima del mínimo aceptable (0.05), representados como puntos rojos en la anterior gráfica. La gran mayoría del resto de obras obtienen valores de correlación muy cercanos a cero, como se observa por su posición en cuanto al eje vertical. Esto significa que las escenas adyacentes no comparten ninguna cantidad significativa de personajes. Curiosamente, en el extremo derecho de la imagen vemos dos obras en que se verifica justo lo contrario: *Bodas de sangre* de Lorca y *La conferencia de Algeciras* de Muñoz Seca. En estas piezas teatrales, el cambio de una escena a otra va asociada con un gran cambio en cuanto a los personajes.

De esta manera, podemos rechazar nuestra primera hipótesis: en la mayoría de los casos los grafos no representan la secuencia de escenas con la cercanía de sus nodos.

4.2. Hipótesis de primera y última escena en periferia

La segunda hipótesis tiene que ver con la pregunta de si las primeras o últimas escenas se separan del resto de las obras en cuanto a su constitución de personajes. Los casos de *Luces de bohemia* y *Juan José* parecen sugerir que las primeras y últimas escenas tienden a compartir escasos personajes con el resto de las obras.

Para formalizar esta segunda hipótesis, iteramos por el corpus observando cada una de las obras. De cada una, extraemos la primera escena y tomamos su valor de grado. Si nuestra hipótesis se sostiene, el grado de la primera escena debería ser notablemente inferior al resto de escenas de esa misma obra. Para realizar esta comparación de manera estadística, calculamos la puntuación estándar (en inglés *z-score*). Si la puntuación estándar de grado de la primera escena es inferior a -2, quiere decir que esa escena tiene un valor significativamente inferior al resto de la obra.

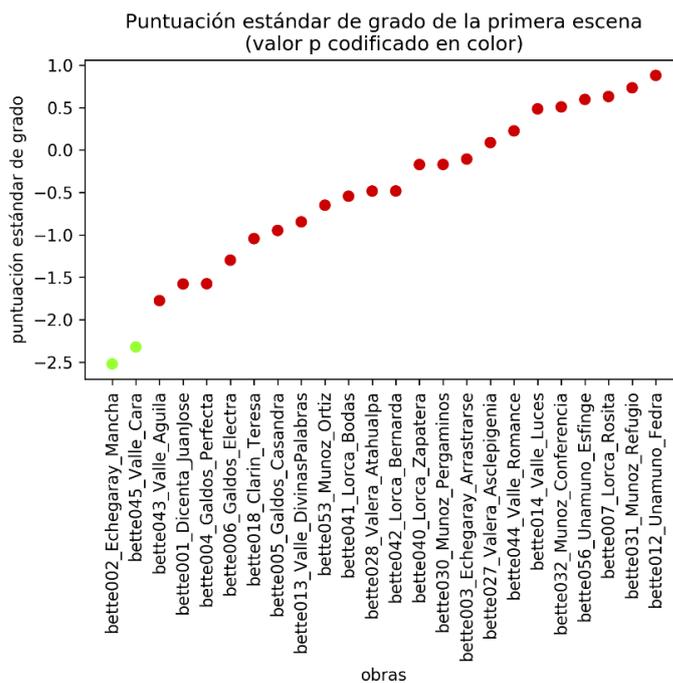


Ilustración 8: Puntuación estándar del grado de la primera escena en comparación con el resto de escenas de cada obra

De manera similar al caso anterior, solo dos obras confirman esta hipótesis: *Mancha que limpia* de Echegaray y *Cara de Plata* de Valle-Inclán, que tienen valores por debajo de -2. Aunque la mayoría de los valores son negativos y el valor medio es -0.48, lo cierto es que la enorme mayoría se encuentran entre los valores -1 y 1, es decir, casos en los que solo hay una desviación estándar de diferencia en relación con el resto de escenas. Es decir, la primera escena

tiende a tener una posición hasta cierto punto periférica en el grafo, pero esta podría ser explicada simplemente por la variabilidad de los datos.

La segunda hipótesis se completa con la pregunta de si la última escena tiene una posición periférica en el grafo. Es decir, si la última escena comparte escasos personajes con el resto de escenas. El procedimiento seguido es idéntico al anterior, analizando la última escena.

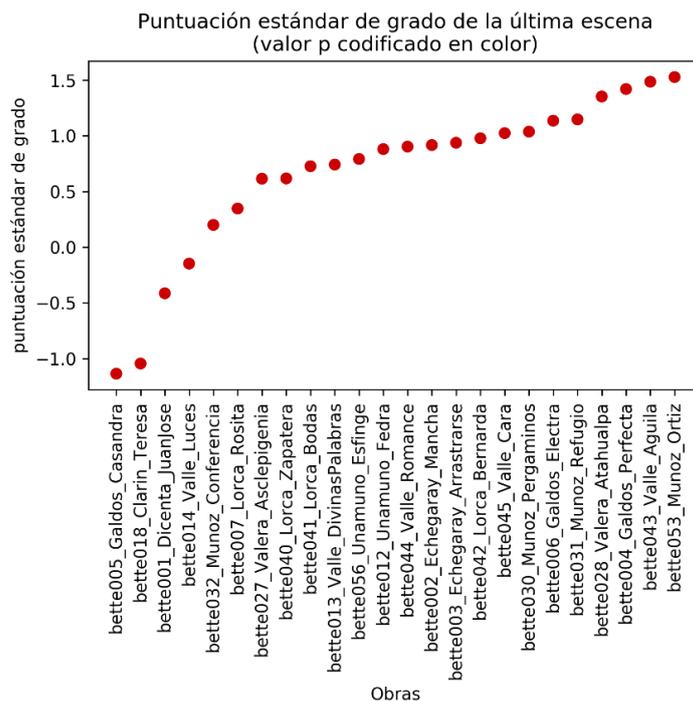


Ilustración 9: Puntuación estándar del grado de la última escena en comparación con el resto de escenas de cada obra

En este caso los resultados son aún más claros: ninguna de las escenas tiene valores notablemente inferiores a -1 (solo los casos de *Casandra* de Galdós, y *Teresa* de Clarín están ligeramente por debajo de ese valor). De hecho, la mayoría tienen valores positivos, con una media de 0.89.

Un aspecto interesante es el hecho de que, si comparamos las primeras y las últimas escenas, sí hay una diferencia estadística. Las primeras escenas tienden a tener un valor tipificado de grado menor que las últimas escenas. En otras palabras, las primeras escenas tienden a compartir menos personajes con el resto de las escenas en comparación con las últimas escenas. Aun así, ambas variantes de la segunda hipótesis pueden ser rechazadas: ni las primeras ni últimas escenas tienden a tener una posición periférica en el grafo aquí propuesto.

4.3. Hipótesis de estructura en actos

La tercera hipótesis que queremos observar está relacionada con los actos. Como hemos señalado, en la figura siete se puede observar que la mayoría de los nodos están distribuidos siguiendo la estructura de actos. De esta manera, las escenas del primer acto se encontrarían en la parte inferior del grafo, pasando en el medio al segundo acto, para terminar en la parte superior con el tercero. No todas las obras fueron originalmente articuladas en actos, por lo que para esta hipótesis trabajamos con 21 piezas teatrales.

Para operacionalizar esta hipótesis, la variable cuantitativa es el peso de la arista en relación al resto de escenas. La variable cualitativa es la diferencia categórica de a qué acto pertenece cada escena. Por cada obra, dividimos las aristas en dos grupos: aquellas que conectan dos escenas que pertenecen al mismo acto (*aristas intra-acto*) y aquellas aristas que conectan dos escenas de diferentes actos (*aristas inter-acto*). Si el grafo estuviese agrupando las escenas por la estructura en actos, los pesos de las *aristas intra-acto* deberían ser significativamente inferiores a los pesos de las *aristas inter-acto*. Si se calcula la media de ambos grupos de aristas de cada obra, efectivamente se observa que las aristas intra-acto tienden a tener pesos menores que las aristas inter-acto. El caso más claro es *Bodas de sangre*, con una diferencia de -0.05. Sin embargo, ese valor extremo es relativamente bajo, y el resto de las obras tiene valores más cercanos a cero. Por ejemplo, en *Juan José* la diferencia media es de -0.02. Para comprobar si esa diferencia es estadísticamente significativa, aplicamos un test de Welch, que aporta una estadística y un valor p por cada obra. Valores p por debajo de 0.05 (que se mostrarían como puntos verdes) señalarían obras donde nuestra hipótesis se sostiene:

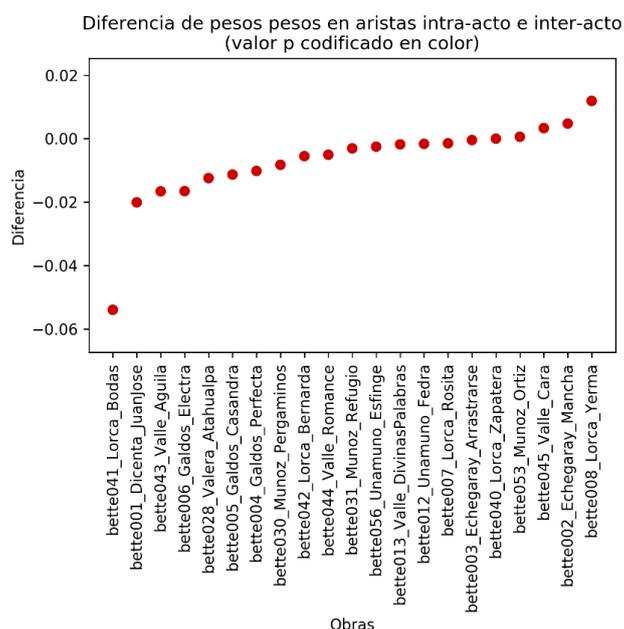


Ilustración 10: Resultados de correlación entre peso entre nodos y la distancia de las escenas en el desarrollo de la obra

Todas las obras están representadas por puntos de color rojo, lo que quiere decir que ninguna de las obras muestra un valor p por debajo de 0.05, siendo el menor 0.43. En la Ilustración 7 hemos observado cómo en la obra *Juan José* de Dicenta la distribución por actos podía observarse en el eje vertical. Lo que la exploración visual daba por aceptable, el test estadístico es capaz de rechazar de manera contundente, ya que tiene un valor p de 0.62, muy por encima del 0.05 necesario para ser estadísticamente significativo. En general los datos rechazan claramente la hipótesis presentada: Aunque es cierto que las aristas intra-actos tienden a mostrar pesos inferiores a las escenas inter-actos, la diferencia nunca es significativa en nuestro corpus. Los resultados son tan claros que somos escépticos de que otras posibles formalizaciones (como la identificación de comunidades) pudiese aportar mejores resultados.

4.4. Hipótesis de metadatos de personajes

La última hipótesis que vamos a explorar en este trabajo utiliza anotaciones cualitativas sobre los personajes. A partir de ellas, construiremos nuestra variable cualitativa para evaluar la hipótesis. Cada obra de BETTE contiene numerosos metadatos. Entre esta información se encuentran diferentes categorizaciones de los personajes, como su género, su papel o su importancia. Este último valor está expresado de manera ordinal, mediante tres posibles valores: personaje principal, secundario o terciario (en inglés *minor*). Esta información fue la base de nuestro trabajo para evaluar las redes de personajes y su importancia (Santa María *et al.*, 2018; Santa María *et al.*, 2020). En este caso esta información nos es útil para operacionalizar la hipótesis de si la distribución de las escenas tiende a relacionarse con la importancia de los personajes. Siguiendo esta hipótesis, las escenas en el centro del grafo estarían dominadas principalmente por personajes principales, frente a las escenas en la periferia, en la que principalmente aparecerían personajes secundarios o terciarios. De esta manera, la centralidad de cada escena debería tener correlación con la importancia de sus personajes.

Concretamente, tomamos como variables cuantitativas diferentes medidas de centralidad como el grado o la centralidad *eigenvector*. Como variable cualitativa de control, tomamos los metadatos de cada personaje anotados de manera manual. En concreto, por cada escena extraemos qué personajes hablan, y comprobamos las anotaciones sobre su importancia. Esta es traducida a un valor numérico ordinal. Finalmente calculamos la media, mediana y desviación estándar de estos valores de importancia. De esta manera obtenemos de cada escena la importancia de los personajes que hablan en ella. La correlación de la variable cuantitativa y cualitativa es medida mediante un test de Spearman por cada obra.

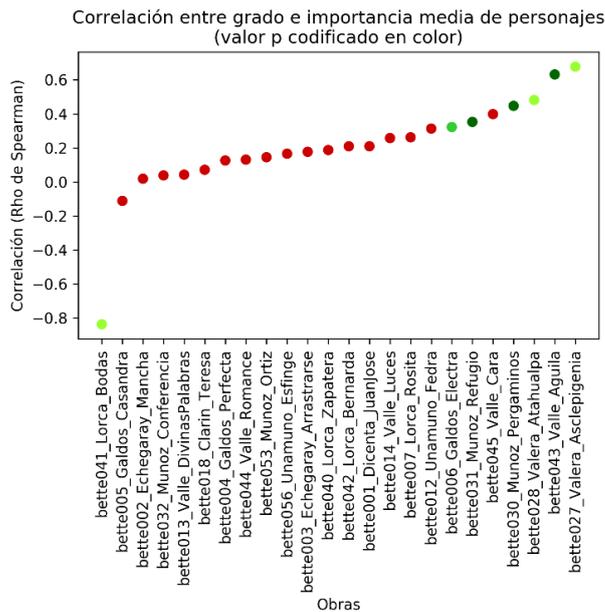


Ilustración 11: Correlaciones entre peso de nodos e importancia de personajes en escena

Los resultados efectivamente muestran que para la mayoría de las obras existe una correlación positiva. Esta es incluso significativa para seis piezas teatrales del corpus, algunas de ellas con valores p por debajo de 0.001 (verde oscuro). Aun así, esta tendencia no puede ser observada en el conjunto del corpus. Incluso esta tendencia es justamente la contraria en *Bodas de Sangre*: una correlación negativa entre el grado de las escenas y la importancia de sus personajes. Es decir, los personajes principales de esa obra tienden a estar localizados en posiciones periféricas del grafo. Las correlaciones con la mediana y la desviación estándar, y la centralidad *eigenvector* producen resultados similares. De esta manera, podemos rechazar la hipótesis de que la importancia de los personajes podría estar explicando la disposición de las escenas en los grafos.

Sin embargo, sí hay una variable que explica estadísticamente la disposición de la mayoría de los grafos. Para obtener esta variable cuantitativa, contamos por cada escena la cantidad de personajes que hablan en ella. Esta variable es utilizada de nuevo en un test de correlación Pearson debido a que en este caso ambos vectores son intervalos a diferencia de los casos anteriores.

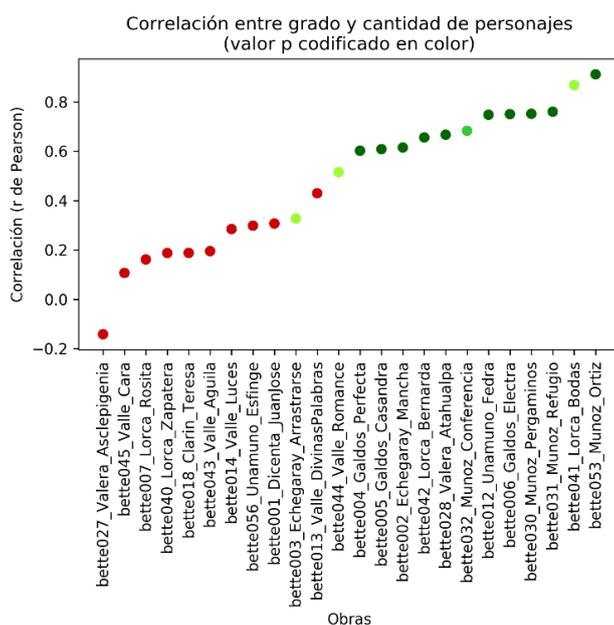


Ilustración 12: Correlaciones entre peso de nodos y cantidad de personajes en escena

En este caso todas las obras obtienen correlaciones positivas entre débiles y muy fuertes, con la excepción de *Asclepienia* de Valera que obtiene un valor negativo. Catorce obras muestran correlaciones estadísticamente significativas. Aunque esta variable no consigue dar explicaciones para todos los casos, muestra una tendencia estadística mayoritaria: Los grafos de escenas tienden a localizar en el centro aquellas escenas que contienen más personajes hablando.

5. Conclusiones

En este trabajo hemos propuesto y evaluado de manera crítica una nueva representación de obras de teatro. Esta está estrechamente relacionada con otras representaciones de novelas y teatro mediante grafos, popularizados desde el trabajo de Moretti en 2005. La mayoría de trabajos de HD con nuevas propuestas metodológicas se limitan a explorar de manera visual exclusivamente ciertas obras. A diferencia de este tipo de trabajos, nuestra investigación contiene una propuesta, una exploración y una evaluación empírica y crítica. Hemos partido de la base de que antes de poder sostener la interpretación de una nueva metodología, el investigador debe demostrar mediante hipótesis que la nueva representación se sustenta. Esta debe ser estadísticamente significativa y observarse en la mayoría de los casos. Esta evaluación estadística con el conjunto de los datos resulta más exacta y exigente que una exploración hermenéutica selectiva. Las diferentes hipótesis para explicar los grafos de escenas efectivamente pueden observarse

hasta cierto punto o en algunas obras. Sin utilizar test estadísticos ni mirar el conjunto de los datos, podríamos haber construido una argumentación a favor de cualquiera de las hipótesis, por ejemplo, quedándonos solo con las obras de *Juan José* de Dicenta, *Luces de Bohemia* y *Cara de Plata* de Valle-Inclán. Sin embargo, al evaluar formalmente el corpus completo, las hipótesis han ido siendo rechazadas una tras otra.

Nuestro trabajo no es una crítica contra la aplicación de grafos en humanidades. Es una propuesta crítica sobre esta formalización de grafos de escenas mediante la evaluación de hipótesis. Otras posibles explicaciones podrían evaluarse, como la importancia que las escenas tienen para diferentes estudios de la historia de la literatura. De cualquier manera, estas continuarían ajustándose al esquema de análisis que hemos planteado y que creemos más fructífero para la fase actual de consolidación de las Humanidades Digitales: la intersección entre datos cuantitativos producidos computacionalmente, datos cualitativos anotados por humanistas y una evaluación estadística sobre el conjunto de los datos.

Bibliografía citada

- ALGEE-HEWITT, MA. (2017): *Distributed Character: Quantitative Models of the English Stage, 1500-1920*, Montréal: McGill University & Université de Montréal.
- ALLINGTON, D.; BROUILLETTE, S. y GOLUMBIA, D. (2016): «Neoliberal Tools (and Archives): A Political History of Digital Humanities», *Los Angeles Review of Books*. <<https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>>, [09/02/2020].
- ARANA, R.; SANTA MARÍA, T. y SIMÓN, M. (2015): «El teatro de la España contemporánea (1868-1936)», *Texto Digital*, 11, 2, 111-127, <<http://dx.doi.org/10.5007/1807-9288.2015v11n2p111>>, [08/02/2020].
- BURROWS, J. (2002): «“Delta”: A Measure of Stylistic Difference and a Guide to Likely Authorship», *Literary and Linguistic Computing*, 17, 267-287.
- CALVO, J. et al. (2015): «Hacia Una Biblioteca Electrónica Textual Del Teatro En Español de 1868-1936 (BETTE)», *Texto digital*, 11, 2, 171-184, <<http://dx.doi.org/10.5007/1807-9288.2015v11n2p171>>, [07/02/2020].
- CEREZO, J. y CALVO, J. (2019): «Autoría y estilo. Una atribución cervantina desde las humanidades digitales. El caso de La conquista de Jerusalén», *Anales Cervantinos*, 51, 231-250, <<https://doi.org/10.3989/anacervantinos.2019.011>>, [07/02/2020].
- CRANENBURGH, A.; DALEN-OSKAM, K. y ZUNDERT, J. (2019): «Vector space explorations of literary language», *Lang Resources & Evaluation* 53, 625-650, <<https://doi.org/10.1007/s10579-018-09442-4>>, [05/02/2020].
- DAN, N.Z. (2019): «The Computational Case Against Computational Literary Studies», *Critical Inquiry* 45, 601-39, <<https://doi.org/10.1086/702594>>, [07/02/2020].
- DE LA ROSA, J. y SUÁREZ, J.L. (2016): «The Life of Lazarillo de Tormes and of his Machine Learning Adversities», *Lemir*, 20, 373-438.
- EDER, M.; RYBICKI, J. y KESTEMONT, M. (2016): «Stylometry with R: a Package for Computational Text Analysis», *R Journal*, 8, 1, 107-121, <<http://journal.r-project.org/archive/2016-1/eder-rybicki-kestemont.pdf>>, [01/02/2020].
- FELLBAUM, C., (ed.) (1998): *WordNet: An Electronic Lexical Database. (Language, Speech, and Communication)*, Cambridge, Mass: MIT Press.
- FIORMONTE, D. (2018): «¿Por qué las Humanidades Digitales necesitan al sur?», en *Humanidades Digitales: construcciones locales en contextos globales: Actas del I Congreso Internacional de la Asociación de Humanidades Digitales – AAHD*.
- GLORIEUX, F. (2016): «Data.bnf.fr, les documents», *J’attends des résultats*, <<https://resultats.hypotheses.org/795>>, [07/02/2020].
- HERMANN, J.B.; et al. (2020): «Response by the Special Interest Group on Digital Literary Stylistics to Nan Z. Da’s Study», *Journal of Cultural Analytics*, 24, <<https://culturalanalytics.org/article/11827-response-by-the-special-interest-group-on-digital-literary-stylistics-to-nan-z-da-s-study>>, [07/02/2020].
- HETTINGER, L.; et al. (2016): «Classification of Literary Subgenres», *DHd Tagung 2016, Leipzig Cultural Analytics*, <<https://culturalanalytics.org/article/11827-response-by-the-special-interest-group-on-digital-literary-stylistics-to-nan-z-da-s-study>>, [07/02/2020].
- HERNÁNDEZ, L. (2017): «Quantitative Syntactic Approaches to Spanish Poetry. A Preliminary Study on Fernando de Herrera’s Poetic Works» en *The Educational Impact of DSE*, 152-154.
- ISASI, J. (2017): «Posibilidades de la minería de datos digital para el análisis del personaje literario en la novela española: El caso de Galdos y los “Episodios Nacionales”», Tesis inédita, University of Nebraska-Lincoln.
- JANNIDIS, F.; et al. (2015): «Automatische Erkennung von Figuren in Deutschsprachigen Romanen», en *DHd-Tagung 2015*, 1-7, <<https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/14333>>, [25/01/2020].
- JANNIDIS, F.; et al. (2016): «Comparison of Methods for the Identification of Main Characters», in *German Novels. DH2016. Krakow: ADHO*, 2016, 578-82, <<https://bit.ly/2TJVAOj>>, [05/02/2020].
- JIMÉNEZ, C.M.; SANTA MARÍA, T. y CALVO, J. (2017): *Biblioteca Electrónica Textual del Teatro en Español (BETTE)*, La Rioja: UNIR, <<https://bit.ly/2SGtwhB>>, [25/01/2020].

- JOCKERS, M.L. (2013): *Macroanalysis: Digital Methods and Literary History*. Illinois: University of Illinois Press.
- JUOLA, P. (2013): «How a Computer Program Helped Reveal J. K. Rowling as Author of A Cuckoo's Calling», *Scientific American*, <<http://www.scientificamerican.com/article/how-a-computer-program-helped-show-jk-rowling-write-a-cuckoos-calling/>>, [28/01/2020].
- KEIM, D.A. (2002): «Information Visualization and visual data mining», *IEEE Transactions on Visualization and Computer Graphics*. January-March, v. 8, n. 1.
- KESTERMONT, M. et al. (2016): «Authenticating the writings of Julius Caesar», *Expert Systems with Applications*, 63, <<https://www.sciencedirect.com/science/article/pii/S0957417416303116>>, [07/02/2020].
- KRAUTTER, B. et al. (2018): «Titelhelden und Protagonisten - Interpretierbare Figurenklassifikation in deutschsprachigen Dramen», *Litlab Pamphets*, 7, <<https://nilsreiter.de/publications/Krautter2018aa>>, [09/02/2020].
- KOOLEN, C.W. (2018): «Reading beyond the female: The relationship between perception of author gender and literary quality», *ILLC Dissertation Series*, Amsterdam, open-access, <<https://dare.uva.nl/search?identifier=cb936704-8215-4f47-9013-0d43d37f1ce7>>, [07/02/2020].
- MARCUS, S. (1973): *Mathematische Poetik*, Mândroui, E. B.(trad.), Frankfurt/Main: Editura Academiei ; Athenäum Verlag.
- MARTÍNEZ CARRO, E. (2019): «Aproximación al teatro lorquiano desde la teoría de las redes sociales: La casa de Bernarda Alba», *Artnodes. Revista de Arte, Ciencia y Tecnología*, 24, 134-141, <<https://artnodes.uoc.edu/articles/10.7238/a.v0i24.3298/metrics/>>, [07/02/2020].
- MARTÍNEZ, E. y SANTA MARÍA, T. (2019): «Biblioteca Electrónica Textual del Teatro Español (1868-1936) e investigación con grafos», *Revista de Humanidades Digitales*, 3, 23-45, <<http://revistas.uned.es/index.php/RHD/index>>, [07/02/2020].
- MEDRANO, J.F.; ALONSO-BERROCAL, J.L. y FIGUEROLA, C. (2011): «Visualización de grafos web», en *Avances en informática y automática: quinto workshop*, 167-190. <www.researchgate.net/profile/Jose_Luis_Berrocal/publication/247936803_Visualizacion_de_Grafos_Web_-_Web_Graph_Visualization/links/5474a3cb0cf245eb436deaeb.pdf>, [07/02/2020].
- MOLINER, M. (1966): *Diccionario de Uso del Español*, Madrid: Gredos.
- MORETTI, F. (2005): *Graphs, Maps, Trees: Abstract Models for a Literary History*, New York: Verso.
- MURRAY, R. y TILLET, B. (2011): «Cataloging Theory in Search of Graph Theory and Other Ivory Towers», *Information Technology and Libraries*, 30(4), 170-184. <<https://doi.org/10.6017/ital.v30i4.1868>>, [07/02/2020].
- NYKAMP, D.: «Graph definition», *From Math Insight*, <<http://mathinsight.org/definition/graph>>, [30/01/2020].
- PALACIOS, A. et al. (2015): «Exploración de patrones de interacción para su uso en la web semántica», *El profesional de la información*, noviembre-diciembre, v. 24, n. 6, 749-758, <<http://dx.doi.org/103145/epi.2015.nov.06>>, [04/02/2020].
- PASCUAL, J.A. y CAMPOS, M. (2013): «Presentación», Nuevo diccionario histórico del español (NDHE), <<http://web.frl.es/DH/org/login/Inicio.view;jsessionid=2EC2BC1E95AC79C4814A647C7DF497CE>>, [04/02/2020].
- PIPER, A. (2020): «Do We Know What We Are Doing?», *Journal of Cultural Analytics*, 23, <<https://culturalanalytics.org/article/11826-do-we-know-what-we-are-doing>>, [05/02/2020].
- POWELL, J. et al. (2011): «Graphs in libraries: A primer», *Information Technology and Libraries*, 30, 4, 157-169, <<https://doi.org/10.6017/ital.v30i4.1867>>, [01/02/2020].
- RIBLER-PIPKA, N. (2018): «Die Digitalisierung des goldenen Zeitalters – Editionsproblematik und stilometrische Autorschaftsattribuion am Beispiel des Quijote», *Zeitschrift für digitale Geisteswissenschaften*, 4, 3.
- ROCHAT, Y. y KAPLAN, F. (2014): «Analyse des réseaux de personnages dans Les Confessions de Jean-Jacques Rousseau», *Les Cahiers du numérique*, 10, 3, 109-133, <www.cairn.info/resume.php?ID_ARTICLE=LCN_103_0109>, [03/02/2020].
- PRIANI, E. (2015): «El texto digital y la disyuntiva de las humanidades digitales», *Palabra Clave*, 18, 4, 1215-1234.
- PUENTES, E. (2018): «La transformación neoliberal de la ciencia: el caso de las Humanidades Digitales», *ArtefaCtoS. Revista de estudios sobre la ciencia y la tecnología*, 7, 1, 17-28.

- SANZ, A. (2018): «Para un comparatismo digital necesario: del relato al dato, del dato al relato» en Julia Ori (ed.), *Nuevos horizontes de la literatura comparada*, Madrid: SELGYC, 33-44, <<http://www.selgyc.com/index.php/es/publicaciones-selgyc/nuevos-horizontes-de-la-literatura-comparada>>, [01/02/2020].
- SANTA MARÍA, M.T.; CALVO TELLO, J. y JIMÉNEZ, C.M. (2018): «¿Existe correlación entre importancia y centralidad? Evaluación de personajes con redes sociales en obras teatrales de la Edad de Plata» en Girón Palau, J. y Galina Russell, I. (eds.), *Digital Humanities 2018 Puentes-Bridges. Book of Abstracts. Libro de resúmenes*, México: Red de Humanidades Digitales, 494-498, <https://dh2018.adho.org/wpcontent/uploads/2018/06/dh2018_abstracts.pdf>, [05/02/2020].
- SANTA MARÍA, M.T.; CALVO TELLO, J. y JIMÉNEZ, C.M. (2020). «¿Existe correlación entre importancia y centralidad? Evaluación de personajes con redes sociales en obras teatrales de la Edad de Plata» en *DSH-Digital Scholarship in the Humanities*, fqa015, <<https://doi.org/10.1093/llc/fqa015>>, [03/02/2020].
- TRILCKE, P.; FISCHER, F. y KAMPKASPAR, D. (2015): «Digitale Netzwerkanalyse dramatischer Texte», en *DHd-Tagung. Graz*, <<http://gams.unigraz.at/o:dhd2015.v.040>>, [05/02/2020].
- TRILCKE, P. et al. (2016): «Theatre Plays as “Small Worlds”? Network Data on the History and Typology of German Drama, 1730–1930» en *Digital Humanities 2016. Conference Abstracts. Jagiellonian University & Pedagogical University, Kraków*, 385-387, <<http://dh2016.adho.org/abstracts/407>>, [29/01/2020].
- ZWEIG, K.A. (2016): *Network analysis literacy*, Viena: Springer-Verlag.