

**Análisis semántico y cuantitativo de *La casa de Bernarda Alba* en el aula con *Voyant Tools*. Una aproximación didáctica al análisis literario**

Semantic and quantitative analysis of *La casa de Bernarda Alba* with the use of *Voyant Tools*. A didactic approach to literary analysis

Monika Dabrowska<sup>1</sup>  
*Universidad Internacional de La Rioja*

Recibido: 02.09.2021

Aceptado: 30.10.2021

## Resumen

El trabajo se propone innovar la enseñanza de la literatura en el aula de Secundaria y Bachillerato, ofreciendo una aproximación al texto literario desde el análisis semántico con las herramientas digitales, para ganar el interés de los alumnos acostumbrados a las aplicaciones tecnológicas y textos en formato digital. En concreto, se presenta un ejemplo de uso de *Voyant Tools*, una herramienta de uso abierto y fácil para realizar estudios cualitativos de textos en línea. En cuanto a la metodología, se muestran actividades que pueden realizar los docentes para analizar *La casa de Bernarda Alba* de Federico García Lorca. Tras una sesión inicial dedicada al autor y su obra, se familiarizará a los alumnos con la herramienta. Después, en grupos, tendrán que analizar nube de palabras, gráfica de frecuencia de los términos y el árbol de palabras, cantidad de palabras únicas, densidad de vocabulario, palabras más utilizadas y su contexto. Finalmente, cada grupo presentará la visualización de los

---

<sup>1</sup> monika.dabrowska@unir.net  
<https://orcid.org/0000-0001-7421-0164>

resultados, sus conclusiones y reflexión final. La propuesta permite enfocar las actividades literarias de forma dinámica y atractiva, estimular la lectura, desarrollar las competencias digitales y dinámicas de colaboración. Además, puede ser una vía para iniciar a los estudiantes en las Humanidades Digitales.

**Palabras clave:** investigación cualitativa, visualización de datos, didáctica, García Lorca, Voyant Tools

### **Abstract**

The work aims to innovate the teaching of literature in the Secondary and Baccalaureate classroom by offering an approach to literary text based on semantic analysis with the use of digital tools to win the interest of students accustomed to technological applications and texts in digital format. Specifically, an example of the use of *Voyant Tools* is presented, as this is an open and easy-to-use tool for conducting qualitative studies of online texts. Regarding the methodology, a few activities that teachers can carry out to analyze *La casa de Bernarda Alba* by Federico García Lorca are shown. After an initial session dedicated to the author and his work, students will be familiarized with the tool. Then, in groups, they will have to analyse word cloud, frequency of terms and the word tree, number of unique words, vocabulary density, most used words and their context. Finally, each group will present the visualization of the results and final conclusions. The proposal allows to run literary activities in a dynamic and attractive way, stimulate reading, develop digital skills and collaboration. In addition, it can be a way to initiate students in the Digital Humanities.

**Keywords:** qualitative research, data visualisation, didactics, García Lorca, Voyant Tools

## Introducción

¿Cómo motivar el interés de la gente joven por la lectura? ¿Cómo convertir las actividades de análisis del texto literario en novedosas y atractivas para los alumnos de la educación secundaria y postsecundaria, acostumbrados al uso de la tecnología, el internet, archivos multimedia e hipertextuales? Una forma de innovar la didáctica de la lectura es incorporar en las clases los textos digitales y abrirse a nuevas herramientas, formatos y lenguajes para su estudio e interpretación, es decir aproximarse a los textos literarios de forma dinámica e innovadora, que puede motivar más a los estudiantes. ¿Qué pasa si examinamos el texto literario prestando atención a las palabras que aparecen con más frecuencia y las relaciones entre ellas? ¿Qué sucede si visualizamos y representamos gráficamente esta información? ¿Qué nueva información y conocimiento sobre la obra estudiada podemos obtener a partir de estas técnicas de la minería de texto (*text mining*)?

En este artículo se revisan las posibilidades de *Voyant Tools* como recurso didáctico para el estudio de textos literarios (Sinclair y Rockwell, 2015; Rockwell y Sinclair, 2016; Hetenyi, et al, 2019). Examinaremos *La casa de Bernarda Alba*, el drama rural de Federico García Lorca escrito en 1936. Por supuesto, la herramienta permite trabajar de la misma manera cualquier tipo de textos, tanto las creaciones poéticas (Huesca, 2021) como narrativas (Boyle y Hall, 2016; Ramsby, 2016; Martín-Martínez 2021). Se pueden investigar conjuntos de textos (*corpora*) más extensos, por ejemplo, de un autor o de varios autores a la vez. Dado el carácter introductorio del ejercicio que proponemos, nos centraremos en una sola obra literaria. Su análisis digital pretende ayudar a los estudiantes a interpretar el texto y a pensar críticamente la lectura tradicional y la mediada por procedimientos computacionales. La práctica de aula será una oportunidad para familiarizar a los alumnos precisamente con los conceptos de la lectura cercana (*close reading*) y la lectura distante (*distant reading*) conceptualizados por Moretti (2015), es decir la habitual que conocen y la que aplica procesamiento informático de datos literarios de *coprus* de gran tamaño.

## 1. Metodología

Se ha optado por el uso de *Voyant Tools* ([www.voyant-tools.org](http://www.voyant-tools.org)) dada la facilidad y gratuidad de su uso. Es un software libre (<https://github.com/sgsinclair/Voyant>), con licencia GPLv3, gratuito y se ejecuta en línea (no necesita registro ni descarga). A diferencia de otras herramientas de análisis semántico del texto (como *NVivo*, *MaxQDA*, *Tableau*) no necesita licencia. El programa fue iniciado por dos académicos canadienses, Stéfán Sinclair (McGill University) y Geoffrey Rockwell (University of Alberta) y sus colaboradores. Actualmente se ejecuta la versión 2.4. El *software* permite procesar y estudiar textos propios o externos, un documento o una colección de textos (*corpus* digital), cargado desde el ordenador o desde la red (puede ser un blog, una página web), en formato de texto plano, MS Word, PDF, RTF, HTML, XM.

Otra razón de peso la sencillez de uso de ese entorno de análisis que pueden explorar los alumnos que no tiene experiencia en análisis computacional. Es suficiente una instrucción básica, además la herramienta es muy intuitiva. La interfaz de inicio está dividida en cinco paneles que muestran diferentes estadísticas y visualizaciones de texto, a saber:

*Cirrus*: nube de palabras, donde el tamaño guarda relación con la frecuencia en el texto (la posición y el color de los términos son aleatorios). Permite visualizar ente 25 y 500 palabras y ver cuántas veces aparece la palabra seleccionada.

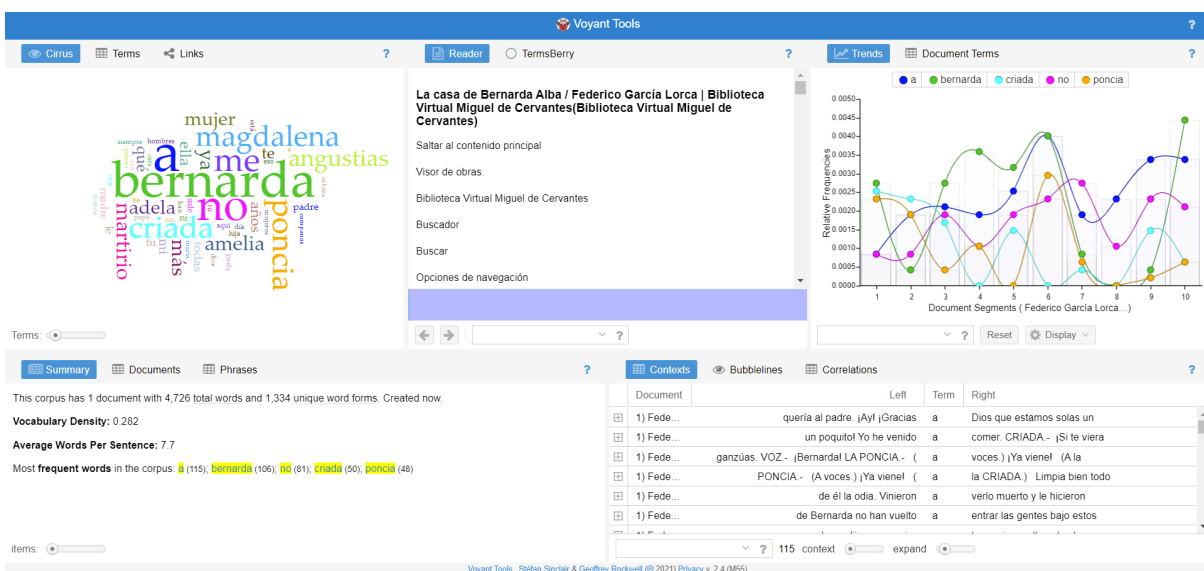
*Reader* (Lector): muestra el texto completo del documento analizado, permite realizar búsquedas por palabras.

*Summary* (Resumen): arroja las estadísticas generales, como la cantidad de palabras, densidad del vocabulario, frecuentes de palabras, palabras únicas y distintivas, promedio de palabras por oración, frases más largas/ cortas, las frases más frecuentes. Automáticamente mostrará las cinco palabras más comunes, pero la opción “ítems” permite ampliar ese rango.

*Trends* (Tendencias): revela la frecuencia absoluta y relativa de palabras en todo el texto (o corpus) examinado. Permite introducir palabras de búsqueda.

*Contexts* (Contextos) muestra palabras anteriores y posteriores al término seleccionado.

Figura 1. Interfaz de *Voyant Tools* con el drama de Lorca cargado para el análisis, donde se pueden apreciar los cinco paneles principales.



Estas ventanas están interrelacionadas, de modo que haciendo *click* en una palabra determinada se actualiza la información en todos los demás paneles. La vista de cada una se puede personalizar, y los resultados guardar en una pestaña nueva, o descargar la visualización. Además, desde estos paneles principales se puede acceder a otras 29 herramientas de análisis y visualización incorporadas en el programa. Su descripción se encuentra en la guía del propio programa <https://voyant-tools.org/docs/#!/guide/tools>. En nuestro caso utilizaremos las cinco funciones básicas.

## 2. Actividades preliminares

Un paso previo imprescindible es una lectura personal y atenta de *La casa de Bernarda Alba* por parte de los alumnos. Es aconsejable que los estudiantes usen el libro en formato impreso.

Incluso sería muy recomendable, además, ver la representación escénica de esta obra, bien en una visita al teatro, bien en una videograbación. Por otro lado, necesitamos que los estudiantes se habitúen con el texto en formato digital (en este caso disponible en la Biblioteca Virtual Miguel de Cervantes (<http://www.cervantesvirtual.com/obra/la-casa-de-bernarda-alba-775125/>) y que comenten las diferencias ambas experiencias lectoras, en papel y en modalidad digital. En la web con la edición digital conviene mostrar la opción “concordancias” (*Ficha/ Contenido de la obra/ Concordancias*) y realizar una exploración inicial, con los ejemplos sugeridos por los alumnos.

Después de introducir la figura de García Lorca y tratar el contexto cultural y literario en el que surge la obra, podemos realizar ejercicios previos de análisis. Se pide a los estudiantes que identifiquen rasgos de los protagonistas, los temas principales del drama, características de estilo. Se trata de que los alumnos determinen cuáles son, a su parecer, los aspectos centrales del drama, elaborando un listado de cinco términos que lo definan. En esta actividad preliminar los estudiantes trabajan en grupos pequeños, donde presentan fragmentos de la obra, discuten y consensuan sus respuestas. Cada grupo presenta sus conclusiones delante de toda la clase, apoyándose en pasajes seleccionados. Es importante que se ponga especial atención a la elección de palabras. Ante las posibles discrepancias conviene que el docente favorezca la discusión, animando a plantear preguntas o inquietudes, y animando a justificar las opiniones con argumentos y citas oportunas. Posteriormente se recogen, por orden de relevancia, los términos más recurrentes que definan las cuestiones planteadas por la obra teatral. Estas respuestas escritas sirven como preparación para una posterior investigación en clase.

En segundo lugar, antes de pasar al análisis digital del texto, tenemos que familiarizar a los estudiantes con la herramienta *Voyant Tools*, mostrando las funciones que vamos a utilizar y dando tiempo para que puedan rastrear las opciones de diferentes módulos y practicar las búsquedas de determinadas palabras. Resulta de utilidad atender a las explicaciones que ofrece el propio software sobre cada una de las herramientas que queramos utilizar (colocándonos en el botón con el signo de interrogación se despliega una información sobre cada función).

## Paso 1. Cargar el texto y observar las primeras estadísticas

El primer paso consiste en cargar la obra introduciendo el link de su versión digital en la herramienta ([http://www.cervantesvirtual.com/obra-visor/la-casa-de-bernarda-alba-775125/html/6aef1659-90e7-443f-bd6c-a8fe62c4c04a\\_2.html](http://www.cervantesvirtual.com/obra-visor/la-casa-de-bernarda-alba-775125/html/6aef1659-90e7-443f-bd6c-a8fe62c4c04a_2.html)) y observar en grupo clase los primeros resultados arrojados por la misma. Comentamos brevemente cada uno de los cinco paneles principales en la vista general, para enseñar su uso y como investigar el texto partiendo de la frecuencia de palabras. Partimos de la pregunta abierta: ¿cuáles son los términos más utilizados en el drama de Lorca? En la ventana *Reader* aparece la información sobre el documento cargado y el texto completo de la obra. El módulo *Summary* destaca los términos que aparecen con más frecuencia, como *bernarda* (106 veces); *criada* (50); *poncia* (48); *magdalena* (42); *angustias* (34); *martirio* (32); *amelia* (28); *adela* (27); *mujer* (27); *madre* (16); *padre* (15); *casa* (11); *hombres* (10), etc. La misma ventana, en la opción *Phrases*, destaca las frases más repetidas (p.ej. “los hombres de aquí no son”). Observamos que esa misma información queda visualizada en el panel *Cirrus*, que además permite comparar la frecuencia de los términos (opción *Therms*), ordenarlos según esa frecuencia o alfabéticamente, así como encortar esos datos para cualquier palabra que introducimos en el campo de búsqueda y su conexión con otras palabras.

Tras una visión global, comprobamos cuáles de ellos coinciden con nuestra lista de términos clave y cuáles no constan el texto dramático (como podrían ser: amor, deseo, libertad, pasión, etc.); cuáles de ellos resultan esperados, p.ej. *bernarda* (106), *mujer/es* (36) y cuáles sorprendentes, p.ej. *criada* (50), *poncia* (48) *no* (87). Indagamos por qué, a que se deben estas estadísticas en cada caso. Cómo podemos interpretar el hecho que las apariciones de Bernarda triplican a las de sus hijas y casi cuadruplican a la de la menor y rebelde Adela. Seguimos prestando atención a la ventana *Trends* donde se puede seguir la frecuencia de la(s) palabra(s) seleccionada(s) en alguno de los módulos anteriores, también las del mismo campo semántico. La herramienta *Context* ofrece más detalles, mostrando la frase y el párrafo donde figura el término que analizamos. En definitiva, se trata de ir interactuando con la herramienta y observar los resultados de este primer análisis en todas las ventanas de la interfaz principal.

Nos fijamos en la correlación entre paneles, escogiendo una palabra y viendo los gráficos en las demás ventanas.

## **Paso 2. *Stop words***

En este primer visionado los alumnos se habían dado cuenta de la presencia de los términos que no resultan significativos para nuestro análisis (artículos, preposiciones, determinantes, pronombres). Pueden señalar las palabras que consideran irrelevantes en la nube de palabras, justificando su opinión. Es el momento oportuno para explicar el concepto de palabras vacías o *stop words* y mostrar cómo podemos depurar el texto filtrando las palabras que consideramos insignificantes para nuestro análisis. La opción correspondiente (*Cirrus/ Definir opciones para esta herramienta*) permite excluir una lista de palabras vacías preestablecida, correspondiente al idioma que seleccionemos, o bien crear una lista propia (se puede usar una lista de palabras existente para *Voyant* en español (disponible en Gutiérrez, 2019). En nuestro caso excluimos los metadatos del texto dramático (como *federico, García, Lorca, biblioteca, virtual, miguel, cervantes*) y las palabras *a, a, y, la, el, un, ni, mí, me, te, se, ya, he, has, que, eso, etc.* Una vez guardada la lista de palabras excluidas hay que aplicarla para la nueva búsqueda. Esta puede introducirse en una herramienta, o en todas las demás a la vez (*apply globally*, opción que aparece por defecto). Hay que ser consciente de estas decisiones y saber que esto afectará a los resultados.

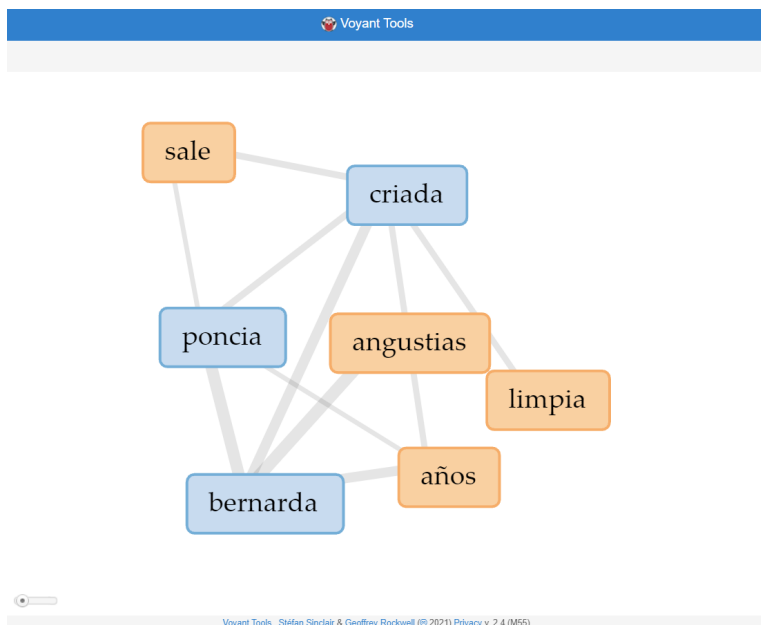
Para este ejercicio los estudiantes pueden, en grupos, crear una lista de *stop words* personal y comparar los resultados con otros grupos. Resultará muy ilustrativo ver como las métricas y los gráficos se quedan influenciados por esas variaciones. Puede ser una ocasión para indagar qué consecuencias pueden tener los términos vacíos para el procesamiento de datos y porqué algunas herramientas no las admiten.



### Paso 3. Procedimientos del análisis computacional

Una vez refinada la lista de palabras comunes (*stop words*) que queremos excluir de la búsqueda, nos centramos en explorar los términos que consideramos significativos para nuestra obra literaria, así como las relaciones entre ellos. Para ello conviene detenerse en la opción *Links* dentro del panel *Cirrus*, para observar el grafo de las relaciones entre palabras más frecuentes (*frequency*) e interconectadas (*proximity*). En la *Figura 2* podemos apreciar un ejemplo, donde las palabras clave aparecen en azul, y las más relacionadas con ellas en color naranja. El grosor de las aristas refleja la densidad de estas relaciones. En este caso se confirma el protagonismo de Bernarda y Poncia, la gobernanta de la casa.

Figura 2. Visualización en forma de grafo de relaciones entre las palabras más frecuentes en el texto lorquiano

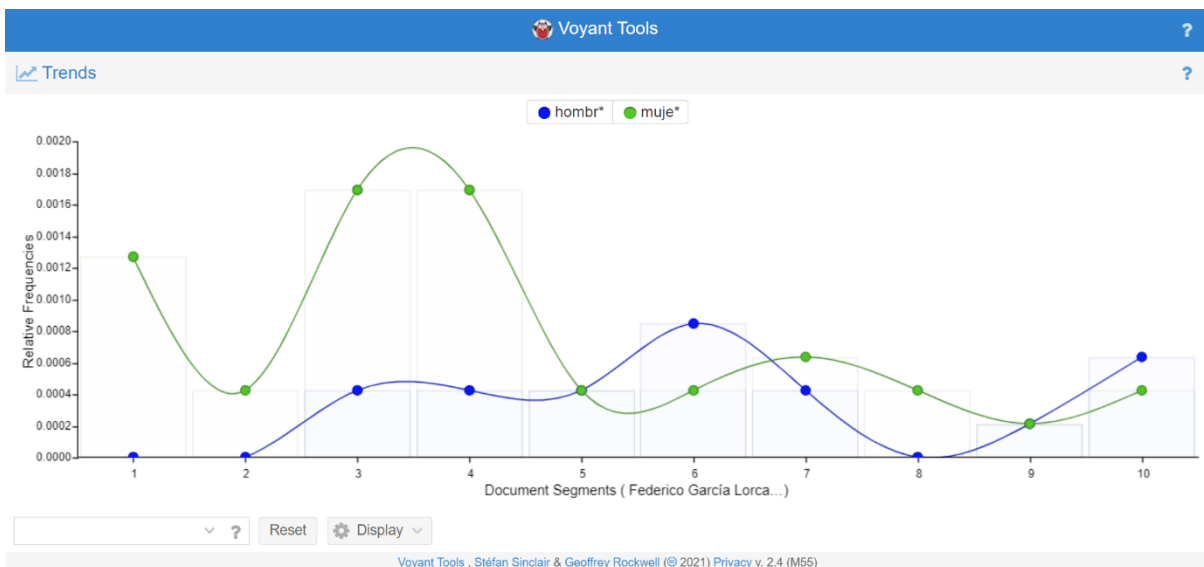


A continuación, realizamos el ejercicio de extracción y análisis de contenido semántico, con el fin de identificar las tendencias y significados más manifiestos que configuran este texto dramático. Teniendo en cuenta la marcada presencia femenina en los dramas lorquianos (y particularmente en las tres tragedias rurales a las que pertenece nuestra lectura), y, en general, las relaciones entre los dos sexos como temática recurrente para el autor granadino (Gil, 2008;

Villazón, 2006; Crovetto, 2010) nos centraremos en el rol femenino y masculino que refleja la obra. De hecho, el autor subtítulo esta pieza tetral “Drama de mujeres en los pueblos de España”. Para ello, pedimos que dos grupos rastreen la palabra “mujer” y otros dos la palabra “hombre”. Los estudiantes tienen que explorar interactivamente las cinco ventanas principales, prestando atención a la frecuencia de uso de cada una de ellas (con sus variantes semánticas y plurales) y los contextos oracionales en los que aparece.

El procedimiento consiste en inspeccionar especialmente el panel *Trends* y *Context*. En el primero cabe reparar en la línea de tendencias y preguntarnos por las funciones crecientes y decrecientes de los dos términos analizados a lo largo de la obra (Figura 3). El otro espacio ayuda a verificar la distribución de palabras alrededor de los términos de interés y ubicarlas dentro del texto de la obra.

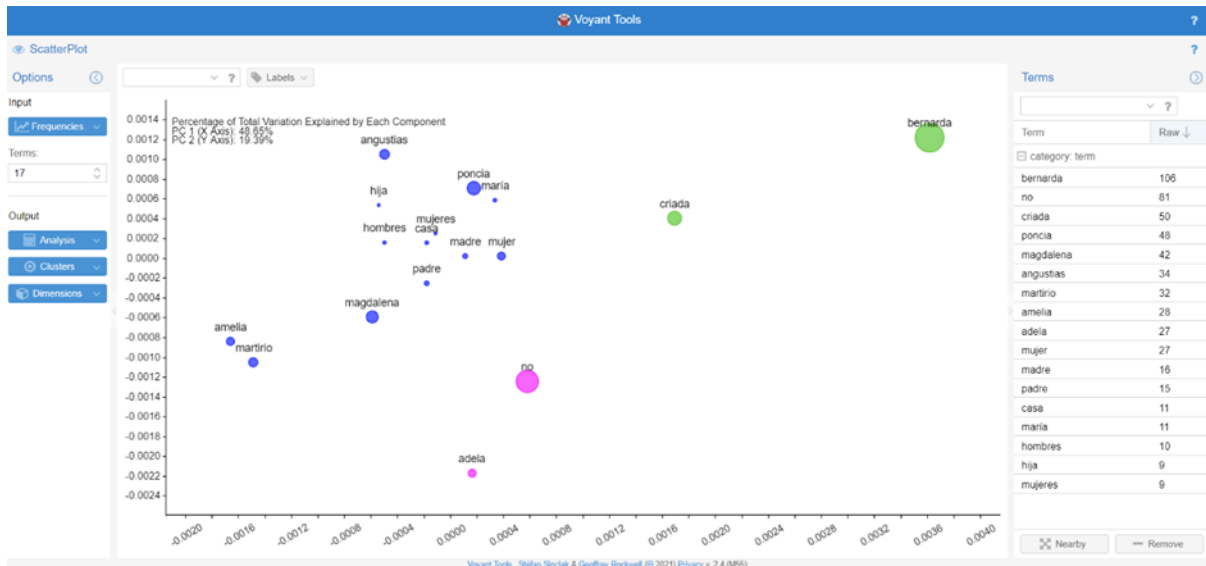
Figura 3. Tendencias de la aparición de la palabra *hombre* y *mujer* en la obra de Lorca.



Si la clase es numerosa, cabe investigar de la misma manera los sustantivos que indican colores (negro, verde, blanco) y averiguar su sentido simbólico dentro del mundo recreado en la escena (Arango, 1998; Murken, 2014). Otro ejercicio posible es rastrear la oposición *dentro /fuera de la casa* (11), incluyendo variaciones del mismo campo semántico (*interior, calle, campo, salir*) y llamando la atención sobre la polisemia de algunos de estos vocablos. Si

disponemos de más tiempo, podemos mostrar, además, otras variantes de visualizaciones (como *Correlations*, *Collocates*, *Mandala*, *ScatterPlot*, *SteamGraph* y otras), introduciendo los términos que queremos explorar. La figura 3 muestra el gráfico de cómo las palabras se agrupan en el documento estudiado.

Figura 4. Visualización de agrupación de palabras con ScatterPlot.



Al acabar esta parte, cada grupo guarda los resultados de su análisis y los comparte con la clase. La comparación de los resultados puede evidenciar algunas discrepancias, creando ocasión propicia para discusión y argumentación por parte de los grupos. El profesor a lo largo de todo ese proceso ha de ofrecer instrucciones detalladas y resolver las dudas de carácter técnico o metodológico.

### 3. Resultados

Los estudiantes han filtrado los términos “mujer” (36 recurrencias) y “hombre” (20 apariciones, incluyendo “varón” (4 apariciones) y han ido anotando las expresiones y contextos en los que aparecen. La primera observación confirma el peso de lo “femenino” en la pieza estudiada. Este dato podría completarse con la cantidad de los nombres propios, también los aludidos, femeninos (Paca la Roseta, Adelaida, La hija de la Librada, la soltera)

y masculinos (Antonio María Benavides, Pepe el Romano, Enrique Humanes, Don Arturo) y cuantificar de nuevo la proporción entre los personajes de ambos sexos. A partir de estos resultados los alumnos revisan las concordancias. El término *hombre* figura vinculado a palabras *fuera, bernarda, flores*, mientras que la palabra *mujer* viene asociada a *casa, abanico, mar, habitación, madre, miedo, maldita, luto* (Figura 4).

Figura 5. Correlaciones léxicas correspondientes a la palabra *mujer* en el drama de Lorca

Term 1	←	→	Term 2	Correlation...	Significanc...
casa			muje*	0.01670791	0.96346164
abanico			muje*	-0.01719797	0.96239054
mar			muje*	-0.01719797	0.96239054
dará			muje*	-0.01970277	0.9569169
habitación			muje*	-0.01970277	0.9569169
madre			muje*	-0.0421262...	0.90801215
miedo			muje*	0.042741287	0.906674
en			muje*	0.048303567	0.89458215
maldita			muje*	-0.04923287	0.89256376
cosas			muje*	-0.0525407...	0.88538396
luto			muje*	0.05874228	0.8719438
gloria			muje*	0.073849306	0.8393328
lenguas			muje*	0.073849306	0.8393328
alta			muje*	0.078811035	0.8286677
año			muje*	0.078811035	0.8286677
buen			muje*	0.078811035	0.8286677
comiendo			muje*	0.078811035	0.8286677
corazón			muje*	0.078811035	0.8286677
cuadros			muje*	0.078811035	0.8286677
cuenta			muje*	0.078811035	0.8286677
das			muje*	0.078811035	0.8286677
debes			muje*	0.078811035	0.8286677
decente			muje*	0.078811035	0.8286677

192 minimum coverage (%50)

Voyant Tools , Stéfan Sinclair & Geoffrey Rockwell ©

La búsqueda ulterior de las oraciones y expresiones ligadas a la masculinidad y feminidad revela las concepciones expresadas por diferentes personajes, siempre femeninos. Los listados de las expresiones más emblemáticas, señalando de boca de qué personaje procede cada una, muestran ideas como:

- *Malditas sean las mujeres* (Magdalena)
- *Las mujeres en la iglesia no deben de mirar más hombre que al oficiante, y ése porque tiene faldas* (Bernarda)
- *Hilo y aguja para las hembras. Látigo y mula para el varón* (Bernarda)
- *A vosotras, que sois solteras, os conviene saber de todos modos que el hombre, a los quince días de boda, deja la cama por la mesa y luego la mesa por la tabernilla, y la que no se conforma se pudre llorando en un rincón* (La Poncia)
- *Los hombres necesitan estas cosas* (La Poncia)

hasta los parlamentos pronunciados por la vieja y alejada madre de Bernarda, María Josefa:

- *Me escapé porque me quiero casar, porque quiero casarme con un varón hermoso de la orilla del mar, ya que aquí los hombres huyen de las mujeres*
- *No quiero ver a estas mujeres solteras, rabiando por la boda, haciéndose polvo el corazón, y yo me quiero ir a mi pueblo. Bernarda, yo quiero un varón para casarme y para tener alegría.*

Se puede apreciar parte de este escrutinio en la Figura 5.

Figura 6. Contextos semánticos de la palabra *mujer* en *La casa de Bernarda Alba*

Document	Left	Term	Right ↓
1) Fede... oscura. BERNARDA - Eso tiene ser oscura.	oscura. BERNARDA - Eso tiene ser <b>mujer</b> .	mujer	. MAGDALENA - Malditas sean las mujeres
1) Fede... los hombres huyen de las	los hombres huyen de las	mujeres	. BERNARDA - ¡Calle usted, madre! MARÍA
1) Fede... mujer. MAGDALENA - Malditas sean las	mujer. MAGDALENA - Malditas sean las	mujeres	. BERNARDA - Aquí se hace lo
1) Fede... haber muerto loca la segunda	haber muerto loca la segunda	mujer	. AMELIA - Y ese infame, ¿por
1) Fede... de Angustias, por Angustias como	de Angustias, por Angustias como	mujer	, yo me alegraría; pero viene
1) Fede... hoy son para mí. MENDIGA -	hoy son para mí. MENDIGA -	mujer	, tú tienes quien te gane
1) Fede... la vida. Le quedan cinco	la vida. Le quedan cinco	mujeres	, cinco hijas feas, que quitando
1) Fede... dos os alegráis. MARTIRIO - ¡Magdalenal	dos os alegráis. MARTIRIO - ¡Magdalenal	mujer	! MAGDALENA - Si viniera por el
1) Fede... que me quedan. CRIADA - ¡Qué	que me quedan. CRIADA - ¡Qué	mujer	! LA PONCIA - Tirana de todos
1) Fede... pinche en los ojos! CRIADA - ¡	pinche en los ojos! CRIADA - ¡	mujer	! LA PONCIA - Pero yo soy
1) Fede... Terminan de entrar las doscientas	Terminan de entrar las doscientas	mujeres	y aparece BERNARDA y sus
1) Fede... No quiero ver a estas	No quiero ver a estas	mujeres	solteras, rabiando por la boda
1) Fede... de esta casa, a una	de esta casa, a una	mujer	que, como su padre, habla
1) Fede... al marido de su primera	al marido de su primera	mujer	para casarse con ella. Luego
1) Fede... cabeza. BERNARDA - Es la única	cabeza. BERNARDA - Es la única	mujer	mala que tenemos en el
1) Fede... madre y de su abuela,	madre y de su abuela,	mujeres	las dos del que la
1) Fede... Lengua de cuchillo! BERNARDA - Las	Lengua de cuchillo! BERNARDA - Las	mujeres	en la iglesia no deben
1) Fede... BERNARDA - ¿Es decente que una	BERNARDA - ¿Es decente que una	mujer	de tu clase vaya con
1) Fede... * MUJER 4 * MUCHACHA,	* MUJER 4 * MUCHACHA,	mujeres	de luto. El poeta advierte

36 context expand

Voyant Tools, Stefan Sinclair & Geoffrey Rockwell (© 2021) f

#### 4. Discusión de resultados

Tras la exploración del texto llevada a cabo en la clase llega el momento de comentar los hallazgos, y quizá de abrir nuevas vías de interpretación del drama lorquiano. También de despertar la discusión sobre las posibilidades y debilidades de análisis cuantitativo, así como problematizar los resultados obtenidos. Prácticamente todo el proceso de interacción de los alumnos con la herramienta crea ocasiones propicias para provocar dudas y discusión. La experiencia de los alumnos, aunque básica, les permite darse cuenta de las posibilidades que abre el análisis informático (procesar numerosos textos a la vez, disminuir el tiempo para obtener información, resúmenes gráficos, etc.) y las cuestiones que deben discutirse críticamente (enfoque cuantitativo, importancia de la interpretación, etc.).

Podemos estimular esta reflexión preguntando: ¿Qué intuiciones se han confirmado, y cuáles han sido los resultados inesperados o sorprendentes? ¿Qué es lo que más les ha llamado la atención en el análisis computacional? ¿Qué beneficios puede aportar en la lectura de otros textos? ¿Si creen que la lectura distante pueda reemplazar la lectura cercana? ¿Qué se pierde y qué puede ganar con procedimientos computacionales? Podemos pedir que un grupo enumere las ventajas y otro las desventajas (o incluso peligros) de ese tipo de lectura. Se trata de que los alumnos se den cuenta como el proceso de recopilación, configuración y visualización de datos influye en los resultados obtenidos. Al final les pedimos posibles soluciones a esta problemática (trasparencia de los procedimientos, contrastar los resultados con interpretaciones tradicionales, etc.). Con este ejercicio nos interesa formar en ellos un acercamiento reflexivo y crítico a los medios tecnológicos y los resultados o visualizaciones que se les presenten como dato objetivo. En la medida de lo posible, se procura hacerles conscientes del sesgo, decisiones o interpretaciones previas.

#### Conclusiones

La ejemplificación de uso de *Voyant tools* lleva a constatar su utilidad como un soporte complementario al análisis tradicional en la clase de lengua y literatura. Más allá de la

gratuidad y facilidad de uso de la herramienta, su empleo permite comprender a los discentes cómo, procesando los datos textuales cuantitativos, se puede detectar estrategias lingüísticas no siempre perceptibles o evidentes en la lectura tradicional. La correlación entre las herramientas y resultados mostrados es una oportunidad para experimentar con palabras de búsqueda, formular preguntas de investigación, proponer hipótesis y realizar descubrimientos absorbentes para los jóvenes. Las diversas estadísticas y visualizaciones aportan de esta forma una visión adicional a la lectura habitual.

En segundo lugar, el ejercicio de este tipo puede potenciar las habilidades de lectura y análisis crítico en los estudiantes. El hecho de ofrecer una nueva experiencia lectora y convertir el acto de leer en una actividad más dinámica y explorativa puede despertar su interés por la lectura de los clásicos y los autores más recientes que conforman los listados de las lecturas escolares. Además, el conocimiento de los procedimientos para la lectura distante, en la época del auge de lecturas digitales, proporciona a los alumnos instrumentos para el tratamiento de cualquier texto o colección de textos (ya no únicamente literarios, sino de cualquier tipo de publicaciones, incluidas las de las webs y redes sociales). Se trata de dotarles de conocimientos y destrezas para que puedan realizar, de forma autónoma, estudios de léxico, estimaciones cuantitativas del contenido y las temáticas abordadas. Hemos presentado aquí las opciones más sencillas de análisis automático, que pueden ir progresando hacia las más complejas.

Respecto a los estudios literarios, el estudio computacional de las obras puede reforzar la interpretación tradicional, pero también descubrir tópicos menos explorados o ausentes en las clásicas guías de lectura. La experimentación con diferentes opciones de búsqueda puede originar descubrimientos de palabras funcionales, nuevos núcleos temáticos, patrones léxicos o estilísticos en los escritos examinados. Esa mirada fresca y abierta a nuevas interpretaciones sin duda es un elemento innovador, también para el propio docente y su labor didáctica.

Por último, todo lo dicho anteriormente, además de aportar una perspectiva nueva para acercarse a los textos literarios, es una excelente oportunidad para desarrollar las competencias

digitales a través del uso de nuevas herramientas informáticas y mostrar su importancia en los estudios humanísticos y literarios. En otras palabras, es una forma de acercar a los alumnos a los conceptos, procedimientos y métodos de trabajo propios de Humanidades Digitales, potenciando su formación no solo literaria, sino también científico-tecnológica. Eso puede abrir una vía para mostrar nuevas perspectivas e incluso salidas profesionales a los alumnos con talento e intereses humanísticos. Aquí está la formación de los futuros filólogos, lingüistas y humanistas digitales.

## Referencias

Arango, M. A. (1998). *Símbolo y simbología en la obra de Federico García Lorca*. Editorial Fundamentos.

Boyle, M; Hall C. (2016). Teaching Don Quixote in the Digital Age: Page and Screen, Visual and Tactile. *Hispania*, 99 (4), 600-614. <https://doi.org/10.1353/hpn.2016.0106>

Crovetto, P. L. (2010). Caballos para montar, sábanas para bordar. La mujer en los dramas rurales y en "Doña Rosita la Soltera" de Federico García Lorca, en: González de Sande, M. y López Criado, F. (ed). *La mujer en la literatura, la sociedad y la historia: identidad, cambio social y progreso en las culturas mediterráneas*, (47-56)

García Lorca, F. (2014). *La casa de Bernarda Alba*. Editorial Vicens Vives.

Gil, W. (2008), La mujer en casa cerrada: Represión y opresión en La Casa de Bernarda Alba de Federico García Lorca y Los Soles Truncos de René Marqués. *Hispanet Journal*, 1, (1-21)

Gutiérrez De la Torre, S. (2019) Análisis de corpus con Voyant Tools. <https://doi.org/10.46430/phes0043>



Hetenyi, G. Lengyel, A. Dr., & Szilasi, M. Dr. (2019). Quantitative analysis of qualitative data: Using Voyant Tools to investigate the Sales-Marketing Interface. *Journal of Industrial Engineering and Management*, 12(3), 393-404. <https://doi.org/10.3926/jiem.2929>

Huesca J. R. (2021) Poetry about the 1968 Mexican Student Movement. An Approach from Testimony, Social Imaginaries, and Digital Humanities.

Martín-Martínez, A. (2021). "Mi poesía es platónica, mi prosa es aristotélica": construcción y verdad en *Estrella distante* de Roberto Bolaño. *Chasqui*, 50(1), 79-94.

Moretti, F. *Lectura distante*. Buenos Aires: Fondo de Cultura Económica. (2015).

Murken L. (2014). *Simbología en "La Casa de Bernarda Alba"*. Grin Publishing

Ramsby, K. (2016). Text-Mining Short Fiction by Zora Neale Hurston and Richard Wright using Voyant Tools. *CLA Journal*, 59(3), 251–258. <http://www.jstor.org/stable/44325917>

Rockwell, G, Sinclair S. (2016). *Hermeneutica: Computer-Assisted Interpretation in the Humanities*. The MIT Press

Rockwell, G., Sinclair, S. *Voyant Tools*. [voyant-tools.org](http://voyant-tools.org).

Sinclair, S., Rockwell, G. (2015). Text Analysis and Visualization. Making Meaning. In Schreibman, S., Siemens, R. and Unsworth, J. (eds). *A New Companion to Digital Humanities* <https://doi.org/10.1002/9781118680605.ch19>

Villazón, N. (2006) Lorca y la mujer. *La Ratonera: Revista asturiana de teatro*, 18 (107-110)

