

SCARPACI, PABLO S.

**MODO DE ACTUALIZACIÓN TÉCNICA Y HERRAMIENTAS DIGITALES**

**MAESTRÍA EN HUMANIDADES DIGITALES**

**FECHA DE ENTREGA: 22-12-2025**

*¿Pueden determinar, con las visualizaciones realizadas, qué temas predominan?*

Creo que se puede determinar claramente que los temas predominantes giran en torno a la crónica de exploración y conquista en la región del Río de la Plata y zonas adyacentes.

Por una parte, la frecuencia con que surgen los términos *indios/yndios*, *gente*, *españoles* y, sobre todo, la aparición de términos como *pueblo* en conjunto con *guaranís* o *chiriguanas*, indica que el contacto con las poblaciones originarias es un eje central. Por otra parte, términos como *río* o *tierra*, más nombres propios asociados como *Paraná*, *Paraguay* y *Plata*, nos dan un fuerte indicio de que se hace hincapié en la descripción del paisaje y color local. Asimismo, palabras como *día*, *tiempo* y *capitán* son usuales en bitácoras y diarios de viajes, por lo que se asume que estos textos pertenecen dicho género narrativo de épocas de la conquista de América.

*¿Pueden evidenciar diferencias entre los textos del corpus?*

Si observamos las tendencias (imagen 5) y las palabras frecuentes y diferenciadas (imagen 6), podemos establecer que el corpus no es homogéneo. Mientras que Azara o Díaz de Guzmán se centran en viajes por el norte de la región, el abordaje de Hudson o Moreno, arroja términos diferentes: *Patagonia*, *mesetas*, *lago*. Esto es indicio de un desplazamiento geográfico hacia el sur de la región. También el enfoque temporal difiere en los textos: en Pigafetta se destacan términos como *maluco*, *rey* y *gentiles*; en Díaz de Guzmán, más rioplatenses (*asumpcion*). En lo que se pueden emparentar todos los autores del corpus es en el uso del registro del lenguaje específico, y el propio de su área de exploración (el *holandeses* en Du Biscay, por ejemplo)

*¿Qué palabras son las más frecuentes? ¿Cuáles luego de eliminar palabras vacías?*

Según las nubes de R y Voyant Tools, la tabla de términos de Voyant y la tabla de palabras repetidas de R, y el gráfico Berry, las palabras más repetidas luego de realizar un filtro son:

- indios (1272 apariciones)

- río (1179)
- tierra (932)
- otros (684)
- día (671)
- gente (628)
- tiempo (601)

Las palabras vacías eliminadas con la función stopwords en R y con el filtro personalizado en Voyant Tools tienen que ver con aquellas clases de palabras *funcionales* (conjunctiones y determinantes), y *clases léxicas* que no repercuten o molestan a la lectura distante: *preposiciones*, *verbos frecuentes* y palabras usadas en *fórmulas literarias* de la época. Dentro de las que no aportan valor semántico al análisis temático, las que más se repiten son **de, la, que, y, el, a**.

Vale destacar que no filtré una de estas palabras, **otro/otros**, porque me interesaba ver el sesgo que tomaba en función de teorías posteriores como la de Tzvetan Todorov, en la cual la aparición de un *Otro* en los textos de la conquista y los diarios de viaje es determinante para la relación conquistador/conquistado. El resultado que arroja esta decisión está desarrollado en la captura 2 (Términos), en donde se realiza un montaje de los términos que más aparecen y en relación con qué otros términos. En el caso de **otros**, vemos que entre las colocaciones principales está *indios* (40), y dentro de las correlaciones, encabeza también *indios* (1272)

Por lo tanto, un término que podría haberse filtrado como "vacío" por ser un determinante, adquiere una trascendencia que podría utilizarse para abordar este corpus desde el *Problema del Otro* de T. Todorov.

*¿Qué hipótesis inicial podrían formular?*

Tomando el gráfico de enlaces de Voyant Tools como aporte ilustrativo a esta idea, se puede esbozar la hipótesis de que este corpus de crónicas coloniales y bitácoras de viajeros del siglo XVI al XIX, tiene como constante la documentación geográfica (principalmente hídrica), y la observación de la organización de poblaciones nativas. Lo que se puede intuir de funciones como la línea de burbujas es que hay una tendencia en los textos a investigar el territorio para el asentamiento español y la explotación de los recursos naturales. En mi opinión, el análisis sugiere al río como vía de comunicación, y al indio como el sujeto de observación, con una vocación evangelizadora o de alianza. Una

visión hispanista sugeriría que no hay ningún término con connotación negativa referida al indio por parte de los españoles que refleje ninguna de las dos herramientas.

*¿Qué pueden decir respecto a este sencillo análisis comparativo entre R y Voyant Tools?*

**Voyant Tools** es muy recomendable para el abordaje de corpus de manera expeditiva e interactiva. Incluso permite abordar relaciones complejas como las redes de enlaces, tendencias temporales y distributivas y contextos de uso. Todo esto nos permite trabajar con lectura distante sin saber programar, y con atractivos patrones visuales. Por otra parte, **R** parece más “estática”, pero en caso de necesitar un análisis estadístico más profundo, sería más precisa (porque trabajamos sobre el código). El gráfico de barras que realicé, por ejemplo, es muy sencillo, pero me llevó semanas poder llegar al mismo, gracias a no tener experiencia en programación. Lo que pude comprobar en este trabajo es que ambas son herramientas complementarias si uno quiere validar con una “segunda opinión” un análisis de lectura distante y minería de datos orientada a las humanidades.