

# Hacia un mapa de discursos: análisis lexicométrico y de contenido de un corpus de la colección editorial Gerardo Molina

## Problema



Estudiar las colecciones editoriales universitarias se vuelve una vía para investigar y problematizar los discursos constituidos en un campo disciplinar y en un espacio institucional. Interrogar una colección editorial universitaria desde un horizonte discursivo debería, entonces, ayudar a revelar las apuestas y características que la han moldeado. Sin embargo, dado que por definición una colección es un objeto extenso, un análisis discursivo sistemático debería preguntarse por metodologías pertinentes que favorezcan el procesamiento de la información y el entrecruzamiento de métodos.

El análisis lexicométrico es un enfoque funcional para resolver inicialmente esta necesidad: es una operación mediada por software que revela datos cuantitativos de un corpus de información digital, evidenciando fenómenos lingüísticos como frecuencias lexicales, coocurrencias, clasificaciones jerárquicas, análisis de similitudes, regiones lexicales, entre otros. Sin embargo, si bien se establecen indicios, es importante interrogarlos desde perspectivas más complejas o relacionales, como sucede con los análisis de contenido y los análisis del discurso.

## Pregunta de investigación

¿Cómo la articulación de un análisis lexicométrico y un análisis de contenido sobre un corpus editorial de la colección Gerardo Molina permite revelar singularidades discursivas en el campo de las ciencias políticas y sociales?

## Marco teórico

**Colección editorial.** Una colección editorial es un objeto constituido a partir de la adscripción de múltiples libros, vinculados bajo una declaración de intención.

**Análisis lexicométrico.** Se entiende como un procedimiento estadístico de análisis de textos para el estudio de un corpus (naturalmente, digital o digitalizado).

**Análisis de contenido.** Es una técnica de investigación que permite describir sistemáticamente contenidos de un corpus textual, por ejemplo, identificando patrones o temas recurrentes que permitan asignar significados. Su propósito es hacer explícitos los contenidos de los mensajes.

**Corpus.** García y García (2011) lo definen como una "colección de textos, almacenados en formato electrónico, seleccionados y organizados de acuerdo con criterios explícitos y para propósitos lingüísticos."

## Resultados



Se parte del supuesto de que los índices lexicales son reveladores de conceptos y categorías disciplinares del campo de las ciencias sociales materializados en el contexto de la colección.

El primer resultado fue el la frecuencia de términos y las nubes de palabras, un dato cuantitativo que permite identificar las unidades lexicales más representativas en el corpus de análisis y, por extensión, un indicio de los contenidos y las materialidades discursivas que informan el corpus construido de la colección.

, un resultado mucho más dicente respecto a las características temáticas y discursivas del corpus es el que se revela a través de un análisis de similitud (ADS), que identifica coocurrencias de palabras y sus conexiones, de forma tal que se perfilan ciertas estructuras y comunidades lexicales.

Usando el mismo software, ahora desde su función más potente: la clasificación jerárquica descendente, o método Reinert, se construye un dendograma a partir de los "mundos lexicales" representativos del corpus.

La potencia de los hallazgos lexicales va más allá de su revelamiento lingüístico: es una enunciación discursiva que revela también representaciones y prácticas que desbordan el plano lingüístico y dan cuenta de apuestas y posicionamientos temáticos y discursivos.

## Conclusiones

• Versus el carácter restringido de los libros universitarios impresos, accesibles a pocos lectores e instituciones por los circuitos ortodoxos, los formatos digitales permiten amplificar no solo el mercado de lectores, sino también dar apertura a otros tipos de lecturas y procesamientos de información valiosos para investigaciones interdisciplinarias mediadas por lo tecnológico. Es el caso de las posibilidades que ofrece en el análisis lexicométrico y de contenido para procesar estadísticamente corpus de información, codificarlos en perspectiva cualitativa y ponerlos en diálogo con apuestas de análisis discursivo más amplio.

• El presente artículo apuesta por una visión marcadamente metodológica en este sentido y sirve de punto de partida para un desarrollo más profundo que articule diversas metodologías de análisis de discurso, cubriendo incluso el horizonte del análisis crítico del discurso, una herramienta pertinente para el campo de las ciencias políticas y sociales, donde apuestas como la resistencia frente al poder, la defensa de los derechos humanos, el ejercicio político o la acción de los movimientos sociales ciertamente están en el centro de sus debates e investigaciones.