

Análisis estilométrico en Python aplicado a epistolarios de la premio nobel de literatura Gabriela Mistral

Stylometric analysis in Python applied to epistolaries of the nobel prize in literature, Gabriela Mistral

Nataly Ruiz-Cornejo^{1*} Mauricio Arriagada-Benítez²

Recibido 06 de enero de 2022, aceptado 08 de mayo de 2022

Received: January 06, 2022 Accepted: May, 08, 2022

RESUMEN

El presente trabajo es un estudio estilométrico que utiliza como corpus una muestra conformada por epistolarios de la poetisa chilena y premio Nobel Gabriela Mistral, dirigidos a Pedro Prado, Manuel Magallanes Moure y Doris Dana. Este estudio es una réplica del ejercicio propuesto por François Dominic Laramée “Introducción a la estilometría con Python” en el cual se utilizan 3 enfoques estilométricos. En este estudio se validan las 3 metodologías y se añade la utilización de una nube de palabras, para reconocer estilos y tendencias de la poetisa a través de la estilometría, donde se puede establecer y determinar la naturaleza de la relación que existía entre Mistral con Pedro Prado, Manuel Magallanes y Doris Dana. Esta investigación contribuye a diferentes áreas del conocimiento que tengan interés de continuar develando la figura e imaginario de Gabriela Mistral, a través de algún análisis automático utilizando herramientas de programación y ciencias de datos.

Palabras clave: Estilometría, Gabriela Mistral, epistolarios, nube de palabras, Python.

ABSTRACT

The present work is a stylometric study that uses a sample of letters from the Chilean poet and Nobel Prize winner Gabriela Mistral addressed to Pedro Prado, Manuel Magallanes Moure, and Doris Dana as a corpus. This study replicates the exercise proposed by François Dominic Laramée, “Introduction to Stylometry with Python,” in which 3 forms of stylometric studies are used. This study validates the 3 methodologies. Additionally, the study complements the analysis by incorporating a word cloud to recognize the styles and tendencies of the poet through stylometry, where the nature of the relationship that existed between Mistral with Pedro Prado, Manuel Magallanes, and Doris Dana can be established and determined. This research contributes to different areas of knowledge that are interested in continuing to reveal the figure and imagination of Gabriela Mistral through some automatic analysis using programming tools and data science.

Keywords: Stylometry, Gabriela Mistral, letters, word cloud, Python.

¹ Pontificia Universidad Católica de Chile. Santiago, Chile. Email: nruizc@uc.cl

² Universidad de Tarapacá. Departamento de Ingeniería en Computación e Informática, Facultad de Ingeniería. Arica, Chile. Email: mauricio.arriagada@academicos.uta.cl

* Autores de correspondencia: nruizc@uc.cl, mauricio.arriagada@academicos.uta.cl

INTRODUCCIÓN

El presente trabajo es un estudio estilométrico en donde se utilizan como objeto de estudio 3 epistolarios escritos por Gabriela Mistral [1], material interesante de investigar en el cual se pueden estudiar sus expresiones afectivas y/o amorosas hacia otro ser en un contexto íntimo.

Las razones por las que se seleccionó a esta autora son su figura relevante como escritora a nivel nacional e internacional, fue la primera mujer latinoamericana en ganar el Nobel de Literatura, además de ser una intelectual de avanzada, ya que se interesó y publicó sobre muchísimos temas vedados en su época.

Los destinatarios de los epistolarios seleccionados son Manuel Magallanes Moure, Pedro Prado y Doris Dana. Fueron escogidos debido a las características de los lazos que Mistral mantenía con ellos, basado en las biografías escritas en torno a su figura y el velo que se buscaba revelar en torno a la naturaleza de la relación que existía entre la Mistral y Doris Dana. Para esto se utilizaron herramientas de ciencias de datos y estilometría para contribuir al debate y la reconstrucción de la biografía e imaginario en torno a Gabriela Mistral.

Analizar textos requiere de una gran cantidad de horas invertidas en lectura. Comparar textos implica un esfuerzo similar o mayor, por cuanto se debe tener un amplio conocimiento del tema y memoria, para inferir conclusiones asertivas. En este sentido, el uso de herramientas de ciencias de datos permite apoyar estos estudios a través de la estilometría [2] automatizando procesos e incorporando métricas que permiten analizar grandes volúmenes de datos en un menor tiempo, apoyando al experto en investigaciones científicas [3].

El presente trabajo y sus resultados, plantean 3 aportes que lo caracterizan para ser útil en diferentes aspectos: Entregar un modelo de uso de estilometría en cartas manuscritas, trabajar con una muestra real de manuscritos de un autor nacional y por último ser el primer modelo estilométrico basado en cartas manuscritas pertenecientes a Gabriela Mistral.

ESTADO DEL ARTE

Un epistolario es un conjunto de cartas, escritas por o para una misma persona, en él es posible observar

aspectos desconocidos del autor respecto a su vida, relaciones personales, emociones, pensamientos, etc. Un ejemplo de esto, ligado a los 3 epistolarios estudiados, es la factibilidad de indagar en la evolución de la identidad de género que se expresa en Mistral a lo largo de sus epístolas y cómo podría cambiar la biografía oficial de esta. Otro aporte que genera la revisión de sus epistolarios, es el acceso a leer sobre sus sentimientos, sus preocupaciones, sus temas de interés, incluso opiniones de la Mistral que podrían servir para complementar la visión que tenemos de ella e incluso generar nuevas publicaciones.

Para estudiar el material escrito, específicamente los epistolarios, se pensó de una manera automatizada de tal propósito, ya que debía ser transcrito y luego analizado. La técnica propuesta para el análisis es a través de estilometría, la cual se resume como una metodología estadística para analizar textos a través de la frecuencia del uso de palabras [4, 5], además, la estilometría analiza ciertos rasgos del estilo del autor y los utiliza para comparar dos o más textos [6]. El punto de base de la estilometría, es que el estilo es algo que nace en el subconsciente, y por esta razón, cada quien tiene su estilo propio, por lo anterior cada autor en su trabajo deja una huella lingüística que lo identifica y diferencia de otros, por lo cual cada estilo es único. La estilometría es una forma de analizar textos en diferentes niveles, algunas de sus aplicaciones son determinar la autoría de una obra, la autenticidad, clasificación de textos, medición de frecuencia de palabras e identificación de lenguas. El plus de la estilometría es el análisis del estilo de un autor a través de la recuperación de datos estadísticos, por lo que aúna diferentes áreas del conocimiento como son las letras y la estadística. Hoy en día, el lenguaje de programación Python, se utiliza en el análisis de texto a través de diversas librerías, siendo una de las más importantes NLTK (Natural Language Toolkit) [7].

Respecto a la característica del material de análisis, el trabajo con epistolarios es un aporte para estudiar al autor desde su espacio privado ya que hay palabras, trazos y conexiones léxicas que van más allá de su obra y que plasma cuando se comunica, en un epistolario podemos ver la huella del estilo del autor.

Middleton [8], indica que “lo valioso de un estilo personal dependerá de que sea o no la expresión

de un auténtico sentimiento individual” y que lo que parezca artificial en la obra del autor será un triunfo del arte, entendiendo la artificialidad como el trabajo del autor quien, a pesar de que utilice un lenguaje natural para expresarse, lo hace de una manera original, permitiéndole comunicar sentimientos y crear mundos a través de sus herramientas únicas e individuales rozando el arte y la creación, a diferencia del concepto artificial que tiene la intención de destacar pero con el anhelo egocéntrico de alcanzar la perfección. Existe una delgada línea entre estos artificios ya que, indica Middleton Murry “Todo estilo es artificial en el sentido de que todos los buenos estilos se logran por artificio”, aún así destacan los buenos estilos cuando se logra la naturalidad en la creación, ya que se entrelaza lo intelectual con lo emotivo.

La poesía es conocida como un medio escrito para expresar las emociones personales, y una carta es un escrito que una persona dirige a otra, para comunicar y/o expresarse.

Este trabajo es un aporte para estudiar a Gabriela Mistral desde un ángulo relativamente nuevo, el cual comenzó oficialmente con los estudios realizados a la muestra tangible del epistolario dirigido a Doris Dana, aquellos tratan de la homosexualidad de Mistral. Este estudio aporta con una mirada estadística a aquellos estudios entregando una visión objetiva (datos), respecto a la relación de Mistral con Doris Dana comparados con los epistolarios de Magallanes Moure y Pedro Prado. El resultado del trabajo puede tomarse para reconstruir su figura

apoyándose de la tecnología, para comprender a la Mistral no solo como la figura en el ambiente literario nacional, quien fue la primera chilena en ganar el Nobel de Literatura, dato relevante para una mujer que por su género no tuvo la relevancia merecida en su país natal, sino también sirve para rescatar y destacar su figura de reformista, de maestra, su figura como ente político, y su legado a la identidad chilena, ya que su patrimonio es un aporte a la historia de Chile.

MATERIALES Y MÉTODOS

Para poder analizar las cartas de los autores referenciados en esta investigación, se siguió el esquema mostrado en la Figura 1.

Se puede apreciar que consta de 4 etapas, las cuales se explican a continuación:

- **Selección de los epistolarios:** Considera las cartas de Gabriela Mistral a Magallanes Moure, Doris Dana y Pedro Prado. Elaboración propia.
- **Creación de los Corpus:** Esta etapa considera los pasos de seleccionar las cartas, revisar el estado de ellas, transcribir y validar el texto con expertos del tema. Elaboración propia.
- **Desarrollo y aplicación de técnicas para análisis de texto:** Se usa el lenguaje de programación Python como principal herramienta de análisis estilométrico, y otras herramientas complementarias para apoyar el estudio. Basado en el estudio de Laramée.
- **Análisis de los resultados:** Esta última etapa considera la interpretación de los resultados

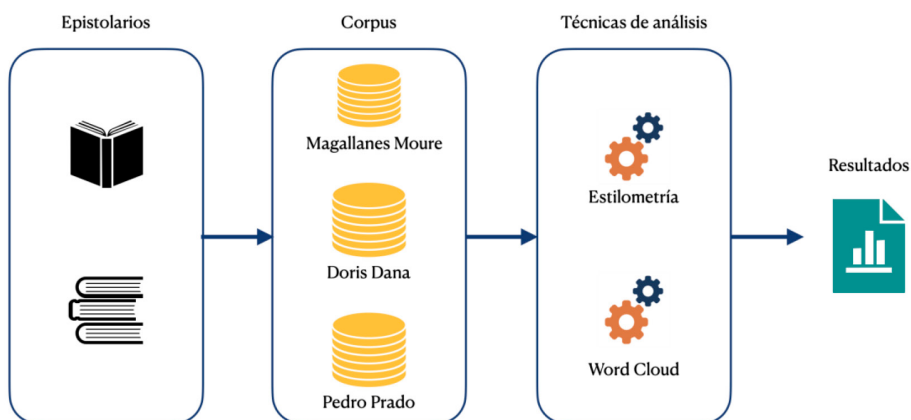


Figura 1. Esquema de la metodología desarrollada.

obtenidos en la etapa de desarrollo, desde dos perspectivas: Análisis de las métricas obtenidas y la interpretación literaria que se deriva del estudio de análisis de texto computacional. Elaboración propia.

A continuación, se describen las etapas presentadas en la metodología (ver Figura 1):

Transcripción de las cartas

Para esta investigación se utilizó una muestra de 84 cartas manuscritas y mecanografiadas por Gabriela Mistral [9], la cual conforman el corpus general (el cual está compuesto por 3 corpus pertenecientes a tres diferentes destinatarios). Parte de la muestra de cartas utilizadas en esta investigación pertenecen a la Colección Patrimonial de la Biblioteca de Humanidades, de la Pontificia Universidad Católica, estos documentos se mantienen resguardados bajo estrictos estándares de preservación. Este material tiene la característica de ser único, ya que en su mayoría está escrito por puño y letra de la poetisa. El resto de cartas se extrajo de los epistolarios publicados por Ediciones UC, Lumen y Dolmen que también son parte de su colección [9, 10, 11].

1. **Corpus 1:** Epistolario de Manuel Magallanes Moure (37 cartas), dirigido al poeta entre los años 1915 y 1923. Las primeras 24 cartas fueron transcritas por Amalia Redondo Magallanes, nieta de Magallanes Moure. La Figura 2 muestra un resumen de las 37 cartas y un total de 21.530 palabras.
2. **Corpus 2:** 17 cartas dirigidas al escritor Pedro Prado Calvo, ganador del Premio Nacional de Literatura de 1949. La Figura 3 muestra un resumen de las 17 cartas y un total de 5.521 palabras.
3. **Corpus 3:** El tercer epistolario es una muestra de 30 cartas dirigidas a Doris Dana que se recopilaron del texto *Niña Errante: Cartas a Doris Dana*. Su transcripción no se realizó a mano, ya que se utilizaron los siguientes recursos detallados paso a paso:
 - a. Digitalización de las páginas del libro a PDF, 300 ppp.
 - b. Se cambiaron a formato de texto (Word) utilizando el software Acrobat Reader.
 - c. Cambio de formato a texto (txt).

La Figura 4 muestra un resumen de las 30 cartas y un total de 12. 114 palabras.

MISTRALMAGALLANES

user/nruizc/mistralmagallanes

MANAGE CORPUS

Cartas de Mistral a Magallanes

GENERAL INFO

Language	Spanish
Tagset	Description
Word sketch grammar	Description

COUNTS

Tokens	25,561
Words	21,530
Sentences	1,443
Documents	37

COMMON TAGS

noun	N.*
verb	V.*
adjective	A.*
adverb	R.*
pronoun	P.*

LEMPOS SUFFIXES

noun	-n
verb	-v
adjective	-j
adverb	-r
pronoun	-p

Figura 2. Cartas de Gabriela Mistral a Magallanes Moure.

MISTRALPRADO

user/nruizc/mistralprado

MANAGE CORPUS

Cartas entre Mistral y Prado / Amistad

GENERAL INFO

Language	Spanish
Tagset	Description
Word sketch grammar	Description

COUNTS

Tokens	6,570
Words	5,521
Sentences	324
Documents	17

COMMON TAGS

noun	N.*
verb	V.*
adjective	A.*
adverb	R.*
pronoun	P.*

LEMPOS SUFFIXES

noun	-n
verb	-v
adjective	-j
adverb	-r
pronoun	-p

Figura 3. Cartas de Gabriela Mistral a Pedro Prado.



The screenshot shows the MISTRALDANA web interface. At the top, it says 'MISTRALDANA' with a user ID 'user/ruizc/mistraldana' and a 'MANAGE CORPUS' button. Below this, it says 'Cartas de Mistral a Doris Dana'. The interface is divided into four main sections: GENERAL INFO, COUNTS, COMMON TAGS, and LEMPOS SUFFIXES.

GENERAL INFO	
Language	Spanish
Tagset	Description
Word sketch grammar	Description

COUNTS	
Tokens	14,460
Words	12,114
Sentences	1,052
Documents	30

COMMON TAGS	
noun	N *
verb	V *
adjective	A *
adverb	R *
pronoun	P *

LEMPPOS SUFFIXES	
noun	-n
verb	-v
adjective	-j
adverb	-r
pronoun	-p

Figura 4. Carta de Gabriela Mistral a Doris Dana.

Finalmente, cuando todas las cartas del corpus se transformaron a formato txt se les agregaron los datos de fecha y lugar desde donde fueron escritas con el fin de rescatar datos georeferenciales que puedan ser utilizados en un trabajo de investigación posterior, los datos se plasmaron en el siguiente orden: Lugar, año, mes. El criterio de ubicar el mes al último fue debido a que lugar y año siempre estaban presentes, a diferencia del mes que solo aparecía en algunas cartas y así se podía obtener la información 0/1/2 (opcional).

Estudio de los datos

Se utilizó el software Sketchengine para probar una herramienta nueva a disposición, y por su facilidad de trabajar los datos para analizar los listados de palabras en los corpus (ver Figura 2, Figura 3 y Figura 4).

El análisis estilométrico se basó en el ejercicio propuesto por François Dominic Laramée, titulado “Introduction to stylometry with Python” [12, 13].

Se selecciona Python como herramienta de programación por la gran cantidad de librerías para aplicar ciencia de datos, entre ellas la librería llamada Natural Language Toolkit (NLTK). Además, se utiliza la librería Matplotlib para visualización.

Para replicar el estudio de Laramée, se trabajó en tres etapas, cada una con un enfoque estilométrico diferente:

1. **Curvas características de composición de Mendenhall** [14]: Técnica con más de cien años de antigüedad que sentó las bases para estudios estilométricos el cual se basa en el conteo del largo de palabras. Hoy en día, no es una técnica que sea considerada precisa para este tipo de

estudios. Sin embargo, se utiliza, tal como dice Laramée, para tener una visualización rápida de análisis de corpus.

2. **Método Chi-cuadrado de Kilgariff** [15]: Técnica que mide la distancia de los vocabularios entre corpus.
3. **Método Delta de John Burrows** [16]: Está diseñado para comparar un texto anónimo (o un conjunto de textos), con las firmas de muchos autores al mismo tiempo. Laramée explica que Delta mide cómo el texto anónimo se diferencia de los conjuntos de textos escritos por un número x de autores.

ANÁLISIS DE RESULTADOS

La primera técnica de análisis es la de Mendenhall para ver las curvas características, tal como lo indica la Figura 5:

En los gráficos se puede apreciar, a pesar de la diferencia de longitud debido a la disparidad de textos por corpus; que, en el sector izquierdo del gráfico, todos los documentos que representan las longitudes de palabras más frecuentes comienzan de manera similar, sin embargo, cabe destacar que las curvas de Dana y Magallanes tienen una curvatura similar en el sector central izquierdo y que además el final de sus colas termina de la misma manera. Esto es consistente con las últimas publicaciones dedicadas a la Mistral, como el documental de María Elena Wood [17], y el texto de Fiol-Matta [18], quienes indican que a través del epistolario de Gabriela Mistral se puede acceder a su inconsciente, se pueden conocer otros ángulos de su personalidad, se pueden encontrar atisbos para cuestionar biografías anteriores, ya que coinciden en el lesbianismo de la Mistral y en la naturaleza de su relación con Doris Dana [19].

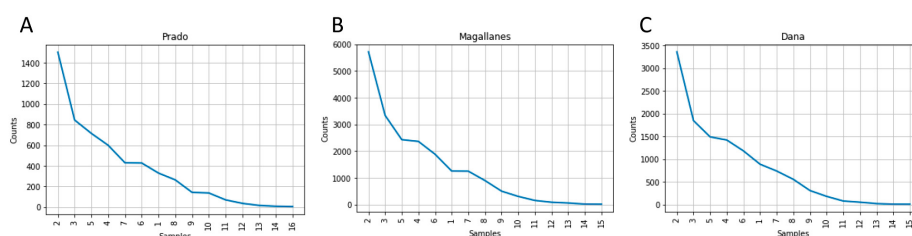


Figura 5. Curvas características de composición de Mendenhall.

La segunda prueba estilométrica consiste en aplicar el método de distancia chi-cuadrado de Kilgariff. Para observar a qué destinatario se asimila más el epistolario dedicado a Dana, se procede a compararlos con los corpus de Pedro Prado y Magallanes Moure dando como resultado la siguiente Tabla 1.

Al estudiar las cartas escritas por Mistral a Dana la prueba de Chi cuadrado nos indica que se asemeja más al set de datos enviado a Pedro Prado para frecuencias de 300 y 500 palabras más comunes en los corpus. Para una frecuencia de 100 palabras más comunes en el corpus, la semejanza de estilo entre los epistolarios escritos a Dana son más cercanos a los epistolarios escritos a Magallanes.

Esta técnica no esclarece, ni nos permite asegurar, la naturaleza de la relación que Mistral sostuvo con sus destinatarios. Es de esperarse que mientras mayor sea la cantidad de palabras más comunes, mejor es el análisis. Sin embargo, no es correcto que exista una cercanía entre los epistolarios enviados a Dana y Prado. Esto se debe a que las palabras con alta frecuencia inciden significativamente en el cálculo de esta métrica, donde se deja de lado la importancia de las palabras menos comunes.

En la tercera prueba de John Burrows se estudiaron 29 conectores, conocidas como palabras de función, propuestos en el ejercicio de Laramée, y traducidos

Tabla 1. Comparación de estilos de escritos entre los epistolarios de Dana y {Magallanes,Prado} usando método Chi-square.

Destinatario	χ^2		
	f = 100	f = 300	f = 500
Pedro Prado	1378,892	2258,501	2794,871
Magallanes Moure	1265,000	2483,248	3331,203

f = frecuencia de palabras más comunes en corpus.

al español. El resultado se muestra en la siguiente Tabla 2.

El valor Delta es menor entre los epistolarios de Magallanes y Dana (3,231), por ende se acerca más al punto de relación, lo que significa que las cartas escritas a estos dos destinatarios se asemejan (este resultado es acorde a lo indicado en las biografías en torno a la naturaleza de la relación que existía entre Mistral y ellos), en cuanto al corpus cuyo destinatario es Pedro Prado, al tener un delta mayor (3.898) nos indica la existencia de una mayor diferencia en el estilo de escritura.

Según lo calculado, podemos ver que el costo computacional del Chi Square es lineal y el del delta score es cuadrático, lo que significa que el computador se demora 1 segundo en analizar las 10.000 primeras palabras, en cambio en el tercer ejercicio el computador se demoraría 10.000 segundos, ya que demoraría 10.0000 veces más, esto también refleja que para este ejemplo la tercera prueba estilométrica, la avanzada, extrajo mejores resultados.

El segundo set de estudio extrajo deltas más bajos, lo que significa que se obtuvieron resultados más acertados que en el 1º set, visibilizando que fue mejor utilizar más cartas en la muestra que solo una carta.

El ejercicio propuesto por Laramée es replicado en su totalidad en este estudio, adaptando los parámetros del algoritmo propuesto. Sin embargo Laramée indica

Tabla 2. Comparación de estilos de escritos entre los epistolarios de Dana y {Magallanes, Prado} usando método Delta.

Destinatario	Δ_c
Pedro Prado	3,898
Magallanes Moure	3,231

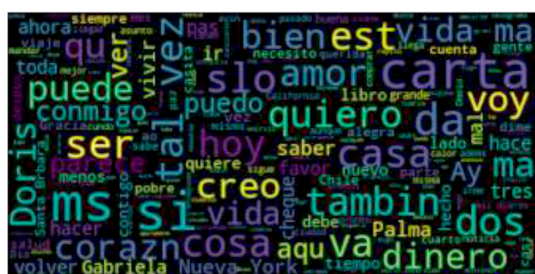


Figura 8. Word cloud del corpus de Doris Dana.

Cabe destacar que en el epistolario dirigido a Doris Dana aparecen términos enfocados a un género masculino, intencionalidad en la redacción que no se observó en los otros epistolarios, un gran ejemplo de esto es la siguiente frase que demuestra un tránsito desde lo femenino a lo masculino en Gabriela Mistral: “Duerme, mi amor, Dios te cure de tu dolencia. Perdona el que te he herido, por no creerme amado, por pensarme postergado en tu corazón. (Sigo mañana.) Te beso, tuyo,” México, 1949 [10].

Análisis extraídos a partir de la nube de palabras

Respecto al contenido de los corpus visualizados en los Word clouds se puede apreciar que a pesar de explicitar en el código y en NLTK que los corpus eran en español, el programa obvió los tildes y es por esto que no aparecen las letras que tuviesen un tilde dentro de la palabra.

En los 3 corpus aparecen los nombres de los destinatarios, lo que indica que Mistral comúnmente se dirigía a sus destinatarios por su nombre de pila.

Podemos observar que en los sub corpus de los epistolarios destinados a Magallanes Moure y Doris Dana se repiten los términos vida y amor, a diferencia del epistolario destinado a Prado en donde no se observan aquellos términos ligados a la emocionalidad o a un lazo más íntimo, pero si la palabra amigo, término no observado en los otros corpus.

Podemos indicar que la nube de palabras apoya lo arrojado en los resultados del ejercicio, en donde relaciona y asimila los corpus de Magallanes y Dana, ya que además de la tendencia de los conectores

estudiados, también nos podemos apoyar por el uso de ciertas palabras claves al comparar los términos utilizados en los corpus.

Análisis general a partir de los resultados

Los métodos propuestos por Laramée, a través de los estudios estilométricos, fueron aplicables y replicables en esta investigación, el método Delta fue la técnica más útil para estudiar similitud de estilo entre subcorpus. Además, se pudo revelar aspectos importantes de los contenidos de los epistolarios de Gabriela Mistral.

También es posible corroborar, gracias a la aplicación de este análisis estilométrico, rasgos estilísticos asociados a dos tópicos literarios concretos: Uno referido al amor pasional (o “Eros”), en donde situamos los epistolarios dirigidos a Manuel Magallanes y Doris Dana y el otro hacia la amistad (el equivalente al concepto griego de “Philos” relacionado al aprecio por alguien), en donde claramente se sitúan las cartas dirigidas a Pedro Prado el “distinguido amigo” de Gabriela Mistral.

CONCLUSIONES

Luego del proceso de revisión del corpus, compuesto por tres subcorpus (epistolarios), de Gabriela Mistral destinados a Pedro Padro, Manuel Magallanes y Doris Dana, utilizando Python y aplicando metodologías de Ciencias de Datos se puede concluir que el método Delta es el más apropiado para analizar la documentación mencionada, este método sirvió para comprobar lo propuesto por Laramée en su artículo Introduction to stylometry with Python. Con este ejercicio se demuestra que su propuesta es replicable para identificar similitud de escritos, y en este trabajo en particular, al realizar los ejercicios de comparación de corpus y ajuste de parámetros (método 2), el método delta es capaz de identificar destinatarios y similitud de estilo entre los subcorpus.

A través de las cartas, se puede apreciar la evolución de Gabriela Mistral como persona y las diferencias de estilo en su literatura epistolar, la cual se modifica acorde al destinatario y al paso del tiempo; no es la misma forma de expresión cuando escribe a Magallanes Moure o Doris Dana que cuando escribe a Pedro Prado, ni su identidad plasmada en las cartas escritas en 1916 comparada con las redactas durante el 1955 en donde se aprecia un cambio en

su identidad sexual el cual no fue visibilizado en las primeras biografías y/o quizás ocultado.

Gracias a la aplicación del análisis estilométrico se pueden reconocer diferencias de estilo en el corpus textual seleccionado de Gabriela Mistral, como lo es el denominar a sus destinatarios por el nombre de pila o dejar muy en claro, a través del término utilizado para dirigirse a su destinatario, la naturaleza de su relación, específicamente a Manuel Magallanes lo trata directamente de Manuel, a Doris Dana la llama por Doris, Dorisin o Cara, y a Prado se dirige como Pedro Prado, amigo o la frase “distinguido amigo”, frase ausente en los dos epistolarios restantes en donde destacan y se repiten los términos: Vida, conmino y amor.

El aporte de este trabajo es que podemos considerarlo como el primer estudio metodológico de análisis estilométrico aplicado a epistolarios de Gabriela Mistral. En donde se estableció y determinó la naturaleza de la relación que existía entre Mistral con Pedro Prado (amistad), contrastándola con la similitud de la relación epistolar mantenida con Manuel Magallanes Moure y Doris Dana, los cuales son epistolarios de carácter amoroso, aportando al debate y a la reconstrucción de la biografía e imaginario en torno a Gabriela Mistral.

Mediante este trabajo se aportan datos para aumentar las cifras de estudios estilométricos latinoamericanos, un aporte chileno para las Humanidades Digitales, el cual se caracteriza por estar enfocado en una figura chilena y latinoamericana, y a su vez ser un modelo de trabajo en español, utilizando epistolarios, basado en la aplicación de tecnologías de estudios estilométricos.

AGRADECIMIENTOS

Extendemos nuestros agradecimientos a la bibliotecóloga Javiera Bravo, quien aportó con sus conocimientos en las transcripciones y letras manuscritas de Gabriela Mistral.

REFERENCIAS

- [1] A.M. Moraga. “Gabriela Mistral: Álbum personal”. Pehuén. Santiago, Chile. 2008.
- [2] U. Kuckartz. “Qualitative Text Analysis: A Guide to Methods, Practice & Using Software”. Sage. Los Ángeles, USA. 2014.
- [3] R. Sarawgi, K. Gajulapalli and Y. Choi. “Gender attribution: Tracing stylometric evidence beyond topic and genre”. 15th Conference on Computational Natural Language Learning, CoNLL, pp. 78-86. Stroudsburg, USA. 2011.
- [4] D.I. Holmes. “The Evolution of Stylometry in Humanities Scholarship”. Literary and Linguistic Computing. Vol. 13 N° 3, pp. 111-117. 1998.
- [5] S.L. Bucklow. “A stylometric analysis of craquelure”. Computer and the Humanities. Vol. 31 N° 6, pp. 503-521. 1998.
- [6] J. Savoy. “Machine Learning Methods for Stylometry”. Springer. Cham, Switzerland. 2020.
- [7] J. Perkins, D. Chopra and N. Hardeniya. “Natural language processing: Python and NLTK”. Packt Publishing. Birmingham, Inglaterra. 2016.
- [8] J. Murry. “El estilo literario”. Fondo de cultura económica. México. 2014.
- [9] G. Mistral. “Manuel, en los labios por mucho tiempo: Epistolario entre Lucila Godoy Alcajaga (Gabriela Mistral) y Manuel Magallanes Moure”. Ediciones UC. Santiago, Chile. 2005.
- [10] G. Mistral. “Niña errante. Cartas a Doris Dana”. Lumen. Santiago, Chile. 2009.
- [11] A.C. Flores, L. Vargas, M.E. Martínez y R. Valdés. “En batalla de sencillez: Epistolario de Gabriela Mistral a Pedro Prado”. Dolmen. Santiago, Chile. 1996.
- [12] F.D. Laramée. “Introduction to stylometry with Python”. Programming Historian. 2018. Fecha de consulta: 05-08-2021. URL: <https://programminghistorian.org/en/lessons/introduction-to-stylometry-with-python>
- [13] E. Loper and S. Bird. “NLTK: The natural language toolkit”. Department of Computer and Information Science. University of Pennsylvania. Philadelphia, USA. DOI: 10.3115/1118108.1118117. 2016.
- [14] D. Wackerly, W. Mendenhall and R.L. Scheaffer. “Mathematical statistics with applications”. Brooks Cole. Boston, USA. 2008.
- [15] A. Kilgariff. “Comparing word frequencies across corpora: Why chi-square doesn’t work, and an improved LOB-Brown comparison”. ALLC-ACH Conference, pp. 169-172. 1996.

- [16] J. Burrows. “‘Delta’: A measure of stylistic difference and a guide to likely authorship”. *Literary and Linguistic Computing*. Vol. 17 N° 3, pp. 267-287. 2002.
- [17] M.E. Wood. “Locas mujeres”. Wood Producciones. Santiago, Chile. 2011.
- [18] L. Fiol-Matta. “A queer mother for the nation: The state and Gabriela Mistral”. University of Minnesota Press. Minneapolis, USA. 2002.
- [19] H.M. Garrido. “Gabriela Mistral y su transgresión poética”. *QVADRATA. Estudios sobre Educación, artes y humanidades*. Vol. 3 N° 5, pp. 65-74. 2021.
- [20] F. Heimerl, S. Lohmann, S. Lange and T. Ertl. “Word cloud explorer: Text analytics based on word clouds”. 2014 47th Hawaii International Conference on System Sciences, pp. 1833-1842. 2014. DOI: 10.1109/HICSS.2014.231
- [21] M.C. Taylor. “Sensibilidad Religiosa de Gabriela Mistral”. Gredos. Madrid, España. 1975.
- [22] P. Zegers. “Yin Yin”. Ediciones Universidad Diego Portales. Santiago, Chile. 2015.