



Anonimato, pseudonimia y delitos en las redes sociales: una propuesta multidimensional de la lingüística forense para la identificación de autoría

Anonymity, pseudonymity and crimes in social networks: a multidimensional proposal of forensic linguistics for the identification of authorship

MARÍA VALENTINA NOBLIA

Universidad de Buenos Aires
Argentina

ALEJANDRO CARLOS RENATO

Universidad de Buenos Aires
Argentina

TOMÁS GERSHANIK

Universidad de Buenos Aires
Argentina

Recibido: 21 de diciembre de 2021 | Aceptado: 2 de marzo de 2022

DOI: 10.35956/v.22.n1.2022.p.122-142

RESUMEN

Una de las primeras apreciaciones del impacto de Internet sobre el mundo social fue la creación de comunidades virtuales, que permitieron congregarse personas dispersas, distantes en el espacio, pero unidas por propósitos comunes (Noblia 2000 a y b). Esta potencialidad de vincular y fomentar grupos de distinta índole, también habilitó la agrupación de individuos aunados por la “cultura del odio” (Zickmund 1997; Dery 1994). Bajo el amparo de la libertad de expresión, del anonimato/pseudonimia, grupos e individuos hacen uso de las redes sociales con el fin de difundir sus discursos radicalizados y sus programas violentos. Las dinámicas de las redes, pautada por “haters”, los ayuda a circular y convocar a nuevos adherentes.

El accionar de estos grupos es un foco de preocupación para los organismos gubernamentales a cargo de fiscalizar delitos de esta índole y prevenirlos, porque, dadas las características y ámbitos de circulación, su detección temprana y su desarticulación es una tarea compleja. En esta labor, la actuación de los lingüistas forenses resulta central, ya que el análisis de los discursos es uno de los recursos clave para la identificación de sus autores.

En el marco de este artículo, y en el marco de la Lingüística Forense (Coulthard 1998; Turrell 2010; Grant y Macleod 2018) nos proponemos presentar una propuesta metodológica multidimensional de análisis de autoría, que resulta productiva para los casos de pseudonimia/anonimato en redes sociales.

PALABRAS CLAVE: *Lingüística forense. Redes sociales. Anonimato. Autoría.*

RESUMO

O impacto inicial da Internet no mundo social foi a criação de comunidades virtuais, que permitiram reunir pessoas dispersas, distantes no espaço, mas unidas por propósitos comuns (Noblia 2000 a e b). Esse potencial de vincular e fomentar grupos de diferentes tipos também possibilitou o agrupamento de indivíduos unidos pela "cultura do ódio" (Zickmund 1997; Dery 1994). Sob a proteção da liberdade de expressão, anonimato ou pseudônimo, grupos e indivíduos fazem uso das redes sociais para divulgar seus discursos radicalizados e programas violentos. A dinâmica das redes, regidas por "haters", os ajuda a circular e convocar novos adeptos. A atuação desses grupos é motivo de preocupação para órgãos governamentais encarregados de fiscalizar crimes dessa natureza e preveni-los, pois, dadas as características e áreas de circulação, sua detecção precoce e desmantelamento é uma tarefa complexa. Para isso, a atuação de linguistas forenses é central, pois a análise de discurso é um dos recursos fundamentais para a identificação de autorias. Neste artigo, no escopo da Lingüística Forense (Coulthard 1998; Turrell 2010; Grant y Macleod 2018), apresentamos uma proposta metodológica multidimensional para análise de autoría, produtiva em casos de pseudonímia/anonimato em redes sociais.

PALAVRAS CHAVE: *Lingüística forense. Redes sociais. Anonimato. Autoría.*

ABSTRACT

One of the first appreciations of the impact of the Internet on the social world was the creation of virtual communities, which made it possible to bring together dispersed people, distant in space but united by common purposes or interests (Noblia 2000 a and b). This potential for linking and generating relationships and groups of different kinds also enabled the grouping of individuals united by the culture of hate (Zickmund 1997; Dery 1994). Thus, under the cover of freedom of expression, and anonymity/pseudonymity, groups and individuals use social networks to spread their radicalized discourses and violent agendas. The dynamics of the networks guided by "haters" help them to circulate and attract new adherents.

The actions of these groups are a source of concern for government agencies in charge of monitoring and preventing crimes of this nature because, given their characteristics and areas of circulation, their early detection and dismantling is a complex task. In this task, the work of forensic linguists is central, since linguistic analysis is one of the key resources for identifying the perpetrators.

Within the framework of Forensic Linguistics (Coulthard 1998, 2004; Turrell 2010; Grant and Macleod 2018), we propose a multidimensional methodological proposal for authorship analysis, which is productive for cases of pseudonymity/anonymity in social networks.

KEYWORDS: *Forensic linguistics. Social networks. Anonymity. Authorship.*

Introducción

Una de las primeras apreciaciones del impacto de Internet sobre el mundo social fue la creación de las comunidades virtuales, que permitieron congregar a personas dispersas, distantes en el espacio, pero unidas por propósitos o intereses comunes (Noblia 2000 a y b). Esta potencialidad de vincular y de generar relaciones y grupos de distinta índole, también habilitó la agrupación de individuos aunados por la “cultura del odio” (Zickmund 1997; Dery 1994; Hawley 2017). Estos grupos, identificados por su clara afiliación a ideologías supremacistas, homofóbicas, sexistas, xenofóbicas, entre otras posiciones radicalizadas, encontraron en la red un ámbito de propagación y difusión que les permitió eludir los controles de los organismos estatales (Hawley 2017; Nagle 2017). Bajo el amparo de la libertad de expresión (uno de los valores más preciados de la cultura digital), del anonimato o de la pseudonimia, estos grupos fueron consolidándose a lo largo del tiempo. En la actualidad, las redes sociales constituyen su hábitat por excelencia, porque les permite solaparse y difundir sus discursos radicalizados, sus símbolos y programas de acción violentos.

El mundo mediado les dio la posibilidad a individuos, que antes de Internet permanecían aislados y anónimos, de cohesionarse y formar grupos extremistas con la inmediatez que posibilitan las tecnologías digitales, en redes que garantizan el anonimato. También, les dio la posibilidad de eludir cualquier fiscalización o rastreo por parte de los organismos de control, que ponga en riesgo sus actividades (Zickmund 1997). Las dinámicas de las redes sociales, pautada por “haters” y debates encendidos, los ayuda a emerger, camuflarse y convocar a nuevos adherentes.

El accionar de estos individuos y grupos radicalizados resulta, en la actualidad, un foco de preocupación internacional, en la medida en que –por la naturaleza misma de las redes– van conformando agrupaciones, que, en algunos casos, pueden trascender la virtualidad para llevar a la práctica sus planes violentos. Por ello, la detección temprana y la desarticulación de estos grupos es una tarea compleja. En esta labor, la actuación de los lingüistas forenses resulta de importancia, en la medida en que los discursos constituyen el principal –en algunos casos el único– recurso para la identificación de los autores.

En el marco de este artículo, a partir de un caso particular, nos proponemos presentar una propuesta metodológica multidimensional de análisis de la autoría que resulta productiva para la identificación o atribución de autoría en casos de pseudonimia/anonimato en redes sociales.

1. Las redes del odio

Para poder explicar el campo de aplicación de la metodología propuesta, es necesario esbozar –aunque sea de una manera muy general y simplificada– el perfil de autor que se busca identificar. Para ello, partimos de dos preguntas: ¿qué caracteriza a los integrantes de grupos radicalizados en los que interactúan las personas a identificar? y ¿qué los diferencia de los usuarios que difunden mensajes violentos y hostigan a otros en las redes sociales públicas como Twitter?

Una de las diferencias más claras de estos grupos es el conocimiento profundo de las posibilidades interactivas de los medios digitales. Su radio de acción son las redes sociales tradicionales como Twitter o Instagram, pero también las plataformas de mensajería como Telegram o WhatsApp, o las de blogs como Wordpress o Blogspot, y hacen un uso diferencial de cada una de ellas.

En las redes abiertas insinúan sus posiciones ideológicas y van testeando, midiendo la tolerancia de la red a la virulencia de sus mensajes. Y cuando son bloqueados por su conducta, se encubren con un nuevo pseudónimo para reingresar, reactivar sus contactos y convocar a nuevos adherentes, con discursos en los que confluyen –sin demasiada uniformidad– posiciones ideológicas de ultraderecha (Traverso 2021; Stefanoni 2021). Estos usuarios asumen un rol evangelizador: entablan diálogos didácticos con sus adherentes y con quienes responden a sus provocaciones.

En los grupos cerrados de las plataformas de mensajería, espacios que –al preservar a ultranza la libertad de expresión y la privacidad de datos de sus usuarios– les garantizan la impunidad más completa, los roles de sus integrantes son claros: por un lado, están los líderes, quienes bajan línea y definen los programas a seguir a través de publicaciones, en las que difunden sus proclamas y programas de acción; por otro lado, sus seguidores, quienes los replican y difunden. Estos grupos cerrados constituyen verdaderas comunidades de prácticas (Scollon 2003), organizadas con objetivos concretos: la segregación y el hostigamiento de los “enemigos” que van configurando de manera aleatoria para justificar su necesidad de defenderse.

El entramado de interacciones, diseminado en diferentes redes y plataformas, los protege y les permite circular sin mayores riesgos, contactar a perfiles similares y sumar voluntades dentro y fuera de la red. Para estos propósitos juega un rol central el trabajo identitario que llevan a cabo de manera programática, tanto sobre el grupo y como de los individuos que los conforman. Estos procesos de identificación, que establecen los límites de pertenencia/exclusión, son poderosos, porque:

- Se glocalizan: retoman las expresiones internacionales de la supremacía blanca, el antisemitismo, las disputas religiosas, para recontextualizarlas en las tradiciones autóctonas. En ese sentido, invocan símbolos y referencias de los principales hitos y próceres de su ámbito local, para sustentar un discurso político alternativo que respalde sus proyectos violentos.
- Retoman algunos referentes de las derechas nacionales, pero rechazan a los referentes políticos que las representan. De este modo, dejan en claro la necesidad de nuevos líderes que cubran ese liderazgo vacante.
- Están abiertos a cualquier expresión que vaya en contra de las luchas por la ampliación de derechos: sea la igualdad de género, los derechos de los inmigrantes, de los homosexuales, etc.
- Reproducen en sus mensajes la clásica fórmula de los discursos discriminatorios (van Dijk 1984), sostenidos en la dicotomía “nosotros” versus “ellos”, en el que el “nosotros” ocupa el lugar de la víctima que debe actuar en defensa de un “ellos”, que representa un peligro inminente para la preservación de la sociedad. Los que se erigen como líderes, mentores o gurús de estos movimientos asumen como su deber la necesidad de organizarse fuera de la red social y pasar a la acción directa, en defensa de un supuesto enemigo que pone en peligro el orden natural, del que ellos forman parte.

Este panorama plantea –de manera muy general– el contexto de circulación y el perfil de la persona a identificar con la metodología propuesta. En otras palabras, hacemos referencia a la identificación de la persona oculta detrás de un pseudónimo, cuyos únicos indicios disponibles son su propia producción discursiva y la información que se filtra o se deduce de ella.

2. ¿Cómo llegar al autor detrás del pseudónimo/anónimo?

Grant y MacLeod (2018) abordan el problema de la autoría, tomando en cuenta los recursos, pero también las limitaciones que presenta la identificación lingüística de las personas online. El aporte de ambos autores se alinea con toda la tradición de los estudios lingüísticos de la autoría (Coulthard 1998; Coulthard y Johnson 2007; Turrel 2010) y también de los discursos digitales (Noblía 2009; 2012).

El fundamento básico de la autoría es concebir al lenguaje como un recurso crucial para la construcción de la identidad social y personal: “Cada texto es un momento de expresión de identidad” (Grant y MacLeod 2018), pero también relativizar su alcance, en la medida en que no es infinito, porque limita tanto como posibilita esa configuración identitaria *on line*.

El supuesto de una identidad personal se asienta en la persistencia en el uso de algunos elementos lingüísticos a lo largo del tiempo. Esta recurrencia de recursos dinámicos y estables permiten relevar aspectos de la identidad de las personas. Grant y MacLeod (2018) clasifican estos recursos en dos grupos: por un lado, los recursos cognitivos, entre los que se encuentran las capacidades psicolingüísticas y la memoria. Y, por el otro, los recursos sociolingüísticos, entre los que encontramos la lengua, las diferentes variedades dialectales, la historia familiar de la persona lingüística, su trayectoria educativa, su historia profesional, entre otros.

Para estos autores, es en la interacción entre las limitaciones y las potencialidades del uso lingüístico en donde se abre la posibilidad de activar protocolos de rastreo de las huellas lingüísticas que los usuarios van dejando involuntariamente, durante el transcurso de su actividad online. Estos indicios constituyen filtraciones involuntarias de sus verdaderas identidades, enmascaradas bajo un pseudónimo.

A partir de la propuesta de estos autores y de la experiencia desarrollada en casos concretos de identificación autoral de seudónimos en redes sociales, hemos elaborado un protocolo de investigación que triangula los métodos de la lingüística forense, en su versión cualitativa (Coulthard 1998; 2005; Turrell 2010; Grant y MacLeod 2018) y estilométrica (Burrows 2002; 2007; Eder, Rybicki y Kestemont 2016), y las técnicas de investigación OSINT (*Open Source Intelligence*).

3. Una propuesta tridimensional

Técnicas Open Source Intelligence (OSINT)

El punto de partida del proceso de identificación y/o atribución de autoría es la construcción de un corpus de textos a analizar a partir de las técnicas de OSINT, metodología cuali y cuantitativa de recolección, análisis y toma de decisiones sobre datos de fuentes disponibles de forma pública para ser utilizados en investigaciones que llevan a cabo los organismos públicos.

Esta etapa es dinámica, cíclica y progresiva: a medida que avanza la investigación se va robusteciendo con todos los elementos textuales y no textuales que se van localizando alrededor del seudónimo a rastrear y toda la red de *links* que remiten a otros autores (que pueden figurar con su identidad civil o bajo seudónimos). La ampliación o circunscripción de potenciales autores requiere volver de manera cíclica al rastreo en las diferentes redes y plataformas, procedimiento que acompaña todo el proceso de la investigación para la identificación del/la autor/a.

La instancia inicial del rastreo planteado permite determinar el tipo de problema de atribución autoral a resolver: si se trata de un problema de atribución autoral *closed-set* (conjunto cerrado), es

decir, si los seudónimos en análisis son parte de un grupo cerrado de potenciales sospechosos con nombre propio previamente determinados, o se plantea un problema de atribución autoral *open-set*, es decir, un verdadero problema de “una aguja en el pajar”.

El desafío principal en una investigación con técnicas OSINT es estar atento a las filtraciones involuntarias que cada persona lingüística tiene en sus producciones textuales o intervenciones en las redes sociales. Muchas veces son los detalles más intrascendentes los que nos permiten llegar a la identificación del autor de un delito o, al menos, a restringir el campo de búsqueda. Es necesario recolectar de manera sistemática los detalles inconexos y disgregados, para luego lograr armar el relato unificador de esa “identidad” en caso de estar ante una identificación positiva.

Análisis cualitativo para el perfilamiento autoral

A partir de la elaboración del corpus, se efectúa el análisis discursivo, con propósitos múltiples:

- La elaboración del perfil autoral, a partir de los indicios que surgen del rastreo de la actividad online de la persona.
- El análisis estilístico/idiolectal (Halliday y Mathiessen 2004; Coulthard 2005; Turrell 2010; Martin y White 2005) del corpus en busca de las elecciones lingüísticas recurrentes y significativas que utiliza el sospechoso.
- El análisis estilométrico (estadístico) por medio del paquete Stylo en R (R Core Team 2020), que, a partir de la frecuencia de ocurrencia de palabras y colocaciones, mide la proximidad estilística, en términos probabilísticos, a fin de establecer si pertenecen a un mismo autor.

A fin de mostrar la operatividad de este método, presentamos un caso, en el que, por razones de confidencialidad se han modificado los datos identificatorios, para poder mostrar el funcionamiento del método, sin violar los requerimientos éticos y procesales de nuestra actuación como peritos.

4. El caso: de XX a MLB

A partir de una denuncia formulada a una fiscalía especializada en temas de discriminación, se inicia la investigación que busca de identificar al autor de los textos de una publicación *on line*. La plataforma de difusión era Telegram, servicio de mensajería conocido por su postura de defensa férrea del anonimato y por el cifrado de datos que hace imposible solicitar, desde el Poder Judicial, información sobre sus usuarios. Los autores que participaban en la revista tenían, a su vez, activa participación en otras redes sociales.

El primer paso fue definir el límite del corpus de textos. Al tratarse de una publicación hipertextual, contábamos con 49 *links* asociados al nombre del autor y 183 a los textos identificados con su seudónimo. Por lo que el primer paso fue determinar si estábamos ante un problema de atribución autoral *closed-set* (conjunto cerrado). Es decir, si los seudónimos podían llegar a ser de alguno de los autores con nombre propio que eran parte del corpus de textos. Luego del análisis

contrastivo, se descartó esta instancia, para plantear el problema en términos de atribución autoral open-set (conjunto abierto), o sea de rastrear “una aguja en el pajar”.

A partir de esta primera decisión, se replanteó la búsqueda desde un enfoque multimodal (Kress 2010), o sea un abordaje que tuviera en cuenta la articulación de recursos verbales, visuales y audiovisuales, y tomando como punto de partida la concepción de estos grupos como comunidad de práctica (Scollon 2003), en el sentido que conforman “un grupo de personas que interactúa con regularidad con el fin de conseguir algún propósito u objetivo común”; interacción que se sostiene, además, por el compromiso y el repertorio de rutinas compartidas por sus integrantes.

Ambas decisiones fueron productivas en la medida en que el rastreo de recursos visuales y verbales permitió precisar el horizonte de la población buscada. Por otro lado, concebirla como comunidad de prácticas (Scollon 2003) permitió sistematizar una búsqueda a ciegas, tomando como eje los símbolos, tópicos, léxico y expresiones recurrentes de los textos publicados en las revistas a través de las otras redes sociales que dieran cuenta de los patrones interactivos del grupo, sus rituales, tópicos de interés, etc.

Perfil de XX

Una vez relevado el corpus inicial, el segundo paso consistió en definir el perfil sociolingüístico del usuario, relevando indicadores de su identidad personal y social, para proponer una hipótesis de trabajo respecto de su nacionalidad, edad, sexo, nivel educativo y formación.

De este análisis, surgieron los primeros datos de la persona detrás del pseudónimo: era argentino, usuario de la variedad rioplatense, género masculino, de entre 20 y 30 años de edad, de formación universitaria, vinculada a las humanidades. Para recabar este último dato, por ejemplo, fueron relevantes las citas que utilizaba en sus textos y las convenciones a las que recurría para mencionar a los autores, propias de una persona con acceso al sistema universitario.

Los tópicos de sus textos y las representaciones discursivas relevadas respecto de los temas tratados, permitieron esbozar un perfil ideológico general asociado al nacionalismo (la invocación de la patria y a la guerra de Malvinas, sus alusiones a la masonería como grupo conspirativo, la referencia a figuras históricas como Rosas, entre otras referencias): “la Argentina de los gauchos y los halcones es el lado correcto de la historia”¹; su rechazo a las reivindicaciones de género y, en especial, su antisemitismo de clara filiación al nazismo: “esto es un mensajito para su colectividad de lacras e hijos de puta de la DAIA, no sigan jodiendo porque les vamos a matar a sus familiares y violar a sus esposas e hijas. Hací patria, maté un judío”².

Una vez definido este perfil general de la persona buscada, se procedió a realizar el primer filtro de los candidatos que conformaban el corpus inicial, seleccionando los que se ajustaran con mayor precisión a este perfil.

1 Ponemos como referencia algunos comentarios del autor, publicados en Twitter.

2 *Idem.* En este caso, se trata de un mensaje en Telegram. Twitter hubiera bloqueado un texto de esta naturaleza.

Tal como lo plantea la Lingüística Forense (Coulthard 2004: 431) “cada hablante nativo tiene su propia versión, distinta e individual, de la lengua que habla y escribe, su propio idiolecto y la suposición de que este idiolecto se manifestará a través de elecciones distintivas e idiosincrásicas en textos”.

A fin de identificar elementos representativos del idiolecto del autor a identificar en el corpus seleccionado, para poder cotejarlo con otros usuarios, se procedió al análisis de los textos. Esta etapa implicó:

- Segmentar todos los textos del corpus relevado.
- Identificar patrones cohesivos, en especial las colocaciones y las cadenas referenciales.
- Rastrear frases hechas, refranes, sentencias, metáforas, hapaxes y citas.
- Reconocer las fórmulas de tratamiento y las claves de contextualización.
- Sistematizar los patrones léxico-gramaticales recurrentes.

En paralelo, se efectuó el análisis estilométrico, que comprendió la compilación de todos los textos de todos los autores inventariados y de todos los textos terciarios que se iban encontrando, a fin de encontrar similitudes entre ellos que permitieran asociarlos al mismo autor o al grupo. Sobre el análisis estilométrico nos detendremos más adelante.

Del análisis efectuado, resultaron de interés algunas frases, cuya combinatoria resultaba poco frecuente y tenían la fuerza de un slogan. En ese sentido, nos resultó útil la propuesta de Woolls (2003) respecto del plagio, para aplicarlo a los casos de autoría anonimizada. Este autor sostiene que para la identificación de un autor no es tan importante el volumen de texto compartido, sino los elementos utilizados “una sola vez” o combinaciones de palabras que remiten a un autor, casi por defecto³. En este caso, se detectaron una serie de frases que proponían un patrón léxico-gramatical recurrente y que se correspondían con el usuario que aquí identificamos como XX. Entre ellas, una fue particularmente significativa, en términos de contenido ideológico y de registro⁴:

“Un empleado de los medios es un adversario de la patria, un hombre armado del establishment. Y ante todo su medio de vida consiste en mentir al pueblo.”

3 Por ejemplo, si se busca en internet la frase "unánime noche" notaremos que todas las referencias giran en torno a Jorge L. Borges y su cuento "El sur". Si bien la palabra "noche" tiene múltiples colocaciones posibles y "unánime" también, la combinación de ambas resulta particular y refiere a un único autor como Borges.

4 Los textos se encuentran levemente modificados, por ser un corpus confidencial, aunque mantienen el sentido y la matriz ideológica de los originales.

Esta frase de XX formaba parte de uno de sus artículos, publicados en la revista que había motivado la denuncia ante el organismo público. Al rastrear la frase en Twitter, se llegó a una versión apenas modificada de esa frase inicial, suscripta, en este caso, por un usuario que aquí designaremos YY, que decía:

“Un empleado de los medios es un adversario de la patria, un hombre armado del orden establecido. Y ante todo su medio de vida consiste en mentir al pueblo.”

A la similitud de la frase se le agregó otro dato relevante: esa frase de XX no tenía más menciones que la de YY. Este dato llevó a la segunda hipótesis: XX e YY podrían llegar a ser la misma persona con pseudónimos diferentes.

De XX a YY: las co-selecciones como indicio clave

Una vez planteada la posibilidad de que el pseudónimo de XX en las redes sociales abiertas fuera YY, se efectuó un relevamiento de toda la constelación de contactos, mediante un análisis de fuentes abiertas de todas sus interacciones en las redes. De este rastreo, se pudieron reconocer tópicos y textos similares, cuya co-selección dio un patrón léxico-sintáctico regular, como el que presentamos a continuación:

Dos grandes *adversarios de la patria*. Los bancos privados y los medios de comunicación privados.

Pasan los años, *los adversarios de la patria* siguen siendo los mismos. Mientras el trabajador queda en la lona, la basura financiera crece.

El macrismo es un *adversario de la patria* y un agente desmalvinizador (...)

La recurrencia en el léxico y la sintaxis permitió reorientar la investigación y enfocar la búsqueda a los dos autores y sus vínculos. En esta etapa cobró relevancia la composición de los pseudónimos (XX/YY), porque pudo detectarse otra recurrencia clave. Los pseudónimos de ambos usuarios estaban compuestos según un patrón regular, que consistía en la combinación de dos palabras, cuya matriz léxica era sinónima. Si, a modo de ejemplo, el pseudónimo de XX era “animal sagrado”, el de YY era “bestia divina”, o sea, los nombres estaban compuestos por la combinación de un sustantivo y un adjetivo que eran sinónimos (“animal”/“bestia”; “sagrado”/“divino”). Cabe aclarar que las palabras utilizadas referían a su universo ideológico, marcadamente nacionalista, vinculado a iconos y referencias a la cultura argentina⁵.

En virtud de esta recurrencia, se inició un rastreo a través de *web.archive.org* de estos usuarios. La búsqueda en este archivo permitió recuperar las diferentes cuentas y sus respectivos pseudónimos que XX había usado a lo largo del tiempo. La composición de los apodos era regular y asentada en la composición aleatoria de esos elementos, por ejemplo, “bestia sagrada”, “animal divino” o “fera sagrada”, etc.

5 Los pseudónimos utilizados no pueden reproducirse por la índole confidencial de los datos.

Archive permitió, también, comprender las razones de los diferentes apodos, ya que el cambio de nombre era sucesivo y motivado por los recurrentes bloqueos del usuario. El patrón compositivo también tenía una razón de ser, en la medida en que operaba como señuelo o clave para poder ser rastreado por sus contactos.

De XX a YY a MLB

El desafío principal en una investigación con técnicas OSINT es estar atento a la filtración de información de las personas, en sus intervenciones en las redes sociales. El conjunto de detalles inconexos y disgregados permiten ir armando un relato unificador de esa “identidad”, una biografía, en caso de estar ante una identificación positiva que dirija al sospechoso.

A las filtraciones involuntarias de XX y sus múltiples seudónimos, se sumaron otros indicios importantes: por ejemplo, una foto de sus antepasados publicada en Twitter que nos permitió establecer una posible nacionalidad de su apellido, conversaciones con otros usuarios, su edad y la fecha de nacimiento.

Como plantean Grant y MacLeod (2018), la identidad de una persona implica una persistencia en el tiempo de algunos recursos que la definen. Aun cuando pensamos que borramos algo en internet, cabe la posibilidad de que haya quedado indexado y, por ende, que pueda ser rastreado. Entre los recursos más relevantes se encuentran las imágenes. En este caso, una de las fotos que utilizaba el autor en Twitter pudo ser rastreada gracias a Google. En esa foto se puede observar la imagen pixelada de su perfil facial. Pero también surgieron otras dos fotos de perfil, en las que se visualizaba lejos a la misma persona en cuerpo entero. En estas fotos lo importante no era sólo la imagen de la persona, sino el paisaje, ya que gracias a la técnica de “*reverse image search*”, se pudo geolocalizar las fotos. Las fotos habían sido tomadas en una ciudad de Europa, y se pudo así recabar evidencia sobre un viaje concreto que adquirió relevancia al analizar los contactos del usuario.

El análisis de las amistades de XX, en las múltiples redes sociales, permitió distinguir sus vínculos en el mundo *online* de los del “*offline*”. Las relaciones online están basadas en el anonimato. Las personas que participan en ellas rara vez se conocen por fuera de ese ámbito. Son personas de diferentes ciudades, provincias y hasta círculos sociales. Estos patrones de socialización online son muy diferentes a los tradicionales, porque no se originan en los espacios de socialización como el club, la escuela, la universidad, el trabajo, el barrio o las afinidades. Esta investigación se concentró, por lo tanto, en un grupo de personas ubicadas en el medio hacia el ápice de la pirámide establecida por McCauley y Moskalkenko (2017)⁶. En base a esta hipótesis

6 Estos autores sistematizan gráficamente el problema de la violencia a través de la imagen de la pirámide: en la base de esta pirámide están los individuos que no se preocupan por una causa política (lo que podría entenderse como una posición neutral); más alto en la pirámide son aquellos que creen en una causa pero no justifican la violencia (simpatizantes); un nivel más arriba están los que justifican la violencia en defensa de la causa (los justificadores); y en la cúspide de la pirámide están los que sienten la obligación moral personal de asumir la violencia en defensa de la causa. Para estos autores, no se trata de un modelo de escalera, o sea, no es progresivo, ya que las personas pueden saltar niveles al moverse hacia arriba y hacia abajo en la pirámide.

de trabajo, se pudieron identificar a aquellas personas que podían ser potenciales contactos de la vida “offline” de XX.

A partir del análisis de las fórmulas de tratamiento, logramos localizar una palabra puntual que resultó fundamental para la identificación de XX: “bro”. Del corpus se analizaron las diferentes formas de invocar a sus contactos en los comentarios. A partir de este cotejo, se reconocieron los destinatarios a los que XX –en sus múltiples seudónimos– se refería de esa manera. De ese grupo, y a través de la búsqueda de cada uno de ellos en Facebook e Instagram, se llegó a la referencia del nombre propio de XX y de las fotos compartidas, entre ellas las del viaje mencionado, a la identificación física y a un nombre propio “MLB” que debíamos confirmar si era XX.

El nombre propio implicó un viraje importante en la investigación, porque fue el inicio de un nuevo rastreo, ya focalizado exclusivamente en MLB, por medio de fuentes abiertas, de todas las apariciones en las redes y Google. Toda la información recabada permitió aportar datos de la identidad real de la persona: la CUIT⁷, fecha de nacimiento, referencias al viaje, otras fotos que permitían corroborar la nacionalidad de sus antepasados, entre otros datos que confirmaron el perfil inicial elaborado.

El análisis estilométrico

Como se planteó en la presentación del método, en las sucesivas etapas de relevamiento de corpus –desde los textos de la revista *on line* y sus *links*, a los textos de los usuarios que iban sumándose como posibles candidatos a ser el autor– se llevó a cabo un análisis exploratorio de los textos, aplicando técnicas estilométricas.

La estilometría es una metodología estadística para analizar textos, a través de sus palabras más frecuentes. Se ha utilizado principalmente para estudiar la atribución de autoría. Se basa en el supuesto idiolectal de que cada persona utiliza los recursos lingüísticos de una manera particular, tanto por las estructuras sintácticas y el vocabulario que selecciona, como el porcentaje de palabras de función utilizadas.

Desde este enfoque, podemos encontrar dos perspectivas de análisis: una exploratoria y otra confirmatoria. El análisis confirmatorio permite hacer una predicción a partir de un conjunto de variables, con un grado de confianza. El análisis exploratorio nos permite explorar los datos para poder extraer conclusiones, hipótesis y relaciones novedosas. En el caso en cuestión, si bien por la dispersión de los géneros discursivos encontrados, la contaminación de citas y préstamos y, algunas veces, la longitud de los textos, no se pudo arribar a un análisis confirmatorio, el análisis exploratorio sirvió como guía para ajustar las hipótesis y orientar la búsqueda.

El objetivo en el caso XX, entonces, fue comparar los estilos de los textos y determinar si pertenecían a un mismo o a distintos autores. Se realizaron varios experimentos, en los cuales se intentó combinar materiales de distinta naturaleza: mensajes de Twitter y Facebook, artículos, *posteos* en foros de discusión, conversaciones. Los resultados obtenidos fueron diferentes según la extensión de los textos.

7 CUIT es la sigla de Clave Única de Identificación Tributaria. Es una clave que se utiliza en el sistema tributario de la República Argentina para poder identificar correctamente a las personas físicas o jurídicas.

En una etapa preliminar, se combinaron textos cortos con otros de mediana longitud, pero los resultados no fueron significativos. Por ello, se resolvió aplicar el análisis estilométrico en textos de mediana longitud: artículos o textos de una extensión mayor a un posteo promedio.

El método utilizado para el cálculo de su proximidad fue Delta de Burrows. La idea en la que se basa Burrows es que la variación de frecuencia de las palabras más frecuentes en un texto permite reconocer a sus autores. Su propia definición de Delta es: “the mean of the absolute differences between the z-scores for a set of word-variables in a given text-group and the z-scores for the same set of word-variables in a target text” (2002: 271). El método de Burrows consiste en extraer las palabras más frecuentes de cada texto y normalizar las frecuencias relativas al conjunto de palabras del texto. Las distancias entre las frecuencias relativas de las palabras más frecuentes normalizadas (distancia de Manhattan) nos permite ver si un texto pertenece a un mismo autor (distancia mínima) o no (distancias mayores).

En este caso, se utilizaron dos tipos de rasgos (*features*): palabras y enagramas, tomando a las palabras como unigramas (frecuencia de aparición de la palabra sola), bigramas (frecuencias de aparición de dos palabras contiguas) y trigramas (frecuencia de aparición de tres palabras contiguas). A estos últimos se los denomina enagramas. Para esta medición, pueden tomarse diferentes números de las primeras palabras o enagramas más frecuentes. Para el caso que presentamos, la variación fue de 200 a 2000 palabras y enagramas.

A modo de síntesis del trabajo realizado en el caso que presentamos, se realizaron los siguientes pasos: a) se extrajeron los contenidos de sus formatos originales y se convirtieron a sólo texto; b) se tokenizaron los textos (divididos en tokens) y se segmentaron por texto y autor; c) a través del método Delta de Burrows, se obtuvo una matriz por cada texto, con la frecuencia de cada palabra o engrama relativa al total y se compararon las matrices de cada texto con el resto. Las distancias entre textos (medidas de similitud o proximidad estilística) fueron visualizadas a través del agrupamiento jerárquico (*Hierarchical Clustering*) y escala multidimensional.

El agrupamiento jerárquico permitió ver las relaciones entre textos a través de una jerarquía; la escala multidimensional fue analizada a través de un plano bidimensional. El análisis estadístico de textos es multidimensional en esencia. Cada palabra en un texto se convierte en una variable a estudiar. Por eso, los textos pueden parecerse o distanciarse de diferentes formas. Los distintos análisis realizados nos permitieron ver las diferentes dimensiones, de acuerdo a los parámetros que se usaron. La diferencia en los resultados tuvo relación con que examinamos los textos desde distintos ángulos, como quien mira un objeto a través de un prisma: el texto en forma individual, agrupado por los autores conocidos, comparando colecciones de textos por autor, época, corriente literaria, etc.

En el presente estudio se probaron diversas formas. Tanto el análisis de los textos, la aplicación del método Delta de Burroughs como la graficación de los resultados de HC (*clustering jerárquico*) y MDS (escala multidimensional) se realizaron con el paquete Stylo (Eder, Rybicki y Kestemon 2016) del software estadístico R (R Core Team).

Debido a la extensión de este trabajo, sólo se presentarán algunos ejemplos, con textos individuales.

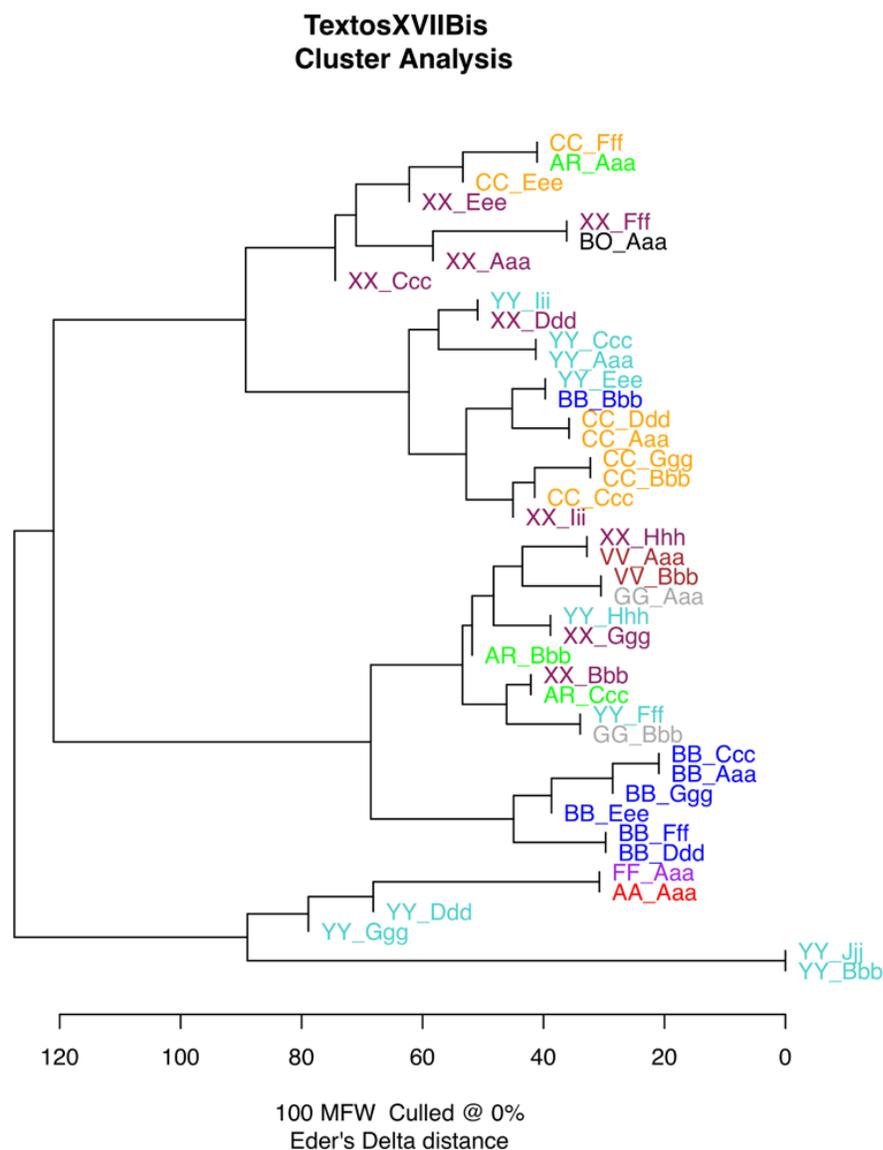
En la primera etapa de la investigación, los textos de XX y YY se compararon con un conjunto de textos de autores próximos a XX. XX y CC son ambos editores de la publicación encontrada en Telegram, a la cual YY no pertenece en principio. GG y VV son autores públicamente conocidos, cuyos textos aparecen publicados en la revista digital, aunque no son originales de dicha revista y

suelen aparecer en foros diversos. BB es un autor próximo a la pertenencia ideológica de la revista y en un análisis se realizó una hipótesis de que podría haber sido XX. Estos autores son citados en la publicación y pertenecen al espectro ideológico próximo a los autores (nacionalistas católicos), aunque no puede determinarse si comparten el antisemitismo ni si son vínculos estrechos del autor. Por otra parte, hay un conjunto de autores de la misma publicación de Telegram, a los que hemos designado con las siglas AA, AR, BO y FF. En esta sección vamos a ver las relaciones entre YY y XX, utilizando autores muy próximos a XX como distractores.

En primer lugar, vamos a mostrar las similitudes encontradas entre los textos utilizando las 100 y las 2000 palabras más frecuentes.

CUADRO 1

Se analiza la similitud entre textos a partir de las primeras 100 palabras. Cada color representa a un autor y cada línea un texto. Abajo se encuentra una regla que va de 0 (más próximos) hasta 120 (más lejanos). Se espera que los textos que caen dentro de cada grupo se asemejen más en términos estilísticos.

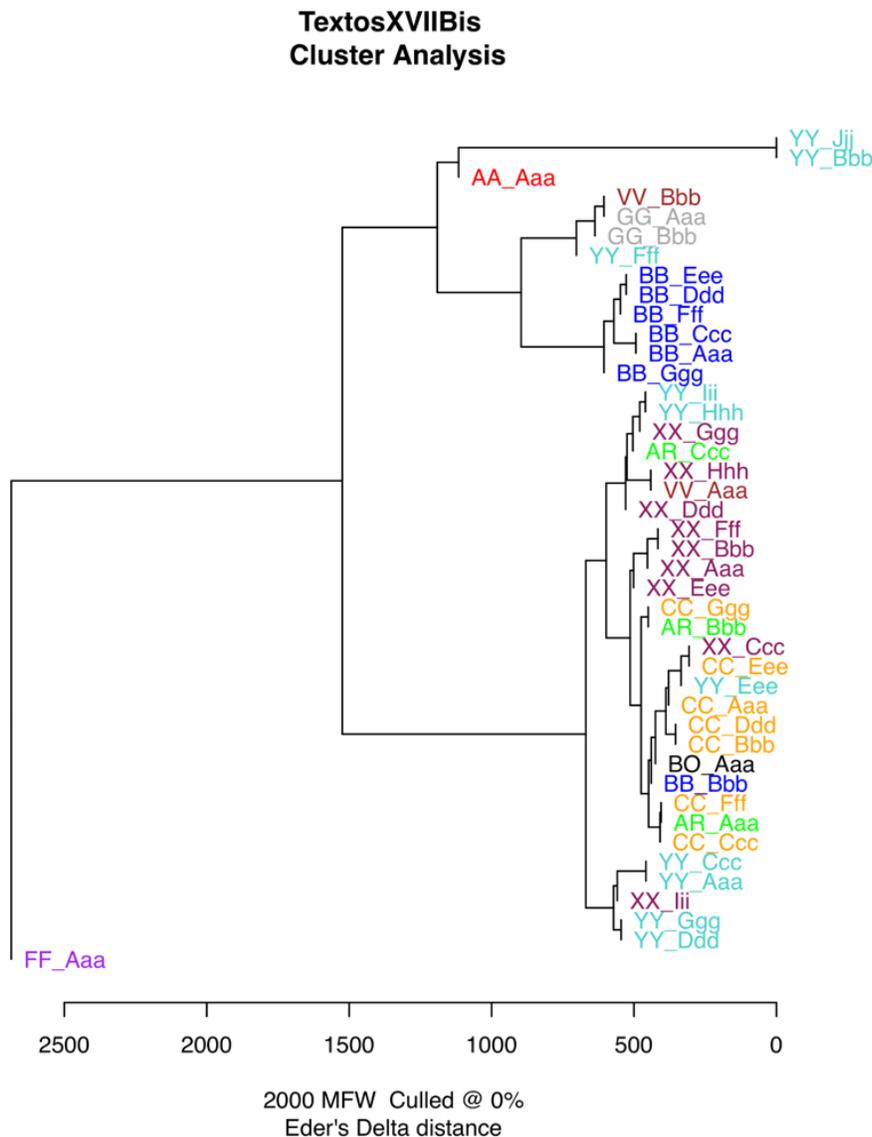


Cuadro 1. Se analiza la similitud entre textos a partir de las primeras 100 palabras. Cada color representa a un autor y cada línea un texto. Abajo se encuentra una regla que va de 0 (más próximos) hasta 120 (más lejanos). Se espera que los textos que caen dentro de cada grupo se asemejen más en términos estilísticos.

En el Cuadro 1 podemos probar que, en el nivel de las primeras 100 palabras, donde aparecen muchos ítems funcionales, podemos diferenciar los textos. Nos interesa ver las relaciones de cercanía entre los textos de XX e YY. Vemos cómo los textos de YY_Iii se relacionan con XX_Ddd y YY_Hhh con XX_Ggg. Otros textos de YY_Jjj y YY_Bbb se relacionan, porque son textos académicos de YY. Aquí podemos ver cómo los textos de BB (profesor de Historia) se relacionan entre sí, por el registro, como cada autor consigo mismo. Hay ciertos cruces entre CC y XX, porque ambos son editores de la misma publicación, y se alternan en ese rol.

CUADRO 2

Realiza el mismo análisis pero a partir de las 2000 palabras más frecuentes.

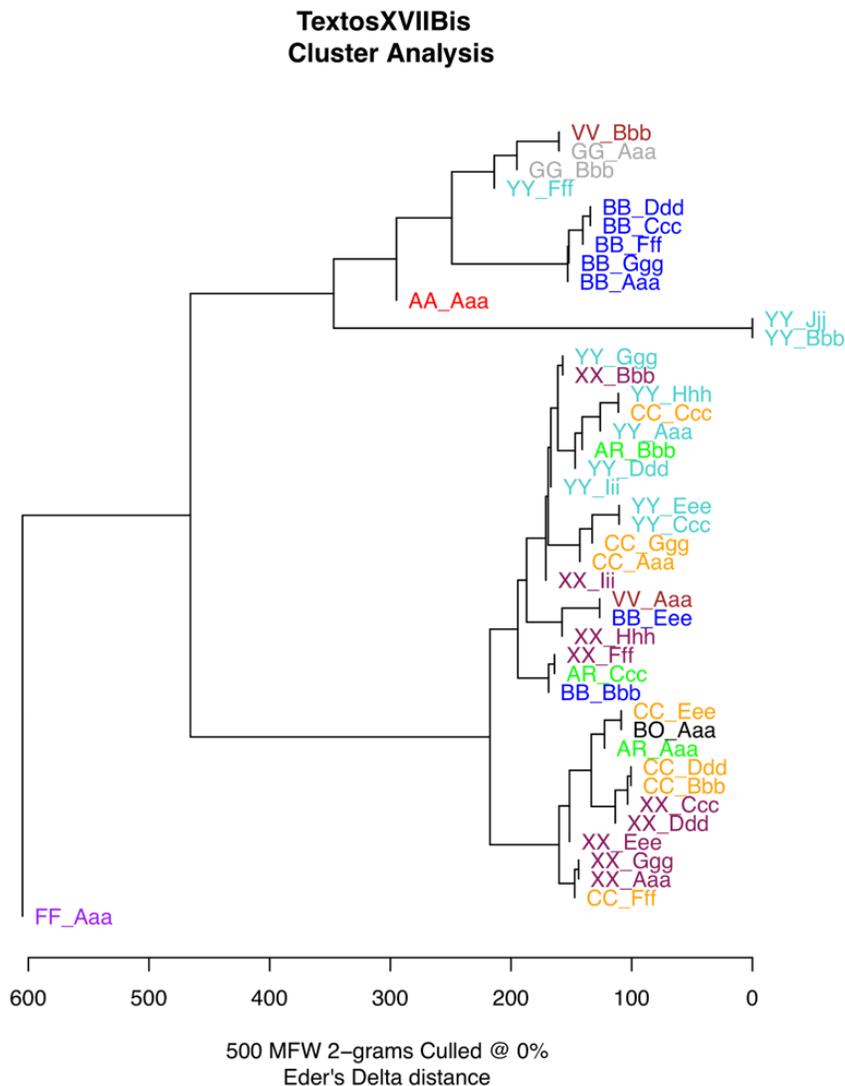


En este dendrograma podemos ver ciertas zonas de confluencia de los textos de YY con XX. Por ejemplo, entre YY_Iii, YY_Hhh y XX_Ggg, entre otros de XX, y entre YY_Ccc, YY_Aaa, YY_Ggg, YY_Ddd y XX_Iii. También entre XX_Ccc y YY_Eee, pensando también en la proximidad entre CC y XX.

En el Cuadro 3 vamos a encontrar el análisis realizado con los 500 bigramas más frecuentes.

CUADRO 3

Análisis de dendrogramas con los 500 bigramas más frecuentes.

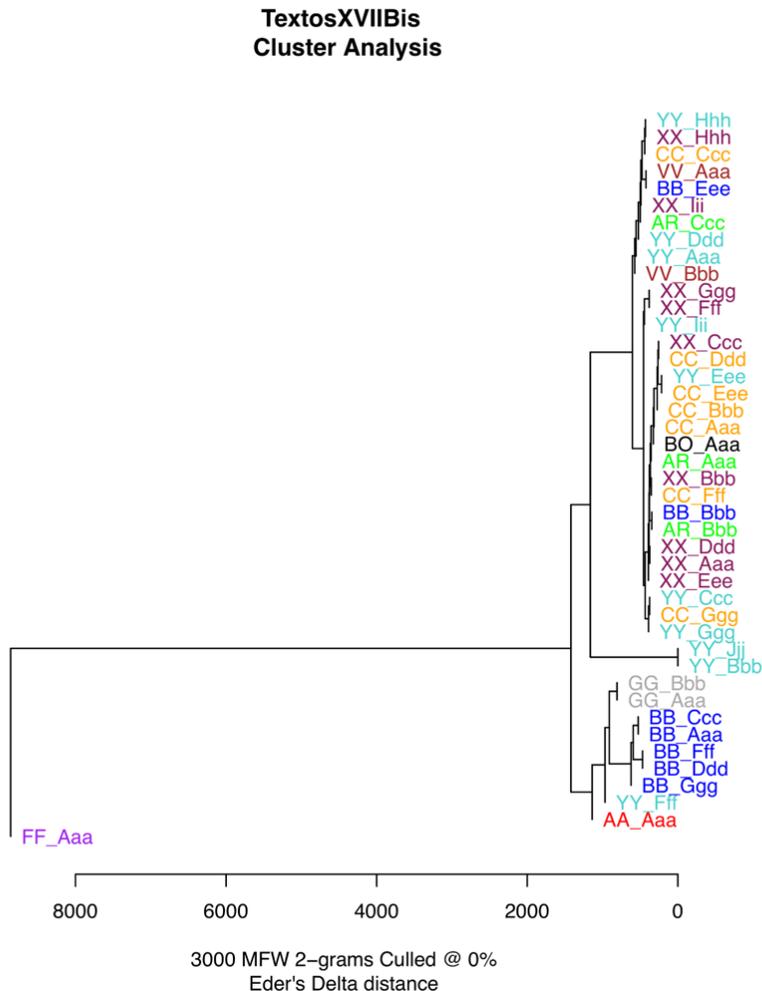


El Cuadro 3 permite visualizar nuevamente relaciones entre CC y XX, por un lado, y de XX y YY por el otro, al mismo tiempo que de YY y CC. En este caso puede notarse que existen varias dimensiones en común entre los autores, a pesar de que YY no pertenezca a la publicación, a la que sí pertenecen XX y CC.

El Cuadro 4 nos muestra los dendrogramas pero con los 3000 bigramas más frecuentes.

CUADRO 4

Dendogramas de las relaciones entre los textos con los 3000 bigramas más frecuentes.



En los gráficos pueden reconocerse aquí grandes agrupaciones, donde puede verse que BB y GG pertenecen a dos grupos diferentes. Sin embargo, estos autores escriben artículos de tipo académico-político. Los textos de YY se mezclan con los autores de la publicación en general: XX, AR, BO y CC, los cuales mantienen relaciones de proximidad. Los trabajos académicos de YY se separan en un grupo aparte > YY_Jjj y YY_Bbb.

En conclusión, podemos encontrar que existen relaciones bastante próximas entre YY y XX para relacionarlos y poder hipotetizar de que se trata del mismo autor. La publicación de Telegram está compuesta por autores que en su mayoría se presentan con seudónimos y podemos suponer también que algunos de ellos puedan ser XX o CC. Por eso, en todos los análisis presentados se encuentran bastante próximos.

Los trabajos académicos presentados por YY, que poseen citas y préstamos de otros textos, aparecen siempre como un conjunto aparte. Por otro lado, algunos artículos de YY aparecen también ligados a trabajos de autores de nombre conocido, lo que supone que también podemos encontrar ahí similitudes de estilo y de contenidos.

5. Resultados

Las tres perspectivas de análisis propuestas en forma integrada (las técnicas OSINT, el análisis textual y el análisis estilométrico) permitieron establecer una metodología para el hallazgo de la identidad del autor de un texto, cuando se trata de rastrear “una aguja en un pajar”.

El análisis de fuentes abiertas permitió rastrear información estratégica y armar el entramado de links que remitieron a otros textos/autores vinculados a la publicación objeto de la denuncia. Con los textos relevados se construyó un corpus inicial, que permitió realizar un perfil del autor buscado, el análisis lingüístico idiolectal de los textos identificados bajo su pseudónimo y los primeros análisis estilométricos, que, si bien permitieron orientar la investigación, no confirmaron las hipótesis preliminares.

Como dijimos más arriba, la aplicación de este triple enfoque metodológico no fue sucesiva, sino circular, recurrente y, por momentos, simultánea: el hallazgo de algún indicio clave en el plano del análisis lingüístico activaba o reorientaba el rastreo en fuentes abiertas y de ese relevamiento se volvía a recurrir al análisis estilométrico para confirmar o rechazar hipótesis. Como pudimos ver en el ejemplo presentado, el relevamiento de algunas frases o palabras clave, vinculadas al espectro ideológico del autor rastreado (por ejemplo, el uso del vocativo “bro”, las colocaciones de ciertos sustantivos y adjetivos (“empleado de los medios”, “adversario de la patria”), los patrones de construcción de sus seudónimos, entre otros recursos) dieron las pistas definitivas que permitieron llegar al autor, a través de la búsqueda en Google, Twitter o Instagram.

El análisis estilométrico permitió encontrar relaciones de proximidad entre los textos de los autores XX y YY y llegar a la hipótesis de que se trataba de un mismo autor, como también la pertenencia de un grupo de autores a una misma publicación. Sin embargo, tanto XX como YY poseían el problema de que ambos eran seudónimos. Por eso un paso adicional consistió en pasar de YY a MLB. La confirmación del autor pudo ser establecida por un exhaustivo rastreo del material verbal, visual y audiovisual por medio del análisis de fuentes abiertas OSINT, que llevó a la identidad pública del autor.

Los aspectos que mayores dificultades plantearon en la investigación tuvieron que ver fundamentalmente con el armado del corpus, en particular, las diferencias de género discursivo entre los textos a analizar, las contaminaciones que algunos de ellos planteaban al introducir citas, traducciones y entrevistas, y el envejecimiento de la persona lingüística, motivada por encontrarse el autor en una etapa de formación académica. No obstante, el encuadre metodológico propuesto fue fructífero en la medida en que permitió, a la vez, descartar posibles candidatos, acotar la búsqueda definiendo un grupo de pertenencia y finalmente la identificación del autor de los textos.

5. Conclusión

El desarrollo de los estudios lingüísticos-forenses sobre la autoría tiene un campo de aplicación especial en las redes sociales, principalmente por el amplio espectro del ciberdelito, que va desde las simples injurias a casos muy complejos, en los que el anonimato opera como pantalla de las acciones criminales. El desafío que afronta la lingüística forense es poder desarrollar herramientas que permitan atribuir la autoría de delitos en espacios en los que los recursos institucionales no están disponibles.

Este trabajo presenta, de manera muy general, los resultados de una investigación que se llevó a cabo en el transcurso de un año y medio, y que arrojó resultados positivos. Es una de las primeras experiencias en nuestro país en la que intervinieron lingüistas forenses, como colaboradores externos de una investigación oficial, por lo que consideramos importante compartir los resultados con el propósito de promover avances en el desarrollo de esta rama de la lingüística aplicada.

Referencias bibliográficas

- BURROWS, J. 2002. 'Delta': A measure of stylistic difference and a guide to likely authorship. *Literary and Linguistic Computing* 17, 3: 267–287.
- BURROWS, J. 2007. All the way through: testing for authorship in different frequency strata. *Literary and Linguistic Computing* 22, 1 :27–48.
- COULTHARD, M. 1998. Making text speak: the work of a forensic linguist. *Studia Anglica Posnaniensia: International Review of English Studies* 33: 117-130.
- COULTHARD, M. 2004. Author identification, idiolect and linguistic uniqueness. *Applied Linguistics* 25, 4: 431–47.
- COULTHARD, M. 2005. Some forensic applications of descriptive linguistics. *Veredas: Revista de estudos linguísticos* 9, 1.
- COULTHARD, M. y JOHNSON, A. 2007. *An Introduction to Forensic Linguistics. Language in Evidence*. New York: Routledge.
- DERY, M. 1994. *Flame wars. The discourse of Cyberculture*. Durham/London: Duke University Press.
- EDER M., RYBICKI J. & KESTEMON, M. 2016. Stylometry with R: a package for computational text analysis. *R Journal*, 8, 1 : 107–121. En: <https://journal.r-project.org/archive/2016/RJ-2016-007/index.html>.
- GRANT, T. & MACLEOD, N. 2018. Resources and constraints in linguistic identity performance – a theory of authorship. *Language and Law = Linguagem e Direito* 5, 1: 80-96.
- HALLIDAY, M. & C. MATHIESSEN. 2004. *An Introduction to Functional Grammar*. London: Routledge.
- HAWLEY, G. 2017. *Making sense of the Alt-Right*. New York: Columbia University Press.
- KRESS, G. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London & New York: Routledge.
- MARTIN, J. & WHITE, P. 2005. *The Language of Evaluation: Appraisal in English*. London: Palgrave/Macmillan.
- MCCAULEY, C. y MOSKALENKO, S. 2017. Understanding Political Radicalization: The Two-Pyramids Model. *American Psychologist* 72, 3: 205–216.

- NAGLE, A. 2017. *Kill All Normies: Online Culture Wars From 4Chan and Tumblr to Trump and the Alt-Right*. Washington, DC: Zero Books.
- NOBLIA, M. V. 2000a. Internet, Comunidad virtual y control. En M. V. Noblia y M. L. Pardo (Eds.). *Globalización y nuevas tecnologías*, pp. 55-58. Buenos Aires: Biblos.
- NOBLIA, M. V. 2000b. Conversación y Comunidad: los chats y la comunidad virtual. *Discurso y Sociedad* 2, 5.
- R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- NOBLIA, M. V. 2009. Modalidad, evaluación e identidad en el chat. *Discurso y Sociedad* 3, 4.
- NOBLIA, M. V. 2012. *La relación interpersonal en el chat. Procesos de construcción y negociación de la identidad*. Tesis doctoral. Buenos Aires: Editorial de la Facultad de Filosofía y Letras. (UBA). E-Book.
- SCOLLON, R. 2003. Acción y texto: para una comprensión conjunta del lugar del texto en la (inter) acción social, el análisis mediato del discurso y el problema de la acción social. En R. Wodak y M. Meyer (Eds.). *Métodos de análisis crítico del discurso*, pp. 220-221. Barcelona: Gedisa.
- STEFANONI, P. 2021. *¿La rebeldía se volvió de derecha?* Buenos Aires: Siglo XXI.
- TRAVERSO, E. 2021. *Las nuevas caras de la Derecha*. Buenos Aires: Siglo XXI.
- TURRELL, M. T. 2010. The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *Speech, language and the law* 17, 2: 211–250.
- VAN DIJK, T. 1984. *Prejudice in Discourse*. New York/Amsterdam: John Benjamins Publishing.
- WOOLLS, D. 2003. Better tools for the trade and how to use them. *Forensic Linguistics* 10, 1: 102-112.
- ZICKMUND, S. 1997. Approaching the radical other: the discursive culture of cyberhate. En S. Jones. *Virtual Culture. Identity & Communication in Cibersociety*, pp. 185-186. London: Sage.

DRA. MARÍA VALENTINA NOBLIA es Doctora en Letras por la Universidad de Buenos Aires. Se ha especializado en el análisis del discurso en medios digitales y en lingüística forense.

En relación con esta última materia, se desempeña desde 2003 como perita en lingüística y forma parte de la nómina del Poder Judicial de Argentina. Es miembro fundadora y actual presidenta de la Asociación de Lingüística Forense Argentina. Dicta cursos y seminarios de especialización en la UBA y otras instituciones del país.

Email: mvnoblia@filo.uba.ar

LIC. ALEJANDRO CARLOS RENATO es Licenciado en Letras por la Universidad de Buenos Aires. Es docente de la Facultad de Medicina (UBA) en la Carrera de Fonoaudiología y en el CEPEL (Universidad Nacional de San Martín). Su área de especialidad es la fonética, el procesamiento del habla y el análisis estadístico del lenguaje y el discurso. Se desempeña en el proyecto PIASHIBA - Hospital Italiano de Buenos Aires, en el área de Procesamiento del Habla y Reconocimiento Automático del Habla.

Email: alejandrorenato@hospitalitaliano.org.ar

LIC. TOMÁS GERSHANIK es Licenciado en Letras por la Universidad de Buenos Aires y miembro del Área de Actividades Interdisciplinarias sobre Conductas Discriminatorias del Ministerio Público Fiscal donde se desempeña como investigador especializado en el fenómeno de la radicalización online.

Email: tgershanik@fiscalias.gob.ar