XV Jornadas de Sociología. Facultad de Ciencias Sociales, Universidad de Buenos Aires, Buenos Aires, 2023.

Análisis de discurso: Similitudes y diferencias en discursos vinculados a educación entre precandidatos a presidente y diarios digitales en Argentina en el primer semestre de 2023.

Joaquín Lovizio Ramos y Ariana Bardauil.

Cita:

Joaquín Lovizio Ramos y Ariana Bardauil (2023). *Análisis de discurso: Similitudes y diferencias en discursos vinculados a educación entre precandidatos a presidente y diarios digitales en Argentina en el primer semestre de 2023. XV Jornadas de Sociología. Facultad de Ciencias Sociales, Universidad de Buenos Aires, Buenos Aires.*

Dirección estable: https://www.aacademica.org/000-025/54



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto*. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: https://www.aacademica.org*.

Análisis de discurso: Similitudes y diferencias en discursos vinculados a educación entre candidatos a presidente y diarios digitales en Argentina entre abril 2022 y abril 2023

Ariana Bardauil, Universidad Nacional de la Matanza (UNLaM)
ariibardauil@gmail.com

Joaquín Lovizio Ramos, Universidad Nacional de Tres de Febrero (UNTREF)
joaquin.lovizio@gmail.com

Opinión Pública, Comunicación y Marketing Político

Nuevas tecnologías, comunicación y participación política

RESUMEN

En el contexto de las próximas elecciones presidenciales y legislativas en Argentina, que se llevarán a cabo en octubre de este año, el análisis de los discursos políticos y mediáticos sobre la educación adquiere especial relevancia. Los comicios representan momentos de generación de agenda política y de debate mediático acerca de las propuestas de los candidatos y los partidos políticos.

El presente trabajo tiene como objetivo indagar los discursos, por un lado, de los principales candidatos a presidente para las elecciones argentinas del año 2023 y, por el otro, aquellos publicados por los portales de noticias más leídos en el país. En este sentido, se buscará encontrar similitudes y diferencias entre ambos actores utilizando como eje temáticas vinculadas a educación tomando como recorte temporal el período comprendido entre enero y junio de 2023. Para tal fin, se llevará a cabo un análisis del discurso de dichos políticos desde una metodología cuantitativa, utilizando como principales fuentes sus intervenciones en redes sociales, entrevistas y discursos en actos políticos y públicos. De esta manera, a partir de técnicas de programación y ciencia de datos, como webscrapping y text mining, reunidas en una herramienta de desarrollo propio creada para tal propósito y titulada “Dicen Los Medios”, se buscará evaluar la similitud entre lo que dicen los políticos y lo que dicen los medios.

1. INTRODUCCIÓN

La política contemporánea se encuentra atravesada por un proceso de hipermediatización, que influye en las prácticas políticas y produce cambios en las formas de la democracia (Carlón, 2016). Al mismo tiempo, la imbricación de la política en las redes sociales permitió que los medios de comunicación ya no sean los únicos intermediarios entre gobernantes y gobernados sino que, la interacción entre los políticos y votantes se pueda dar de manera directa reforzando la personalización y personalismo de la política (Slimovich, 2019)

Esta transformación ha llevado a un cambio de paradigma, pasando de una política mediatizada a una política mediática, donde los medios de comunicación ya no son un actor más dentro de la esfera pública, que media entre la ciudadanía y la política brindando información auténtica e imparcial como en épocas pre digitales (Castells, 2008), sino que se han convertido en actores con capacidad para construir y moldear esta esfera pública. Así, la transmisión del contenido de los medios a las audiencias no es lineal sino que contiene una compleja trama de significaciones sobre los mismos. De modo que "los medios de comunicación construyen la realidad política y, así, dan forma al retrato que el público arma en su mente para entender lo que ocurre a su alrededor" (Aruguete 2015, p.30).

Por otro lado, los medios de comunicación se constituyen como actores con intereses específicos que "se mueven en un campo atravesado por relaciones de poder" (Califano 2015, p.63), a través del cual, otorgan prioridad a ciertos acontecimientos, omitiendo otros y/u otorgándoles mayor o menor jerarquía desde un encuadre particular. Actualmente, poseer incidencia en los medios es tener influencia sobre la opinión pública, se trata del "espacio donde se crea el poder" (Castells 2009, p.262). Es así como el vínculo entre medios de comunicación y discursos políticos presenta especial interés para las ciencias sociales, particularmente en relación con el estudio de la representación política y la opinión pública en la actualidad.

En este contexto, es esencial comprender cómo las declaraciones de políticos se entrelazan con la discursividad mediática, particularmente en temas clave como la educación. El análisis del discurso[1], se presenta como una herramienta fundamental para explorar las estrategias, enunciativas y argumentativas, utilizadas tanto por los políticos como por los medios de comunicación en la construcción de sentido y representaciones sobre la educación. En este sentido, los conceptos y enfoques de Eliseo Verón y Van Dijk resultan relevantes.

[1] La teoría discursiva afirma que, la "razón" concebida por la tradición filosófica como razón moral y como razón política, puede ser reformulada con una racionalidad práctico-comunicativa, que está incorporada en los procesos de diálogo y en los procedimientos argumentativos. (Habermas, 1987).

Eliseo Verón (1998), desde una perspectiva semiótica y sociológica, enfatiza la importancia de comprender la materialidad del sentido, en otras palabras, cómo el significado adquiere forma y se manifiesta en la comunicación. Este enfoque reconoce que el sentido no es algo abstracto sino que se encuentra arraigado en prácticas y discursos específicos que se desarrollan en la interacción social. El estudio sobre superficies textuales a través de la aplicación de métodos cuantitativos, como el análisis estadístico o el análisis de redes semánticas, puede brindar una dimensión adicional al análisis cualitativo proporcionando datos numéricos y patrones que respaldan las observaciones y conclusiones de la investigación.

El estudio sincrónico de los discursos en redes sociales, medios de comunicación audiovisual y prensa digital nos permite visualizar estos cuerpos discursivos como un corte en el proceso de la semiosis social[2]. Por su parte, Van Dijk (2009), en sus postulados sobre la teoría del discurso, destaca la influencia del contexto social y las relaciones de poder en la producción y procesamiento del discurso. Ambos elementos cobran muchísima importancia en momentos electorales, por lo que cabe pensar que los discursos producidos por diferentes actores que disputan en la arena política presentan configuraciones merecedoras de ser estudiadas por las ciencias sociales. Esta teoría permite analizar la manera en que las estrategias discursivas utilizadas se reflejan en la cobertura mediática o viceversa. Gracias al uso de herramientas informáticas que permiten procesar extensos cuerpos textuales, podemos establecer relaciones entre tópicos y sintagmas de interés en una cuantía mucho mayor.

Los avances en la tecnología y la constante producción de información motivan la incorporación de diferentes métodos de procesamiento de texto de manera automatizada permitiendo identificar patrones, tendencias y similitudes en los grandes volúmenes de información con los que se cuenta actualmente. Las técnicas de *text mining* (minería de texto en inglés) para el análisis de texto emergen como herramientas fundamentales que permiten extraer, procesar y categorizar información relevante de extensos cuerpos textuales[3] (Barrera, 2014). Dichas técnicas permiten explorar las complejas estructuras de producción de sentido presentes en la comunicación social, descubriendo conocimiento en cantidades considerables de información no estructurada.

La elección de la educación como temática central para el presente trabajo se fundamenta en su relevancia en la sociedad contemporánea y su papel fundamental en la configuración del

[2] Manifestación material de toda producción de sentido. Dimensión significante de fenómenos sociales (Verón, 1993)

[3] En este trabajo se entiende a la *minería de texto* como el *"estudio de la información digital y en particular de los documentos textuales, con el objetivo de descubrir tendencias, patrones, desviaciones y asociaciones de una colección de textos, para –en última instancia– pasar al descubrimiento de conocimiento en considerables cantidades de información no estructurada"* (Barrera, 2014, p.131)

futuro de una nación. La educación no solo es un pilar fundamental en el desarrollo individual y colectivo, sino que también representa un aspecto crítico de la política pública en cualquier país.

Durante los periodos electorales, los discursos sobre educación adquieren un papel protagónico al reflejar no solo las prioridades políticas de los candidatos y partidos, sino también sus visiones sobre la igualdad de oportunidades, la equidad, la innovación y la competitividad a nivel global. La educación es un tema que resuena en la conciencia de los votantes, ya que afecta directamente a las futuras generaciones y a la formación de ciudadanos informados y críticos.

Además, los medios de comunicación desempeñan un rol clave en la construcción de discursos sobre educación, lo que a su vez influye en la percepción pública de este tema vital ¿Qué educación queremos? ¿Cómo están las instituciones educativas públicas y privadas hoy en día? ¿A qué escuela llevaría a mis hijos? Son preguntas que obtienen respuesta a partir de lo que reflejan, en parte, los medios de comunicación y las redes sociales. Por lo tanto, este análisis busca comprender no sólo las estrategias comunicativas y enunciativas utilizadas, sino también la visión de los actores políticos y mediáticos de la política educativa, el desarrollo y la calidad de la educación en nuestras sociedades contemporáneas

Por estas razones, en el contexto de las próximas elecciones presidenciales y legislativas en Argentina, que se llevarán a cabo en octubre de este año (2023), el análisis de los discursos políticos y mediáticos sobre educación adquiere especial relevancia. Los comicios representan momentos de generación de agenda política y de debate mediático acerca de las propuestas de los candidatos y los partidos políticos. Al mismo tiempo, se realzan a través de las construcciones discursivas diferentes representaciones de los panoramas sociales, económicos y políticos (Acosta, Demirdjian y Lassi, 2021).

De esta manera, los procesos electorales se presentan como una instancia de generación e imbricación entre la agenda política que tratan de configurar los referentes políticos y, la agenda mediática, que intentan constituir los medios de comunicación. La forma en que los medios realizan la cobertura de los procesos electorales tendrá “relevancia en la conformación de una opinión pública más o menos activa, en particular a través del tratamiento que realicen de la información; es decir, de los encuadres empleados” (Acosta, Demirdjian y Lassi, 2021). En este sentido, la manera en que los políticos y los medios de comunicación presentan y discuten temas educativos puede tener un impacto significativo en la construcción de la realidad social y en la percepción de la educación como pilar fundamental en la sociedad. Al mismo tiempo, a pesar de su relevancia, según la exploración realizada, no se han encontrado trabajos investigativos que aborden la similitud de discursos sobre educación

entre candidatos a presidente y medios digitales, en el actual periodo electoral de la República Argentina. Esta ponencia se enmarca en un proyecto más amplio que abarca el trabajo continuo del Laboratorio de medios de El Ateneo A.C llamado "Dicen los medios" que tiene como misión analizar la manera en que los medios de comunicación informan

El presente trabajo, tendrá como objetivo general analizar los discursos sobre educación de los candidatos a presidente y los portales de noticias más leídos en Argentina - Infobae, La Nación, Clarín, Infobae, Página 12, Perfil, Ámbito Financiero, Todo Noticias y el Destape -, durante el período comprendido entre abril 2022 y abril 2023, con el fin de identificar similitudes y diferencias en la construcción del discurso político y mediático y al mismo tiempo tendrá como objetivo secundario poner a prueba una metodología utilizada en anteriores producciones.

A tal efecto, se buscará:

> Identificar los discursos sobre educación de los candidatos a presidente en redes sociales, entrevistas y discursos en actos políticos y públicos.
>
> Identificar los discursos presentes en los portales de noticias más leídos en Argentina, en relación a la temática de educación.
>
> Comparar y contrastar los discursos de los candidatos a presidente y los portales de noticias en cuanto a la temática de educación, buscando identificar las similitudes, diferencias y posibles influencias entre ambos actores discursivos.
>
> Monitorear la performance de la metodología utilizada en anteriores producciones.

La hipótesis que se sostiene en este trabajo es que, si bien los discursos entre políticos son diferentes entre sí, y a su vez los diarios poseen líneas editoriales disímiles; existen similitudes en los discursos sobre educación entre los candidatos a presidente y los portales de noticias más leídos en Argentina, lo que podría mostrar algún tipo de influencia y/o retroalimentación entre ambos actores vislumbrando alianzas no intencionales.

Para ello, se utilizará una metodología mixta basada en fuentes como las intervenciones en redes sociales, entrevistas y discursos en actos políticos y públicos, analizando las similitudes y diferencias entre estos discursos y las noticias de los diferentes medios. En este marco, a partir de técnicas de programación y ciencia de datos, como *webscrapping* y *text mining,* reunidas en una herramienta de desarrollo propio creada para tal fin titulada "Dicen Los Medios", se buscará contribuir al análisis de la construcción discursiva en el ámbito electoral argentino

El presente trabajo se propone analizar la similitud de discursos sobre educación entre políticos y medios. Para ello, en primer lugar, se describe la metodología utilizada para la comparación de los discursos, así como las técnicas aplicadas y las fuentes abordadas en esta ponencia. En segundo lugar, se presenta una descripción de los corpus de discursos de los candidatos a presidente y de las noticias de los portales digitales, junto con los resultados de la investigación. Por último, en las consideraciones finales, se abordan las conclusiones del trabajo, así como los desafíos y aprendizajes para futuras líneas de investigación.

2. METODOLOGÍA

Para abordar esta ponencia se emplea una metodología mixta, es decir una combinación de métodos cuantitativos y cualitativos. El análisis se concentra en buscar el grado de similitud entre los discursos sobre educación entre los políticos seleccionados y los diarios digitales entre abril de 2022 y abril de 2023.

Tabla 1: Políticos Seleccionados

Político	**Fuerza**
Sergio Tomás Massa	Unión por la Patria
Patricia Bullrich	Juntos por el Cambio
Javier Milei	La Libertad Avanza
Myriam Bregman	Frente de Izquierda
Juan Schiaretti	Hacemos por Nuestro País

Fuente: Elaboración propia

Los medios digitales se seleccionaron en base a trabajos realizados previamente por este equipo, como forma de lograr algún tipo de continuidad investigativa y teniendo en cuenta la

información ofrecida por el Digital News Report (2022)[4]. Para generar los corpus de textos con los que se trabajó en el presente documento, se utilizaron diferentes técnicas.

En primer lugar, la base de noticias de los diferentes diarios se construyó a través de la herramienta de scrapeo web de elaboración propia y denominada Dicen Los Medios. El scrapeo web o *scrapping* como se lo denomina en inglés, también conocido como *web extraction* o *harvesting,* es una técnica que permite "rascar" datos de la World Wide Web (WWW), lo que se puede traducir como localizar elementos en el código HTML de las diferentes páginas de un sitio web y ordenar esa información en forma de tabla (Zhao, 2022). A partir de esta herramienta, se construyó un programa que se ejecuta diariamente y recolecta elementos tales como el título, la sección o segmento, el texto, la fecha, entre otros, de las noticias publicadas en los principales diarios digitales del país. Sobre esta base se realizó una consulta para filtrar las noticias publicadas entre el 01 de abril del 2022 y el 30 de abril del 2023.

Posteriormente se desarrolló lo que se conoce como 'bolsa de palabras' con términos relativos al tema de investigación, es decir, una lista de términos referidos a la temática educación. Esta bolsa fue construida por el equipo investigador incluyendo palabras tales como educación, docentes, profesores, estudiantes, becas, escuela, entre otras. De esta manera, se filtraron noticias que mencionaran al menos 4 palabras distintas de la bolsa de palabras y con una extensión no mayor a 500 caracteres. Esta decisión recayó en la idea de que, a mayor cantidad de palabras en una noticia, mayor probabilidad de encontrar diversos términos y por consiguiente mayor probabilidad de que algunos términos de la bolsa de palabras fueran mencionados. Si bien esta preselección pudo haberse utilizado tal y como fue construida, se decidió realizar un filtrado más, teniendo en cuenta que algunas noticias podían referirse a cuestiones no vinculadas directamente con la temática educación pero que sí utilizaran términos de la bolsa de palabras.

En este punto, se realizó una comparación de discursos entre la bolsa de palabras y el corpus de cada noticia, a través del elemento estadístico *similitud coseno*. Este elemento mide el ángulo de coseno entre dos vectores o cadenas de texto que se desean comparar en una escala normalizada. Cuanto menor es el ángulo entre los dos vectores, mayor es la similitud; de la misma manera, si el ángulo entre los dos vectores es grande, entonces las cadenas son muy diferentes una de la otra (Singh, Maurya, Tripathi, Narula, & Srivastav, 2020). Respecto a la interpretación, los valores están comprendidos entre 0 y 1, entre más cercano a uno sea

[4] Información disponible en https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf

el valor resultante mayor será la similitud entre ambos discursos (Durán Chinchilla, García Quintero y Rosado Gómez, 2022)

Una vez obtenido el número, a partir de esta técnica que relaciona el nivel de distancia entre una noticia y el corpus referido a la temática de educación, se procedió a generar una muestra de 970 casos que fueron clasificadas manualmente según estuvieran relacionadas realmente a la temática educación o no. Pudo constatarse que algunas noticias de la preselección trataban sobre hechos de inseguridad en escuelas (particularmente tiroteos en colegios de Estados Unidos), denuncias sobre robos en inmediaciones de escuelas, protestas sindicales que incluían al gremio docente, entre otras. Sobre esta muestra se entrenó un modelo de clasificación logística para establecer el threshold que determinara qué valor de similitud coseno podía funcionar para clasificar las noticias como pertenecientes a la temática educación o no, logrando una precisión del 74%. Este modelo fue aplicado al conjunto restante de noticias, lo que permitió reducir significativamente la cantidad de noticias que inicialmente formaban parte del corpus pero que no deberían haber sido consideradas, reduciendo el ruido en la muestra inicial. De esta manera, quedó construida la base de discursos de cada diario digital referidos a la temática educación.

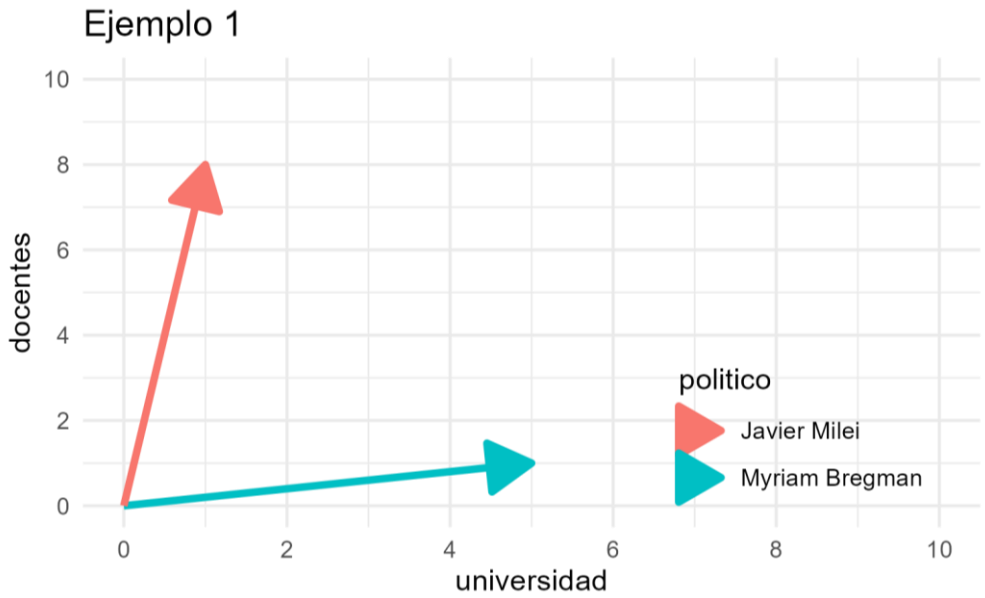
En segundo lugar, respecto a los discursos de los candidatos a presidente, para generar los corpus de textos se decidió trabajar con dos fuentes de información. Por un lado, los tweets publicados por cada político durante el período objetivo. Para obtener esta información se realizó una conexión a la API de Twitter. Una API es un protocolo de intercambio de información entre diferentes plataformas. A diferencia de técnicas de scrapeo web, al consultar la API se obtiene información ordenada y de manera oficial, ya que es la provista por la misma plataforma. Debido a que la extensión máxima de un tweet es de 280 caracteres, mucho menor si se lo compara con las noticias, el filtro aplicado para determinar si un tweet se refería a la temática educación se limitó a encontrar al menos una mención a algún término de la bolsa de palabras. Por otro lado, se consultaron entrevistas de los candidatos presidenciales relacionadas a la educación. Este paso se realizó de una manera cualitativa, buscando entrevistas tanto en la web como en la plataforma Youtube y seleccionando aquellos fragmentos donde los políticos hablaran expresamente sobre la temática. Para obtener las desgrabaciones de estas entrevistas, se utilizaron librerías especializadas en extracción de subtítulos de videos de la aplicación, permitiendo una optimización de tareas de manera considerable.

Como resultado, se obtuvieron las tres fuentes de texto deseadas a través de tres técnicas diferentes: scrapeo web, consultas a una API y descarga con librerías especializadas. Además, se realizaron diferentes técnicas de filtrado para conseguir corpus más acabados.

En el caso de las noticias, consistió en la comparación con una bolsa de palabras utilizando la medida de similitud coseno y un modelo logístico de clasificación. En relación a los discursos de los políticos seleccionados, se recolectaron tweets que mencionaran términos de la bolsa de palabras de educación y se seleccionaron entrevistas de manera más bien cualitativa.

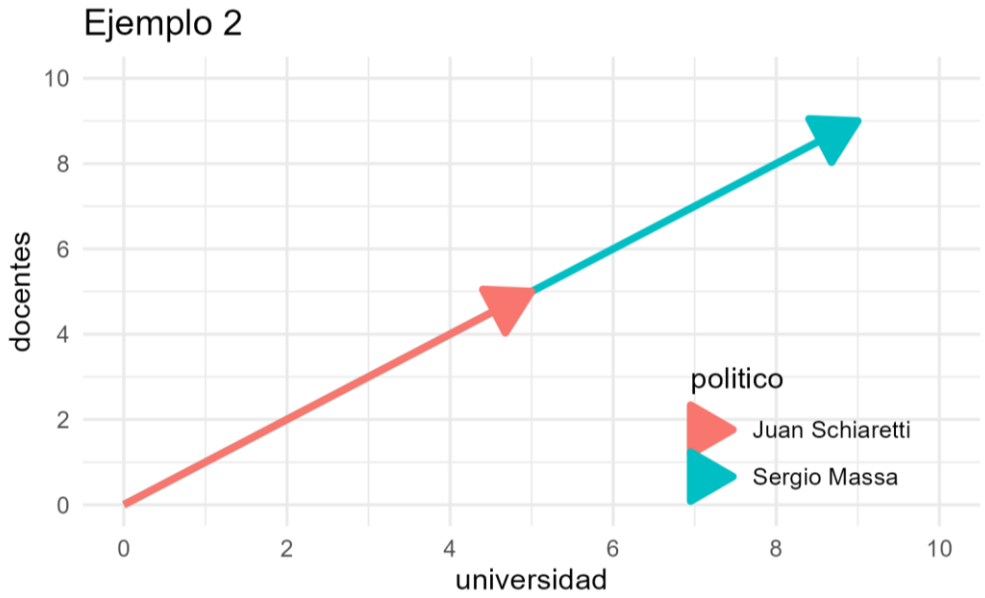
Una vez recolectada toda la información, se procedió a comparar los discursos de cada agente a través del elemento estadístico similitud coseno. Como se mencionó anteriormente, esta técnica se utiliza para medir la distancia entre dos vectores a través del ángulo que forman cada uno en un espacio de grandes dimensiones, proporcionando un valor entre 0 y 1, siendo 0 mínima coincidencia entre los vectores y 1 máxima coincidencia. Para demostrar el funcionamiento de este elemento, se ejemplifica con un espacio de dos dimensiones. A modo de ejemplo, se toman los términos “docentes” y “universidad” y se graficara un caso hipotético de comparación entre los discursos de Javier Milei y Myriam Bregman

Para los discursos de Javier Milei, la palabra “docentes” tiene una frecuencia de 1 mención y la palabra “universidad” tiene una frecuencia de 5 menciones. Por otro lado, para Myriam Bregman la palabra “docentes” tiene una frecuencia de 7 menciones y “universidad” una frecuencia de 1 mención. Al aplicar similitud coseno, se podría decir que los discursos tienen poca similitud, ya que las frecuencias de las palabras utilizadas tienen cierta diferencia. En este caso, el ángulo que se forma entre ambos vectores es más cercano a 90°, por lo que el coseno de ese ángulo se acercará al valor 0.



Ahora bien, supongamos un segundo caso de comparación entre Juan Schiaretti y Sergio Massa donde la frecuencia de las palabras nos arrojara 5 menciones de la palabra “docentes”

y 5 de la palabra “universidad” para el primer político, y 10 menciones de la palabra “docentes” y 10 de la palabra “universidad” para Massa. En este caso, si bien la frecuencia es más alta para uno de los políticos, la coincidencia en los discursos es máxima ya que ambos vectores forman un ángulo de 0° y el coseno de 0° es 1.



Los vectores de todos los discursos de cada político y cada diario hacen que sea imposible graficarlos y siquiera imaginarlos, pero utilizando nociones de álgebra lineal se pueden calcular los cosenos de estos vectores y así obtener un valor entre 0 y 1 que indique la similitud entre discursos.

3. RESULTADOS

En este apartado, se presentan los resultados obtenidos a través del análisis realizado y la aplicación de las técnicas descritas en la sección anterior. El análisis presentado refleja la relación entre los discursos de los candidatos a presidente y las noticias de los portales digitales, y ofrece un punto de partida para mejorar y ajustar la metodología para futuras líneas de investigación.

De esta manera, en primera instancia, se expondrá la proporción de discursos que abordan la temática educación en ambos corpus. A continuación, se llevará a cabo un análisis exploratorio de la composición de los discursos, considerando los términos más utilizados por los medios y, posteriormente por los/las candidatos/as, así como la diversidad léxica de los diferentes actores involucrados. Por último, se presentarán los resultados obtenidos mediante la aplicación de las técnicas de comparación de discursos, específicamente la similitud entre los discursos sobre inseguridad de los medios y los candidatos/as.

3.1 Corpus de discursos

En primer lugar, resulta relevante presentar cómo terminó conformado el corpus en base a los relevamientos y filtros realizados. El corpus de entrevistas y tweets de los candidatos a presidente se conformó según se muestra en la *Tabla 2.*

Tabla 2: Cantidad de discursos por candidato

Fuente: Elaboración propia en base a Dicen los Medios

Precandidatos a presidente	Discursos sobre educacion	Tweets sobre educación	Tweets totales	% de tweets sobre educación
Juan Schiaretti	5	119	1.426,0	8,3%
Myriam Bregman	5	369	5.781,0	6,4%
Patricia Bullrich	2	44	1.301,0	3,4%
Sergio Massa	4	39	1.163,0	3,4%
Javier Milei	14	631	43.435,0	1,5%

Como puede observarse en la Tabla 2, el político con mayor cantidad de referencias al tema educación es Javier Milei. Sin embargo, cuando se toman en cuenta las frecuencias relativas, los candidatos que admiten mayor espacio en sus discursos para temas sobre educación son Juan Schiaretti y Myriam Bregman.

Tabla 3: Cantidad de noticias por diario

Fuente: Elaboración propia en base a Dicen los Medios

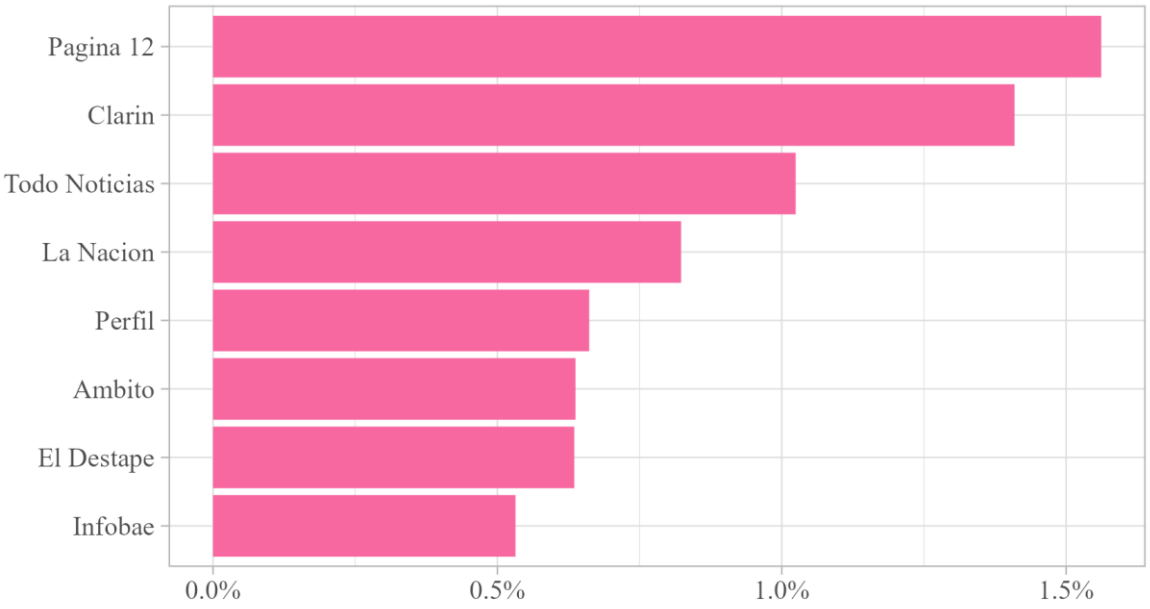
Diarios	Noticias sobre educación	Noticias totales	%
Ambito	209	32.793,0	0,6%
Clarin	501	35.550,0	1,4%
El Destape	272	42.821,0	0,6%
Infobae	1.322	248.588,0	0,5%
La Nacion	446	54.195,0	0,8%
Pagina 12	337	21.579,0	1,6%
Perfil	173	26.157,0	0,7%
Todo Noticias	300	29.284,0	1,0%

Respecto al corpus de noticias, en la *Tabla 3* se muestra la distribución de las noticias relacionadas con la temática educación en los diferentes medios. En este punto, se observa

una situación similar en cuanto a la diferencia en la cantidad de noticias entre los diferentes diarios, pero que, de alguna manera, se encuentra relacionada con la frecuencia de publicación de cada portal digital, por ejemplo, Infobae es por lejos el mayor productor de noticias sobre educación entre los diarios comparados, sin embargo, es de los que menos espacio le otorga a la temática en la producción general de contenido.

Gráfico 1: Noticias sobre Educación

Porcentaje por diario



Fuente: Elaboración propia en base a Dicen los Medios

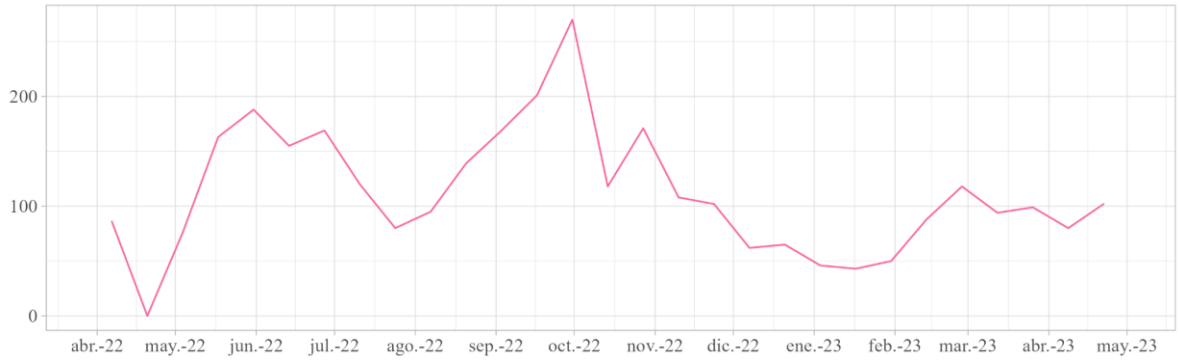
En este sentido, el *Gráfico 1* muestra el porcentaje de noticias sobre educación por cada portal de noticias. Se podría dividir la gráfica en dos grupos, uno que atiende la temática de forma más constante (Página 12, Clarín y Todo Noticias) y otro que lo aborda menos asiduamente y de manera más particular (La Nación, Perfil, El Destape, Ámbito Financiero e Infobae).

En el *Gráfico 2* se observa la evolución de la cantidad total de noticias que abordan la temática educación en el período analizado. Se podría interpretar la tendencia de la cobertura como un comportamiento cíclico con temporadas de menor y mayor representación en la agenda mediática. Entre septiembre y octubre de 2022, se observa un pico de noticias que puede corresponder no sólo al dia del maestro y al día del estudiante (11 y 21 de septiembre) sino también a que durante esos meses se sucedieron las paritarias docentes y la toma de 15

colegios de la Ciudad de Buenos Aires incluido el Mariano Acosta[5], que resultó en un protocolo "anti toma de colegios" donde las familias debían hacerse cargo de los costos y gastos que se desprendieron de los daños de los establecimientos[6].

Grafico 2: Noticias sobre educacion

Desde el 1 de abril de 2022 hasta el 30 de abril de 2023



Fuente: Elaboración propia en base a Dicen los Medios

3.2 Composición discursiva

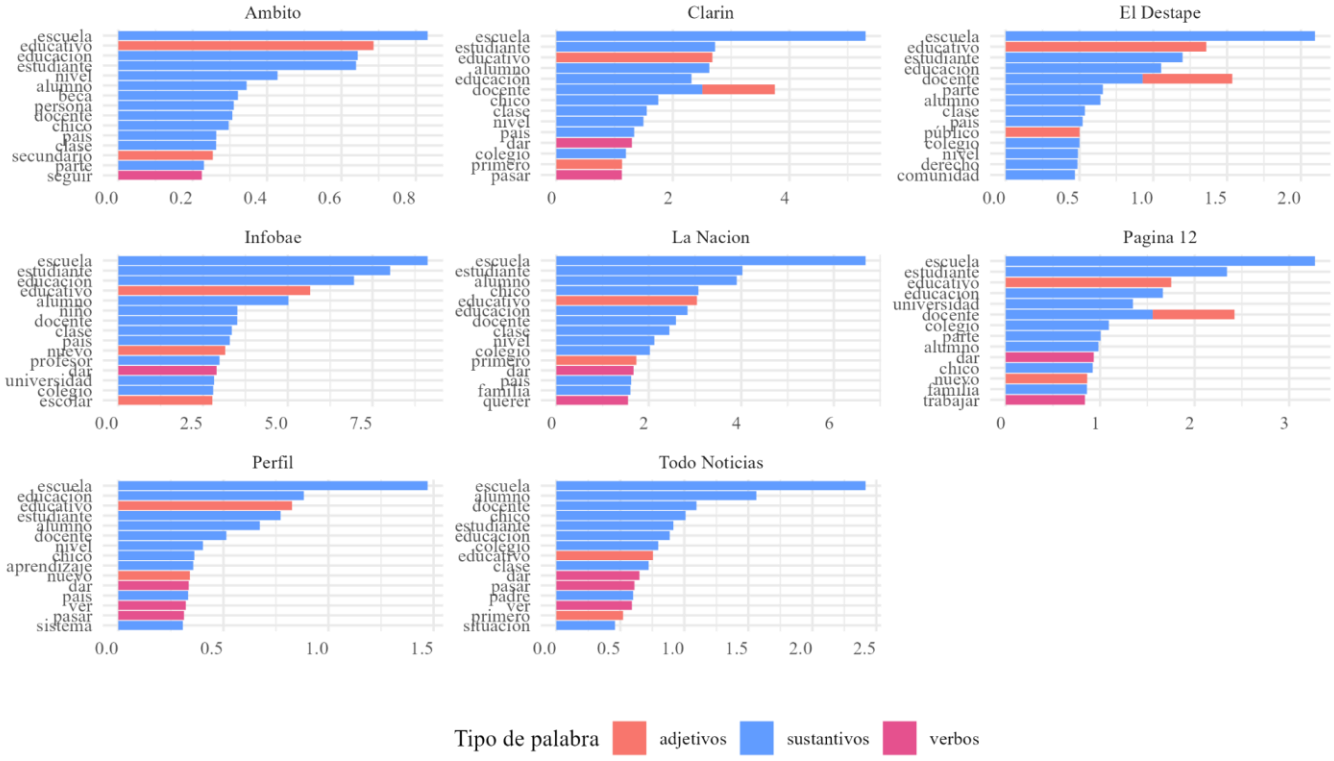
Respecto a los términos utilizados por cada corpus, en los *Gráficos 3* y *4* se presentan los sustantivos, adjetivos y verbos más mencionados por cada candidato y medio digital.

[5] https://www.infobae.com/educacion/2022/09/23/tomas-de-colegios-en-caba-el-mariano-acosta-definio-la-medida-y-podria-extenderse-a-otras-20-escuelas

[6] https://www.infobae.com/educacion/2022/09/26/el-gobierno-porteno-califico-la-toma-del-colegio-mariano-acosta-como-una-medida-violenta-y-responsabilizo-al-kirchnerismo/

Gráfico 3: Palabras más utilizadas por cada portal de noticias

En millones. Del 01 de abril de 2022 al 30 de abril de 2023

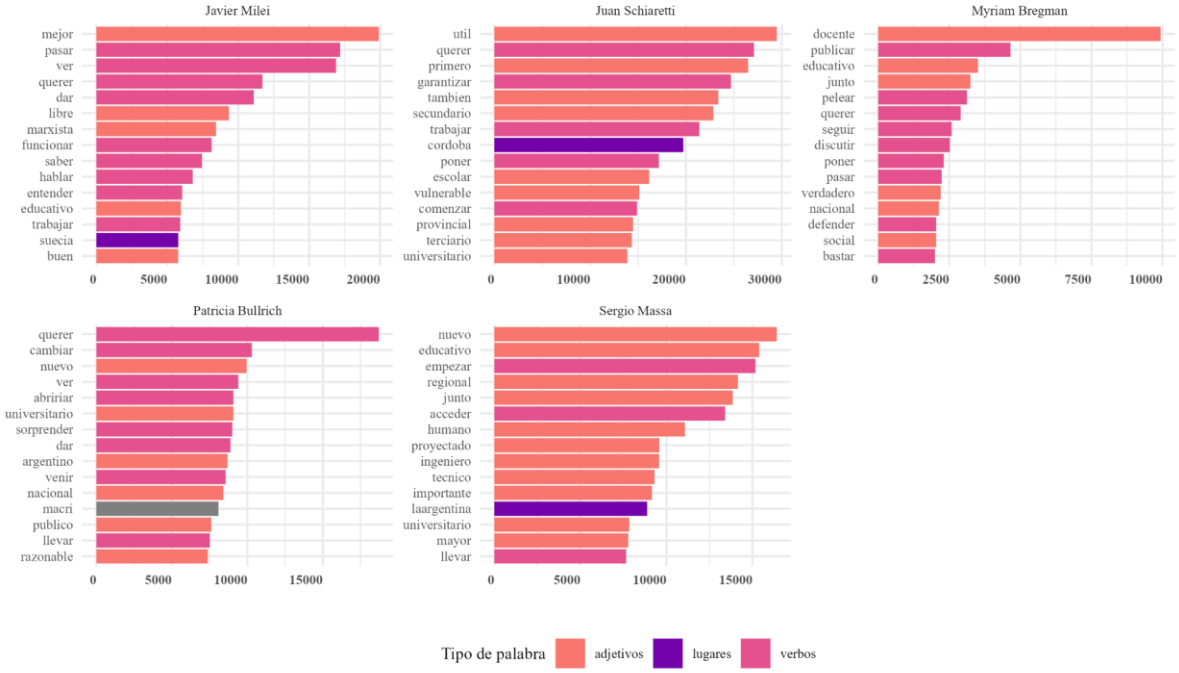


Fuente: Elaboración propia en base a Dicen los medios

Del *Gráfico 3* podemos observar que las palabras más mencionadas por los medios digitales son: escuela, educativo, educación, alumno, beca, estudiante. En líneas generales, los términos utilizados por los diferentes medios coinciden entre sí.

Gráfico 4: Palabras más utilizadas por cada precandidato

Del 01 de abril de 2022 al 30 de abril de 2023

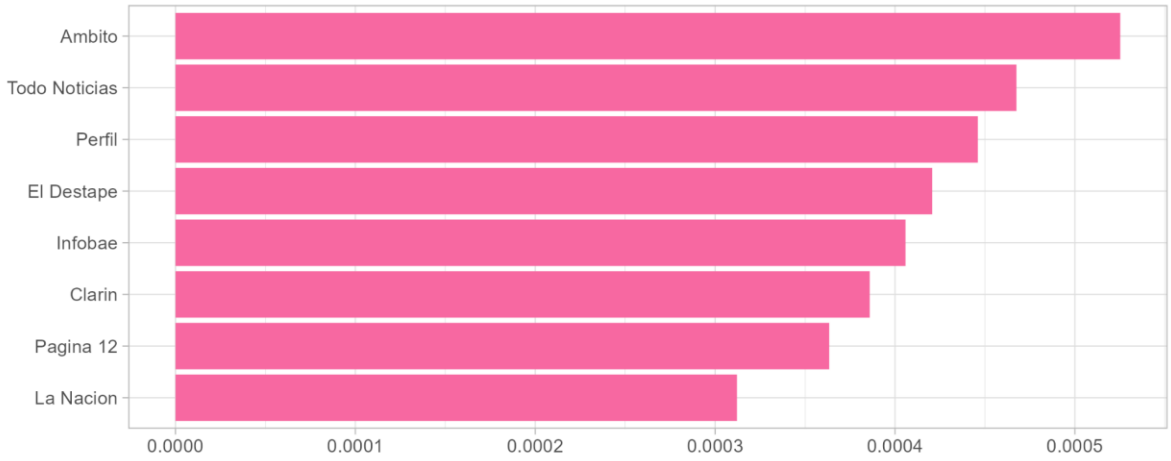


Fuente: Elaboración propia en base a Dicen los medios

El *Gráfico 4* refleja los principales términos utilizados por los candidatos a presidente, se destacan términos que agregan dimensiones en comparación con las palabras utilizadas por los portales digitales. Las referencias a la *protesta* en relación a la *defensa de la educación* con palabras como *"pelear" "defender" "seguir"* aparecen en las enunciaciones de la candidata del Frente de Izquierda (Myriam Bregman). Javier Milei tiene menciones a "Suecia", "libre" y a "marxismo" como parte de su retórica sobre la educación. Juan Schiaretti y Sergio Massa poseen referencias más relacionadas con el proyecto educativo como son los diferentes niveles, palabras como "garantizar" "poner" "acceder" "empezar" "nuevo". Por último, Patricia Bullrich tiene palabras referidas a la propuesta de "Cambio", al "macrismo" y a "sorprender".

Gráfico 5: Diversidad de léxico por portal de noticias

En noticias sobre educación (Abril 2022 - Abril 2023)



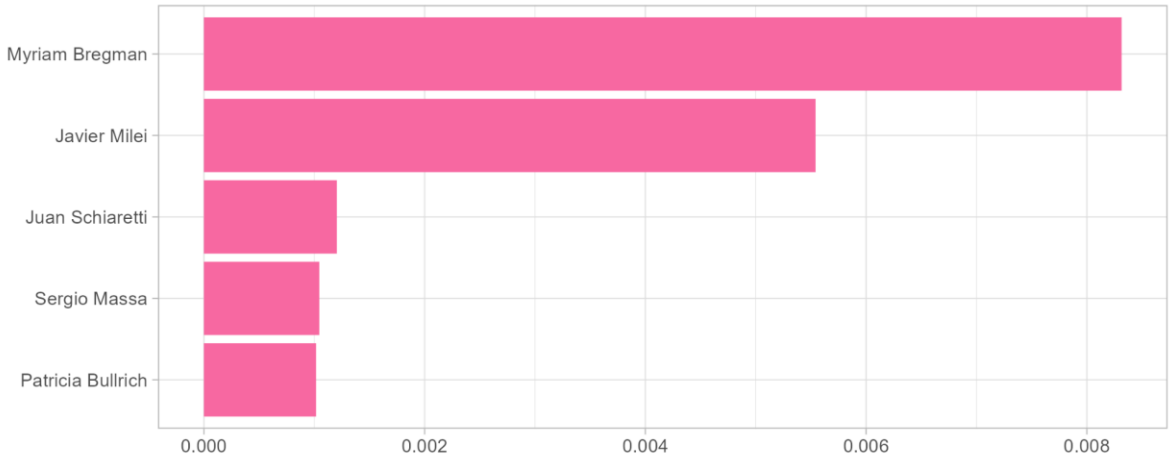
Fuente: Elaboración Propia en base a Dicen los medios

Los *Gráficos 5* y *6* muestran el grado de diversidad léxica de los portales digitales y candidatos a presidente. Este indicador, puede ser entendido como el cociente entre los términos únicos mencionados por cada actor y la totalidad de palabras expresadas. El mismo toma valores de 0 a 1 y puede ser interpretado de la siguiente manera: si el resultado es cercano a 1, existe una alta diversidad léxica, es decir, se utilizan muchas palabras diferentes en relación con el tamaño total del texto. En cambio, cuando la diversidad léxica es cercana a 0, significa que existe una menor variedad de palabras utilizadas.

En el Gráfico 5 se observa que, si bien en términos porcentuales el grado de diversidad léxica es pequeña en la generalidad de los diarios, son Ámbito, Todo Noticias y Perfil los medios con mayores palabras diferentes expresadas en las noticias.

Gráfico 6: Diversidad de léxico por precandidato

En discursos sobre educación (Abril 2022 - Abril 2023)

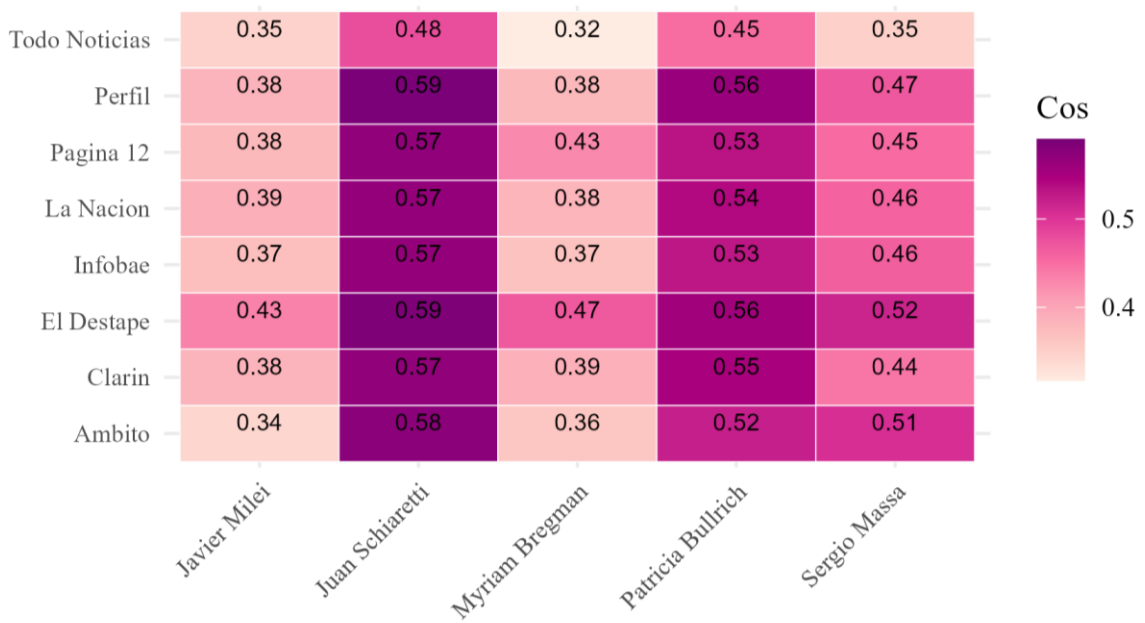


Fuente: Elaboración Propia en base a Dicen los medios

En el *Gráfico 6* se muestra el grado de diversidad léxica por cada candidato a presidente en discursos sobre educación. Al respecto, Myriam Bregman es por lejos la candidata que mayor diversidad de léxico utiliza al referirse al tema educación, en segundo lugar, se ubica Javier Milei. En el caso de Juan Schiaretti, Sergio Massa y Patrici Bullrich casi no hay diferencias en el grado de diversidad léxica utilizado. Podría decirse que los candidatos más extremos dentro del espectro político nacional son los que mayor verborragia poseen al tratar el tema educación en sus discursos y quizás que aquellos con cierto recorrido en el ámbito público tienen mayor moderación u homogeneidad al referirse a temas sobre educación.

3.3 Similitud de discursos

Gráfico 7: Similitud de discursos sobre educación entre candidatos a presidente y diarios digitales



Fuente: Elaboración propia en base a Dicen los Medios

Finalmente, en el *Gráfico 7* se puede observar el grado de similitud de discursos entre los portales de noticias digitales y los candidatos a presidente. En primer lugar, se observa que Juan Schiaretti es quién representa el mayor grado de similitud discursiva sobre educación con casi todos los medios, siendo el diario Perfil y El Destape los que poseen mayor valor; al igual que Patricia Bullrich en una proporción menor. Por el contrario, Javier Milei y Myriam Bregman son los candidatos que menor similitud posee con respecto a los diarios digitales. Esto puede deberse a que ambos candidatos hablan sobre la temática con terminologías que,

si bien son relevantes para su plataforma ideológica, no son parte de la conversación común en la educación (Ej. Vouchers y lucha social)

Esta divergencia en la terminología utilizada por estos grupos políticos sugiere una brecha significativa entre su enfoque y las discusiones educativas que suelen presentarse en los medios de comunicación convencionales, lo que plantea preguntas sobre cómo estos candidatos se comunican y cómo son percibidos por el público en un contexto electoral.

Finalmente, Sergio Massa presenta mayor similitud con los diarios El Destape, Ámbito y Perfil y mucho menor con Todo Noticias. Esto podría deberse a su rol en el Ministerio de Economía y a las concepciones económicas y de gestión que involucra en sus discursos

4. CONSIDERACIONES FINALES

En esta sección se presentan reflexiones sobre el análisis realizado y la metodología elaborada y empleada para abordar la ponencia, si bien los hallazgos obtenidos en el presente trabajo contienen sustento teórico y metodológico, algunos cuestionamientos o dudas fueron surgiendo a medida que se avanzaba con la investigación, elementos que podrían fortalecer futuras producciones.

En primer lugar, cabe preguntarse si la cantidad de unidades de análisis relevadas, principalmente por el lado de los candidatos presidenciales, podría influir en el valor obtenido sobre la similitud de discursos. A mayor cantidad de entrevistas y tweets, los vectores formados por las palabras de los discursos encuentran valores más altos, por el contrario, a menor cantidad de unidades relevadas, los vectores se fuerzan a tomar como valor 0 (cero) para las palabras que no se encuentran presentes. De esta manera, si para el candidato A no se encuentran los términos “escuela”, “estudiante” y “docente” en su discurso, mientras que los mismos sí se encuentran en los discursos de algunos diarios, el vector que se forma para representar el discurso del político toma valores 0 (cero) para esos términos. Esto podría producir que para los políticos con mayor cantidad de unidades de análisis conseguidas, los vectores tomen valores más semejantes a los discursos de los diarios.

En segundo lugar, existen enfoques mucho más complejos para el análisis de términos, tal como lo demuestran desarrollos informáticos al estilo de Chat-GPT y Bard. Estos desarrollos utilizan algoritmos basados en el contexto de los términos (como *word embeddings)* y utilizan tecnologías como Transformers y Multi-Head Attention. Un ejemplo que se pudo observar en el presente trabajo consiste en las palabras “primaria”, “secundaria” o “terciaria”, que al ser utilizadas por los medios o los políticos con referencia a conceptos como “elección primaria” cobran significados diferentes al que se busca en esta investigación. Si bien en el apartado metodológico se describe la búsqueda de una manera de excluir este tipo de noticias, un acercamiento desde técnicas que utilizan mecánicas de word embeddings podría arrojar resultados mucho más sólidos. Por otro lado, en el presente trabajo se realizó un análisis y comparación de discursos con cierta efectividad, basándose en una sóla dimensión de los corpus discursivos: el léxico utilizado por cada agente. Al respecto, agregar una dimensión vinculada con el contexto de los léxicos podría mejorar y enriquecer los resultados obtenidos. La oración *"las escuelas primarias no aportan al desarrollo económico de manera positiva."* resulta muy similar a la oración *"el desarrollo económico es dependiente de la inversión positiva en las escuelas primarias."* y sin embargo ambas oraciones se orientan a significados

totalmente opuestos. En este sentido, no podría afirmarse que ambos discursos sean similares, aunque pudieran haber obtenido un alto valor de similitud coseno.

En otro orden de ideas, el presente análisis abre nuevos interrogantes que bien pudieran volverse futuras líneas de investigación. Uno de ellos se centra en cómo medir la transferencia de elementos discursivos de un agente a otro, es decir, la transferencia de ideas, conceptos, palabras desde medios de comunicación a políticos o viceversa. Como puntapié para esta posible línea, el análisis debería considerar la variable temporal para medir el momento en que se empezó a hablar por primera vez de un tema en alguno de los dos espacios, y luego identificar ese tema y sus características en el otro espacio. Este interrogante podría encontrar sustento en trabajos sobre el 4to nivel de Agenda Setting, que se enfoca en la pregunta sobre quién fija la agenda de temas de los medios de comunicación. Si bien los principales desarrollos realizados bajo el amparo de este marco teórico se centran en la relación entre los propios medios (en sus diferentes formatos: prensa papel, televisión, radio, etc) y las agencias de noticias, es posible encontrar trabajos referidos a discursos presidenciales como fijadores de agenda[7].

Además, otro interrogante que surge, en un nivel más abstracto al de este trabajo, gira en torno a la posibilidad recientemente abierta para las ciencias sociales sobre abordajes más cuantitativos a fenómenos sociales anteriormente estudiados en claves teóricas. Los actuales desarrollos informáticos, junto a la aceleración creciente de producción de datos y el ingenio para acceder a los mismos permiten la operacionalización y la incorporación de enfoques reservados a otras ciencias comúnmente llamadas “duras”. Esta pregunta tiene el riesgo de caer o abrir falsas dicotomías, quizás las preguntas más concretas recaen en sí estos nuevos abordajes son consistentes y cómo pueden plantearse en una clave complementaria para los enfoques teóricos, con el fin de aumentar su riqueza cognitiva.

Finalmente, en trabajos anteriores se dejó abierta la posibilidad de que la metodología diseñada pudiera ser replicada en futuras investigaciones o no y en el presente trabajo se tomó este desafío como objetivo secundario. Al respecto podemos concluir que la metodología utilizada es consistente y útil para el abordaje de la similitud de discursos entre diferentes corpus y ofrece como ventajas una alta flexibilidad (traducida en la alta capacidad de adaptación a diferentes temáticas discursivas), un bajo costo de mejoramiento y rapidez para obtener resultados.

[7] Gilberg, S.; Eyal, Ch.; McCombs, M.; Nicholas, D. “The State of the Union Address and the Press Agenda”, en Protess, D.; McCombs, M. (1991). Op. cit., pp. 223-236 y Wanta, W.; Stephenson, M.A.; Turk, J.; McCombs, M. “How president’s State of Union Talk Influenced News Media Agenda”, Journalism Quarterly, N° 66, 1992, pp. 537-541.

REFERENCIAS BIBLIOGRÁFICAS

Acosta, M., Demirdjian, S., y Lassi, A. (2021). El framing de campaña. Las elecciones presidenciales de 2015 en la prensa de referencia argentina. *RIHUMSO,* *9*(18), 3-27. Recuperado de https://repositoriocyt.unlam.edu.ar/bitstream/123456789/697/1/RiHumSo%209-18-1.pdf

Aruguete, N. (2015). *El poder de la agenda: política, medios y público*, Buenos Aires, Argentina: Biblos.

Aruguete, N. (2009). Estableciendo la agenda. Los orígenes y la evolución de la teoría de la Agenda Setting, *Ecos de la comunicación, 2*(2). Recuperado de https://repositorio.uca.edu.ar/bitstream/123456789/7119/1/estableciendo-la-agenda.pdf.

Barrera, M. C. (2014). Minería de texto: una visión actual. *Biblioteca Universitaria*, *17*(2), 129-138. Recuperado de https://www.redalyc.org/pdf/285/28540279005.pdf

Califano, B. (2015). Los medios de comunicación, las noticias y su influencia sobre el sistema político. *Revista Mexicana de Opinión Pública, 19,* 61-78. Recuperado de https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-49112015000200007

Casermeiro, A. (2003). *Los medios en las elecciones: la agenda setting en la Ciudad de Buenos Aires*. Buenos Aires, Argentina: EDUCA.

Castells, M. (2008). Comunicación, poder y contrapoder en la sociedad red (I). Los medios y la política. *Telos: cuadernos de Comunicación e innovación,* 74, 13-24.

Castells, M. (2009). *Comunicación y poder*. Madrid, España: Alianza Editorial.

Durán Chinchilla, C. M. ., García Quintero, C. L. ., y Rosado Gómez, A. A. . (2022). Aplicación del procesamiento del lenguaje natural como técnica de análisis en la producción textual, caso estudiantes de Ingeniería de Sistemas UFPSO. *Revista Boletín Redipe*, *11*(1), 489–497. https://doi.org/10.36260/rbr.v11i1.1656.

Habermas, J. (1987), *Teoría de la acción comunicativa*, Madrid, España: Santillana.

Singh, R., Maurya, S., Tripathi, T., Narula, T., & Srivastav, G. (2020). Movie Recommendation System using Cosine Similarity and KNN, *International Journal of Engineering and Advanced Technology (IJEAT)*, 9(5), 556-559. Recuperado de https://www.ijeat.org/wp-content/uploads/papers/v9i5/E9666069520.pdf

Slimovich, A. (2019). La mediatización contemporánea de la política en Instagram. Un análisis desde la circulación hipermediática de los discursos de los preprecandidatos argentinos. *Revista Sociedad*, (39), 31-45.

van Dijk, T. A. (1999). El análisis crítico del discurso. *Anthropos,* (186), 23-36. Recuperado de http://www.discursos.org/oldarticles/El%20an%E1lisis%20cr%EDtico%20del%20discurso.pdf

van Dijk, T. A. (2009). *Society and Discourse: How Social Contexts Influence Text and Talk.* Nueva York, Estado Unidos: Cambridge University Press.

Verón, E. (Ed.).(1993). *La semiosis social: fragmentos de una teoría de la discursividad*. Barcelona, España: Editorial Gedisa.

Zhao, B. (2017). *Web Scraping*. Recuperado de https://www.researchgate.net/profile/Bo-Zhao-3/publication/317177787_Web_Scraping/links/5c293f85a6fdccfc7073192f/Web-Scraping.pdf