



Aportes de la lexicometría a las percepciones sobre la transición energética

Marcela F. Albornoz, Juan M. Paris

Resumen: En el marco de un proyecto de investigación, perteneciente a una Universidad, que aborda problemáticas vinculadas a la industria de los hidrocarburos, se realiza una encuesta durante el mes de septiembre de 2021 para indagar la percepción de los ciudadanos respecto de diferentes temas relacionados con la industria. La transición energética surge como una solución al cambio climático que es comúnmente asociado a la quema de combustibles fósiles. En este trabajo se pretende investigar a través del vocabulario cómo están instaladas estas ideas en las personas. Se analizan las respuestas a las preguntas abiertas: "Explique con sus palabras que entiende por transición energética", "¿Cuál es su opinión sobre la transición energética?" y "¿Quiénes deberían liderar este cambio que lleve a la transición energética?", utilizando el análisis estadístico de datos textuales. Estas preguntas forman parte de una encuesta más extensa. El empleo de la estadística textual en el análisis de respuestas a preguntas abiertas aporta información acerca de lo que los sujetos investigados piensan ya que las ideas se manifiestan a través del vocabulario. Adicionalmente, este tipo de análisis permite diferenciar y agrupar a los sujetos estableciéndose categorías que prescinden de la subjetividad del investigador. Se utiliza el lenguaje R complementándolo con el software SPSS Statistics. Se analizan 188 encuestas, con el fin de identificar diferencias y similitudes léxicas entre grupos etarios. Mediante Lexicometría se contextualizan las palabras construyendo tablas de contingencia de los diferentes grupos. Estas características identificadas en el léxico son estudiadas mediante el Análisis Factorial de Correspondencia aplicado a la tabla léxica agregada. Durante este proceso, las respuestas no son modificadas ni categorizadas. Este trabajo permite determinar que palabras están instaladas en la mente de las personas respecto de la transición energética y corrobora que existe una diferencia léxica en las percepciones entre los diferentes grupos etarios.

Palabras clave: Lexicometría; Transición Energética; Preguntas Abiertas.

Contributions of Lexicometry to Perceptions about the Energy Transition

Abstract: Within a research project, belonging to a University, that addresses problems related to the hydrocarbon industry; A survey is carried out during the month of September 2021, to investigate the perception of citizens regarding different issues related to the industry. The energy transition emerges as a solution to climate change, which is commonly associated with the use of fossil fuels. This research aims to investigate through vocabulary how these ideas are installed in people's minds. The answers to the open questions are analyzed: "Explain in your own words what you understand by energy transition", "What is your opinion about the energy transition?" and "Who should lead this change that leads to the energy transition?", using the statistical analysis of textual data. These questions are part of a larger survey. The use of textual statistics in the analysis of answers to open questions provides information about what people think about it, since ideas are manifested through vocabulary. Additionally, this type of analysis allows differentiating and grouping the subjects, establishing categories that leave aside the subjectivity of the researcher. The R language is used, complementing it with the SPSS Statistics software. 188 surveys are analyzed in order to identify lexical differences and similarities between age groups. Through the use of Lexicometry, the words are contextualized by building contingency tables of the different groups. These characteristics identified in the lexicon are studied by Correspondence Factor Analysis applied to the aggregated lexical table. During this process, responses are not modified or categorized. This research allows us to determine what words are installed in people's minds regarding the energy transition and corroborates that there is a lexical difference in perceptions between different age groups.

Keywords: Lexicometry; Energy Transition; Open Questions.

1. Introducción y Objetivo

En el marco de un proyecto de investigación que aborda problemáticas vinculadas a la industria de los hidrocarburos, perteneciente a la Universidad Nacional del Comahue, se realiza una encuesta durante el mes de septiembre de 2021 para explorar las percepciones de los ciudadanos respecto de diferentes temas relacionados con la industria.

Una de las problemáticas abordadas es la transición energética. Dado que la transición energética surge como una solución o mitigación al cambio climático, que es comúnmente asociado a la quema de combustibles fósiles, el tema tiene una relación directa con la industria de los hidrocarburos.

Para indagar las ideas, juicios y percepciones de la población, existen diversas posibilidades. Pueden ser mediante la implementación de encuestas con preguntas cerradas, de interpretación sencilla, o mediante preguntas abiertas. El método de indagación más adecuado, son las encuestas con preguntas abiertas o bien entrevistas semiestructuradas, pues lo que expresan los individuos encuestados con respecto a un determinado tema es reflejo de las ideas o concepciones en ellos subyacentes.

El objetivo del presente estudio es indagar a través del vocabulario como están instaladas estas ideas en las personas con respecto a la transición energética. Vale mencionar que no se trata de analizar la veracidad de las respuestas, sino de visualizar las diferentes percepciones y puntos de vista. No es objetivo del presente estudio analizar la transición energética “per se”, por lo cual no se ahondarán temas técnicos vinculados a la misma.

Para ello, se utilizan técnicas de la estadística descriptiva multivariada, más específicamente las desarrolladas para el tratamiento de variables cualitativas o categóricas que constituyen una herramienta básica del análisis de datos textuales o lexicometría.

El análisis estadístico de datos textuales o lexicometría (Lebart y Salem, 1988,1994) es un área de la estadística que se desarrolla ante la necesidad de poseer una herramienta para analizar las preguntas abiertas realizadas en encuestas. La lexicometría trata de alejar la mirada subjetiva del investigador (Becue, et al., 1995) y analizar los datos textuales después de diversas codificaciones a fin de obtener información sobre frecuencia de palabras, contexto en el que se hallan, frecuencia de dichos contextos, riqueza del vocabulario, etc. Estas frecuencias son posteriormente analizadas utilizando técnicas estadísticas multivariadas.

El desarrollo de estas técnicas ha hecho que el análisis estadístico de textos se constituyera en una herramienta interdisciplinaria, integrada por la estadística, el análisis del discurso, la lingüística, la informática, el procesamiento de encuestas, la investigación documental y es cada vez más utilizada en diversos campos de las ciencias sociales: historia, política, economía, sociología, psicología, entre otros (Albornoz, 2018).

El empleo de la Estadística Textual o Lexicometría (Bécue, 1991; Lebart y Salem, 1994) en estas investigaciones, puede facilitar el conocimiento de las ideas que poseen los sujetos investigados; ya que las mismas se manifiestan en el manejo del vocabulario, concretamente en el uso predominante de ciertas palabras y en las frecuencias de su empleo (Baccalá y de la Cruz, 1995, 2000).

La Estadística Textual trabaja con tablas de contingencia que pueden ser de respuestas (individuos) por palabras distintas, que se denomina Tabla Léxica, o agrupaciones de respuestas (textos) por palabras distintas, que es denominada Tabla Léxica Agregada.

El análisis de esta información se realiza utilizando el análisis factorial de correspondencias simple (AFC).

El AFC aplicado a las tablas léxicas da una visualización de las proximidades entre individuos y entre formas, permitiendo observar que formas y/o expresiones diferencian a los individuos. Procede por comparación de perfiles léxicos. Con la utilización de este método no se trata de saber qué dicen los individuos, pero sí saber si dicen lo mismo (Bécue, 1991).

En síntesis, el AFC permite poner en evidencia los grandes rasgos estructurales relativos a ambos conjuntos (formas gráficas e individuos o grupos de individuos) analizados, mediante proyecciones sobre subespacios de dimensión reducida, pero manteniendo la máxima dispersión de los datos originales.

2. Metodología

Se realiza una encuesta durante el mes de septiembre de 2021 para explorar las percepciones de los ciudadanos respecto de diferentes temas relacionados a la transición energética. Son analizadas en este estudio las respuestas a tres preguntas abiertas:

1. Explique con sus palabras que entiende por transición energética.
2. ¿Cuál es su opinión sobre la transición energética?
3. ¿Quiénes deberían liderar este cambio que lleve a la transición energética?

Se analizan 188 encuestas, utilizando el lenguaje R versión 4.1.1 (2021-08-10) y complementado el mismo con el software SPSS Statistics versión 25, para el procesamiento de la información.

La población es segmentada en cinco grupos etarios, según los siguientes rangos: 25 o menor, 26 a 35, 36 a 45, 46 a 55, 56 o más años.

En el análisis del corpus se aplican las técnicas de la estadística textual o lexicometría. Se eliminan las palabras conectoras (de, y, que, etc.) y los artículos (una, un, la, el), seleccionándose para el estudio las formas gráficas que por sí solas poseen significado, como ser sustantivos, verbos, adverbios y adjetivos. Durante este proceso, las respuestas no son modificadas ni categorizadas, honrando de esta forma lo que cada individuo ha expresado. Para que las palabras tengan un aporte significativo se considera las que presentan una frecuencia mayor e igual a 7, con longitud mayor o igual a 4 caracteres o letras y que están presentes al menos en 4 documentos.

Para evaluar si existe entre los grupos etarios una diferencia léxica, se construyen tablas de contingencia, en las cuales los diferentes grupos etarios ocupan las filas y las palabras ocupan las columnas.

3. Análisis de los Resultados

Teniendo en cuenta el objetivo propuesto para esta investigación se fijan los umbrales: cantidad mínima de documentos, cantidad mínima de caracteres y frecuencia mínima.

Se detectan en el léxico determinadas palabras que presentan una asociación natural en el lenguaje, es decir, cadenas de palabras que se repiten en el mismo orden. Por lo tanto, se consideran como palabras a las estructuras tales como “combustiblesfósiles”, “matrizenergética” y “nosé”.

Cabe mencionar que en todas las preguntas analizadas el porcentaje de individuos que manifestaron no tener conocimientos sobre el tema (respuestas “No sé”) es inferior al 10%. Esto indica que los individuos poseen, en su mayoría, conocimientos sobre los temas tratados.

Para una mejor interpretación se analizan las respuestas a las preguntas de manera individual. A continuación, se presentan los resultados de cada una de ellas.

3.1 Pregunta 1: Explique con sus Palabras que Entiende por Transición Energética

La figura 1 muestra el primer plano factorial del análisis Factorial de Correspondencias Simples aplicado a la tabla léxica. En el mismo se observa que el primer eje explica el 42.98% de la variabilidad y el segundo 25.10%.

En el mismo se observa que no existe una asociación de palabras con los diferentes grupos etarios en particular¹, lo cual significa que el léxico es mayormente compartido por todos los grupos. En la figura 1 sólo son visualizadas las palabras.

¹ Por esta razón no se incluyen los grupos etarios en el plano factorial de la figura 2.

Las palabras utilizadas para describir que entienden por transición energética (figura 1), se estructuran en general según variantes del siguiente discurso:

“La transición energética es *cambiar*² de una *matriz energética* basada en la quema de *combustibles fósiles* (hidrocarburos como el *petróleo* o *carbón*), las cuales producen *emisiones contaminantes* para el *medio ambiente*, por energías *renovables, limpias y sustentables* tales como la *eólica y solar*.”

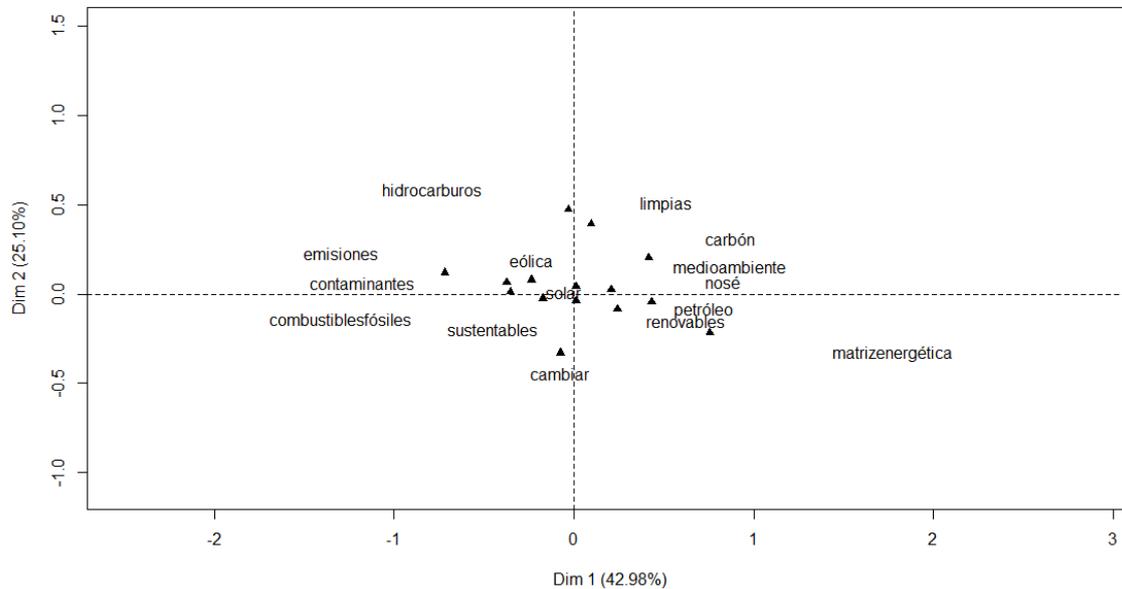


Figura 1. Plano factorial correspondiente a la pregunta 1. Sólo son graficadas las palabras.

Cabe mencionar que dentro de las energías sustentables solo están instaladas en el lenguaje la energía eólica y solar; no siendo mencionadas las restantes energías limpias (hidráulica, geotérmica, mareomotriz, biomasa, etc.).

Si bien se ha estructurado una definición utilizando todas las palabras analizadas, en general, los individuos encuestados utilizan sólo alguna de las palabras mencionadas, brindando definiciones en algunos casos parciales sobre el tema.

Otro aspecto a resaltar, es que, si comparamos la definición estructurada con la definición que figura en Wikipedia, ambas presentan un alto grado de similitud:

“La transición energética actual surge por la necesidad de acciones climáticas para mitigar el calentamiento global.⁵ Para mantenerse dentro de los 1.5°C que propone el Acuerdo de París, es necesario detener las emisiones de gases de efecto invernadero para 2040 o 2050.

Esto implica descarbonizar los sistemas energéticos, es decir, abandonar los combustibles fósiles como el petróleo, el gas natural, el lignito y el carbón, y reemplazarlos por fuentes que no generen o generen bajas emisiones como el combustible nuclear (uranio) y las fuentes de energía renovable eólica, hidroeléctrica, solar, geotérmica, marina, undimotriz, entre otras”³.

Si bien no es objetivo del presente estudio, evaluar la veracidad de las afirmaciones de los individuos, sino más bien, intentar visualizar sus percepciones, resulta evidente que los individuos encuestados tienen una percepción del significado de transición energética mayormente correcta.

² La tipografía *italica* es utilizada para indicar las palabras analizadas (las cuales figuran en el plano factorial).

³ Fuente: https://es.wikipedia.org/wiki/Transici%C3%B3n_energ%C3%A9tica

3.2 Pregunta 2: ¿Cuál es su Opinión Sobre la Transición Energética?

La figura 2 muestra el primer plano factorial del análisis Factorial de Correspondencias Simples aplicado a la tabla léxica. En el mismo se observa que el primer eje explica el 37.24% de la variabilidad y el segundo 25.30%. El primer eje permite diferenciar a los grupos de individuos menores a 45 años, de los de 46 a 55 años; mientras que el segundo eje diferencia los individuos mayores de 56 años con respecto del resto de los grupos.

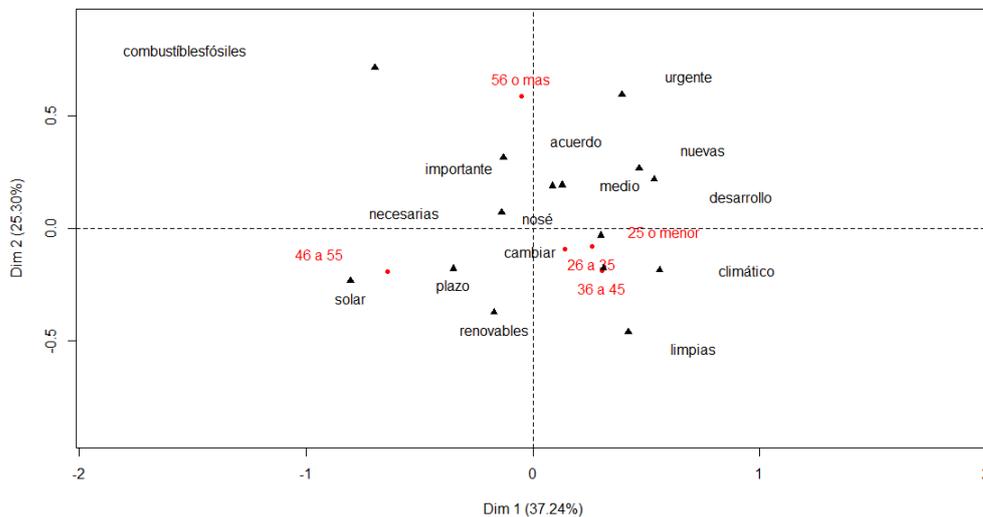


Figura 2. Plano factorial correspondiente a la pregunta 2. Las palabras son graficadas en color negro y los grupos etarios en color rojo.

Al analizar los planos factoriales debe observarse con especial cuidado las distancias entre palabras, las distancias entre grupos etarios y las distancias entre palabras y grupos etarios.

La cercanía entre tres grupos etarios menores a 45 años, indican que comparten mucho del contenido léxico, mientras que la lejanía entre estos grupos y los mayores de 56 y los que tienen entre 46 a 55 años, es índice que estos grupos utilizan diferentes palabras como respuesta a la pregunta planteada. Existe un vocabulario compartido por los grupos (palabras ubicadas en general en el sector central del gráfico y equidistantes de los grupos etarios) y vocablos característicos de grupos etarios en particular.

En el siguiente listado las palabras son expresadas dentro del contexto en que han sido utilizadas.

Vocabulario compartido por todos los grupos:

- *Cambiar* de energías fósiles a energías renovables / estar de *acuerdo* / *no sé* / implementación de medidas y políticas *necesarias* / *importante*.

Los grupos con individuos menores a 45 años hacen hincapié en el desarrollo de *nuevas* energías *limpias* para frenar o mitigar el cambio *climático*.

Adicionalmente los individuos entre 46 y 55 años mencionan a la energía *solar* como alternativa y los mayores de 56 años la disminución en la quema de *combustibles fósiles*.

Una de las palabras más utilizadas por todos los grupos ha sido *cambiar*. Los mayores de 56 muestran una tendencia a considerarlo *urgente*, mientras que el resto de los grupos consideran que el cambio hacia energías *renovables* se producirá a largo *plazo*.

3.3 Pregunta 3: ¿Quiénes Deberían Liderar este Cambio que Lleve a la Transición Energética?

La figura 3 muestra el primer plano factorial del análisis Factorial de Correspondencias Simples aplicado a la tabla léxica. En el mismo se observa que el primer eje explica el 50.78% de la variabilidad y el segundo 33.21%.

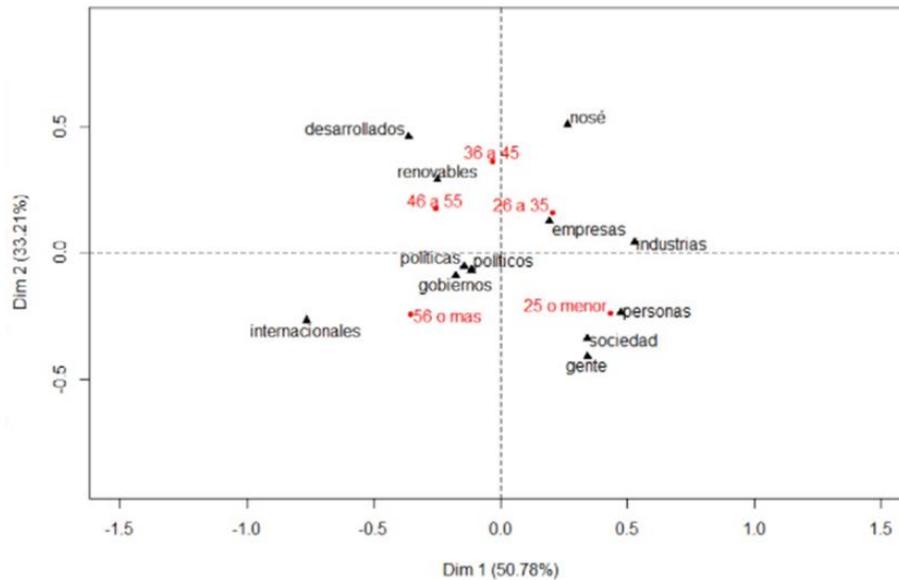


Figura 3. Plano factorial correspondiente a la pregunta 3. Las palabras son graficadas en color negro y los grupos etarios en color rojo.

Para los más jóvenes (25 años o menos) la responsabilidad de liderar la transición energética recae en las *personas*, la *sociedad*, la *gente* y en las *industrias*.

Por otro lado, para el grupo etario de mayor edad (56 años o más) la transición energética debe ser liderada por: los organismos u organizaciones *internacionales*, los líderes *políticos*, las *políticas* públicas y los *gobiernos*. Estas tres últimas palabras mencionadas (*políticos*, *políticas* y *gobiernos*) son también léxicos compartidos por con el resto de los grupos.

Forma parte también del vocabulario compartido: La implementación de energías *renovables* y las *empresas*.

Los países *desarrollados*. Siendo este último vocabulario compartido por los grupos etarios principalmente de edades intermedias, pero no representando una opción por los de 25 años o menos (observar que se encuentran distanciados en el plano factorial).

4. Conclusiones

El análisis de las características del léxico permite aproximar que ideas asocian los individuos con respecto a la transición energética a través de diferentes grupos etarios. Los resultados que emergen de los análisis muestran que los diferentes grupos manejan un vocabulario específico, pues es posible distinguir palabras que son propias de cada uno.

En todas las preguntas analizadas el porcentaje de individuos que manifestaron no tener conocimientos sobre el tema es inferior al 10%. Esto indica que los individuos poseen, en su mayoría, conocimientos sobre los temas tratados.

Respecto a la definición que proporcionan los individuos sobre la transición energética, si bien no es objeto de este trabajo evaluar el grado de veracidad de la información, esta se ajusta a las definiciones formales que existen sobre la misma; lo que indica que el común de los individuos maneja información correcta sobre el tema.

En cuanto a las opiniones personales sobre la transición energética, todos los grupos etarios la consideran importante. Los mayores de 56 muestran una tendencia a considerarla urgente, mientras que el resto de los grupos consideran que el cambio hacia energías renovables se producirá a largo plazo.

Por último, los más jóvenes (25 años o menos) consideran que la responsabilidad de liderar la transición energética debe recaer en las personas, la sociedad, la gente y en las industrias, mientras que para el grupo de mayor edad (56 años o más) la transición energética debe ser liderada por los organismos u organizaciones internacionales.

5. Referencias

- Albornoz, M (2018). Aportes de la Lexicometría en Indagaciones sobre la Enseñanza de la Matemática. Publicia, ISBN 978-620-2-43054-8. Dusseldorf, Alemania.
- Baccalá, N.; de la Cruz, M., (2000). La importancia de la estadística textual aplicada al estudio de las concepciones de enseñanza. In M. Rajman and J-C. Chappelier, editors, Proc. Of 5es Journées internationales d'Analyse statistique des Données Textuelles, Volume II. Copy Service Pillet, Martigny, pag. 519-524. Lausana , Suiza
- Bécue Bertaut, M. (2018). Textual Data Science with R. Chapman & Hall/CRC Computer Science & Data Analysis) 1st Edición, ISBN-13: 978-1138626911. Florida, EEUU.
- Bécue-Bertaut, M. (1991). Análisis de datos textuales, Métodos estadísticos y Algoritmos, CISIA, París
- Bécue, M.; Pagès, J. (1999). Intra-Set Multiple Factor Análisis. Application to textual data. In: proa. of the 9th International Symposium on Applied Stochastic models and Data Analysis, J. Jansen et al.(Eds.). Universidad de Lisboa. Lisboa, pp.51.60.
- Bécue, M.; Pagès, J. (2000). Analyse Factorielle Multiple intra-tableaux. Application à l'analyse simultanée de plusieurs questions ouvertes. In JADT 2000, 5ème Journées Internationales d'Analyse statistique de Donnée Textuelles, Rajman M. et Chappelier J.C.(Eds.)EPFL, Lausanne, pp. 425-432.
- Bécue, M.; Pagès, J. (2003). Análisis factorial múltiple para tablas de contingencia: Estudio de mortalidad en las comunidades autónomas de España. 27 Congreso Nacional de Estadística e Investigación Operativa, Lleida, 8-11 de abril de 2003
- Kiers, H. A. L. (1991). Hierarchical relations among three-way methods. Psychometrika, 56: 449-470.
- Lébart, L.; Salem, A. (1994). Statistique textuelles, Dunod, Paris.

Marcela F. Albornoz

Universidad Nacional del Comahue, Argentina

 <https://orcid.org/0000-0002-4299-1987>

✉ mfalbornoz.mat@gmail.com

Juan M. Paris

Universidad Nacional del Comahue, Argentina

 <https://orcid.org/0000-0001-7059-7197>

✉ mp.geol@gmail.com

Fecha de envío: 02/2022

Fecha de evaluación: 04/2022

Fecha de publicación: 07/2022