

Análisis lingüístico en contenidos sobre el tema “lazos creativos”: tratamiento semántico y representación de terminologías

Adilson Luiz Pinto (PGCIN/UFSC)

adilson.pinto@ufsc.br

Gabriel Ângelo da Silva Gomes (PGCIN/UFSC; Polícia Federal de Brasil)

gabriel.gasg@pf.gov.br

Raphael Baggio de Luca (PGCIN/UFSC; Polícia Federal de Brasil)

raphael.rbl@pf.gov.br

Rodrigo Gonçalves Teixeira (PGCIN/UFSC; Polícia Federal de Brasil)

teixeira.rgt@pf.gov.br

Ricardo Henrique Pereira (PGCIN/UFSC; Polícia Federal de Brasil)

henrique.rhp@pf.gov.br

1. Tema y Problema de estudio

El término lazo creativo no tiene autoría de funcionalidad[□], significado de actuación, ni de aplicación. Lo que hay, son tentativas de relacionar prácticas artísticas (Saura, 2014), de la administración (Lee, 2014), la política (Lesgart, 2013), las ciencias de la información (Teigland; McLure–Wasko, 2000), la salud mental (DeLaney, 2015), el análisis de redes (Daskalaki, 2010), la comunicación (Rehman; Ibrahim, 2011) y las ciencias de la computación

(Leichik; Shchelov, 2009) con el término, como algo que fusiona conjuntos, relaciones y comportamientos.

A pesar de la complejidad de la definición del término “lazo creativo”, intentaremos explicarlo a partir del punto de vista del que fue recuperado de datos e informaciones en sus innumerables aplicaciones, citadas en el párrafo anterior.

Para esta investigación, entenderemos el término “lazo creativo” como el proceso de creación dentro de un área, campo y hasta una ciencia, en el cual existen relaciones de acción con agentes, personas y sistemas. También creemos que se relaciona con el desarrollo continuo y, en cierta forma sin complicaciones, de interconectar la creatividad.

En relación a la creatividad, partimos del concepto de Ghiselin (1952), que la define como “el proceso de cambio, de desarrollo, de evolución en la organización de la vida subjetiva”. Sin embargo, hay complementos en este significado y significancia representados por las simbologías, sean éstas escritas, numéricas o figurativas.

Ante los elementos y la forma de cómo pensamos el entendimiento de los lazos creativos, este tema tiene un problema principal: el de no tener una terminología fundada ni conocer el universo literario sobre el tema, así como no saber las capilaridades de análisis del tema. Cuando presentamos nuestra problemática, estamos tomando, de cierta forma, el universo científico, que hasta la actualidad aún no acuñó el término con significado y significancia. También vale aclarar que no tenemos la pretensión de acuñar una definición, sino de explorar cuáles son las áreas y representaciones semánticas que están por detrás de este proceso.

Así, este estudio busca analizar los títulos publicados sobre el término “lazos creativos” en los idiomas español, inglés, francés y portugués para poder trazar algunas reflexiones semánticas al tema. El objetivo es mapear la literatura sobre el tema agregando significado semántico significativo, a partir del descubrimiento del grado de términos relevantes (Ley de Zipf), identificar contenido semejante, como documentos, hapax, identificar clases de palabras y aplicar un análisis métrico para identificar autoridades (Ferrazza; Pinto, 2017).

2. Análisis Lingüístico para nuestro tema

La lingüística saussuriana distingue las relaciones semánticas secuenciales por estructuras jerárquicas y asociativas, así como la estructura de la propia frase, bien similar a las clases y subclases de los sistemas ontológicos.

Complementariamente, la lingüística estructuralista propone la distinción entre el eje sintagmático (horizontalmente, las relaciones de sentido entre las unidades de la cadena hablada, que se dan en presencia) y el eje paradigmático (verticalmente, las relaciones virtuales entre las unidades conmutables, que se dan en ausencia). En el primer eje, se abren las relaciones que pertenecen al dominio del habla, y, en el segundo, al dominio de la lengua. Estas relaciones sintagmáticas y paradigmáticas, por lo tanto, no se limitan al nivel lexical o gramatical del signo, sino que abarcan también al nivel fonológico (Ceia, 2010).

La descripción sintagmática es expuesta en dos ejes:

el paradigmático y el asociativo, el primero es equivalente a la concepción de una unidad sintáctica para una clase (o categoría), y son jerárquicas. El segundo, se refiere a la relación entre las unidades sintácticas, y son asociaciones lineales de concurrencias y subordinadas jerárquicas.

Relacionada a esa perspectiva, existe la estandarización del esquema por *Mind Mapping*, que es una técnica de representación gráfica de ideas, conceptos o informaciones. El *Mind Mapping* forma una estructura de árbol a partir de una idea central que se irradia para otros datos, y muchas veces es presentado con colores, gráficos y símbolos. La estructura en forma de árbol sirve para identificar las relaciones de un concepto central, buscando averiguar los pares. La idea central de esa representación por *Mind Mapping* es visualizar los resultados como un mapa mental, vislumbrando el pensamiento asociativo destacando las vinculaciones entre un concepto o una idea y las informaciones asociadas a ellos.

Además, tenemos dentro de este tipo de representación la estructura de un mapa heurístico, que es un diagrama que muestra la organización de relaciones semánticas entre diferentes ideas o relaciones de subordinación entre los diferentes conceptos.

No obstante, al contrario del esquema conceptual, este tipo de representación ofrece una representación en árbol de los datos imitando el progreso y el desarrollo del pensamiento.

Lo que puede ser complementario en estas representaciones es la cuestión cognitiva, que es discursiva sobre un determinado objeto (Cosette, 1994), el mapa con la visión

cognitiva normalmente es utilizado para generar opciones para profundizar el pensamiento.

Frente al contexto, las representaciones semánticas que vamos a presentar en este estudio tienden a traer todas estas características para una análisis semántico de los contenidos lingüísticos.

3. Metodología

Este estudio tuvo dos focos de análisis de redes: 1) trabajando con la mentalidad métrica de co-ocurrencias de autoridades, medios de publicaciones y visibilidad de citas, y; 2) con Análisis semántico de los contenidos de las obras, siendo categorizado por el título de las publicaciones, explorando análisis lexicográfico, análisis factorial, clasificación de términos, análisis de similitud y word cloud. Como herramienta para el desarrollo realizamos búsquedas en un sistema que extrae los términos y registros de Google Académico. Este software extrae indicadores de las búsquedas y del contenido recuperado y genera índices de citas (Harzing, 2010), más allá de reproducir los campos principales de las obras, como autoría, link de acceso, citas y año.

Para la cuestión semántica, utilizamos otro software que perfeccionó todo el contenido (Ratinaud; Marchand, 2012), seguido del análisis basado en redes sociales utilizando software lingüístico, donde vislumbramos: 1) relaciones de las clases, documentos, términos y el dendrograma de las clases; 2) términos semejantes por clusters; 3) agrupamiento de textos, contenidos y clases de palabras por

grupos semánticos, y; 4) la relación de tags de palabras.

De los contenidos recuperados en Google Académicos a partir del software *Publish or Perish*, fueron identificados un total de n=529 registros, siendo n=50 en Español, n=16 en Portugués, n=393 en Inglés y n=70 en Francés.

La estrategia de búsqueda utilizada fue por el campo *Keywords* de *Publish or Perish* y la expresión fue exclusiva para cada idioma de estudio, siendo: Inglés = creative ties; francés = liens créatifs; español = lazos creativos, y; portugués = laços criativos.

La generación de los indicadores fue ejecutada a partir de un programa de hojas de cálculos (Excel) y la generación de relaciones semánticas fue ejecutada por el software *Iramuteq*.

4. Resultados

La representación de los resultados se dio en la formación de los indicadores por análisis de conjunto de datos recuperados y del análisis semántico de los datos trabajados.

Para la primera parte, obtuvimos el índice-h del conjunto de datos, las relaciones de autoridades, las tipologías informacionales y el índice medio de citación. Para la segunda parte, obtuvimos el diagrama de Zipf, agrupamientos de términos, especificidad por agrupamiento temático, grado de interrelación, agrupamiento de relevancia, distribución por clases y grupos textuales, distribución por clases de términos, clases de stop words, dendogramas de clases, simetría de términos y word cloud.

4.1 Análisis métrico

Sobre análisis métricos, hay varias aplicaciones que pueden ser utilizadas en este tipo de estudio, como visibilidad, impacto, representación (Ferrazza; Pinto, 2017) y hasta la estadística descriptiva, que no necesariamente son aplicaciones de estudios métricos de la información.

Sobre impacto, se encuentra el índice-h que representa la cantidad de citas que un autor puede concentrar con la misma carga de publicación. En este sentido, priorizamos identificar las obras relevantes, totalizando $h=22$.

Tabla 1: Índice-h de los registros recuperados sobre el tema
Fuente: elaboración propia, 2022.

Citas	Publicaciones
2377	B Uzzi, J Spiro. Collaboration and creativity: The small world problem. 2005.
343	F Tustin. Les états autistiques chez l'enfant. 2021.
224	A Pande. "It may be her eggs but it's my blood": Surrogates and everyday forms of kinship in India. 2009.
143	JH Kirkland. The relational determinants of legislative outcomes: Strong and weak ties between legislators. 2011.
133	AF Wysocki. Impossibly distinct: On form/content and word/image in two pieces of computer-based interactive multimedia. 2001.
125	M Daskalaki. Building 'bonds' and 'bridges': Linking tie evolution and network identity in the creative industries. 2010.
114	I Rechberg, J Syed. Ethical issues in knowledge management: conflict of knowledge ownership. 2013.
110	C Hah, AJ Wennerstrom. Three-dimensional flowfields inside a transonic compressor with swept blades. 1991.
106	EF Koku, B Wellman. Scholarly networks las learning communities. 2004.
76	J Wells, RM Barry, la Spence. Using video tutorials as a carrot-and-stick approach to learning. 2012.
65	MS Ibrahim. Integrated marketing communication and promotion. 2011.

65	M Mayfield, J Mayfield. Leader talk and the creative spark: a research note on how leader motivating language use influences follower creative environment perceptions. 2017.
62	N Blanc. Les nouvelles esthétiques urbaines. 2012.
50	A Sannino, V Ellis. Learning and collective creativity. 2015.
47	C Sorensen, M Kakahara. Knowledge discourses and interaction technology. 2002.
41	M Huysman. Communities of practice: Facilitating social learning while frustrating organizational learning. 2004.
34	J Collado-Ruano. Educación y desarrollo sostenible: la creatividad de la naturaleza para innovar en la formación humana. 2017.
31	R Teigland, M McLure-Wasko. Creative ties and ties that bind: examining the impact of weak ties on individual performance. 2000.
31	S Lee, C Lee. Creative interaction and multiplexity in intraorganizational networks. 2015.
30	AS Mikhaylov. Socio-spatial dynamics, networks and modelling of regional milieu. 2018.
27	HJ Watson, RD Sousa, I Junglas. Business school deans assess the current state of the IS academic field. 2000.
23	C Hah, AJ Wennerstrom. Three-dimensional flowfields inside a transonic compressor with swept blades. 1990.

De este universo hay $n=19$ textos en inglés, $n=2$ en francés y $n=1$ en español, limitando el tema para el idioma anglosajón.

En relación al universo de colaboraciones científicas sobre el tema, obtuvimos un total de 245 nodos en un formato edges, de modalidad 1. Los autores con mayor tamaño de relacionamiento y grado de centralidad en la red fueron: CB Nko'o, seguido por A Douline y Crété, IA Sleiman, O Moles, T Joffroy, BD Stevens, el Moles, S Hawkins, S Niblock, la Podoksik, P Garnier, CR Smit, DF Lin, K Hammou, K Curnutt, K Hammou.

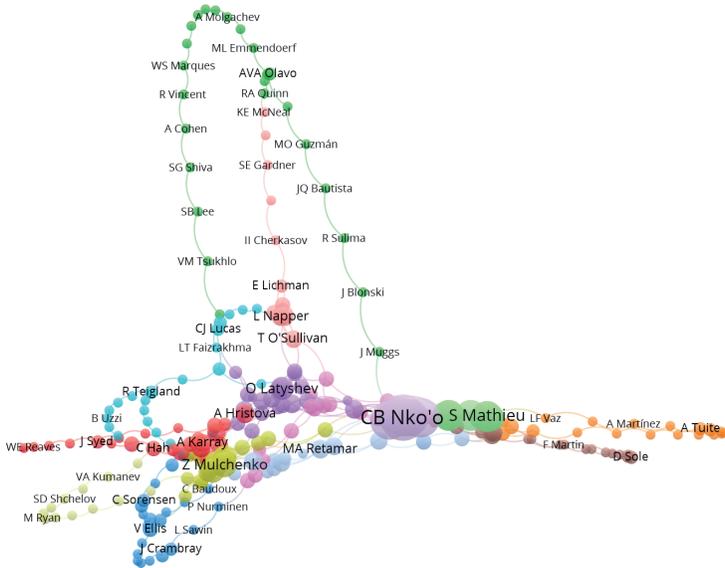


Figura 1: Red de colaboración de publicaciones
Fuente: elaboración propia, 2022.

En términos de cantidad de producción, que es otro tipo de poder en la red, los principales son: B Zaleski $n=10$, GD Ramos $n=5$, FL Lin $n=4$, otros 10 autores con $n=3$ (A Malaev–Babel, A Douline, M Laskina, O Moles, S Missonnier, S Turcotte, T Joffroy, GF Chiou, M Huysman) y otros 38 autores con $n=2$ (V Bakirov, M Ryan, JW Murry, la Khan, R Teigland, la Mela, ВУ Коваль, la Sannino, L Mansilla, la Toldo, el Latyshev, AC Aman, S Karlyuk, AJ Wennerstrom, VS Gribov, AVA Olavo, CP Туманова, C Hah, K Clark, CB Nko'ò, LA Reis, CJ Lucas, N Demidov, D Sole, la Enrici, y Aston, RSM Honorato, y Crété, SG Feldman, y Kovalev, V Ellis, EB Cortés, BB Квіцинська, I Yan, ОЮ Пономаренко, J Zychowicz, JP Matot, ТГ Мариупольская).

Sobre las cuestiones de relaciones entre autores, hay algunos aspectos destacados en (i) *Betweenness* (Intermediación) de la red, con énfasis de poder para J Syed $n=1345$, I Rechberg $n=1020$, LJ Reaves $n=768$, WE Reaves $n=514$ y M Suárez-Tamayo $n=258$; (ii) *Closeness* (Proximidad) de la red, con los siguientes investigadores: NAM Nordan $n=4339$, IA Sleiman $n=4019$, Z Mulchenko $n=3947$, y Crété $n=3740$ y P Garnier $n=3645$; Para (iii) Eigenvector (Vector propio) está S Gibson $n=4132$, JN Tranter $n=4086$, LL Denisova $n=4065$, MS Poole $n=4025$ y L Deng $n=4020$. Con relación al grado de la red, como fue formada una red pequeña no obtuvimos un grupo específico de centralizadores (260 destacados), por lo tanto no consideramos este cuantitativo para este Análisis.

Sobre las tipologías informacionales encontradas en las publicaciones analizadas, se obtuvo un total de $n=297$ artículos de revistas científicas, seguida por $n=129$ libros/capítulo de libros, $n=53$ tesis/disertación, $n=21$ trabajos presentados en eventos científicos, $n=15$ páginas personales de investigadores, $n=7$ informes gubernamentales y técnicos, $n=4$ artículos publicados en periódicos (magazines) y $n=3$ proyectos de investigación.

4.2. Análisis Semántico

Pensando en una aplicación semántica de análisis, empleamos algunas variaciones de críticas, pensando en: a) análisis estadístico-lexicográfico; b) especificidades y análisis factorial de correspondencia; c) método de la clasificación o método de Reinert; d) análisis de similitud, y; e) nubes de palabras. El desarrollo de representación de los

datos se dio por escalas, gráficos, diagramas, dendrograma y Word Cloud.

Nuestro primer análisis se volvió hacia cuestiones lexicales, en especial en la división del corpus estructurado de nivel macro y micro textual, con el objetivo de generar clases, subclases, atributos y relaciones. La lexicografía utilizada se basa en la generación de enciclopedias (diccionarios), sistemas de clasificación, glosario y vocabulario con términos relevantes, hapax, artículos y preposiciones.

En la estructuración de los datos en este tipo de análisis fue posible identificar los términos que fueran presentados con la frecuencia de solamente una aparición en el conjunto de los datos. Este tipo de datos son llamados de hapax, y se obtuvo un total de $n=1280$ términos aislados con frecuencia $n=1$.

Otros datos identificados fueron la cantidad de adjetivos ($n=51$ términos), sustantivos ($n=927$ términos), verbos ($n=123$ términos) y términos en los que no se consiguió reconocer la clase de las palabras ($n=642$). Complementariamente, fue posible identificar las stop words, que son las palabras vacías, sin sentido en el texto, como artículos y preposiciones.

Entre las stop words las principales en frecuencia son: the ($n=509$), of ($n=476$), and ($n=276$), in ($n=221$), la ($n=86$), for ($n=51$), on ($n=50$), to ($n=46$), las ($n=32$), s ($n=30$), an ($n=27$), new ($n=19$), from ($n=15$), between ($n=15$), with ($n=14$), at ($n=14$), by ($n=12$), is ($n=11$), through ($n=8$), two ($n=7$), some ($n=7$), my ($n=7$) y first ($n=7$).

Como final del análisis lexical, se encuentra la representación de la ley de Zipf, que también puede ser una

ecuación para métricas de las terminologías más recurrentes, en el cual destacamos los 7 términos más referenciados, siendo 6 sustantivos y un adjetivo, como puede ser vislumbrado en la figura 2.

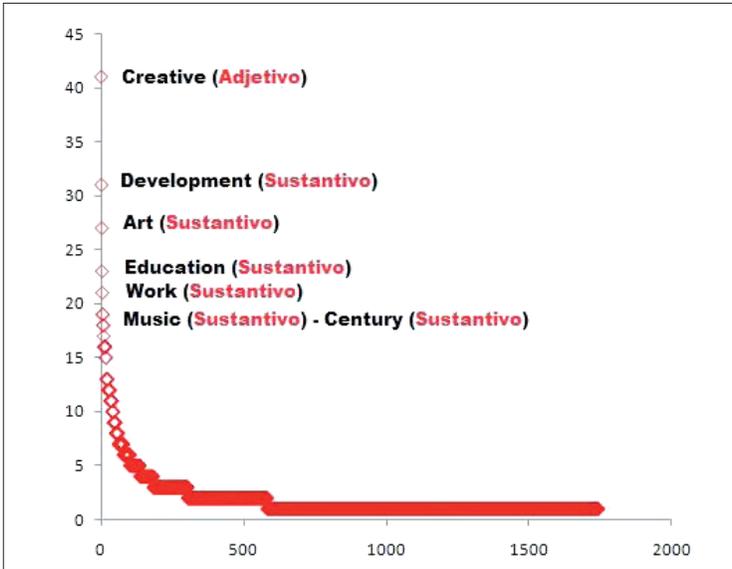


Figura 2: Escala de Zif: similitud y agrupamiento temático
Fuente: elaboración propia, 2022.

Siguiendo los patrones de análisis, hay también uno por clasificación, que es jerárquico en la forma descendente y tiene por objetivo generar un ordenamiento en el grupo de textos presentados en el Análisis (Reinert, 1990). Las finalidades de aplicación son: representación de las clases por dendrograma, representante de cuadrantes por términos centrales y stop words.

Al eliminar una parte del conjunto de datos, llamados términos suplementarios, se encuentra el agrupamiento de

cuatro clases, que actúan con jerarquías y agregan subclases y atributos.

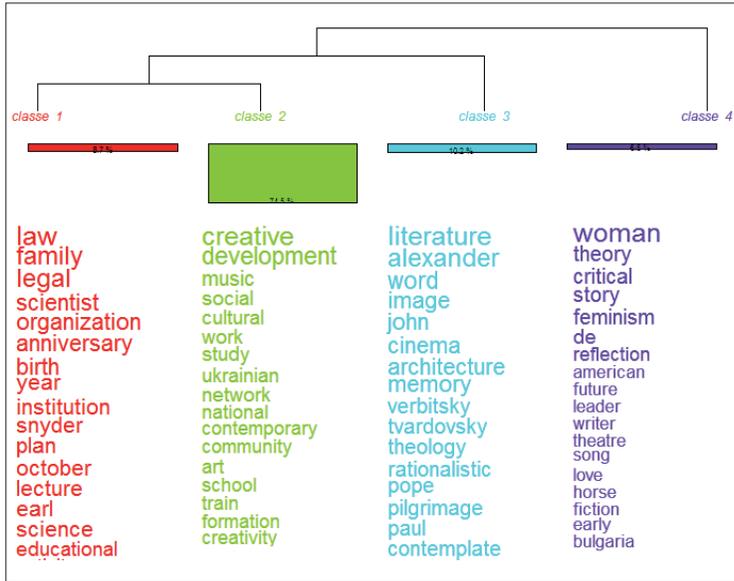


Figura 3: Clases en forma de Dendrograma
Fuente: elaboración propia, 2022.

En esta primera visualización, se encuentra el dendrograma de las categorías, que realiza la clasificación de los textos con la clase 1 (en rojo), clase 2 (en verde), clase 3 (celeste) y clase 4 (en violeta), en el cual la clase verde concentra el 74,5% de los términos y agrupa, semánticamente, el norte de los textos sobre lazos creativos del conjunto de datos.

Lo curioso es que este conjunto (Clase 2) es prácticamente el mismo que es representado en la ley de Zipf para identificar las terminologías más relevantes del Análisis.

Con respecto al agrupamiento por centralidad de la

Otro análisis que realizamos semánticamente fue la similitud de los términos, que pueden ser vislumbrados por la matriz de Zipf, con la selección de pocos términos, y la versión extendida de los principales términos de análisis, con recorte de las hapax (los términos que aparecen solamente una vez en todo el análisis).

El análisis de similitud presentada por Marchand y Ratinaud (2012) es la técnica basada en la teoría de los grafos, que busca verificar el análisis de correspondencia y clasificación lexical de los textos, identificando las correlaciones por medio de la Lexometría. Romero-Pérez, Alarcón-Vásquez y García-Jiménez (2018) enfatizan que la lexicometría es una herramienta aplicada para identificar expresiones conceptuales y categorizar temáticas, por medio del análisis de patrones ocultos en datos textuales.

Para que quede más llamativo el análisis, utilizamos la escala de Zipf para poder trabajar con pocos datos, con el objetivo de agrupar y categorizar los textos por temáticas, verificando la frecuencia de uso de cada término, distinguiéndolos de los suplementarios, con esto fue posible identificar las temáticas comunes y divergentes para cada una de las redes que el sistema indica.

La simetría de red de la escala de Zipf mostró que las temáticas clave de cada red fueron: development, creative, century, music, education, international, social y potential. Las palabras que siguen, todas tienen relación directa con los aspectos semánticos, siendo que los hilos de cada urdimbre están interconectados entre sí, indicando el punto de relevancia de cada uno, en la muestra analizada.

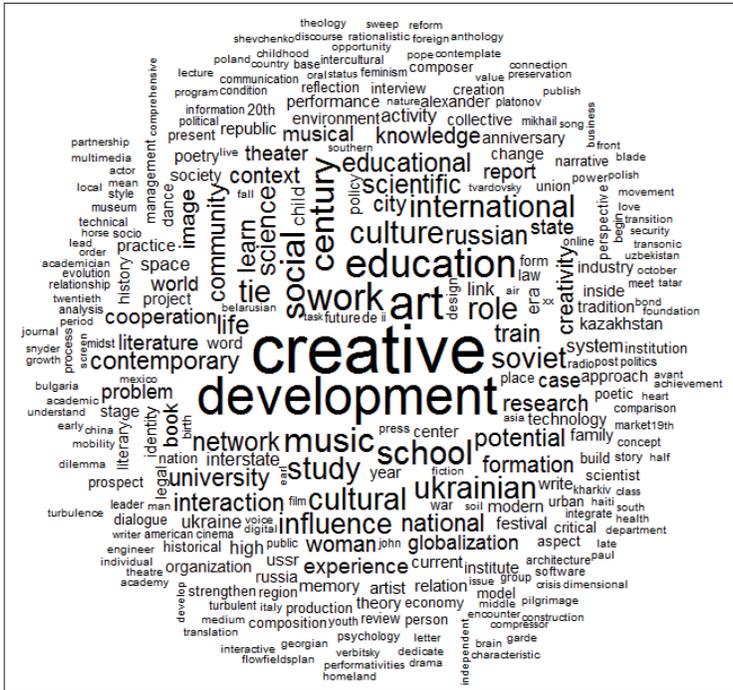


Figura 7: Word Cloud de los términos obtenidos

Fuente: elaboración propia, 2022.

5. Consideraciones Finales

Por tratarse de una misma muestra y emplear la escala de Zipf en algunas representaciones, para los términos obtenidos en el análisis de la frecuencia de palabras, los resultados son los mismos que fueron presentados en la simetría, de clases y de la similitud de la red. Lo que difiere en estos casos es la forma de presentación.

Lo que los modelos de representación dicen es que la temática “lazos creativos” está relacionada a seis términos claves, que son Adjetivos y Substantivos.

Para concluir el Análisis de los temas de aplicación de “lazos creativos”, se presenta la relación de los temas que los investigadores demostraron tener más preocupación: Creativo, Desarrollo, Arte, Educación, Trabajo, Música y Siglo.

Por otro lado, la cuestión bibliométrica muestra una vertiente dirigida al Arte, Ciencia de la Computación, Ciencia de la Información, Administración, Psicología y Sociología, siendo las áreas que envuelven los estudios del término investigado.

La investigación tiene algunas limitaciones, al transformar contenidos del idioma español, portugués y francés al idioma inglés, ya que puede ser que algunos términos hayan sido perdidos o no adaptados de forma correcta. Es posible que hayan ocurrido pequeños errores, en este sentido. Sin embargo, no creemos que eso pueda afectar de forma significativa al contexto general del estudio.

Otro detalle de limitación fue que algunos textos en el idioma inglés tenían términos referenciales de otras localidades, como Rusia, Hungría, Bulgaria y Suecia. Esos textos traen consigo términos no reconocidos por Iramuteq.

Así, no fue posible realizar algunos estudios de semántica de documentos por identificación de relación entre los textos, que podrían mostrar que existe relación ontológica de los términos de los textos estudiados. Infelizmente, el sistema solamente consigue identificar este segmento en un mismo idioma, sin relacionar un idioma con otro.

De todos modos, los signos aquí identificados y correlacionados a partir de la muestra multidisciplinar demuestran que los investigadores de variadas áreas están

abordando el tema “lazos creativos”<<<<, el que evidencia la existencia de preocupaciones emergentes sobre el tema en diversas áreas del conocimiento.

6. Referencias

- Ceia, C. (2010). Eixo sintagmático/eixo paradigmático. *E-dicionário de termos literários*. <http://www.edtl.com.pt/business-directory/6857/eixosintagmaticoeixo-paradigmatico/>
- Cossette, P. (1994), “La carte cognitive idiosyncrasique. Étude exploratoire des schèmes personnels de propriétaires-dirigeants de PME”, dans P. Cossette, *Cartes cognitives et organisations*, capítulo 6, p. 122. Québec, Presses de l’Université Laval; Paris, Eska.
- Daskalaki, M. (2010) Building ‘Bonds’ and ‘Bridges’: Linking Tie Evolution and Network Identity in the Creative Industries. *Organization Studies*, 31(12), pp. 1649–1666. <https://doi.org/10.1177/0170840610380805>
- DeLaney, B. (2015) *Youth center for improving health and wellness*. California State University, Long Beach [Disertación de maestría].
- Ferrazza, A. y Pinto, A. (2017). *Métodos quantitativos aplicados à Biblioteconomia e à Documentação*. Chapecó: Argos.
- Ghiselin, B. (1952) *The creative process*. California: Mass Market Edition.
- Harzing, A. (2010) *The Publish or Perish Book: Your guide to effective and responsible citation analysis*. Melbourne: Tarma Software Research Pty.
- Lee, S. y Lee, C. (2015) Creative Interaction and Multiplexity in Intraorganizational Networks. *Management Communication Quarterly*, 29(1), pp. 56–83. <https://doi.org/10.1177/0893318914553120>
- Leichik, V.M. y Shchelov, S.D. (2009) The activities of the terminological center of the RAS (a Brief Review). *Scientific and Technical Information Processing*, 36(2), pp. 96–98. <https://doi.org/10.3103/S01476888209020063>

- Lesgart, C. (2013) Pasado y presente de la Ciencia Política producida en Argentina. Apuntes para un debate de su porvenir. *Temas Y Debates*, 14, pp. 119–157. <https://doi.org/10.35305/tyd.voi14.108>
- Ratinaud, P. y Marchand, P. (2012) Application de la méthode ALCESTE à de “gros” corpus et stabilité des “mondes lexicaux”: analyse du “CableGate” avec IraMuTeQ. In: *Actes des 11^{eme} Journées internationales d'Analyse statistique des Données Textuelles* (835–844). Liège.
- Rehman, S. y Ibrahim, M. (2011) Integrated Marketing Communication and Promotion. *Researchers World, Journal of Arts, Science & Commerce, International Refereed Research Journal*, 2 (4), pp. 187–191. <https://ssrn.com/abstract=2383065>
- Saura, D. (2014) La fuerza de los lazos creativos: las redes sociales de un ensayo de danza. *REDES–Revista hispana para el análisis de redes sociales*, 25(2), pp. 86–105.
- Teigland, R. y McLure–Wasko, M. (2000) “Creative Ties and Ties That Bind: Examining the Impact of Weak Ties on Individual Performance”. In: *ICIS 2000 Proceedings*. 29. <https://aisel.laisnet.org/icis2000/29>