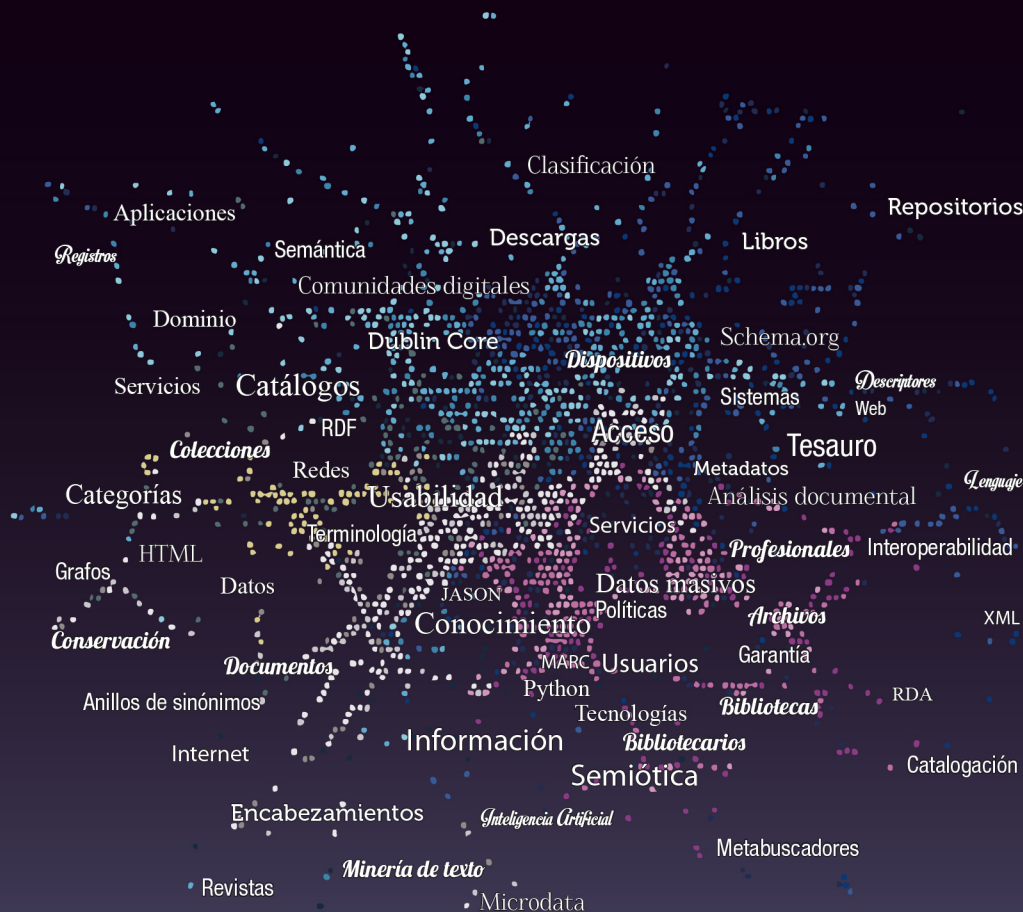


Organización del conocimiento

Bibliotecología, Terminología e Información



Catalina Naumis Peña
Adriana Suárez Sánchez
Coordinadoras



Z1006
O743

Organización del conocimiento : bibliotecología, terminología e información /
coordinadoras Catalina Naumis Peña, Adriana Suárez Sánchez. –Primera edición.
– Ciudad de México : Universidad Nacional Autónoma de México, Instituto de
Investigaciones Bibliotecológicas y de la Información, 2025.
xi, 240 páginas. – (Sistematización de la información documental)

Este libro fue producido en el marco del Proyecto PAPIIT IT 400122 Sistema
de información terminológica sobre bibliotecología, archivología e información.

ISBN: 978-607-587-131-8

1. Organización del conocimiento. 2. Terminología bibliotecológica.
I. Naumis Peña, Catalina, editor. II. Suárez Sánchez, Adriana, editor. III. serie.

Este libro fue producido en el marco del Proyecto PAPIIT IT 400122 *Sistema de
información terminológica sobre bibliotecología, archivología e información.*

Diseño de portada: Carlos Papaqui Landeros

Primera edición: febrero de 2025

D. R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
Instituto de Investigaciones Bibliotecológicas y de la Información
Circuito Interior s/n, Torre II de Humanidades,
pisos 11, 12 y 13, Ciudad Universitaria, C. P. 04510,
Alcaldía Coyoacán, Ciudad de México

ISBN: 978-607-587-131-8

Esta edición y sus características son propiedad de la Universidad Nacional Autónoma
de México. Prohibida la reproducción total o parcial por cualquier medio sin la
autorización escrita del titular de los derechos patrimoniales.

Publicación dictaminada
Hecho en México

Contenido

Presentación	vii
<i>Birger Hjørland</i>	

Prefacio	ix
<i>Catalina Naumis Peña</i>	

I.

PERSPECTIVAS DE EXPERTOS

La transformación conceptual y terminológica sobre catalogación en el siglo XXI	3
<i>Catalina Naumis Peña</i>	
Distribución de relaciones léxicas como indicador de propiedades estructurales en vocabularios controlados: el caso de la sinonimia	21
<i>Diego Ferreyra</i>	
La enseñanza de la construcción de tesauros: experiencias y reflexiones	39
<i>Francisco Javier García Marco</i>	

II.

APLICACIONES PRÁCTICAS

La representación temática facetada: análisis y propuesta de conversión y representación de encabezamientos de materia a FAST	63
<i>Noé Ángeles Escobar</i>	
Los sistemas de anotación como instrumento de organización del conocimiento	87
<i>Jesús Tramullas</i>	

Hacia una sistematización de la terminología de las ediciones: una contribución desde la organización del conocimiento a la industria editorial.	103
----------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Mario Barité

Carolina Saravia Rebollo

Consolidando un ecosistema de documentación del patrimonio histórico y cultural de la UNAM: propuesta metodológica <i>ad hoc</i> para el desarrollo de un microtesauro.	121
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----

Claudio Molina Salinas

III.

PERSPECTIVAS INNOVADORAS

Gestión de datos digitales: desafíos de la organización de información.	141
---------------------------------------------------------------------------------	-----

Lourdes Feria Basurto

Elizabeth Nashielli Márquez Sánchez

Repositorios institucionales, redes y organización del conocimiento: un estado de la cuestión en Hispanoamérica.	163
--------------------------------------------------------------------------------------------------------------------------	-----

Gema Bueno-de-la-Fuente

María del Carmen Agustín-Lacruz

IV.

ORGANIZACIÓN DEL CONOCIMIENTO DIGITAL

Mapas tópicos y ontologías terminológicas: características y tendencias en la organización temática de colecciones digitales.	195
---------------------------------------------------------------------------------------------------------------------------------------	-----

Adriana Suárez Sánchez

Las relaciones terminológicas entre los sistemas de información y los sistemas informáticos.	211
------------------------------------------------------------------------------------------------------	-----

Hugo Alberto Guadarrama Sánchez

La organización de la información en la era digital.	227
--------------------------------------------------------------	-----

Blanca Estela Sánchez Luna

Distribución de relaciones léxicas como indicador de propiedades estructurales en vocabularios controlados: el caso de la sinonimia

Distribution of lexical relations as an indicator of structural properties in controlled vocabularies: the case of synonymy

Diego Ferreyra
Facultad de Filosofía y Letras
Universidad de Buenos Aires

Los vocabularios controlados constituyen un tipo específico de sistema de organización del conocimiento (soc). Delimitan y formalizan un dominio estableciendo los términos que lo conforman, el alcance conceptual de cada uno de ellos y las relaciones que los articulan. Se trata de dispositivos de mediación social que basan su productividad y agencia en su capacidad para designar de manera no ambigua un conjunto de referentes en el contexto de un dominio y articularlos a través de una red coherente de relaciones léxico-semánticas. Entre las relaciones léxicas contempladas se encuentra la relación de sinonimia, entendida como la presencia en el lexicón de diversas formas lingüísticas con un significado similar o equivalente en el contexto de un dominio específico. Este artículo presenta un modelo de análisis y evaluación que permite definir una proporción óptima entre términos alternativos y preferentes en el contexto global de un vocabulario controlado.

Estudiaremos modelos que abordan el léxico como una red autoorganizada y se presentará una serie de modelos aplicables al contexto de vocabularios controlados. Finalmente, explicaremos la implementación de estos criterios en la herramienta de gestión de vocabularios controlados TemaTres.

Nocetti y Figueiredo (1978) relevan los puntos de contacto entre los vocabularios controlados y los lenguajes naturales. En su trabajo identifican, entre otros aspectos, el caso de relaciones léxicas entre términos como uno de los fenómenos de estudio compartidos. La sinonimia –y la polisemia– constituyen un aspecto estructural en los lenguajes naturales y un componente funcional

de los vocabularios controlados. En los lenguajes naturales estos procesos forman parte de las dinámicas que proveen elasticidad, economía y polivalencia como dispositivo sociocomunicacional. En el caso de los vocabularios controlados, la sinonimia forma parte de los fenómenos que pretenden corregir y remediar, la diversidad de términos referidos a un mismo concepto es abordada como una falla que debe ser subsanada a través del control (ANSI/NISO Z39.19:2005 [R2010]; ISO 25964-1:2011). Observando este particular punto de articulación, creo que es posible realizar un acercamiento, definir aspectos distintivos, formular algunas preguntas e intentar explorar herramientas para el análisis y la elaboración de sistemas de organización del conocimiento basados en vocabularios controlados.

LA SINONIMIA COMO RELACIÓN LÉXICA

La sinonimia es una relación entre unidades léxicas caracterizada por la verificación de la convergencia de significados entre palabras. Los procesos de sinonimia establecen una relación de equivalencia entre los significados léxicos de dos o más formas léxicas. Por tratarse de un proceso omnipresente en todos los actos de habla, todos los hablantes disponen de una aproximación nocional de la sinonimia.

Existen diversos enfoques para determinar la sinonimia como relación léxica. Desde la perspectiva de la semántica composicional, la sinonimia puede ser explicada como la coincidencia entre los rasgos de sentido composicionales de dos unidades léxicas. Desde una perspectiva funcional, la sinonimia se verifica y estudia en los usos y la posibilidad de sustitución entre ítems léxicos en contextos oracionales. Desde una perspectiva contextualista, las relaciones de sinonimia se analizan con base en la recurrencia de ítems léxicos en corpus textuales. Más allá de los marcos teóricos, es posible establecer un acuerdo en la consideración de la sinonimia como una relación caracterizada por la convergencia de diversas formas verbales sobre un significado. Este punto resulta central en la presente indagación: entender la sinonimia como la red de formas o unidades léxicas que se organizan alrededor de un mismo significado o concepto. Es pues una relación que agrupa una diversidad de entidades que en el contexto del lenguaje conforman la dimensión más concreta y material posible —como lo son los significantes o las formas verbales— en vinculación con un significado o concepto. Es quizás la dimensión más material de los términos al estar basada en relaciones entre diversos significantes y un significado y no estrictamente en las relaciones entre significados (Casas Gómez 1999; García Jurado 2003), como es el caso de otras relaciones léxicas (hiperonimia-

hiponimia, meronimia, términos asociados). Si la polisemia opera a través de la acumulación de significados en un mismo significante, la sinonimia, como proceso, acumula significantes en torno a un significado.

La onomasiología es el estudio del significado con respecto a los distintos significantes que pueden designarlo (García Jurado 2003). En términos de actividad terminológica, podemos considerarla como un derrotero que parte de un significado establecido en un dominio específico y que busca, selecciona y asigna un significante puntual. Si la sinonimia explora la posibilidad de convergencia entre diferentes términos y un único significado, es pues la relación onomasiológica por excelencia.

Nos interesa concebir la relación sinonímica como un proceso, es decir, como una relación dinámica en el tiempo y en el contexto de los conceptos y significados que organizan un dominio o una situación comunicativa.

En términos dialógicos, es decir, a través del tiempo, tal como Casas Gómez (1999), se analizan diferentes estados sinonímicos en una determinada lengua siguiendo diversos procesos (semantización/desemantización, irradiación o derivación semántica, etcétera) e incluso es posible identificar momentos en los que una palabra es utilizada para expresar significados opuestos (antónimos) o para reunir diferentes hipónimos (es decir, términos más específicos) bajo un mismo término. Es así que en determinados momentos una serie de palabras comparten un mismo significado y, eventualmente, a lo largo del tiempo pueden ir diferenciándose hasta incluso conformar series con significados contrarios (Abel 1988).

La cantidad de sinónimos posibles para un mismo concepto o significado varía según sea la lengua, el dominio o la situación comunicativa (Casas Gómez 1999; Strori, Bombaci y Bingol 2007; Makaruk y Owczarek 2008; ISO 25964-1:2011).

Desde una perspectiva psicolingüística, la sinonimia se explica como una relación gradual y continua entre significados y definida en términos de semejanzas en el contexto de una red de ítems léxicos. Esto implica que se verifica experimentalmente (Horno-Chéliz, Timor y Sarasa 2017) que el lexicón mental se encuentra organizado como una red en la que todas las piezas léxicas difieren —ya sea a través de matices, contextos de aplicación, registro de uso u otros aspectos— y donde la relación de sinonimia es de naturaleza gradual en términos de gradientes de semejanza y no de identidad. Eventualmente estas aproximaciones establecen que un término en el contexto del lenguaje natural no opera como una etiqueta léxica que se adosa a un concepto, sino que se desprende de la red de relaciones que organiza un léxico o vocabulario (Zapico y Vivas 2014). Esto presenta un conjunto de problemas y promesas para el campo de los vocabularios controlados y los sistemas de organización

del conocimiento, problemas que no son soslayados pero que serán abordados en futuros trabajos.

Se observa, pues, que la sinonimia se presenta como un proceso fluido, dinámico y situado ya sea que sea abordado como un fenómeno a través del tiempo, como un aspecto de la lengua en tanto sistema o desde un enfoque psicolingüístico. Corresponde analizar la función de esta relación léxica y su operacionalización en el contexto de los vocabularios controlados como componentes de un sistema de organización del conocimiento (soc).

Los soc son dispositivos de mediación social singularizados por su carácter formal, sistemático y orientado a objetivos específicos (Ferreyra 2021; Glushko 2014). En ese contexto, los vocabularios controlados constituyen un componente articular que organiza un dominio o espacio semántico a través de una red de relaciones léxico-semánticas coherente, es decir, relaciones no contradictorias entre sí según un régimen conceptual definido. En ese contexto, las relaciones de equivalencia léxica –como la sinonimia– constituyen un recurso central que permite diversificar el espectro de ítems léxicos que refieren un mismo significado, aportando mayor asertividad léxica y, por lo tanto, disminuyendo el nivel de ambigüedad y mejorando, por lo tanto, las capacidades para operar en un dominio dado.

La ambigüedad es un problema en cualquier contexto comunicativo. El lenguaje natural recurre a la redundancia –en registros diversos y concurrentes– para resolver ambigüedades y proveer eficacia comunicativa. Los vocabularios controlados, en cambio, se caracterizan por recurrir a un esquema de coherencia estructural en la red de relaciones léxicas. Esta coherencia estructural permite definir y diferenciar cada valor léxico de manera interdependiente. Los lenguajes artificiales en general tienden a un ideal modélico en el que cada término delimita con otros, conformando conjuntos diferenciados e interdependientes (Geckeler 1976). Sin embargo, es importante remarcar que se trata de una orientación, un modelo ideal que como tal se encuentra en permanente tensión y torsión con los procesos comunicativos basados en el lenguaje.

Aun así, esta es una de las fortalezas distintivas de los vocabularios controlados y los léxicos en general: la portabilidad operativa que le permite intervenir y asignar significados con relativa autonomía con respecto a diversos contextos. Un vocabulario controlado no varía en cada situación comunicativa, en cada registro o según los interlocutores y –en los casos óptimos– resuelve procesos de asignación de significados con eficacia. Es aquí donde las relaciones de equivalencia tienen un rol central en la consolidación de la eficacia de un vocabulario controlado o léxico. La relación entre términos preferentes y términos alternativos define el territorio total del campo semántico entendido como inventario léxico del dominio. A través de las relaciones de equivalencia

léxica se establece y delimita la totalidad de lo que puede ser denominado en el contexto del dominio. Las otras relaciones léxicas (relaciones jerárquicas de inclusión, relaciones asociativas) estructuran el léxico, organizan el territorio pero no amplían el contorno lexical. Aunque parece ser una relación léxica desjerarquizada y considerada de menor complejidad, la relación léxica de sinonimia define el inventario total de material léxico. Las relaciones de sinonimia resultan en ocasiones subvaluadas como instrumentos menores, considerados estamentos primarios en la jerarquía de complejidad de los vocabularios controlados (ANSI/NISO Z39.19:2005 [R2010]; Zeng 2008) y, sin embargo, tienen un rol preeminente y son el común denominador de todas las tipologías de léxicos y vocabularios controlados.

Consideremos que el mundo en general o un dominio en particular como destino referencial constituyen entidades infinitas, mientras que el repertorio terminológico consolidado en un vocabulario controlado es un fenómeno finito. En tal sentido, la capacidad de un vocabulario controlado para operativizar un dominio o proyecto comunicacional es justamente establecer un espacio finito de significados a través de un léxico puntual y discreto.

El fenómeno del proceso de sinonimia expresa como ningún otro la materialidad del lenguaje. Diversas formas pueden expresar un mismo significado y eventualmente denotar un mismo referente, pudiendo variar según el dominio, el lenguaje o el contexto comunicacional.

Se nos presenta entonces una pregunta importante para el campo de los vocabularios controlados y los soc: ¿existe una distribución de sinónimos característica en el contexto de un léxico? Si todos los términos pueden ser expresados por uno o más sinónimos, ¿sería deseable o correcto que todo término disponga de un sinónimo?, ¿cada término debería disponer de más de un sinónimo?, ¿cuántos? Estas preguntas giran en torno de una pregunta subyacente: ¿existe un patrón o modelo ideal de distribución de sinónimos en el léxico? Y si existiera, ¿qué tipo de indicador es y cómo debería interpretarse?

Los sinónimos no se distribuyen de manera uniforme en el léxico, este fenómeno se verifica tanto en el lenguaje natural como en los lenguajes controlados (Casas Gómez 1999; Strori, Bombaci y Bingol 2007; Makaruk y Owczarek 2008). Quizás la noción misma de clasificación como actividad pueda ofrecer un modelo explicativo de este fenómeno. La clasificación es, en los términos más generales posibles, el establecimiento de una discontinuidad sobre el continuo de la realidad, implica obviar ciertas diferencias para priorizar determinados rasgos salientes que nos permitan establecer divisiones (Bowker y Star 2000; Bauman 1996). Los procesos de sinonimia implican establecer una indiferenciación entre significantes para un determinado contexto comunicacional o dominio, crear una continuidad material entre varios

ítems léxicos en relación con un concepto. El proceso sinonímico generaliza y obvia diferencias entre formas léxicas en torno a un referente o concepto. Articula pues dos conjuntos no congruentes: agrupa una diversidad de significados como una unidad bajo una variedad de ítems léxicos.

Es así que la sinonimia expresa un aspecto central del conocimiento y del lenguaje en un dominio o contexto comunicacional: identifica los espacios en los que un dominio prefiere ampliar el rango de amplitud léxica para un significado que no debería ser segmentado. Es un indicador de la necesidad de una comunidad de prácticas de diversificar y extender el material verbal disponible para abarcar un campo semántico puntual. La presencia de sinónimos resulta ser un indicador útil para detectar la densificación del material léxico en una determinada zona del conocimiento de una comunidad o proyecto. Es un analizador que permite reconocer las zonas en las que una comunidad de prácticas o dominio opone una continuidad semántica frente a la evidencia de discontinuidades léxicas en las prácticas sociales en el seno de un dominio dado.

Hemos podido observar la importancia de las relaciones de sinonimia en la conformación del universo material y léxico de un vocabulario controlado. Reconocer los componentes de la dimensión material del léxico en nuestro campo resulta central, ya que constituyen el cuerpo y la geografía sobre las cuales habremos de diseñar y elaborar las operaciones de identificación, representación, búsqueda y organización de los artefactos culturales que conforman nuestros entornos de trabajo. Hemos observado también que los procesos de sinonimia exponen un campo de tensiones en el que es posible reconocer tanto concentraciones como vacíos. Ha llegado el momento de explorar la estructura de estas afinidades.

REDES DE RELACIONES LÉXICAS

Es posible analizar el léxico como una red de relaciones en la que cada palabra es un nodo y cada relación léxica un arco que la vincula a través de una relación específica con otra palabra (Cancho y Solé 2001; Sigman y Cecchi 2001; Zapico y Vivas 2014). Si en el lenguaje natural estas relaciones pueden analizarse a partir de la interacción recurrente entre palabras, en el caso de los léxicos y vocabularios controlados el análisis resulta más sencillo. En nuestro caso los términos –no las palabras– son los nodos y sus relaciones los vínculos que los conectan. Es así que los vocabularios controlados pueden analizarse como un espacio de nodos y conexiones: cada término/nodo puede establecer más de tipo de relación léxica y cada relación dispone de una direccionalidad (Holanda, Torres Pisa, Kinouchi, Souto Martínez y Seron Ruiz 2004).

Para avanzar en este análisis resulta necesario recurrir a ciertos conceptos propios de la teoría de grafos. Esta teoría parte de la consideración de que un grafo es una representación que se compone de una serie de nodos –también llamados vértices– y vínculos –también llamados arcos– que relacionan pares de nodos. Si las relaciones entre nodos expresan una direccionalidad, será un grafo dirigido, caso contrario será un grafo no dirigido. Con estos elementos es posible caracterizar una red según aspectos como ser el promedio de pasos entre cualesquiera dos nodos (distancia promedio), el diámetro máximo de la red (la distancia máxima entre dos nodos) o coeficiente de agrupamiento –puede ser entendido como el promedio de relaciones de los nodos de una red (Steyvers y Tenenbaum, 2005).

Los procesos de sinonimia en lenguajes naturales se analizan como un grafo no dirigido en el que un conjunto de palabras forman parte de una red vinculada por un significado similar. Es así que reciben a veces la denominación de synset o de anillos de sinónimos (ISO 25964-1:2011; ANSI/NISO Z39.19:2005 [R2010]). En una red de este tipo, los nodos establecen relaciones directas y simétricas entre sí (Strori, Bombaci y Bingol 2007). Un grafo no dirigido de sinónimos podría graficarse de la siguiente manera (Figura 1):

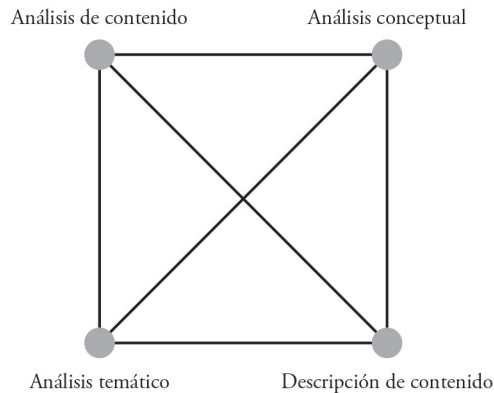


Figura 1.

Por el contrario, en vocabularios controlados orientados a procesos de indicación, se trata de un grafo dirigido, ya que se establece un término preferente y un conjunto de posibles términos equivalentes alternativos. En una red de este tipo un nodo –el término preferente– establece relaciones directas y asi-

métricas hacia otros nodos –los términos alternativos. Un grafo dirigido de sinónimos podría graficarse de la siguiente manera (Figura 2):

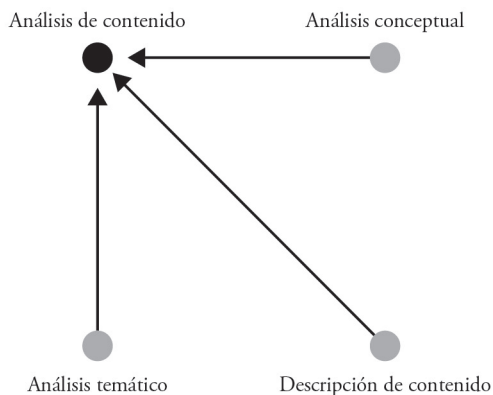


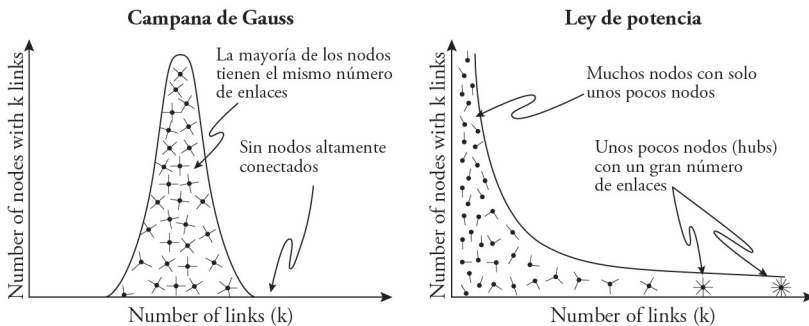
Figura 2.

En este sentido, un vocabulario controlado puede ser visto como la documentación formal de la red léxica que caracteriza un dominio.

Más allá de variaciones en los métodos, corpus o lenguas, se verifica en las investigaciones relevadas que las redes de relaciones léxicas en el lenguaje presentan propiedades compatibles con el modelo de redes libres de escala (Budél *et al.* 2023; Cancho y Solé 2001; Csermely 2006; Fóris 2013; Makaruk y Owczarek 2008; Sigman y Cecchi 2001; Steyvers y Tenenbaum 2005; Storri, Bombaci y Bingol 2007; Zapico y Vivas 2014). Las redes libres de escala formalizan un modelo presentado inicialmente por Barabási y Albert (1999) en el que no existe un patrón homogéneo en la distribución de enlaces por nodo, sino que un número acotado de nodos concentra proporcionalmente la mayor cantidad de vínculos, en tanto la mayoría de los nodos establecen pocos vínculos. Es un modelo representado gráficamente como una larga cola donde no es posible establecer un número característico o predominante de nodos según cantidad de relaciones. Este patrón, en términos de representación gráfica, resulta compatible con la distribución resultante de una relación de potencia entre 2 magnitudes (ley de potencia) o con la distribución de Pareto (Newman 2005). Este tipo de redes se verifica en un sinnúmero de fenómenos y actividades, como ser las citaciones científicas, la frecuencia de palabras, frecuencia y duración de llamadas telefónicas, diámetros de cráteres o el tamaño de las ciudades (Barabási 2003; Newman 2005; Solé 2009). Y en

el análisis de esta recurrencia se identifican una serie de aspectos característicos como una singularidad de las redes libres de escala (Figura 3):

- Conforman redes caracterizadas por la robustez como capacidad de tolerancia frente errores o variaciones en la cantidad de nodos o las conexiones (Fóris 2003).
- Se presenta siempre el fenómeno de *mundos pequeños*, esto significa que es posible acceder en pocos pasos entre nodos aparentemente distantes entre sí y entre nodos poco conectados; pese a que la mayoría de los nodos dispone de pocos enlaces, la existencia de hubs o concentradores permite disminuir drásticamente la distancia media entre dos nodos cualesquiera en la red (Watts y Strogatz 1998).
- Se autoensamblan y evolucionan nodo a nodo según el principio de conexión preferencial (*preferential attachment*); este principio considera que cada nodo que se suma a la red tiende a seleccionar nodos más densamente conectados; resulta un principio relevante por su capacidad explicativa de fenómenos evolutivos en grandes estructuras no coordinadas hasta la emergencia de patrones organizados robustos y sistémicamente estables (Klemm y Eguiluz 2001; Barabási 2003; Solé 2009; Fóris 2013).
- Permite describir globalmente una estructura compleja con una distribución proporcional consistente con la denominada ley de Pareto –también conocida como principio 80-20– según la cual el 80% de los recursos o eventos de algún tipo se concentra en el 20% de actores o factores (Barabási 2003; Newman 2005; Makaruk y Owczarek 2008).



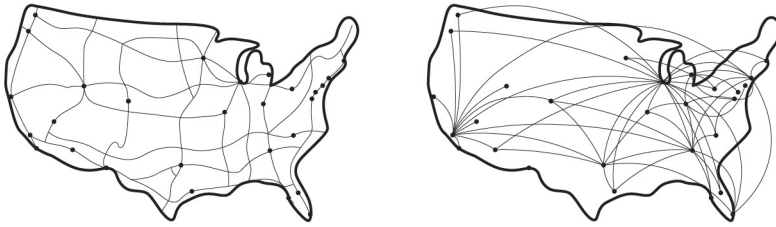


Figura 3. La distribución de enlaces en redes aleatorias se refleja como una curva de campana donde la mayor parte de los nodos tienen similar cantidad de enlaces. Las redes libres de escala –con una distribución acorde con la ley de potencia– concentran en pocos nodos la mayor cantidad de enlaces. Fuente: Barabási 2003.

Reconocer la conformación de redes compatibles con el modelo de redes sin escala en el lenguaje es relevante, especialmente por los trabajos que verifican este fenómeno en particular en el caso de las relaciones de sinonimia (Strori, Bombaci y Bingol 2007; Makaruk y Owczarek 2008).

Determinar distribuciones de relaciones léxicas que se verifican en cualquier lengua natural permite suponer que se trata de rasgos sistémicos o adaptativos que, a través de un sinnúmero de reiteraciones y variaciones, ofrecen condiciones de tolerancia, robustez y eficiencia. Si se pudieran identificar los aspectos que caracterizan la red de relaciones léxicas resultante, sería eventualmente posible aplicar los mismos patrones estructurales en la elaboración de vocabularios controlados y léxicos en general. Se espera que, a diferencia de otros aspectos del lenguaje en general y de los vocabularios controlados en particular, estas características resulten independientes de la escala o magnitud, el dominio, la lengua o el contexto comunicacional. Esto permite especular con la posibilidad de establecer un indicador regular e invariante en todos los escenarios y casos (Figuras 4 y 5).

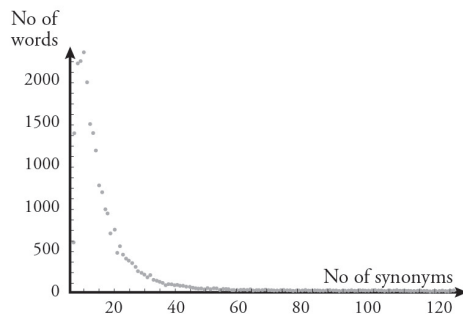


Figura 4. Número de sinónimos frente al número de palabras en polaco. Fuente: Makaruk y Owczarek, 2008.

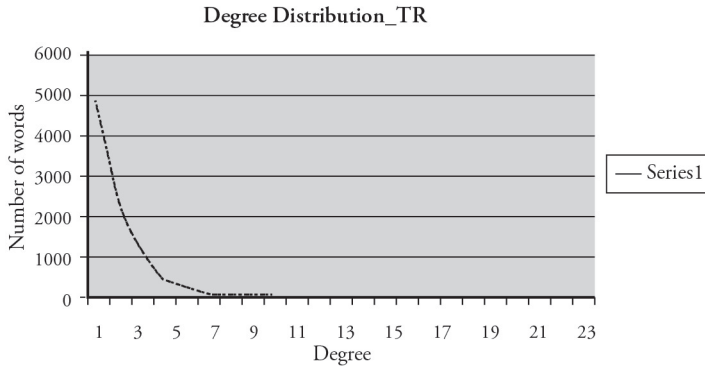


Figura 5. Número de sinónimos frente al número de palabras en turco. Fuente: Strori, Bombaci y Bingol, 2007.

INDICADORES DE DISTRIBUCIÓN DE RELACIONES LÉXICAS INCORPORADOS A TEMATRES

TemaTres es una aplicación web de gestión de vocabularios controlados.¹ Existen cientos de implementaciones, una familia de herramientas web asociadas a la explotación y reúso de servicios terminológicos² y dispone de una agenda de desarrollo activa y participativa. Tal como se señaló anteriormente, el interés de la presente exploración es proponer la consideración de indicadores asociados a la gestión de vocabularios controlados y léxicos que sean capaces de dar cuenta de propiedades estructurales y globales con base en las relaciones léxicas formalizadas en los mismos. Como se ha podido comprobar a lo largo de nuestra indagación, en el caso de las relaciones léxicas de sinonimia es posible verificar la emergencia sistémica de patrones de distribución compatible con el modelo de redes libres de escala. Es así que en la última versión disponible de TemaTres se han incorporado tres componentes informativos orientados a caracterizar la distribución global de relaciones léxicas de sinonimia y ofrecer en tal sentido indicadores que permitan evaluar el grado de adecuación de cada vocabulario, con una distribución compatible de las redes libres de escala, a saber:

1 <https://github.com/tematres/TemaTres-Vocabulary-Server/releases>

2 <https://github.com/tematres>

1. Un gráfico que describe la relación de términos preferentes según la cantidad de términos alternativos.
2. Un gráfico que describe la relación porcentual sobre el total de términos de los términos preferentes y los términos alternativos.
3. El coeficiente de agrupamiento (*clustering*) correspondiente a la red de relaciones léxicas resultante entre términos preferentes y los términos alternativos.

El primero de los indicadores permite establecer si la red de relaciones léxicas resultante describe una función compatible con la ley de potencia y, por lo tanto, podría caracterizarse como una red libre de escala. Debería verificarse que un conjunto minoritario de términos concentra la mayor cantidad de términos alternativos mientras que la mayoría de los términos dispone de un número drásticamente menor de términos alternativos.

El segundo gráfico indicador permite establecer si la distribución de relaciones de equivalencia léxica es compatible con la ley de Pareto. Dicha ley establece que el universo de términos se compone de un 80% de términos preferentes y un 20% de términos alternativos. Este gráfico resulta en realidad un orientador intuitivo, un indicador riguroso debería establecer la concentración del 80% de las relaciones de equivalencia en el 20% de los términos preferentes. Siendo que esa proporción se encuentra informada detalladamente en el indicador gráfico anterior, el presente gráfico espera orientar un criterio aproximado relativo a la necesidad de incrementar o decrementar la cantidad global de términos equivalentes.

Finalmente, el tercer indicador ofrece un coeficiente de agrupamiento, es el promedio de relaciones léxicas de equivalencia establecidas en el vocabulario. Estima en qué medida los términos del vocabulario, entendidos como una red léxica, están agrupados o conectados entre sí. El coeficiente permite establecer un valor numérico que describa la tendencia a formar grupos densamente interconectados a través de relaciones de equivalencia léxica entre términos en el contexto del vocabulario. Es un indicador cuya mayor utilidad se encuentra orientada a establecer métricas comparativas entre vocabularios controlados o indicadores evolutivos en el contexto de un mismo vocabulario controlado a través del tiempo.

Se presentan a modo de ejemplo los indicadores resultantes para el vocabulario el Tesauro Spines (Figura 6), el Tesauro latinoamericano en ciencia bi-

bliotecológica y de la información (Figura 7) y el Tesauro brasileiro de ciência da informação (Figura 8).

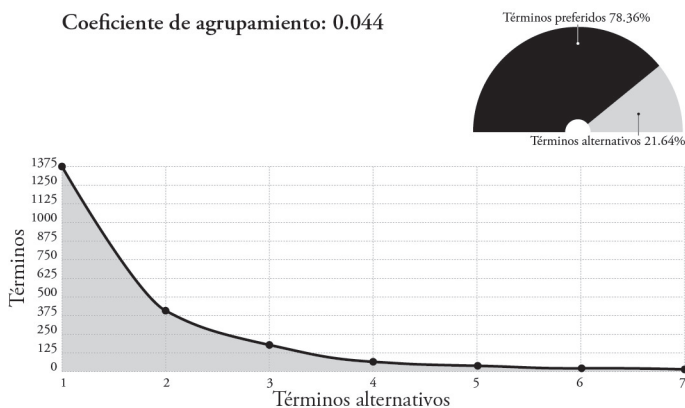


Figura 6. Términos alternativos según término. Tesauro SPINES, <https://vocabularyserver.com/spines/sobre.php>

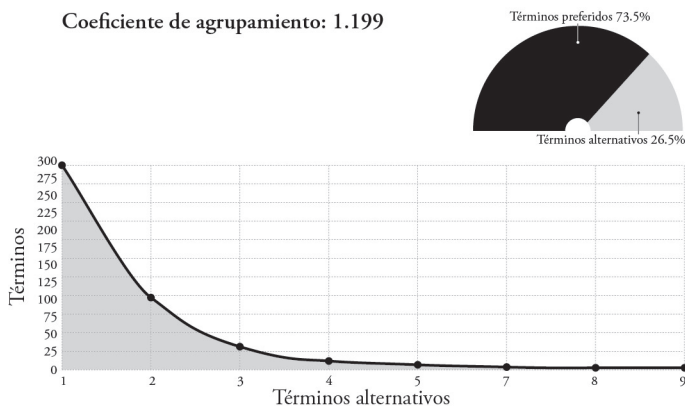


Figura 7. Términos alternativos según término. Tesauro latinoamericano en ciencia bibliotecológica y de la información, <https://vocabularyserver.com/telacibin/sobre.php>

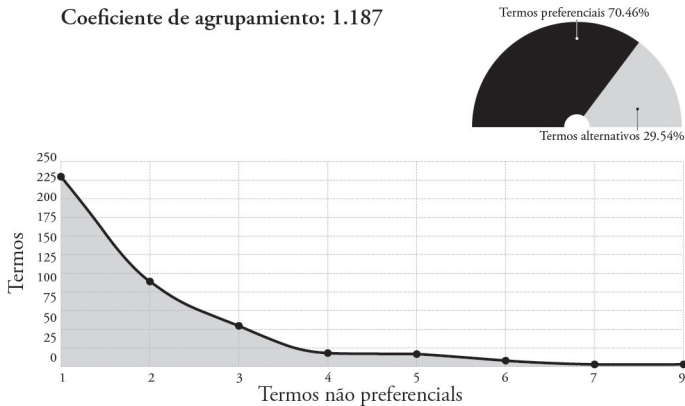


Figura 8. Términos alternativos según término. Tesauro brasileiro de ciência da informaçã, <https://vocabularyserver.com/tbci/sobre.php>

CONCLUSIONES

La construcción de una red coherente y robusta de relaciones léxicas promueve la persistencia estructural de un paradigma de sentido. En este trabajo se ha observado que las relaciones de sinonimia no se distribuyen de manera aleatoria ni homogénea. La conformación de una red de relaciones léxicas compatible con el modelo de redes libres de escala parece ofrecer mejores posibilidades para sobrellevar variaciones de todo tipo. En el caso de la sinonimia, fortalece las capacidades de un léxico para absorber las variaciones léxicas que se concentran en determinadas zonas de un dominio. El presente análisis sostiene que el análisis de las redes de relaciones léxicas de sinonimia permite identificar los *centros de atracción* en el contexto de un dominio. Establece un espectro de indiferencias entre variantes léxicas asociadas a un mismo concepto, es una forma de ampliar el alcance asertivo de un conjunto de ítems léxicos sobre un mismo vértice conceptual. Analizar la distribución de las relaciones de sinonimia permite evaluar la robustez de la red de ítems léxicos en su estamento más material posible y menos abstracto. En segundo término, facilita la diferenciación de espacios concentrados de significado en el contexto de un dominio al identificar el conjunto minoritario de términos que reúne la mayor cantidad de variaciones lexicales. En tercer lugar, permite establecer comparativas entre vocabularios controlados. Por último, considero que la disponibilidad de indicadores capaces de caracterizar vocabularios controlados como redes de relaciones léxicas abre importantes posibilidades

para nuevos interrogantes: ¿es posible establecer un análisis similar en otras relaciones léxicas adoptadas en los vocabularios controlados (relaciones de inclusión, relaciones asociativas)?, ¿existen efectos o factores articulares entre los modelos de red de cada tipo de relación léxica?, ¿la verificación de patrones de distribución de relaciones de sinonimia en diversos dominios, escalas y lenguas permitiría considerar la existencia de universalidades estructurales y variaciones contextuales según sean las zonas en las que se concentran los casos de sinonimia? El modelo permite comparar variaciones diferenciales según cada lengua en la composición de las redes de sinonimia. En el caso de los vocabularios multilingües ¿qué implicancias conlleva para el enfoque centrado en conceptos (web semántica, ISO 25964-1:2011) las diferencias entre las redes de sinonimia –que varían según cada lengua– en el contexto de un mismo vocabulario multilingüe? (Makaruk y Owczarek 2008).

REFERENCIAS BIBLIOGRÁFICAS

- Karl Abel, "Acerca del sentido antitético de las palabras primitivas", Karl Abel, Wilhelm Von Humboldt y Hans Sperber, *El psicoanálisis y las teorías del lenguaje*, Catálogos, 1998.
- Baldinger, Kurt, *Teoría Semántica: hacia una semántica moderna*, Madrid: Alcalá, 1977.
- Barabási, Albert-László y Réka Albert, "Emergence of Scaling in Random Networks", *Science* 286 (5439), 509-12, <https://doi.org/10.1126/science.286.5439.509>
- Barabási, Albert-László, *Linked: The New Science of Networks*, Cambridge: Perseus Publishing, 2003.
- Bauman, Zigmunt, "Modernidad y ambivalencia", Josetxo Beriaín, comp., *Las consecuencias perversas de la modernidad: modernidad, contingencia y riesgo*, Anthropos Editorial, 1996.
- Bowker, Geoffrey C., y Susan Leigh Star, *Sorting Things Out*, Cambridge: The MIT Press, 1999, <https://doi.org/10.7551/mitpress/6352.001.0001>
- Budel, Gabriel, Ying Jin, Piet Van Mieghem, y Maksim Kitsak, "Topological Properties and Organizing Principles of Semantic Networks", *Scientific Reports* 13 (1), <https://doi.org/10.1038/s41598-023-37294-8>
- Cancho, Ramon Ferrer I. y Ricardo Vicente Solé, 2001, "The Small World of Human Language", *Proceedings of the Royal Society B. Biological Sciences* 268 (1482), 2261-65, <https://doi.org/10.1098/rspb.2001.1800>
- Casas Gómez, Miguel, *Las relaciones léxicas*, Walter de Gruyter, 2015.
- Csermely, Peter, *Weak Links: The Universal Key to the Stability of Networks and Complex Systems*. Springer, 2009.
- Ferreira, Diego, "Sistemas de organización del conocimiento: propuesta de un modelo unificado de definición", *Palabra Clave* 11 (1), e141, <https://doi.org/10.24215/18539912e141>
- Geckeler, Horst, *Semántica estructural y teoría del campo léxico*, Barcelona: Gredos, 1976.
- Glushko, Robert J. *The Discipline of Organizing: Professional Edition*, O'Reilly Media, Inc., 2016.
- Holanda, Adriano de Jesus, Ivan Torres Pisa, Osame Kinouchi, Alexandre Souto Martinez y Evandro Eduardo

- Seron Ruiz, "Thesaurus as a Complex Network", *Physica, A* 344 (3-4), 530-36, <https://doi.org/10.1016/j.physa.2004.06.025>
- Horno-Chéliz, María del Carmen, Raquel Timor y Antonio Sarasa Cabezuelo, "¿Qué ocurre cuando comparamos dos unidades léxicas sinónimas? Un estudio psicolingüístico sobre la naturaleza de la sinonimia", *RLA, Revista de lingüística teórica y aplicada* 55 (1), 149-68, <https://doi.org/10.4067/s0718-48832017000100149>
- ISO, ISO 25964-1:2011 Information and documentation - *Thesauri and interoperability with other vocabularies* - Part 1: Thesauri for information retrieval, 2011.
- Jurado, Francisco García, *Introducción a la semántica latina: de la semántica tradicional al cognitivismo*, Madrid: Universidad Complutense de Madrid, 2003.
- Klemm, Konstantin y Víctor M. Eguíluz, "Highly Clustered Scale-free Networks", *Physical Review. E, Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 65 (3), <https://doi.org/10.1103/physreve.65.036123>
- Makaruk, Hanna E. y Robert Owczarek, "Hubs in Languages: Scale Free Networks of Synonyms", arXiv.Org, February 28, 2008, <http://arxiv.org/abs/0802.4112>
- National Information Standards Organization, ANSI/NISO Z39.19:2005 [R2010], *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*, 2005.
- Newman, M. E. J., "Power Laws, Pareto Distributions and Zipf's Law", *Contemporary Physics* 46 (5), 323-51, <https://doi.org/10.1080/00107510500052444>
- Nocetti, Milton A. y Regina Célia Figueredo, "Línguas naturais e linguagens documentárias: traços inerentes e ocorrências de interação", *R. Bibliotecon, Brasília* 6 (1); 23-27, jan./jun.1978.
- Sigman, Mariano y Guillermo A. Cecchi, "Global Organization of the Wordnet Lexicon", *Proceedings of the National Academy of Sciences of the United States of America* 99 (3), 1742-47, <https://doi.org/10.1073/pnas.022341799>
- Solé, Ricard, *Redes complejas: del genoma a Internet*, Tusquets Editores, 2016.

- Steyvers, Mark y Joshua B. Tenenbaum, "The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth", *Cognitive Science* 29 (1), 41-78, https://doi.org/10.1207/s15516709cog2901_3
- Strori, Dorina, Ahmet Bombaci y Haluk Bingol, "Cross Comparison of Synonym Graphs in a Multi Linguistic Context", arXiv.Org, September 16, 2007, <https://arxiv.org/abs/0709.2476>
- Watts, Duncan J. y Steven H. Strogatz, "Collective Dynamics of 'Small-world' Networks", *Nature* 393 (6684), 440-42, <https://doi.org/10.1038/30918>
- Zapico, Martín y Jorge Vivas, "La sinonimia como caso particular de distancia semántica", *Encontros Bibli* 19 (40), 253, <https://doi.org/10.5007/1518-2924.2014v19n40p253>
- Zeng, Marcia Lei, "Knowledge Organization Systems (kos)", *Knowledge Organization* 35 (2-3), 160-82, <https://doi.org/10.5771/0943-7444-2008-2-3-160>

Organización del conocimiento. Bibliotecología, Terminología e Información. Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. Edición digital. Coordinación editorial: Sergio Javier Sepúlveda H. y Angélica Valenzuela; revisión especializada: Diana Martínez Aguilar; corrección de pruebas: Carlos Ceballos Sosa y Marcos Emilio Bustos Flores; formación editorial y portada: Carlos Papaqui Landeros. Se terminó en febrero de 2025.