



Delimitación temática de la investigación en Lingüística y Comunicación mediante análisis de co-palabras

Carmen Gálvez¹

Recibido: 23 de enero de 2018/ Aceptado: 17 de diciembre de 2018

Resumen. Comprender las complejas interrelaciones semánticas de los conceptos científicos, principalmente en las áreas interdisciplinarias, sigue siendo un problema para los investigadores. En este trabajo se exponen los principales resultados de un estudio cuyo principal objetivo fue identificar y visualizar los focos de investigación de dos categorías *Web of Science (WoS)* correspondientes a las áreas de la *Lingüística* y la *Comunicación*, durante el periodo comprendido entre 2007 y 2017. Primero, las palabras-clave, extraídas de la producción científica indexada en *WoS*, se usaron para identificar los conceptos y temas de investigación de forma estadística. A continuación, se realizó un análisis de co-palabras basado en matrices de co-ocurrencia en el que, por un lado, se agruparon las palabras-clave para identificar conglomerados temáticos y, por el otro, se adoptó el método del análisis de redes para identificar la estructura jerárquica de ambos dominios científicos. Los resultados mostraron que los principales focos de investigación se sintetizaron en: *Análisis del discurso* y *Géneros discursivos*, *Etnografía de la comunicación*, *Identidad lingüística*, *Pragmática* y *Análisis de la conversación*, *Lingüística textual*, *Teoría de la argumentación*, *Retórica*, *Traducción*, *Intertextualidad* y *Sociolingüística*. La conclusión general fue que los mapas y redes obtenidas permitieron alcanzar un mayor nivel de comprensión del dominio de conocimiento interdisciplinario de la *Lingüística* y la *Comunicación*.

Palabras clave: Lingüística, comunicación, análisis de co-palabras, métodos de visualización de dominios de conocimiento.

[en] Thematic delimitation of the research in Linguistics and Communication through co-word analysis

Abstract. Understanding the complex semantic interrelationships of scientific concepts, mainly in interdisciplinary areas, remains a problem for researchers. This paper presents the main results of a study whose main objective was to identify and visualize research focuses of two categories *Web of Science (WoS)* corresponding to the areas of *Linguistics* and *Communication*, during the period between 2007 and 2017. First, the keywords, extracted from the scientific production indexed in *WoS*, were used to identify the concepts and research topics in a statistical way. Next, a co-word analysis was carried out based on co-occurrence matrices in which, on the one hand, keywords were grouped to identify thematic clusters and, on the other hand, the method of network analysis was adopted to identify the hierarchical structure of both scientific domains. The results showed that the main research focuses were synthesized in: *Discourse Analysis* and *Discursive Genres*, *Ethnography of Communication*, *Linguistic Identity*, *Pragmatics* and *Conversation Analysis*, *Textual Linguistics*, *Theory of Argumentation*, *Rhetoric*, *Translation*, *Intertextuality* and *Sociolinguistics*. The general

¹ Universidad de Granada (España). Correo electrónico: cgalvez@ugr.es

conclusion was that the maps and networks obtained allowed reaching a higher level of understanding of the interdisciplinary knowledge domain of *Linguistics* and *Communication*.

Keywords: Linguistics, communication, co-word analysis, knowledge domain visualization methods.

Índice. 1. Introducción. 2. El análisis de co-palabras. 3. Metodología. 4. Resultados y discusión. 4.1. Resultado del análisis de las relaciones de asociación. 4.2. Resultado del análisis de la estructura jerárquica. 5. Conclusiones. Referencias bibliográficas.

Cómo citar: Gálvez, C. (2019). Delimitación temática de la investigación en Lingüística y Comunicación a través de análisis de co-palabras. *Círculo de Lingüística Aplicada a la Comunicación* 77, pp. 187-200. <http://dx.doi.org/10.5209/CLAC.63283> <http://webs.ucm.es/info/circulo/no77/Galvez.pdf>

1. Introducción

A medida que la ciencia avanza los investigadores tienen más dificultades para comprender las complejas conexiones semánticas de los conceptos científicos con los que trabajan. Con respecto a las disciplinas fuertemente interrelacionadas, como es el caso de la *Lingüística* y la *Comunicación*, este problema se agrava. Los límites de ambas disciplinas son muchas veces difíciles de delimitar. En el dominio científico de la *Lingüística*, intervienen diversos campos vinculados con la *Comunicación lingüística*, como la *Pragmática*, la *Didáctica de las lenguas*, la *Lingüística aplicada a la enseñanza de Lenguas Extranjeras*, la *Psicolingüística* o la *Sociolingüística*. En el caso concreto de la *Lingüística aplicada*, debido a su carácter interdisciplinario, sigue siendo un problema común entre los lingüistas delimitar su ámbito de estudio, dentro de las ciencias del lenguaje (Luque Agulló, 2004-2005).

Para identificar las relaciones entre conceptos y temas en los diferentes dominios de conocimiento se han desarrollado métodos cuantitativos, tales como el análisis de co-palabras (Callon *et al.*, 1986), que consideran que los documentos científicos se pueden reducir a apariciones conjuntas de las palabras que lo componen. El análisis de co-palabras constituye un instrumento valioso para descubrir la estructura conceptual y temática de un campo de investigación (Van Raan, 2005). Por otra parte, este método entroncaría, de alguna forma, con la corriente lingüística del distribucionalismo de Zellig Harris (1951), en la que se consideraba que el estudio de una lengua determinada consiste en la reunión de un conjunto amplio de enunciados (o corpus lingüístico) utilizados por los hablantes y, posteriormente, descomponer los enunciados de ese corpus para lograr una descripción ordenada y sistemática de los constituyentes inmediatos que tengan la misma distribución.

Por otra parte, dentro de las disciplinas que estudian la gestión del conocimiento, como son las *Ciencia de la Documentación*, el análisis de co-palabras toma, como base del análisis, datos de naturaleza bibliográfica. El origen de este planteamiento surgió de la posibilidad de medir las relaciones científicas por medio del grado de co-ocurrencia de palabras (Small y Griffith, 1974). Así, las palabras que aparecen juntas en el campo «título», «resumen» o «descriptor» del mismo registro de las bases de datos se consideran una indicación de la conexión entre los conceptos y temas que representan.

El uso del análisis de co-palabras en combinación con técnicas de representación gráfica consigue no sólo una adecuada visualización de la información sino transferir conocimiento. En este sentido, la visualización de la información tiene por objeto transformar datos abstractos y relaciones complejas en mensajes visibles, dando lugar a una nueva ciencia de la comunicación visual denominada la *esquemática* (Costa, 1998). Se trata de un área que estudia el lenguaje gráfico de los esquemas como mensajes y la esquematización como procedimiento para la visualización de la información que no es visible en la realidad. Pero los esquemas poseen asimismo sus condiciones propias como lenguaje de comunicación: la capacidad de transmitir conocimientos (Costa, 1998). Dentro de estos esquemas se hallarían: *i*) los mapas de la ciencia, también denominados mapas del conocimiento o *cienciogramas* (Small, 2006), que reflejan las relaciones entre disciplinas y sus etiquetas las conexiones semánticas (Garfield, 1994); y *ii*) las redes que representan el conocimiento en forma de grafos, en los que hay nodos y aristas que interactúan, permitiendo un análisis de características estructurales y topológicas de las redes por medio de diferentes métricas.

El objetivo de este trabajo fue aplicar la metodología del análisis de co-palabras a la producción científica indexada en *Web of Science (WoS)* dentro de las categorías *Lingüística* y la *Comunicación*, durante el periodo comprendido entre 2007 y 2017, para identificar y visualizar, en forma de mapas del conocimiento y redes de palabras, los principales focos de interés en los que convergen ambas disciplinas.

2. El análisis de co-palabras

El análisis de co-palabras, también denominado análisis de palabras asociadas, se define como una técnica de análisis de contenido que usa patrones de co-ocurrencia de pares de elementos (es decir, palabras o sintagmas nominales) cuando aparecen simultáneamente en un mismo *corpus* de textos, en un mismo documento o en un conjunto de documentos representativos de la producción de un área (Callon *et al.*, 1986). Este método se incluye dentro de la clasificación de los indicadores bibliométricos relacionales (Leydesdorff y Welbers, 2011), que se basan en el conteo de matrices de datos, donde cada uno de los componentes representa la co-ocurrencia de las variables comunes. En relación con esto, la *Bibliometría* se ocupa de estudiar los documentos científicos por medio de la aplicación de indicadores cuantitativos a la producción científica (Van Leeuwen, 2004). Los instrumentos de la *Bibliometría* para la evaluación de dicha producción son los indicadores bibliométricos. Una clasificación ampliamente aceptada distingue (Callon *et al.*, 1995): *i*) indicadores de actividad, que proporcionan datos acerca del volumen y del impacto de las actividades de investigación; y *ii*) indicadores de relación (o *reticulares*), que proporcionan datos acerca de los vínculos entre campos científicos. Los indicadores de relación se basan precisamente en el conteo de matrices de datos, donde cada uno de los componentes representa la co-ocurrencia de las variables comunes.

A su vez, los resultados del análisis de co-palabras se representan en mapas y redes de palabras, donde se identifican los grupos (o *clusters*) de palabras estrechamente unidas unas a otras, así como las relaciones existentes entre ellas. Estos grupos serían asimilables a los focos y tendencias temáticas o conceptuales que componen un dominio científico. Así, a partir de un conjunto de documentos

representativos de la producción científica de un área, y después de seleccionar los términos que se van a procesar, en el método del análisis de co-palabras se construyen matrices de co-ocurrencias y redes de co-palabras a través de las cuales se desarrollan diferentes tipos de análisis, de tal forma que la medida del enlace entre dos palabras será proporcional a la co-ocurrencia de esas dos palabras en el conjunto de documentos que se ha tomado como muestra. Por tanto, examinar las características de las matrices de co-ocurrencias y de las redes generadas constituye el paso más importante del análisis de co-palabras. En este trabajo se van a considerar dos tipos principales de características:

- Relaciones de asociación a través de análisis de *clustering*, o agrupamiento, consistente en descomponer la red de co-palabras en grupos similares e interconectados. Las asociaciones entre las palabras obtenidas serían asimilables a las líneas temáticas o frentes de investigación de los campos científicos analizados.
- Estructuras jerárquicas a través de *Análisis de Redes Sociales (ARS)*, consistente en identificar subredes o conjuntos de nodos altamente conectados. Según los valores de conexión de las palabras, la estructura de la red de co-palabras se descompone en núcleos (o capas jerárquicas) desde las capas externas a los más centrales, que se representan gráficamente en cajas que se contienen entre ellas. Los núcleos de palabras resultantes se corresponderían con la estructura de los diferentes bloques temáticos.

El análisis de *clustering* se ha aplicado con éxito a diversas materias y disciplinas (Ding *et al.*, 2001; Kostoff *et al.*, 2006; Li *et al.*, 2008; Leydesdorff y Hellsten, 2006; Cho, 2014; Hu y Zhang, 2015; Ravikumar *et al.*, 2015). Por su parte, el análisis de la estructura jerárquica ha demostrado ser adecuado para aplicaciones que se dirigen a la detección de las comunidades de nodos más altamente interconectados entre sí (Fortunato, 2010; Yang y Dia, 2008) y a la identificación de nuevas tendencias de investigación en las disciplinas científicas (Du *et al.*, 2007).

3. Metodología

Para llevar a cabo el análisis de co-palabras se siguieron varias etapas: *i)* extracción de los registros de la plataforma *WoS* correspondientes a las áreas de investigación de la Lingüística y la Comunicación; *ii)* conteo de las frecuencias de apariciones de las palabras-clave; *iii)* selección de las palabras-clave con una frecuencia alta; *iv)* construcción de una matriz de co-ocurrencia con las palabras-clave seleccionadas (*palabra-clave x palabra-clave*); *v)* creación de un mapa de palabras-clave agrupadas según sus relaciones de asociación; *vi)* creación de una red de palabras-clave descompuesta según su estructura jerárquica; y *vii)* interpretación de los resultados obtenidos.

El primer paso del análisis de co-palabras consistió en la obtención de los documentos representativos que se van a examinar y mapear. Para la recuperación de la producción científica se utilizó la plataforma *WoS*, a través de la cual se accedió a las bases de datos *Science Citation Index Expanded (SCIE)*, *Social Sciences Citation Index (SSCI)* y *Arts & Humanities Citation Index (A&HCI)*. La estrategia de búsqueda empleada consistió en buscar en el campo «Categoría de *Web of Science*» los términos

‘WC=*Linguistics and Communication*’, en el campo «tipo de documento» el término ‘TP=*Article*’ y limitada el periodo temporal 2007-2017. En total se recuperaron 2353 documentos. Los documentos recuperados fueron descargados directamente en formato de texto plano para ser analizados posteriormente.

A continuación, se seleccionaron las palabras-clave con las que está indexada la producción científica en *WoS*. En esta plataforma, la producción científica está indexada por medio de dos tipos de palabras-clave: a) palabras-clave de autor, *Author Keywords* (KWA), proporcionadas por los propios autores; y b) *KeyWords Plus* (KW+), extraídas de forma automática a partir de la frecuencia de aparición de las palabras en los títulos de las referencias de los artículos citados. En este trabajo, se optó por seleccionar las palabras-clave de autor, debido a que las palabras-clave extraídas de forma automática son menos específicas y comprensibles que las palabras-clave aportadas por los autores (Zhang *et al.*, 2016). En total, se identificaron 6906 palabras-clave de autor en la muestra seleccionada. Dentro de la distribución obtenida, se tuvo en cuenta las frecuencias de aparición más altas para obtener una representación suficientemente significativa de las áreas de investigación analizadas, por esta razón se decidió fijar un umbral de 4 ocurrencias mínimas (esto es, cuya frecuencia fuera ≥ 4). Después de una etapa de pre-procesamiento para eliminar duplicados y normalizar los ítems, se obtuvo una lista de 448 palabras-clave.

Con las 448 palabras-clave seleccionadas, el paso siguiente consistió en construir una matriz de co-ocurrencia. Para ello, se calculó la frecuencia con la que dos palabras-clave aparecen simultáneamente en el mismo documento. Actualmente, hay diversas aplicaciones informáticas para la construcción de estas matrices (tales como, *BibExcel*, *Citespace* o *Matlab*). Utilizando la herramienta *BibExcel* se construyó una matriz de co-ocurrencia de 448 x 448 palabras-clave. A partir de la matriz de co-ocurrencia, se generaron redes de palabras-clave, que se sometieron a un análisis de agrupamiento y a un análisis de su estructura jerárquica.

El análisis de agrupamiento se realizó para la búsqueda de grupos homogéneos en función de las similitudes entre las palabras-clave. Con el propósito de normalizar los valores de co-ocurrencia e identificar parejas de palabras-clave que representaran de forma adecuada el corpus analizado y facilitaran su agrupamiento fue necesario aplicar un índice de similitud, en este estudio se aplicó el índice *Fuerza de Asociación* (FA), *Association Strength* (Van Eck y Walkman, 2009). Con la aplicación del índice de similitud se consiguió, entre otras cosas, dar mayor valor a las palabras con una frecuencia baja pero con una gran frecuencia de co-ocurrencia, y, lo contrario, dar menor valor a las palabras con una alta frecuencia pero con una baja frecuencia de co-ocurrencia. La red de palabras-clave obtenida se visualizó en dos mapas utilizando la aplicación *VOSviewer* (Van Eck y Waltman, 2010): *i*) mapa etiquetado con palabras-clave, donde cada palabra estuvo representada por una etiqueta y cuyo tamaño fue proporcional a su peso o índice de similitud; y *ii*) mapa de densidad (también denominado *mapa de calor*) en el que cada nodo estuvo representado con un color que fue del rojo al azul, reflejando la densidad de las relaciones entre las palabras-clave.

Por su parte, el ARS estuvo dirigido a proporcionar información sobre la estructura jerárquica de la red de palabras-clave. A partir de la matriz de co-ocurrencias y utilizando las aplicaciones computacionales basadas en análisis de

redes, *UciNet* y *NetDraw* (Borgatti, 2002), se construyó un grafo, en el que cada palabra-clave se describió por un nodo (o vértice), y las líneas entre los nodos representaron las relaciones entre dos palabras-clave. Hay una serie de métricas para estudiar la importancia de los nodos dentro de las estructuras de redes complejas. Entre estas métricas se encuentra la detección de comunidades y subredes dentro de una red más amplia, estas medidas se basan en el principio de la identificación de los nodos más densamente conectados, tales como *cliques*, *n-cliques* y *k-cores* (Wasserman y Faust, 1994). En este trabajo, se decidió utilizar la métrica *k-core* (*k-núcleo* o *núcleo de grado k*) (Seidman, 1983), consistente en descomponer la red de nodos en núcleos cada vez más cohesionados, a través de un proceso de sucesivos anidamientos. El *k-core* de una red de nodos se configura como la subred más grande en la que cada nodo tiene al menos *k* interconexiones (Dorogovtsev *et al.*, 2006). De esta forma, variando el valor del parámetro *k*, se obtuvieron diferentes imágenes de la red de palabras-clave (esto es, a medida que aumentamos el valor de *k*, la relación entre las palabras-clave se hizo más densa).

4. Resultado y discusión

Con las 448 palabras-clave obtenidas se calculó sus frecuencias de aparición, se construyó la matriz de co-ocurrencia, y se generó una red de palabras-claves en la que se identificaron las relaciones de asociación y las estructuras jerárquicas. En la Tabla 1 se muestra una selección de las palabras-clave con las frecuencias de aparición más altas en la producción científica analizada.

Palabras-clave	Frecuencias	Palabras-clave	Frecuencias
<i>Narrative</i>	99	<i>Translation</i>	32
<i>Conversation analysis</i>	80	<i>Persuasion</i>	30
<i>Identity</i>	73	<i>Rhetoric</i>	29
<i>Argumentation</i>	57	<i>Metaphor</i>	29
<i>Discourse analysis</i>	56	<i>Semiotics</i>	29
<i>Gender</i>	52	<i>Interaction</i>	29
<i>Discourse</i>	46	<i>Culture</i>	28
<i>Multimodality</i>	42	<i>Communication</i>	26
<i>Human-robot interaction</i>	37	<i>Language ideology</i>	26
<i>Language</i>	37	<i>Conversation</i>	23
<i>Pragmatics</i>	33	<i>Computer-mediated communication</i>	21
<i>Ideology</i>	33	<i>Critical discourse analysis</i>	21

Tabla 1: Selección de una muestra de las palabras-clave con las frecuencias de aparición más altas

3.1. Resultado del análisis de las relaciones de asociación

A partir de la aplicación de técnicas de *clustering* se analizaron las relaciones de asociación entre las palabras-clave. Utilizando la herramienta *VOSviewer* (Van Eck y Waltman, 2010), se realizaron dos tipos de visualizaciones: *i*) un mapa etiquetados (Figura 1), donde cada palabra-clave estuvo representada por una etiqueta y un círculo (cuyo tamaño fue proporcional a su peso o índice de similitud), además se distinguieron con un color aleatorio los diferentes grupos

temáticos al que se adscribió cada palabra-clave; y *ii*) un mapa de densidad (Figura 2), en el que cada zona estuvo representada con una tonalidad que fue del rojo al verde, reflejando la mayor o menor densidad de las relaciones entre las palabras-clave (las zonas de mayor grado de interconexión entre palabras-clave aparecieron en una tonalidad próxima al roja, y las zonas de menor grado de interconexión aparecieron en una tonalidad próxima al amarillo y verde).

En la visualización del mapa etiquetado, las 448 palabras-clave se dividieron en 5 grupos, o *cluster*, distinguidos con diferentes colores aleatorios. Si las palabras-clave se agrupan en un mismo conglomerado se consideró que era probable que reflejaran temas relacionados. En la Tabla 2 se muestra una selección de las cinco palabras-clave con mayor peso dentro del correspondiente grupo. En cada grupo (*G*) se asociaron diferentes números de palabras-clave vinculadas a un campo temático:

- **G1:** agrupó a 116 palabras-clave vinculadas con los campos del *Análisis del discurso* y los *Géneros discursivos*. En estos campos temáticos se posicionaron las palabras-clave referidas al estudio del discurso escrito y hablado como una forma del uso de la lengua, como hecho de comunicación y de interacción, en sus contextos cognitivos, sociales, políticos, históricos y culturales.
- **G2:** agrupó a 116 palabras-clave relacionadas con los campos de la *Etnografía de la comunicación* y la *Identidad lingüística*. En este foco temático se situaron las palabras-clave relacionadas con el uso del lenguaje por parte de los miembros de un determinado grupo, las situaciones en las que se produce ese uso, reglas que utilizan los interlocutores, o las variaciones que se observan entre diversos grupos.
- **G3:** agrupó a 111 palabras-clave concernientes con los campos de la *Pragmática* y el *Análisis de la conversación*. En este conglomerado se situaron las palabras-clave vinculadas a cómo los hablantes producen e interpretan enunciados en contexto, los factores extralingüísticos que determinan el uso del lenguaje, tales como los interlocutores, la intención comunicativa, el contexto o el conocimiento del mundo.
- **G4:** agrupó a 46 palabras-clave relacionadas con la *Lingüística textual*, la *Teoría de la argumentación* y la *Retórica*. En este grupo temático se situaron las palabras-clave vinculadas al estudio del texto, entendido como acontecimiento comunicativo con plenitud de sentido, a los procesos de producción e interpretación de los textos y al trabajo con determinados géneros de discurso: fundamentalmente discursos públicos, tales como el discurso político o periodístico, los discursos persuasivos o los discursos publicitarios.
- **G5:** agrupó a 59 palabras-clave relacionadas con los campos de la *Traducción*, la *Intertextualidad* y la *Sociolingüística*. En este foco temático se posicionaron las palabras-clave referidas a la socialización del lenguaje, la influencia que tienen en una lengua los factores derivados de las diversas situaciones de uso en las que produce la comunicación lingüística y la interrelación de los procesos lingüísticos y culturales.

Grupo temático	Selección de palabras-clave	Peso
G1	<i>Narrative</i>	181
	<i>Gender</i>	108
	<i>Discourse analysis</i>	88
	<i>Media</i>	49
	<i>Human-robot interaction</i>	35
G2	<i>Identity</i>	164
	<i>Culture</i>	67
	<i>Multimodality</i>	67
	<i>Ideology</i>	59
	<i>Language ideology</i>	48
G3	<i>Conversation analysis</i>	138
	<i>Discourse</i>	99
	<i>Pragmatics</i>	82
	<i>Interaction</i>	78
	<i>Communication</i>	57
G4	<i>Argumentation</i>	138
	<i>Rhetoric</i>	64
	<i>Persuasion</i>	56
	<i>Pragma-dialectics</i>	46
	<i>Interpretation</i>	32
G5	<i>Translation</i>	56
	<i>Intertextuality</i>	41
	<i>Language socialization</i>	38
	<i>Indexicality</i>	32
	<i>Sociolinguistics</i>	32

Tabla 2: Detalle de los grupos temáticos identificados en la red de palabras-clave

El resultado más relevante fue que los cinco focos temáticos estuvieron fuertemente interconectados. Los grupos más importantes se asociaron a los campos del *Análisis del discurso*, la *Etnografía de la comunicación* y el *Análisis de la conversación*. Además, los conglomerados **G1**, **G2** y **G3** aglutinaron el mayor número de palabras-clave. Por tanto, estos grupos temáticos fueron los que recibieron más interés por parte de los investigadores. Otro resultado significativo fue que en el centro de los mapas se situó el **G1**, referido al *Análisis del discurso*, cuyas palabras-clave establecieron más conexiones con el resto de los grupos temáticos. En consecuencia, el *Análisis del discurso* se configuró como el campo más relevante al estar vinculado con otros campos temáticos, como la *Etnografía de la comunicación*, el *Análisis de la conversación*, la *Pragmática*, o la *Lingüística textual*.

3.2. Resultado del análisis de la estructura jerárquica

Con la aplicación del método *k-core* se obtuvo la segunda red de palabras-clave, para ello se utilizaron las herramientas de ARS, *UciNet* y *NetDraw* (Borgatti, 2002). Mediante al análisis de *k-core* la red de palabras-clave se descompuso en diez núcleos (o capas) desde los nodos externos a los más centrales. En el caso del núcleo central $k=10$, las palabras-clave de esta subred estuvieron interconectadas con, al menos, 10 palabras-clave, lo que se interpretó como la subred en la que situaron los conceptos y temas fuertemente relacionados y relevantes. Por el contrario, en el caso de $k=1$, las palabras-clave de esta sub-red estuvieron conectados con, al menos, 1 palabras-clave, lo que se interpretó como la sub-red en la que estuvieron representadas los conceptos y temas menos relacionados e irrelevantes. Para facilitar la distinción de los núcleos obtenidos, cada capa se marcó con diferentes formas geométricas (Figura 3): nodos circulares ($k=10$) en la capa central; nodos en forma de triángulos ascendentes ($k=9$, $k=8$ y $k=7$), en la capa media; nodos cuadrados ($k=6$, $k=5$ y $k=4$) en la capa lateral; y nodos en forma de rombo ($k=3$, $k=2$ y $k=1$), en la capa de borde. Cada capa integró diferentes números de palabras-clave correspondientes a los diferentes conceptos y temas de investigación (Tabla 3).

- Capa central: incluyó 120 palabras-clave que indicaron los conceptos y temas principales e influyentes.
- Capa intermedia: incluyó 215 palabras-clave que representaron los temas especializados relacionados con los temas principales y los focos emergentes.
- Capa lateral: incluyó 95 palabras-clave que representaron los temas específicos, con pocas conexiones con el resto.
- Capa de borde: incluyó 18 palabras-clave que representaron los temas tangenciales.

Núcleo	Capa	Selección de palabras-clave
$k=10$	Central	<i>Narrative; Identity; Argumentation; Conversation analysis; Gender; Discourse; Discourse analysis; Pragmatics; Interaction; Rhetoric; Multimodality; Cultura; Ideology; Communication; Persuasion; Translation; Media; Language ideology; Semiotics; Metaphor; Intertextuality.</i>
$k=9$ $k=8$ $k=7$	Intermedia	<i>Pragma-dialectics; Human-robot interaction; Aristotle; Discrimination; Autism; Strategic maneuvering; Linguistic intergroup bias; Impression formation; Ethics; Critical discussion; Strategic manoeuvring; Internet; Twitter.</i>
$k=6$ $k=5$ $k=4$	Lateral	<i>Repair; Assessments; Perelman; Formulations; News interviews; Universal audience; Argumentative pattern; Pragmatic argumentation; Semantics; Fieldwork; Institutional interaction; Political speeches; Semiotic resources.</i>

$k=3$ $k=2$ $k=1$	Borde	<i>Endangered languages; New rhetoric; Argumentation scheme; Artificial intelligence; Crime fiction; National allegory; Narrative practice; Oral history; Social robots; Argument evaluation.</i>
-------------------------	-------	---

Tabla 3. Selección de palabras-clave incluidas en cada capa jerárquica

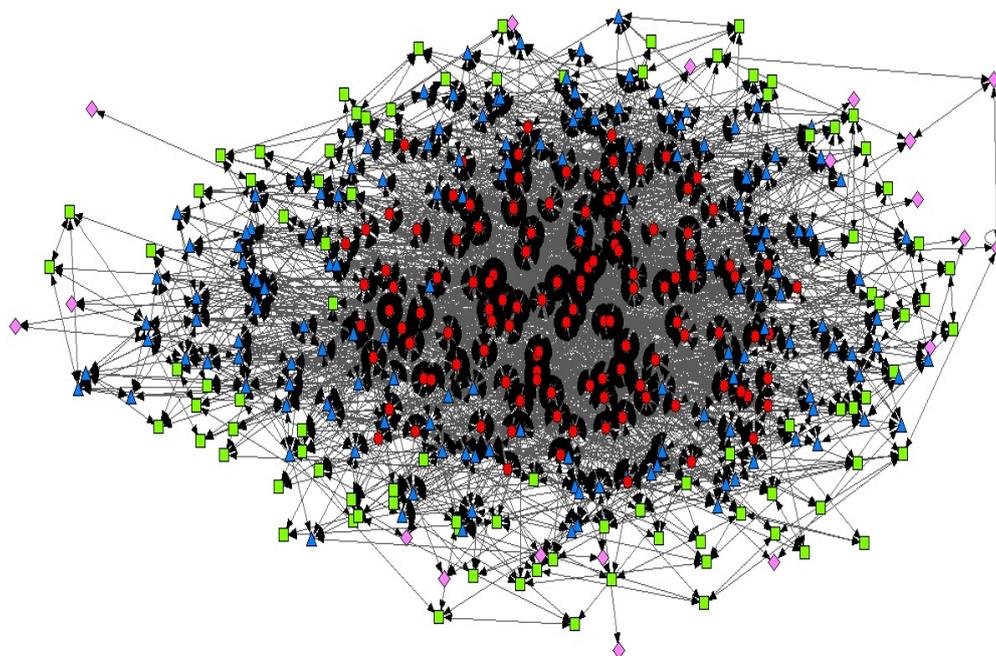


Figura 3: Red de palabras-clave dividida en capas jerárquicas

Las 120 palabras-clave de la capa central representaron la base de conocimiento esencial de la investigación interdisciplinar de las áreas de la Lingüística y la *Comunicación* (Figura 4), en este núcleo se situaron los siguientes subcampos de trabajo: *Etnografía de la comunicación, Identidad lingüística, Análisis del discurso, Teoría de la argumentación, Retórica, Pragmática, Análisis de la conversación, Traducción, Intertextualidad y Sociolingüística*. Las 215 palabras-clave de la capa intermedia representaron los conceptos y temas estrechamente vinculados con el núcleo central, que en algunos casos simbolizaron la evolución de los temas principales, tales como la *Teoría pragma-dialéctica de la argumentación*, la *Interacción humanos-robots*, la *Teoría de la acomodación lingüística*, o los estudios sobre el discurso en la red *Twitter*. Las 98 palabras-clave de la capa lateral representaron los temas específicos, tales como la *Comunicación de la ciencia*, la *Argumentación jurídica*, el *Auditorio universal* o la *Teoría de la argumentación jurídica de Perelman* (Perelman y Olbrechts-Tyteca, 1989). Y, por último, las 18 palabras-clave de la capa de borde representaron los temas periféricos sin apenas conexión con el resto de los temas analizados, tales como las *Lenguas en peligro de extinción*, *Género de la ficción criminal*, *Inteligencia artificial* o los estudios sobre *Robots sociales* capaces de comunicarse con las personas.

fundamentales, estrechamente interconectados, que formando la base de conocimiento constante de los campos de actuación de la *Lingüística y Comunicación*. Esta capa básica incluyó los siguientes subcampos principales: *Etnografía de la comunicación y Identidad lingüística*, *Análisis del discurso*, que estuvo fuertemente asociado a la *Teoría de la argumentación* y la *Retórica*, los subcampos de la *Pragmática* y el *Análisis de la conversación* y los subcampos de la *Traducción*, *Intertextualidad* y *Sociolingüística*. La capa intermedia contuvo 215 palabras-clave, que se interpretaron como los conceptos y temas especializados y muy relacionados con el núcleo central. En esta capa intermedia, los focos de investigación se dirigieron esencialmente a la *Teoría pragma-dialéctica de la argumentación*, la *Interacción humanos-robots* o la *Teoría de la acomodación lingüística*. En el resto de las capas se situaron las palabras-clave aisladas, con escasas conexiones entre ellas, que representaron los temas menos relevantes, tales como la *Inteligencia artificial* o los *Robots sociales*, que se interpretaron como los campos inmaduros y las nuevas tendencias de investigación, que tal vez necesitarían ser más estudiados.

En general, el doble planteamiento propuesto y desarrollado en este trabajo, al contrastar los resultados de la agrupación de palabras-clave con la perspectiva de su estructura jerarquía, permitió mapear y explorar la base de conocimiento de los campos en los que convergen la *Lingüística y Comunicación*. Las representaciones reducidas de la realidad obtenidas se podrían utilizar para alcanzar una mejor comprensión de las interrelaciones de ambas áreas, o simplemente completar una imagen instantánea que ya se tenía de ellas, por parte de los expertos en la materia. De cualquier forma, se ha logrado destacar los subcampos más importantes y se han revelado las relaciones entre subcampos que de otra forma serían difíciles de apreciar.

Referencias bibliográficas

- Borgatti, S. P. (2002). *NetDraw [computer software]*. Harvard, MA: Analytic Technologies
- Borner K., Chen, C. M. y Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37,179-255.
- Callon M., Courtial, J-P. y Penan, H. (1995). *Cienciometría. La medición de la actividad científica: de la bibliometría a la vigilancia tecnológica*. Gijón: Trea.
- Callon, M., Courtial, J-P., Turner, W. A. y Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22, 191-235.
- Callon, M., Rip, A. y Law, J. (1986). *Mapping the Dynamics of Science and Technology*. London: The Macmillan Press Ltd.
- Cho, J. (2014). Intellectual structure of the institutional repository field: A co-word analysis. *Journal of Information Science*, 40, 386-397.
- Costa, J. (1998). *La esquemática: visualizar la información*. Barcelona: Paidós.
- Courtial, J-P., y Law, J. (1989). A co-word study of artificial intelligence. *Social Studies of Science*, 19, 301-311.
- Ding, Y., Chowdhury, G. G. y Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing and Management*, s37, 817-842.
- Dorogovtsev, S. N., Goltsev, A. V. y Mendes, J. F. F. (2006). K-core organization of complex networks. *Physical Review Letters*, 96, 040601-040604.
- Du, N., Wu, B., Pei, X., Wang, B. y Xu, L. (2007). Community detection in large-scale social networks. En *Proceedings of the 9th WebKDD workshop*, 16-25. San Jose, California, USA: ACM.

- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486, 75-174.
- Freeman, L. (1978). Centrality in social networks: conceptual clarification. *Social Networks*, 1, 215-239.
- Garfield, E. (1994). Cienciography: mapping the tracks of science. *Current contents: social & behavioral sciences*, 7, 5-10.
- Harris, Z. (1951). *Methods in structural linguistics*. University of Chicago Press, Chicago.
- Hu, J., Zhang, Y. (2015). Research patterns and trends of recommendation system in China using co-word analysis. *Information Processing and Management*, 51, 329-339.
- Kostoff, R. N. Stump, J. A., Johnson, D., Murday J.S., Lau, C. G. Y. y Tolles, W. M. (2006). The structure and infrastructure of the global nanotechnology literature. *Journal of Nanoparticle Research*, 8, 301-321.
- Leydesdorff, L. y Hellsten, I. (2006). Measuring the meaning of words in contexts: an automated analysis of controversies about 'Monarch butterflies', 'Frankenfoods', and 'stem cells'. *Scientometrics*, 67, 231-258.
- Leydesdorff, L. y Welbers, K. (2011). The semantic mapping of words and co-words in contexts. *Journal of Informetrics*, 5, 469-475.
- Li, T., Ho, Y.S. y Li C.Y. (2008). Bibliometric analysis on global Parkinson's disease research trends during 1991-2006. *Neuroscience Letter*, 41, 248-252.
- Luque Agulló, G. (2004-2005). El dominio de la Lingüística Aplicada. *Revista Española de Lingüística Aplicada*, 17-18, 157-173.
- Perelman, Ch. y Olbrechts-Tyteca, L. (1989). Tratado de la argumentación. Madrid: Gredos.
- Ravikumar, S., Agrahari, A. y Singh, S. N. (2015). Document mapping the intellectual structure of scientometrics: A co-word analysis of the journal scientometrics (2005–2010). *Scientometrics*, 102, 929-955.
- Seidman, S. (1983). Network Structure and Minimum Degree. *Social Networks*, 5, 269-287.
- Small, H. (2006). Tracking and predicting growth areas in science. *Scientometrics*, 68, 595-610.
- Small, H. y Griffith, B. C. (1974). The structure of the scientific literatures: identifying and graphing specialities. *Science studies*, 4, 17-40.
- Van Eck, N. J. y Waltman, L. (2009). How to normalize co-occurrence data? An analysis of some well-known similarity measures. *Journal of the Association for Information Science and Technology*, 60, 1635-1651.
- Van Eck, N.J. y Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84, 523-538.
- Van Leeuwen, T. (2004). Descriptive versus Evaluative Bibliometrics. Moed, H. F., Glänzel, W., Wolfgang, S., eds., *Handbook of Quantitative Science and Technology Research*, 373-388. The Netherlands: Kluwer Academic Publishers.
- Van Raan, A. F. J. (2005). Measurement of Central Aspects of Scientific Research: Performance, Interdisciplinarity. *Measurement*, 3, 1-19.
- Waltman, L., Van Eck, N. J. y Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4, 629-635.
- Wasserman, S., y Faust, K. (1998). *Social network analysis. methods and applications*. Cambridge: Cambridge University Press.
- Yang, W.-S. y Dia, J. B. (2008). Discovering cohesive subgroups from social networks for targeted advertising. *Expert Systems with Applications*, 34, 2029-2038.
- Zhang, J., Yu, Q., Zheng, F., Long, C., Lu, Z. y Duan, Z. (2016). Comparing keywords plus of WOS and author keywords: a case study of patient adherence research. *Journal of the Association for Information Science and Technology*, 67, 967-972.