



**UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO
FACULTAD DE PSICOLOGÍA**

**REDES SEMÁNTICAS NATURALES:
DESARROLLO Y MODELOS DE ANÁLISIS**

TESIS

**QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN PSICOLOGÍA**

**PRESENTA:
JORGE ANDRADE VIDAL**

**DIRECTOR: Mtro. Francisco Javier Urbina Soria
REVISOR: Dr. Cesáreo Estrada Rodríguez**



Ciudad Universitaria, CD. MX. 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

Resumen	3
Introducción	4
Memoria humana	7
Sistemas de memoria	10
Memoria semántica	12
<i>Semantic priming</i>	16
Modelos y teorías de redes semánticas	19
Modelo Collins & Quillian	20
Modelo “HAM” de Anderson y Bower	23
Teoría de la Comparación de Rasgos	27
Técnicas de análisis semántico	30
Redes semánticas y procesamiento de lenguaje natural	30
Modelo del Hiperespacio Análogo al Lenguaje HAL	32
Modelo de Análisis Semántico Latente LSA	34
Redes semánticas naturales	36
Redes semánticas naturales modificadas	40
Redes semánticas con escalamiento multidimensional	47
Análisis estructural de redes semánticas naturales	51
Métricas del análisis	52
Análisis comparativos entre aproximaciones	56
Discusión y conclusiones	63
Referencias	67

Resumen

La realidad se concibe y construye de diversas formas; una de ellas mediante los llamados significados psicológicos. Por este medio el humano abstrae el conocimiento que utiliza para desempeñarse en su medio. En las últimas décadas se han aplicado diversas metodologías para evaluar cómo es que las personas hacen esa abstracción. Quillian y Collins fueron de los primeros en estudiar y proponer uno de estos sistemas estructurados. Utilizando diversas teorías y conceptos pertenecientes a la psicología cognoscitiva, construyeron lo que se llama una red semántica, que tiene como fin conocer el significado de un concepto y entender la estructura jerárquica que le confiere sentido en una red de subconceptos relacionados. Figueroa y colaboradores propusieron un enfoque distinto, haciendo a los participantes preguntas directas de asociación libre con el concepto en cuestión y utilizando un método de jerarquización proporcionado también por los mismos colaboradores; por este motivo es que se le nombró técnica de redes semánticas naturales (RSN). El presente trabajo se aborda la evolución de las RSN, desde los conceptos que le dieron origen hasta los diferentes métodos que se emplean hoy en día y se hace una comparación entre ellos, presentando las diferencias y cualidades entre estas diferentes técnicas aplicadas al análisis de la estructura cognoscitiva de un concepto y su significado.

Palabras clave: redes semánticas naturales, metodología de redes semánticas naturales, análisis estructural.

Introducción

“Por la experiencia, progresan la ciencia y el arte en el hombre” (Aristóteles)

Ser niño es ser curioso. ¿Por qué el cielo es azul?, ¿qué animal es ese?, ¿qué es un vagabundo? Por qué y qué, son las formas prácticas que tiene un infante de satisfacer esa curiosidad. Así es como adquieren significado objetos y fenómenos de todo tipo, no sólo en la infancia sino a cualquier edad. Así es como se inicia el aprendizaje semántico (significado) y se consolida el lenguaje. El aprendizaje de conceptos, en todos los ámbitos del quehacer cotidiano y científico, es la herramienta más elaborada que tiene el ser humano para operar y cooperar con sus semejantes. La construcción de este sistema de conocimiento, ya sea en las etapas tempranas o la edad adulta, ha fascinado desde siempre a los filósofos, psicólogos cognoscitivos; lingüistas, sociólogos, y muchos otros profesionistas.

La memoria y la formación de conocimiento son temas estrechamente relacionados, ampliamente investigados y discutidos en la psicología. La memoria tiene una raíz profunda desde la historia de la filosofía y ahora, en nuestros días, con base en teorías computacionales. El cognoscitivism, tiene como materia de estudio la mente humana y la naturaleza de los conocimientos. Para explicar la función de la mente y los procesos que le subyacen, la ciencia cognoscitiva, sienta sus bases en la época de la Grecia clásica. En ella, Platón menciona que el conocimiento es innato y nunca se aprende algo totalmente nuevo. La mente evoca recuerdos que le permiten al ser humano operar en el mundo. Por otro lado, Aristóteles plantea una teoría que se contrapone a la de su maestro Platón. La humanidad tiene como característica especial el intelecto o la razón, las cuales actúan sobre la experiencia sensorial para abstraer los conceptos. El intelecto es innato, pero las representaciones de las categorías (conceptos) son producto de lo experimentado sensorialmente (Valdez, 2010).

Dichos planteamientos dan pie al surgimiento del racionalismo y el empirismo, en los cuales se mantienen diferentes perspectivas en relación con la adquisición del conocimiento. Mientras Platón establece por medio del racionalismo, que el conocimiento es formulado mediante un proceso de razonamiento y que los sentidos no intervienen de ninguna forma en ello; el empirismo de Aristóteles pone como base la experiencia humana con el entorno como fuente principal para la obtención del conocimiento, introduciendo la noción de asociaciones mentales, que se van registrando y almacenando en la mente (Valdez, 2010).

Descartes (1637) , siguiendo los principios del racionalismo, añadió un método (la duda metódica), para conocer el mundo con certeza, poniendo a prueba todo el conocimiento, yendo desde lo simple a lo complejo y compara la forma del saber con un árbol cuyas raíces son la metafísica, el tronco es la física y las ramas son la demás ciencias (Hernández, 2005). Separando con su idea de la metafísica, la mente y el cuerpo, argumentando que el mundo se conforma de un espacio y un pensamiento, siendo esta última la principal cualidad que Dios nos da; dejando al cuerpo como simples sistemas de engranajes para interactuar con el mundo.

Más tarde, el asociacionismo inglés propone definitivamente que el conocimiento no es innato, sino que se organiza por medio de las percepciones, la memoria y el pensamiento. Locke (1690) plantea que el mundo se va conociendo a partir de la asociación de ideas (teoría muy similar a la que propone Aristóteles). Mientras Hume (1748), el ilustre filósofo escocés, menciona que las impresiones del mundo exterior (la experiencia), son el conocimiento original y que el pensamiento copia estas ideas. Sin embargo, no es sino hasta que Kant (1781), que se habla de representaciones: sensaciones, intuiciones o categorías que se agrupan de forma ordenada. Este desfile de pensadores sienta las bases del movimiento cognoscitivista, que se interesa en conocer cómo se obtiene, almacena, clasifica, ordena y recupera la información que

proviene de la experiencia en la mente humana. Cómo es que se representa y constituye en mapas o “redes” cognoscitivas en constante construcción y reconstrucción, lo cual permite a las personas interactuar con su entorno (Valdez, 2010).

El presente trabajo describe las teorías actuales acerca de cómo se adquiere el conocimiento y se forman los conceptos que afectan las posturas, ideas, afecciones, conducta entre otras. Asimismo, se comparan las técnicas contemporáneas más útiles en la extracción y análisis de estos conceptos en la memoria humana. El objetivo general es el de comprender y contrastar las diferentes técnicas de las redes semánticas naturales. En primer lugar, se describe el sistema de memoria, lugar donde se procesa, almacena, clasifica y se recupera la información de todo tipo: sensorial, factual y episódica. En seguida se presentan los principales modelos de redes semánticas, elemento principal de esta tesis. En tercer lugar, se presentarán las principales técnicas que se usan actualmente para medir y evaluar el conocimiento, así como la organización de estas representaciones o conceptos mentales y se hace un análisis comparativo entre esas técnicas, esbozando los actuales usos y alcances que tienen las mismas. Por último, se hace una breve reflexión sobre líneas de investigación que podrían dar continuidad al desarrollo de redes semánticas naturales; por ejemplo, su inserción en un contexto más amplio como lo es el procesamiento del lenguaje natural.

Memoria humana

La memoria adquiere y selecciona cualquier información para darle significado, acorde a las relaciones, necesidades e interacciones del individuo con el entorno. Éste interioriza del entorno elementos que, cuando son registrados en la memoria, pasan a formar parte de su conocimiento de la realidad: “la experiencia presente está inexorablemente entrelazada con los recuerdos, y el significado de las personas, de las cosas y de los eventos del presente depende de algún modo de la experiencia previa, es decir, de la memoria, de modo que el presente pierde su realce sin el pasado” (Mora, 2010).

Ebbinghaus, en 1885, es el primer investigador en realizar estudios experimentales sobre la memoria (Goldstain, 2015). Este científico decidió poner a prueba la memoria realizando un estudio experimental que consistía en la repetición verbal de sílabas carentes de sentido. La intención era analizar cuántas de esas sílabas se podían evocar de una lista sin equivocarse. En este simple estudio, describe por primera vez la mejoría progresiva del rendimiento durante la adquisición de nuevas tareas o curva de aprendizaje (Montealegre, 2003). Bartlett (1916) experimentó igualmente con esta cualidad mental, encontrando una diferencia en relación con el estudio de Ebbinghaus, que radica en que si el contenido contiene un significado o el contenido es familiar para el sujeto, era menos probable que los sujetos distorsionaran, omitan detalles, abreviaran o transformaran los hechos. Este hallazgo da sustento a la hipótesis de que el conocimiento es esquemático y que se alimenta de la experiencia. De esta manera, cuando se presenta un nuevo material, éste se ajusta a los diseños ya establecidos, pero sometiéndose a un proceso dinámico de reconstrucción y construcción.

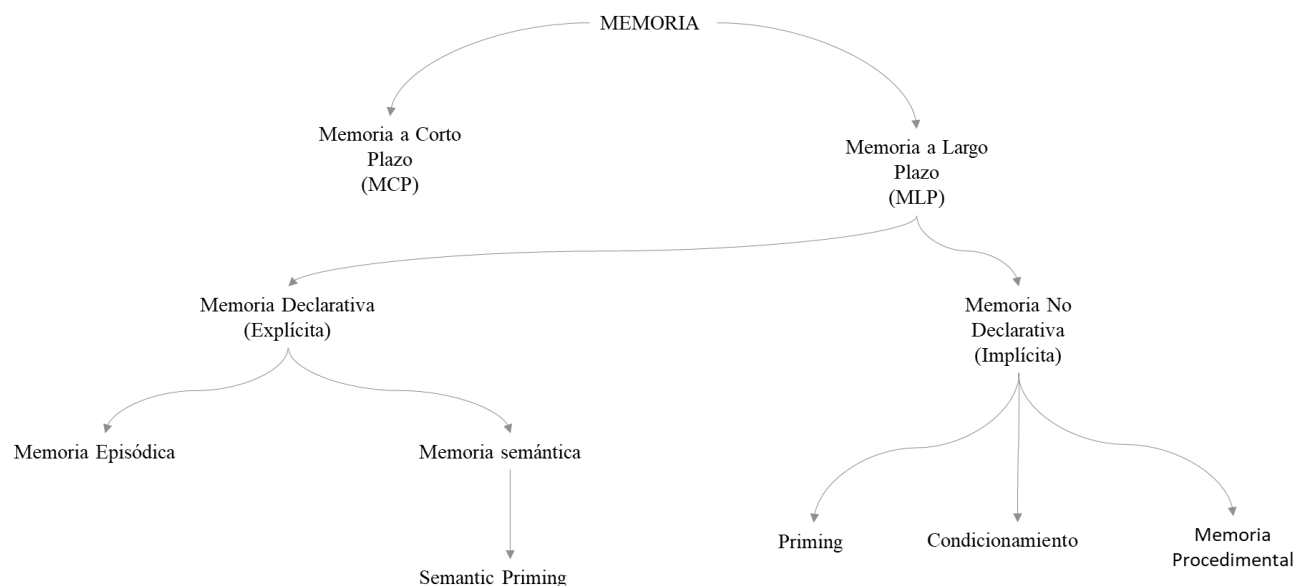
James (1890) fue el primero en hacer una distinción en los procesos de la memoria; la divide en memoria de corto plazo (MCP) y memoria de largo plazo (MLP). La MCP mantiene la información recién percibida por un periodo corto de tiempo y coordina los flujos de información (mediador para MLP), generados por los estímulos y cognición del sujeto. Mientras que la MLP conserva de forma perdurable representaciones realizadas por la persona, implicando selección, organización y consolidación de información de acuerdo con las cualidades del conocimiento. La recuperación requiere generalmente de la creación de imágenes mentales que son evocadas en el momento y cuyo recuerdo puede incluir una gran variedad de tipos de información: visual, auditiva, olfativa, somestésica, temporal, espacial, emotiva, entre otras. Este proceso puede llegar a recrear el evento pasado de una forma muy vívida en el presente (Aguado, 2001).

Varias corrientes de la psicología han postulado las diferentes formas en las que se maneja el conocimiento. La teoría de la Gestalt alude a que la experiencia psicológica nunca es estática, comprende un campo de eventos bien organizado y dinámico en la que los elementos interactúan y se afectan mutuamente. Las actividades mentales no dan lugar a una copia idéntica del mundo percibido, comprende los procesos de extracción y selección de información relevante encargándose de generar un estado de claridad y lucidez que permita el desempeño dentro del mayor grado de racionalidad y coherencia posibles con el mundo circundante. El efecto producido por la influencia material es la generación de las ideas y procesos de orden categorial (Oviedo, 2004). La psicología cognoscitiva expresa que las asociaciones que se tienen en la mente se pueden ir construyendo y reconstruyendo, son dinámicas y conforman mapas o redes que son parte del procesamiento humano de la información. La mente es considerada como un sistema de manipulación simbólica que tiene su base en la adquisición, manejo y procesamiento de la información.

La complejidad de la memoria es un tema siempre actual entre científicos de la mente e investigadores cognoscitivistas. Las definiciones y alcances de sus investigaciones y teorías son variables y en ocasiones muestran imprecisiones o contraposición. Pero, aun así, dejan ver que el estudio de los procesos puede hacerse desde distintas aproximaciones. Los diferentes sistemas de memoria trabajan en paralelo, ocasionalmente de forma independiente e inclusive de forma competitiva. Este trabajo se centra en una parte de la memoria a largo plazo (ver Figura 1) un subsistema de ésta, que es la memoria semántica y sus teorías, como el *semantic priming* (diferente al *priming* en la memoria no declarativa).

Figura 1

Representación de los sistemas en la memoria



Nota. La estructura fue hecha con base en el trabajo de Squire (2004).

Sistemas de memoria

La memoria comprende varios sistemas, comparten varios elementos y cualidades entre sí. Esta propiedad de la mente es de vital importancia en el actuar y debido a esto es tema frecuente de los investigadores, aún en nuestros días va aumentando los elementos que sabemos de esta.

La memoria a corto plazo permite almacenar cierta cantidad de información durante un periodo corto de tiempo. Para ejemplificar, imagine que alguien a quien aprecia mucho le da su número telefónico, si no tiene algo en dónde anotarlo y/o al no haber repetición, al pasar un tiempo aquellos dígitos se pueden olvidar irremediablemente. De acuerdo con Cerdán (2011), este tipo de memoria almacena impresiones momentáneas, tiene una persistencia limitada y en situaciones experimentales puede retener de 7 elementos distintos en promedio; este fenómeno es discutido por Miller (1956), argumenta que si estos fueran más, crece la posibilidad de olvidar o confundir elementos. Si la información es repetida o retenida, pasa a formar parte de la memoria a largo plazo, la cual administra y evoca estos datos.

La existencia de la memoria hace posible el aprendizaje. Existen varios modelos que explican cómo es su funcionamiento y uno de ellos es el Modelo de Propagación de la Activación también llamado ACT por sus siglas en inglés (*Adaptative Character of Thought*) el cuál considera que el sistema de la memoria maneja dos tipos de información: conocimiento declarativo y conocimiento no declarativo.

La memoria no declarativa, también llamada *memoria implícita*, la conforman varias subcategorías a) el conocimiento procedimental, el cuál es el saber hacer, es la ejecución de procedimientos, estrategias, técnicas, habilidades, destrezas o métodos; depende continuamente de la práctica y es difícil expresarlo en palabras. B) El *priming* que es un tipo de memoria

implícita que no requiere de ninguna recolección consciente de experiencias previas, y que comparte algunas características con la memoria procedimental, pero también con la memoria semántica. Al igual que la memoria procedimental, el *priming* (diferente al *semantic priming*) significa un incremento de habilidades, afectado por condiciones de codificación y recuperación de la información. Por otro lado, también se asemeja a la memoria semántica en que involucra representaciones cognoscitivas del mundo exterior y su expresión es más cognoscitiva que conductual (Mora, 2010). El condicionamiento deriva de la asociación repetida entre un estímulo y una respuesta (condicionamiento clásico) o entre un estímulo y una conducta (condicionamiento operante).

La memoria declarativa es conocimiento de datos, hechos, conceptos y principios, pudiéndose comunicar verbalmente y se adquiere repentinamente. La memoria episódica es dependiente del contexto, evoca experiencias del mundo gracias a un indicio temporal o espacial en el que se imprimió ese recuerdo. En particular, el presente trabajo se centra en la memoria semántica, porque representa el fenómeno organizador de la memoria semántica que da fuerza y sustento a todo un desarrollo e investigación dentro de la psicología cognoscitiva con impactos teóricos, pero sobre todo prácticos.

Memoria semántica

El sistema de memoria que codifica, organiza y almacena de una forma abstracta las experiencias se denomina memoria semántica. Ésta no discrimina entre recuerdos agradables o desagradables; opera con ellos de la misma forma, pero con una condición, que tengan significado.

Efectivamente, la memoria no es un ente exclusivamente archivista, sino que también tiene voz y voto en la conducta humana al interiorizar y sintetizar toda esa información, a menudo auxiliada por la memoria episódica. En la memoria semántica los conceptos se van relacionando entre sí, pero también dan lugar a otros que se reconstruyen con base en información nueva (Carrillo-Mora, 2010).

El conocimiento semántico se organiza en dominios específicos. Por ejemplo, animado (animales, plantas, entre otros) e inanimado (herramientas, muebles y demás). Como resultado de limitaciones evolutivas, el conocimiento semántico se desarrolló a lo que se conoce como la *cuenta específica del dominio*, que consiste en diversos circuitos neuronales especializados para facilitar una ventaja de reconocimiento de ciertas categorías necesarias para la supervivencia (Caramazza, 1998). Estas características dan lugar a almacenes de conocimiento separados para categorías específicas. En 1984, por ejemplo, Warrington y McCarthy describieron pacientes con recuperación parcial de encefalitis por herpes simple. Un paciente con un déficit de memoria semántica era inferior para clasificar los conceptos del dominio de objetos inanimados que para los de tipo natural; otros pacientes mostraron un patrón contrario de deterioro semántico; un incorrecto desempeño en seres con vida que en objetos no vivos. Tales déficits semánticos fueron referidos como "categoría específica". Estas observaciones sugirieron que el sistema de memoria semántica se subdivide en diferentes componentes en función del contenido de los conceptos (Warrington, 1984).

De manera continua, las asociaciones de características sensoriales y motoras de los conceptos de un objeto constituyen el principio organizador primario en la memoria semántica. En este punto de vista sensorio-motor, los conceptos de objeto están compuestos, en parte, de asociaciones de características sensoriales y motoras que son críticas para su significado y estas características se almacenan cerca de las regiones sensoriales y motoras correspondientes del cerebro. Según esta teoría, las representaciones semánticas se basan en redes distribuidas de características y estas redes son paralelas a las características anatómicas. Aunque el sistema semántico parece ser paralelo a la distribución de los sistemas sensoriales y motores, este sistema puede ser distinto a un nivel neural de grano fino (Kranjec, 2010).

La corteza prefrontal es otra región cerebral implicada en el procesamiento semántico y es constantemente activada en tareas de memoria semántica; se cree que las funciones de la corteza prefrontal son distintas de los mecanismos que admiten representaciones en memoria semántica y a menudo se caracterizan por ser "procesos" de dominio general que actúan sobre conceptos semánticos y otras representaciones mnemotécnicas y perceptivas. Por ejemplo, se cree que la corteza prefrontal es compatible con un número de procesos ejecutivos de alto nivel en la memoria semántica, como recuperar información específica de la memoria semántica y seleccionar la representación apropiada de una serie de alternativas semánticas. Otros autores sugieren que esa corteza prefrontal juega un papel activo en la organización de la información representada en la memoria semántica (Price, 2015). Aunque se requiere más trabajo para especificar las muchas funciones de la corteza prefrontal en la memoria semántica, esta región parece ser crítica para la organización, recuperación y uso de representaciones semánticas.

La actividad en la corteza prefrontal ventrolateral izquierda (VLPFC) se ha asociado fuertemente con el control de la memoria semántica; específicamente, como guía de

recuperación y selección posterior a la recuperación de información conceptual almacenada en áreas temporales posteriores y quizás en otras áreas corticales. Una evidencia importante de la idea de que la información sobre la forma del objeto visual se almacena en la corteza temporal ventral proviene de estudios que muestran que las regiones activas cuando los objetos se reconocen visualmente también lo están cuando los sujetos generan imágenes visuales de esos objetos (Martin, 2007). Estos hallazgos, que respaldan la idea de que la información de objetos visuales se almacena en esta región del cerebro, también ofrecen una alternativa a las interpretaciones basadas en apelaciones al procesamiento conceptual: específicamente, que la actividad temporal posterior simplemente refleja la recuperación explícita de las imágenes de objetos visuales que acompañan el desempeño de habilidades. Los estímulos que denotan objetos concretos, ya sean imágenes o palabras, desencadenan imágenes visuales, que luego recluta indirectamente la corteza temporal ventral posterior, la cual no hace ningún trabajo conceptual.

Es evidente que los sistemas de memoria episódica y semántica dependen de redes cerebrales a gran escala, parcialmente superpuestas, que incluyen no solamente los lóbulos temporales medial y lateral, sino también porciones de los lóbulos frontal y parietal. Así, además de que la distinción teórica entre memoria episódica y semántica ha demostrado ser útil de muchas maneras, hay mucha evidencia de grandes diferencias anatómicas entre los dos sistemas.

La mayoría de los conceptos son representaciones del mundo exterior, sin embargo, nosotros podemos tener términos de conceptos abstractos los cuales no corresponden a una entidad física, como cuando nos mencionan “el tiempo es invaluable” o “no puedes comprar el amor” (por mucho que duela), debido a que en ningún lugar material conseguiremos tiempo o amor ya que no es algo físico; aun así, estos términos igual se componen en la memoria semántica. La diferencia a nivel de organización es que los conceptos abstractos se basan

principalmente en un sistema de asociaciones verbales, mientras que los conceptos concretos (de entes corpóreos, no abstracciones) se basan en asociaciones de características verbales y sensoriales (Paivio, 1971). Esta hipótesis cognoscitiva predeciría correlatos neuronales distintos para estos dos procesos y, de hecho, parece ser el caso de que provocan la activación en regiones parcialmente distintas. Los diferentes tipos de memoria llegan a operar juntos o en competición; sin embargo, gracias a esto es posible que se refuerce un aprendizaje con experiencias al integrarlas a nuestro conocimiento en general. Al tratar de definir ya sea un objeto animado o inanimado el cerebro activa diferentes regiones dependiendo del dominio de la palabra, así como sus características y atributos: si es parte de una conducta motora, emocional, o abstracta, así como los procesos que permiten su recuperación de forma eficiente para su utilización en el pensamiento y el lenguaje. Parte de este procedimiento que utiliza la memoria se explica por la teoría del *semantic priming*.

Semantic priming

Si se pide a distintas personas que piensen en comida y se mencionan países al azar, la probabilidad de obtener respuestas semejantes para cada país es muy alta. Por ejemplo, Japón recuerda al sushi, Italia la pizza, México a los tacos. No es difícil llegar a tales relaciones. El responsable de este fenómeno es el *semantic priming*, el cual es un mecanismo perceptual que da un punto de partida, origina un contexto, derivado de una experiencia o aprendizaje, optimizando la velocidad o exactitud de responder ante un estímulo, ya sea una imagen o palabra relacionada semánticamente. Esto se debe a que existe verdadera relación entre las palabras y su significado; típicamente el concepto contiene una mezcla de asociaciones semánticas, según la circunstancia (McNamara, 2004).

La memoria humana organiza la información por el papel que desempeñan, las junta por sus semejanzas, características y cercanía entre conceptos. Así es como se da lugar a grupos dependiendo de los atributos antes mencionados. Es más fácil relacionar perros y gatos que perros con halcones, ya que, aunque ambos son animales, no pertenecen al mismo grupo de asociaciones mentales. El primero puede referirse a mascotas domésticas mientras que el segundo a animales en general, que es un grupo mucho más extenso. Se puede llegar eventualmente a la asociación perros-halcones, pero resulta más rápido llegar a la asociación perros-gatos, debido a que tienen una mejor relación semántica. Identificar si una palabra tiene conexión con la otra, será simple si ésta guarda un vínculo, a diferencia de si no existiera ese lazo (Meyer David E, 1971). En general, se acepta la idea de que la preparación semántica (enfocarse en una categoría semántica) puede amplificarse mediante la atención consciente, tal vez debido a las expectativas para seguir una secuencia con la palabra objetivo. Es probable que las palabras

estímulo-objetivo relacionadas estén representados en la memoria debido a un historial de co-ocurrencia, por lo que tienen un alto grado de familiaridad y permite una decisión léxica relativamente rápida. Los pares no relacionados tienen menor familiaridad y se necesita más tiempo para tomar una decisión léxica.

Un ejemplo sencillo y práctico; si se menciona la frase “trastornos psicológicos”, algunas posibles respuestas son: estrés, ansiedad, depresión; llegará un punto en el que se hará difícil el aportar un nuevo término que no se haya dicho antes, hasta se podrían repetir palabras e inclusive caer en un número limitado de conceptos. Lo interesante de este ejemplo es que se aprecia de primera mano cómo funciona el *semantic priming*, cómo la mente empieza hacer relaciones semánticas de acuerdo a un contexto, además la interacción que hay entre conceptos. Mientras más se esté en la dinámica, es mayor la posibilidad de equivocarse, ya que cada vez se deben hacer más subdivisiones semánticas del estímulo base; además, si se pregunta a diferentes personas con distintos enfoques de la psicología, se darán términos característicos del área; si se pregunta a quien trabaja en el desarrollo del aprendizaje, sus listas de trastornos diferirían de otra persona que nunca ha estado en ese campo.

Hay situaciones en las que se requiere controlar las respuestas habituales relacionadas con un estímulo, o en las que las respuestas contienen secuencias nuevas de acciones que no están bien aprendidas; tal es el caso de las tareas tipo *Stroop*. Éstas consisten en nombrar el color de la tinta en que está escrita una palabra que a su vez denota algún color. Éste puede ser el mismo de la tinta, ROJO escrito en color rojo (situación congruente), o un color diferente, por ejemplo, ROJO escrito en color verde (situación incongruente). Cuando se junta con el *semantic priming* causa fenómenos interesantes los cuales pueden ejemplificar los modos en los que se hacen decisiones semánticas. Tomando como situación el efecto *stroop*, al nombrar el color de la

tinta de un estímulo neutral (por ejemplo, cuando la tinta del color y la palabra no interfieren uno con el otro) el procesamiento es más rápido que en condiciones incongruentes, ya que se presenta una facilitación semántica del procesamiento léxico-semántico del estímulo objetivo, ya sea a través de un mecanismo de expansión automática de la activación de las representaciones de los conceptos almacenados en la memoria semántica o mediante la creación de una expectativa. Por el contrario, al nombrar estímulos incongruentes el procesamiento es más lento (cuando la tinta del color y la palabra no concuerdan), ya que se da una interferencia semántica (MacLeod Colin M., 2000). Una explicación alternativa del *semantic priming* y su relación con estos fenómenos se deriva de modelos de redes neuronales de identificación de palabras. En muchos de estos modelos, el conocimiento sobre las palabras está representado por un patrón de activación en una colección de unidades de procesamiento. Las palabras relacionadas tienen patrones similares de activación. La presentación de un estímulo relacionado mueve la red a un patrón similar al asociado con el objetivo, lo que le da a la red una ventaja para instar el patrón de activación del objetivo, creando así un efecto de *priming*. Aunque el *semantic priming* puede ser interrumpido por una palabra no relacionada que interviene, este efecto supone que el evento principal modifica los pesos de conexión en la red para favorecer las palabras relacionadas, lo que equivale a modificar la memoria a largo plazo (Masson, 2001). Esta intervención que se puede hacer en la memoria semántica (memoria a largo plazo) es sin duda una de las razones principales por las que interesa a los investigadores desarrollar modelos de redes semánticas, tema del siguiente capítulo.

Modelos y teorías de redes semánticas

El estudio del *semantic priming* ha llevado a diferentes investigadores a construir modelos para desentrañar el proceso en que la mente va categorizando, organizando y extrayendo los significados, denotando las cualidades de cada concepto al formar sus grupos, instituyendo un nivel jerárquico.

Diversos autores han propuesto métodos y visiones para construir e interpretar las redes semánticas. El término “red semántica” deriva de la tesis de doctorado en psicología cognoscitiva de Collins & Quillian (1969), quien lo introdujo por primera vez como modo de expresar la organización de la memoria semántica humana. Es decir, un orden jerárquico dentro de la memoria en el cual los conceptos son expresados en palabras (Forte, 2005). Además, estas palabras se ilustran con un gráfico en el cual los nodos o vértices etiquetados representan conceptos o características específicas, mientras que los arcos, también etiquetados, representan vínculos de diversas clases entre conceptos (Vivas J. R., 2007). Las relaciones entre conceptos se analizan por medio del arco, también llamado arista, esta relación descubre la dinámica en la que el concepto forma las conexiones.

Los conceptos, por sí solos, no dicen nada; el contenido de un concepto no depende de su significado lexicográfico sino de la red de relaciones que establece con los otros conceptos, conectados a la red cognoscitiva (memoria semántica) del participante. Esto considera que los conceptos son representaciones mentales de una clase de objetos; representan el significado de conceptos en la memoria semántica. Estas representaciones están profundamente relacionadas entre sí, por lo cual forman una estructura con base en sus interrelaciones que reproduce el significado del concepto dentro de la memoria. La densidad conectiva de un concepto define su

relevancia semántica; más aún, con los nuevos modelos de agrupamiento o modularidad se extraen subconceptos derivados de dicha densidad de conexiones. Esto sólo es posible verlo aplicando el análisis estructural (Freeman, 2012). A continuación, se abordan algunos de los modelos más representativos de redes semánticas.

Modelo Collins & Quillian

Collins & Quillian diseñó una red semántica tal como hipotetizaba que funciona la memoria semántica (Collins, 1969). Esta red tuvo su origen con el diseño de un programa de computadora que fuera capaz de comprender enunciados en lenguaje natural y responder a los mismos de forma aceptable. Se podría decir que preconizaba el valor del procesamiento del lenguaje natural, el cual es actualmente un área de tremendos alcances tanto para la industria de la información (redes sociales, buscadores, inteligencia artificial) como para la ciencia (Vásquez, Vega Huerta, & Pa, 2009). El TCL (*Teachable Language Comprehender*) o CLE (Comprendedor de Lenguaje Enseñable) tenía diversos objetivos como ser capaz de analizar sintácticamente los enunciados y de realizar algún tipo de interpretación semántica de ellos (Rodríguez, 2010). Al tratar de resolver el problema de la comprensión por computador, Quillian dio con el problema de la memoria semántica. Es decir, que el problema de la comprensión no se resolvía introduciendo en el computador un diccionario de términos, sino que es necesario disponer de conceptos y relaciones entre conceptos formados por grupo de términos y sus relaciones o aristas.

Este método dividió los vínculos en dos clases: los vínculos ISA, llamado así debido a que se refiere a “*is a*” (“es un” ó “es una” en inglés) que representan relaciones jerárquicas o categóricas entre conceptos y los vínculos simples, propiedades que representan características específicas que se asocian a conceptos particulares. Ambos tipos de vínculos son direccionales;

los vínculos de clases conceptuales (“perro”, p ej.), que representan relaciones jerárquicas o categóricas entre conceptos (“cánido”, “animal”) y los vínculos de propiedades (“tiene pelo, ladra”, p ej.) que representan características específicas que se asocian a conceptos particulares. La red semántica se organiza por líneas de similitud semántica. Cuantas más propiedades (vínculos) tienen dos conceptos en común más próximamente relacionados se hallarán y por lo tanto más rápido se propagará la asociación semántica entre ellos (Fig.2).

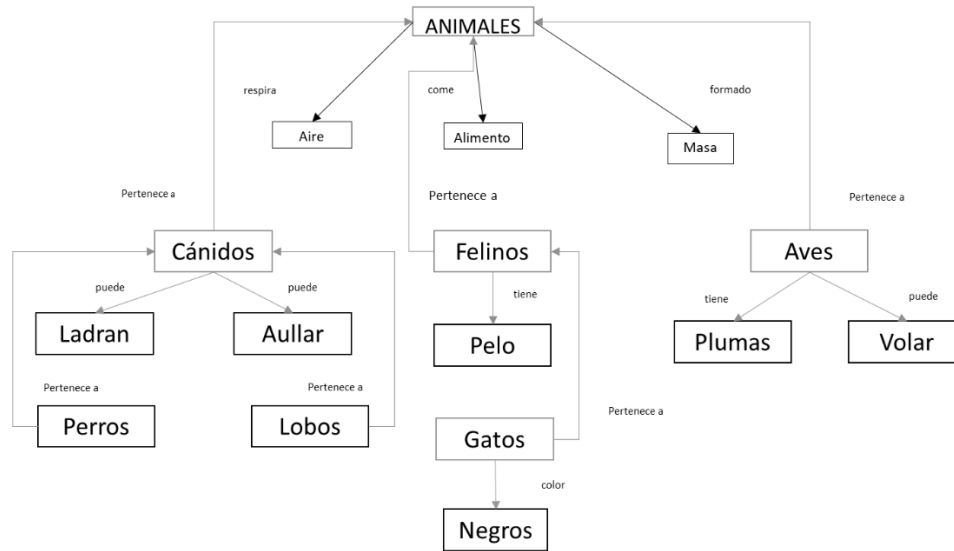
La jerarquía o red semántica tiene un par de restricciones:

1. Cada nodo sólo puede tener un vínculo ISA que llega a él.
2. Cada vínculo de propiedades debe aparecer tan alto en la jerarquía como sea posible.

De este modo, el modelo permite, de un modo rápido y sencillo, hacer predicciones específicas. Para hacerlo, basta con contar el número de “operaciones” de diferentes clases que una tarea particular requiere de acuerdo con la teoría y verificar si los datos emergentes de la ejecución de la tarea arrojan resultados consistentes con dichas predicciones.

Figura 2

Ejemplo de red semántica de Quillian y Collins



Nota. Los nodos son conceptos y las relaciones son vínculos con algunas características entre los conceptos. Figura basada en Collins 1969.

A partir, del modelo de Quillian se formularon nuevas teorías, como la teoría de la Propagación de Activación, la cual trabaja sobre la base de un modelo de taxonomía jerárquica (categorías jerárquicas). Posteriormente se modificó con base en distancias semánticas, formulando la teoría Extendida de la Propagación de la Acción (Forte, 2005).

Este proceso de difusión o propagación de la activación es el fundamento de los recorridos semánticos, aquello que marca la trayectoria de la circulación de ejes semánticos a través de la cadena de nodos, identificando los ejes semánticos del argumento. El recorrido semántico es un encadenamiento directo de nodos relacionados que forman un encadenamiento de unidades semánticas (proposiciones) que permiten identificar los múltiples ejes semánticos del discurso sobre el tópico. Supongamos que se escucha una frase, los nodos correspondientes a los conceptos presentes en la red se activan, propagándose a todos los nodos que están conectados. La propagación es más rápida si el lazo que une a los nodos es más “cercana”. El

ciclo se detiene hasta que el sentido de la activación es diferente. Esto “imita” la forma en que el pensamiento busca la información en la memoria a largo plazo.

La red semántica se organiza por líneas de similitud semántica. Cuánto más propiedades en común tienen dos conceptos, más próximamente relacionados se hallan. Los nombres de los conceptos son guardados en una red léxica (diccionario) organizado por similitud fonológica y ortográfica (Vivas J. R., 2008).

Sin embargo, en el modelo se encontró sistemáticamente una falla que se dio en llamar “efecto típico”. En ciertos casos, dos nodos subordinados a la misma distancia del nodo superior mostraron diferencias significativas en el tiempo de respuesta, cuando el tiempo debería haber sido igual (Forte, 2005). Si se toma la Figura 2 añadiendo un nodo en el apartado “cánidos”, llamado tanuki (perro mapache) ambos se hallarían a distancia uno del nodo cánidos. Sin embargo, aunque la predicción del modelo es que ambos deberían tomar el mismo tiempo en la tarea de verificación, las personas demoran sustantivamente más cuando deben verificar tanuki. La facilidad de acceso a un concepto, calculada por el tiempo de respuesta, depende más de la frecuencia con que dos conceptos aparecen juntos que de la posición que cada uno ocupa en la jerarquía semántica. En realidad, se trataba más de un orden discreto que jerárquico.

Modelo “HAM” de Anderson y Bower

Gracias al modelo de las redes semánticas los investigadores pudieron formular distintas propuestas para estudiar los procesos de propagación de la activación que el sistema cognoscitivo utiliza para manejar información conceptual. Al igual que en el modelo anterior, en el modelo HAM de Anderson y Bower (Vivas J. R., 2008) la unidad cognoscitiva primordial de

la red son los nodos, conceptos con significado, aunque en este modelo se considera que existen dos tipos de información: conceptual y contextual. La información conceptual provee aquellas palabras definitorias, mientras que la información contextual comprende las características físicas en el momento en que se presentó el concepto, las asociaciones implícitas con ese elemento. La recuperación en la memoria semántica deriva de la activación de su representación interna, la activación se propaga desde un concepto a los otros conceptos vinculados, y la activación residual derivada del anterior proceso es acumulada facilitando una posterior recuperación (Vivas J. R., 2008). Por esta razón y no por otra, la técnica del análisis estructural señala la importancia de contextualizar cualquier red semántica natural; es decir, ubicar al sujeto en la categoría o concepto principal (Andrade, 2017).

Las proposiciones son especificaciones de las relaciones semánticas entre conceptos teniendo un valor de verdad y son abstractas, cada proposición define el contexto en términos de cómo fue vivida. Además, se trata de una organización jerárquica donde se distinguen como primer elemento jerárquico las ideas simples, creciendo hacia ideas complejas. Todo esto se articula en un sistema reticular binario en donde cada nodo se bifurcaría en otros dos, formando una compleja red asociativa, donde distinguen entre ocho tipos de conexiones posibles: contexto-hecho, tiempo-lugar, sujeto-predicado y relación-objeto (Manzanero, 2006).

Manzanero menciona un experimento en el que se muestra el proceso de memorización. En una primera fase se asocian los reactivos a una lista de marcadores que constituyen el contexto; lo cual comprende el cómo fue que se presenta la información. Las asociaciones de ambos elementos forman las huellas de memoria, que activan la red de conocimientos relacionados conceptualmente, que a su vez activan otros nodos relacionados de acuerdo con ambas categorías (propagación de la activación). Los nodos activados que tienen una relación

sensorial y conceptual con los elementos que se presentan al sujeto, afectan también a la composición de la huella de memoria a través de la lista de marcadores, ya que influirán en el significado. Cuando se presenta al sujeto una lista de palabras para ser memorizada, el sistema la trata como una unidad simple donde los reactivos que la forman están sujetos a una asociación todo-o-nada. La información sobre el contexto donde se presentó la lista es para el sistema de aprendizaje un conjunto de elementos independientes de la lista, pero que servirán para identificarla. Esta información sobre el contexto constituye los marcadores de lista, únicos para cada elemento. Para que pueda haber recuperación deben darse dos condiciones: que exista en el sistema un prototipo (contexto) con el cual el sistema haya construido una representación del estímulo original y que será el desencadenante del recuerdo; y que se presenten en el momento de la recuperación los conceptos adecuados que servirán de activadores de los nodos que representan la información.

Un conjunto de elementos contextuales puede servir para identificar con éxito cada reactivo de una lista de palabras, aunque los elementos aparezcan por separado o conjuntamente. El elemento más importante de la propuesta de Anderson y Bower es que las diferentes tareas de recuperación implican procesos distintos. Distinguen entre tareas de reconocimiento y tareas de recuerdo. Reconocimiento, cuando se presenta una palabra a un sujeto para ser reconocida, el nodo que la representa se conectaría con otros nodos mediante “marcadores” que contienen información del contexto donde se presentó la palabra, llevándose entonces a cabo dos procesos: el de emparejamiento, conexiones de nodos de memoria que establecen correspondencias entre patrones en los procesos de codificación en la memoria, y el de identificación, que analiza qué información de la emparejada es útil para codificar, distinguiendo qué información debe ser etiquetada como nueva (Manzanero, 2006). Las tareas de recuerdo abarcan un proceso de

búsqueda que estará guiado por principios de organización y asociación del material almacenado. Es decir, la búsqueda no es al azar, sino que necesita de indicios contextuales que servirán de activadores de los nodos que representan los conceptos a recordar.

El paso del tiempo provoca un debilitamiento de las huellas de memoria, ya que las asociaciones entre nodos que forman la red proposicional se debilitan. Igualmente, la presentación de nueva información afecta a la información almacenada con el paso del tiempo, dependiendo de si las nuevas entradas están relacionadas con las huellas de memoria existentes o no. Se distingue entre entradas relacionadas conceptualmente y entradas relacionadas contextualmente. En el primer caso (relacionadas conceptualmente) se encuentra la aparición múltiple de un mismo concepto en contextos diferentes, dándose un efecto de generalización. En las entradas relacionadas contextualmente puede hablarse de interferencia producida de unas estructuras proposicionales sobre otras similares. La interferencia (debilitamiento en las asociaciones de la memoria) puede darse de forma retroactiva, del material viejo sobre el que se está aprendiendo; o proactiva, del material nuevo sobre el que ya estaba almacenado. También se produce interferencia cuando las nuevas entradas no están relacionadas. En términos generales, las adquisiciones de nuevas asociaciones tienden a “enterrar” las viejas asociaciones haciéndolas inaccesibles, aunque la interferencia será mayor cuando los estímulos sean similares. (Manzanero, 2006).

Teoría de la Comparación de Rasgos

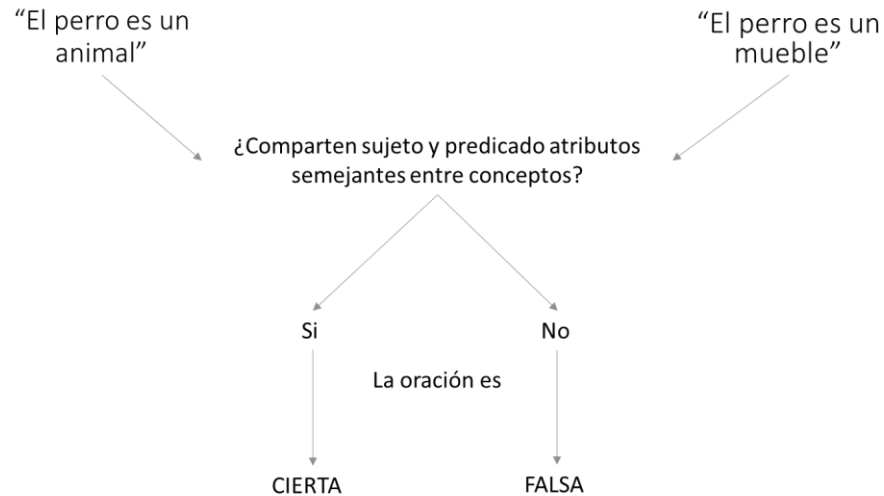
Tomando en cuenta las características de los anteriores modelos se puede notar que el hecho de recuperar un concepto puede ser difícil, ya que se tienen que comprobar diversas situaciones, como jerarquías semánticas o contextuales; sin embargo, Smith, Shoben y Rips (Valdez 2010), postularon que hay conceptos que son más representativos que otros en una categoría, debido a que tienen más atributos típicos y significativos del campo semántico en el que se encuentran.

El significado de una palabra está representado por una lista de rasgos o características semánticas que tienen valores como tamaño, forma, color. Algunos de estos elementos son típicos en una categoría, y mientras estas características se alejen de ese prototipo, el concepto será más difícil de recuperarse dentro de la categoría. Gracias a esto es posible tener la “definibilidad” de un concepto; en un extremo de la escala están los rasgos esenciales para la definición del significado de una palabra, mientras que en el otro extremo están los rasgos que son sólo atributos particulares de ese elemento. En el caso de la red semántica propuesta por Collins & Quillian, perro, lobo y tanuki pertenecen a los cánidos, y como todos ellos se encuentran a la misma distancia jerárquica todos deberían tener la misma importancia para la categoría en general. Sin embargo, esto no es así, ya que unas palabras son mejores que otras al representar, ya que poseen rasgos más definitorios y significativos del grupo cánidos.

Este proceso de recuperación se da en dos etapas (ver Figura 3), si alguien dice “el perro es un animal” primero se aparean todos los rasgos del sujeto y del predicado, si existe una semejanza entre los dos tipos de rasgos, se produce una toma de decisión rápida por la positiva: sí, si es cierto. Por el contrario, cuando la oración contiene poca o nula semejanza, se toma una decisión negativa.

Figura 3

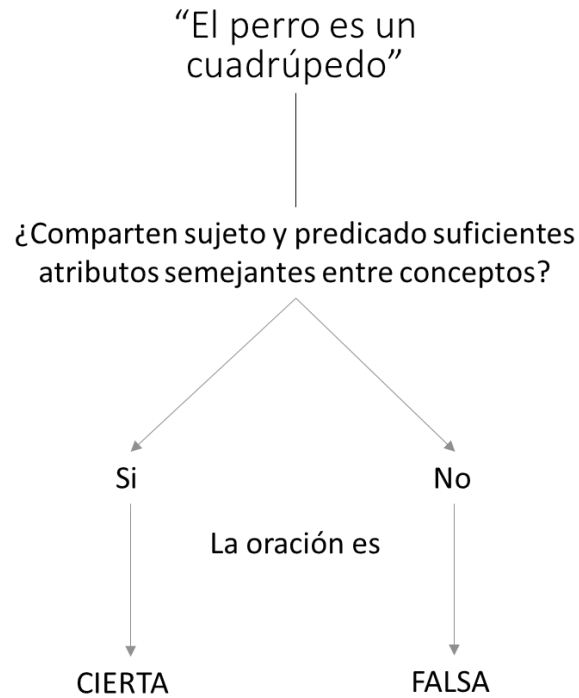
Representación de la etapa 1 del proceso de recuperación



La segunda etapa (ver en Figura 4) se refiere a si se presenta una semejanza entre las proposiciones “el perro es un animal” y “el perro es un cuadrúpedo”. Entonces se comparan los rasgos definitorios del sujeto con los del predicado. Si los rasgos del predicado están incluidos en el conjunto del sujeto, se toma una decisión de “verdadero” de lo contrario la premisa se declara falsa (Vivas J. R., 2008).

Figura 4

Representación de la etapa 2 en la comparación de rasgos



Nota. La etapa 2 ocurre cuando una oración o palabra puede resultar ambigua o no hay suficientes elementos discriminatorios entre sujeto y predicado en una oración.

El modelo predice que las oraciones verdaderas serán verificadas más rápidamente si el sujeto y el predicado comparten más atributos pertenecientes a la categoría; por ejemplo, “un gorrión es un pájaro” contra “un pingüino es un pájaro”. El significado de una palabra puede ser representado como un conjunto de rasgos semánticos. Los rasgos asociados con una categoría dada varían en el grado de definición de esa categoría, rasgos más definatorios o aspectos más esenciales del significado.

Técnicas de análisis semántico

Redes semánticas y procesamiento de lenguaje natural

Del medio que nos rodea proviene una gran cantidad de conocimiento; al compartirlo, esos saberes se transforman y evolucionan; además, la forma de documentar y comunicar el entendimiento se ha ido mejorando, desde simples vocalizaciones hasta la actual era de la información. La manera en la que se interpreta la información es de suma importancia para el avance de la ciencia, por lo cual las técnicas han ido progresando para optimizar y desenmarañar el proceso por el que se transmiten los pensamientos y las ideas. En apartados anteriores se vio que los diferentes preceptos que conforman el intelecto son parte de un rompecabezas en donde la memoria y en general la mente se encarga de darles un sentido mediante el uso del lenguaje.

La lengua no es función del sujeto hablante, sino el producto que el individuo registra pasivamente, nunca supone premeditación y la reflexión no interviene en ella más que para la actividad de clasificar. El habla es el acto individual de voluntad y de inteligencia, ya que supone composición premeditada haciendo uso de la lengua (Vásquez, 2009). Cuando se evalúa el procesamiento del lenguaje natural, se enfoca la construcción de sistemas que almacenan explícitamente los hechos lingüísticos mediante esquemas de representación del conocimiento. Figueroa y colaboradores (1976), notaron que la forma óptima de recuperar el conocimiento semántico era sin modelos preestablecidos. De esta forma hacen una propuesta de redes semánticas naturales creando un método que da validez de contenido, en otras palabras, el método evalúa eficazmente al concepto la tarea que se pretende investigar.

La importancia del modelo de Collins & Quillian radica en que la recuperación de la información, evocación de un significado, tópico, argumento o tema se realiza mediante la

propagación de la activación, asociaciones semánticas entre conceptos. Esa fue una manera muy interesante de ejemplificar y explicar el funcionamiento del conocimiento en la memoria a largo plazo. Este fenómeno se explicaba por la teoría del *semantic priming*. Desafortunadamente, en las siguientes décadas los investigadores se centrarían en esas características; olvidando por completo que, si la memoria funciona así, es debido a un fenómeno natural y como tal había que estudiarlo; es decir, de frente a individuos de carne y hueso. Las redes semánticas naturales es justamente lo que proponen. Eso pone un punto final a los modelos de memoria semántica computacionales y abre una nueva forma de investigar más empírica y con ventajas más que obvias sobre los modelos anteriores (Andrade, 2017). Sin embargo, de nueva cuenta, el desarrollo de los modelos fue lento, debido a la deficiencia que tenía este modelo de redes semánticas naturales para comprobar la estructura latente, que habían propuesto de manera temprana los primeros investigadores (Quillian y Collins). No es sino hasta que con la ayuda de las súper computadoras actuales y el desarrollo de la teoría de grafos –que se aborda en el apartado de fundamentos del análisis estructural-, retoman el principio estructural y natural, conciliando ambos enfoques, lo que da validez y certeza a muchas de las teorías que hay en torno a la memoria semántica.

El procesamiento de lenguaje natural (PLN) parte de un enfoque estadístico que aplica técnicas matemáticas sobre extensos corpus (documentos de texto) informatizados con el fin de clasificar y estudiar los tópicos y conceptos que hay en una gran cantidad de documentos literarios. Es un campo del conocimiento al que han contribuido disciplinas como la lingüística, la informática, la ciencia cognoscitiva y la ingeniería electrónica, esta última más estrechamente relacionada con las tecnologías del habla (Periñán, 2012). Esto se hace con el fin de la utilización de un lenguaje natural para la comunicación con la computadora, debiendo ésta entender las

oraciones que le sean proporcionadas. El uso de estos lenguajes naturales facilita el desarrollo de programas que realicen tareas relacionadas con el lenguaje o bien, se usa para desarrollar modelos como el modelo Hiperespacio Análogo al Lenguaje (HAL) y el Análisis Semántico Latente (LSA) que ayudan a comprender los mecanismos humanos relacionados con el lenguaje.

Modelo del Hiperespacio Análogo al Lenguaje HAL

El modelo del Hiperespacio Análogo al Lenguaje (HAL), crea una simulación que muestra características de la memoria semántica humana al hacer un análisis de un texto sintetizado por la experiencia del sujeto, los conceptos se representan como puntos en un espacio; la posición de cada punto a lo largo de cada eje está relacionada de alguna manera con el significado de la palabra.

El modelo HAL, que es léxico semántico, utiliza la coocurrencia global de la palabra desde un texto, calculando así la distancia entre las palabras que coocurren en dicho espacio. La distancia entre los puntos correspondientes en el espacio representa la semejanza semántica entre conceptos, esto es útil para examinar las relaciones entre palabras o conceptos porque, las relaciones pueden cuantificarse aplicando distancias métricas a los puntos dentro del espacio.

Operativamente, la metodología que se utiliza involucra seguir las coocurrencias lexicales dentro de una ventana móvil de 10 palabras, que se desliza a través de un corpus del texto. El corpus incluye aproximadamente 300 millones de palabras tomadas de los USENET Newgroups (un sistema de discusión en Internet distribuido mundialmente), que contienen textos en inglés. El vocabulario de HAL consiste en los 70.000 símbolos más frecuentemente usados en ese corpus. Alrededor de la mitad de estos símbolos tiene entradas en el diccionario estándar del sistema operativo Unix, el resto incluye no-palabras, faltas de ortografía, nombres propios y el

argot. La metodología, por lo tanto, produce una matriz de 70.000 por 70.000 valores de coocurrencia (Manzanero, 2006).

Si dos palabras son presentadas sucesivamente, se crea una combinación usada para ser comparada con otras combinaciones que antes fueron almacenadas en la memoria a largo plazo, creando ciertas claves y haciendo que palabras que coocurrieron frecuentemente tengan un vínculo más fuerte que las palabras no relacionadas, lo que aumenta la velocidad de procesamiento causado por la familiaridad. Debido a esto cuando se presentan dos palabras que aparezcan en contextos similares, éstas poseerán vectores similares, al tener una alta similitud semántica. Para calcular los vectores semánticos se establece una matriz de coocurrencia entre las palabras; las palabras más cercanas (con más coocurrencia) obtienen los mayores pesos. Como esta metodología se basa en un corpus de texto, las palabras contiguas reciben un peso de 10, las palabras separadas por una palabra reciben un peso de 9 y así sucesivamente. El significado de una palabra se captura en un vector de 140.000 elementos, obtenido concatenando el vector fila y el vector columna para esa palabra. Cada vector se puede pensar como un punto en un espacio dimensional de 140.000 elementos; la semejanza entre el significado de dos palabras se define como la distancia euclidiana entre sus puntos correspondientes en el espacio. (Lund Kevin, 1996).

HAL es un modelo estructural de explicación del significado, pero no posee una arquitectura particular de proceso, aunque es una metodología capaz de capturar información sobre significados de palabras a través del análisis de texto sin supervisión.

Promete ser una técnica generalmente útil para el análisis de semántica de palabras, además de proporcionar una alternativa a los métodos tradicionales de generar vectores de

características para palabras. Las demostraciones de este modelo se basan en evidencias cualitativas o en las correlaciones entre índices generados por el modelo y la conducta humana.

Modelo de Análisis Semántico Latente LSA

El Análisis Semántico Latente (LSA) extrae y representa el significado del contexto de uso de las palabras a partir de cálculos estadísticos aplicados a un corpus grande de textos; este modelo sostiene que puede emular y explicar cómo los seres humanos aprenden, organizan y posteriormente utilizan el significado de las palabras y el conocimiento en general, presentándolo, en consecuencia como una teoría psicológica del significado y de la representación del conocimiento en la mente humana (Gutiérrez, 2005).

La idea subyacente es que si se dispone de la totalidad de la información sobre todos los contextos en los que una palabra aparece y no aparece, esta información proporciona un sistema de restricciones mutuas que determina en gran medida la semejanza entre los significados de las palabras y de las oraciones (Manzanero, 2006).

Los conceptos en el LSA son representados por vectores en un espacio de aproximadamente 300 dimensiones, algo muy parecido al modelo HAL. El LSA es un modelo espacial de alta dimensión de representación del significado, usando las semejanzas entre los significados de los conceptos. El LSA se diferencia de otros abordajes estadísticos en dos sentidos significativos ya que usan como datos de entrada no sólo las coocurrencias de palabras, sino los patrones detallados de ocurrencias de palabras en una cantidad muy grande de contextos, tales como oraciones o párrafos, tratados como conjuntos unitarios. Además, subraya la dimensionalidad en la cual todas las relaciones locales de palabra-contexto se representan

conjuntamente y asume que es posible reducir las dimensiones de los datos observados desde el número de contextos iniciales a uno mucho más pequeño.

La entrada al LSA es una matriz en la cual las filas representan tipos de acontecimientos y las columnas representan los contextos en los cuales los tipos de acontecimientos ocurren. En muchas aplicaciones, las filas corresponden a tipos de palabras y las columnas corresponden a muestras de texto (por ejemplo, párrafos) en las cuales las palabras aparecen. Cada celda en la matriz contiene el número de veces que un tipo particular de palabra aparece en un contexto particular. Esta matriz se analiza usando descomposición en valores singulares (SVD, por sus siglas en inglés), que es una importante herramienta del álgebra lineal utilizada para la factorización de una matriz rectangular real o de una matriz compleja (Landauer, 2009). En el algoritmo es que un fragmento textual (contexto o documento) puede ser representado como una ecuación lineal, cuyo significado correspondería a la frecuencia en que concurren los significados de las palabras que lo conforman y el significado de las palabras en ese fragmento. Este modelo es fundamentalmente utilizado como una herramienta para evaluar la similitud semántica de piezas de información verbal. En pocas palabras, el LSA es un sistema computacional que permite inducir y hacer patente las relaciones semánticas que piezas textuales pertenecientes a un mismo contexto mantienen entre sí (Gutiérrez, 2005).

Redes semánticas naturales

Cada cabeza es un mundo, expresa el dicho popular, que podría parafrasearse como cada mente es un mundo. Quizás por ello, Figueroa y colaboradores (1976) propusieron una manera de estudiar el universo de los significados evaluando a éstos de una forma natural. La técnica que desarrollaron (Redes Semánticas Naturales, RSN) permitiría comparar a dos o más participantes, de acuerdo con el significado que le asignan los grupos a ciertos conceptos, que sean del interés para alguna investigación. Dice Andrade (2017, p. 51) “La red semántica de un elemento se forma por los conceptos organizados en la memoria. Esta red no se debe exclusivamente a vínculos de asociación; las redes semánticas se dan por un proceso de memoria que organiza. La selección de las abstracciones de una red no se basa en la fuerza de la asociación sino más bien en la clase y propiedades de estas”.

El método para obtener el significado de una palabra es la siguiente: se seleccionan una o más palabras-estímulo de las cuales se quiere saber el significado y se pide al participante que mencione todas las asociaciones que le vengan a la mente. Para que el participante entienda la tarea, se usa, a modo de ejemplo la palabra-estímulo “manzana”. Entonces se pide al entrevistado definir la palabra-estímulo mediante un mínimo de cinco palabras sueltas, pudiendo ser verbos, adverbios, adjetivos, sustantivos, nombres o pronombres, sin utilizar artículos ni proposiciones. Una vez concluida esta tarea, se le pide que las ordene, asignando el número 1 a la palabra que mejor defina el concepto o esté más cerca de él, el 2 a la siguiente palabra en cercanía, el 3 a la palabra que sigue en cercanía y así sucesivamente. En una base de datos se vacían los datos relacionando todas las palabras dichas por los sujetos (ver por ejemplo la Tabla 1). Cada palabra tiene a su derecha diez casillas para registrar las frecuencias y jerarquía con que las personas asignan cada uno de los elementos. Una vez obtenida la información, se debe revisar la base de

datos para corregir errores de captura e integrar los términos proporcionado por los sujetos (palabras definidoras), utilizando un solo término, el más común para los sinónimos. Este paso es necesario para que el conteo de palabras sea preciso.

Tabla 1.

Ejemplo del vaciado de datos de la red semántica de "manzana"

Definidoras	Jerarquía									
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
Roja	3	2		1	1	2		1		1
Fruta	2	2			5					1
Jugo	1	2		1		2		1	1	

Nota. En este ejemplo se puede apreciar el vaciado de datos que se lleva a cabo en las redes semánticas naturales. Las definidoras son los conceptos que dieron los participantes, mientras que la Jerarquía es el conteo de la cantidad de veces que las personas clasificaron cada palabra definidora en cada jerarquía, dependiendo de qué tan cercana la consideraron de la palabra-estímulo.

Después de construir una tabla con todas las palabras definidoras que fueron usadas, se calculan los dos principales valores de la red semántica:

A) El valor J, que es el total de palabras definidoras generadas por los sujetos, fundiendo los términos que son sinónimos, usando el más usual de ellos: es un indicador de la riqueza semántica de la red: a mayor cantidad de palabras, mayor riqueza semántica.

B) El valor M, que indicaría el peso semántico de cada palabra definidora. Para calcularlo (ver Tabla 2), se asigna un valor semántico donde se invierten las escalas de las jerarquías haciendo que del valor 10 a la jerarquía 1, valor de 9 a la jerarquía 2, y así sucesivamente, hasta darle un valor semántico de 1 a la jerarquía 10 (ver Tabla 2). Se multiplica la frecuencia de aparición en cada lugar de la jerarquía por el valor semántico que se le da a esa jerarquía y se suman los resultados obtenidos para las diez posiciones jerárquicas.

Tabla 2.*Ejemplo del valor semántico, valor "M"*

Valor Semántico	10	9	8	7	6	5	4	3	2	1	
Definidoras	Jerarquía										
	1	2	3	4	5	6	7	8	9	10	Valor M
Roja	30	18	0	7	6	10	0	3	0	1	74
Fruta	20	27	8	0	25	0	0	0	0	1	81
Jugo	10	18	0	7	0	10	0	3	2	0	50

Nota. En la tabla puede verse un ejemplo del proceso en la que se establece el valor "M" en las redes semánticas naturales.

Para obtener el índice de cercanía de una palabra definidora con respecto a la palabra-estímulo, se divide el valor M de las palabras definidoras entre el puntaje total de la palabra con el valor M más alto, en este caso se tomará el valor M de Roja (74) y se dividirá con el valor M de Fruta (81), el índice de cercanía para Roja es de 91% ($74/81 = .91$); el mecanismo es el mismo para cualquier otra palabra. Al conjunto de las 10 palabras con los puntajes más altos (puntajes M) le pusieron Conjunto SAM (*Semantic Association Memory*). Los conjuntos SAM son la base para comparaciones y correlaciones posteriores intra e inter grupos; ya que este análisis permite hacer comparaciones por diversas características del grupo (sexo, edad, nivel socio económico entre otros)

Por último, el valor FMG representa la distancia semántica en términos de porcentajes entre las diferentes palabras definidoras que conforman el conjunto SAM. Se obtiene a través de una sencilla regla de tres, partiendo de la palabra definidora con el valor M más alto de la red, misma que representa el 100%. Si se toma la Tabla 2 este valor sería para fruta el 100%, roja 91% y jugo 61%.

De acuerdo con Hinojosa (2008), este método tiene varios puntos en contra, sobre todo desde una perspectiva estadística, como la poca solidez que toman los valores J y M. La riqueza de la red (el valor J) depende del tamaño del grupo de participantes, no se puede saber de qué tamaño debe ser el grupo para que J alcance el valor asintótico; además, basta con que un participante dé respuestas atípicas para que J aumente. La arbitrariedad de asignar un valor semántico de 10 a la jerarquía 1 en el valor M, y el manejar valores ordinales (la jerarquización de las palabras), como si fueran valores intervalares que se pueden multiplicar y sumar.

El conjunto SAM, por su parte, depende del valor arbitrario que se dé a los pesos semánticos. Al definir manzana usando las posiciones jerárquicas, que son datos ordinales, se multiplican por un escalar, el peso semántico, el resultado sigue siendo ordinal, en el mejor de los casos. Ahora bien, si los datos ordinales se suman, el resultado es un número sin sentido. Ordenar las palabras definidoras según su valor M puede ser sólo un reflejo muy deformado de la red de significados del grupo. Por la misma razón, usar técnicas estadísticas como la correlación de Pearson, la prueba t o el análisis de varianza con valores M es un ejercicio con valor estadístico dudoso (Hinojosa, 2008).

Finalmente, los valores M dependen del tamaño del grupo. Si el grupo crece indefinidamente, los valores M crecerán también indefinidamente. Aun salvando las dificultades mencionadas, no es posible comparar los valores M obtenidos en una investigación con los obtenidos en otra, a menos que los grupos sean del mismo tamaño.

Redes semánticas naturales modificadas

Reyes-Lagunes (1993), analizando los problemas o inconvenientes de las redes semánticas propuestas por Figueroa y colaboradores, formula redefiniciones en los conceptos para que éstos realmente expongan el cálculo que se realiza mediante el método; así, propuso un procedimiento para efectuar la captura y análisis usando programas de computadora, obteniendo los valores principales y un gráfico radial que representa la red. Para determinar la cantidad de palabras definidoras que han de conformar el núcleo de la red, genera un gráfico lineal con los pesos semánticos de las definidoras y, en dicha gráfica, ubica el punto de quiebre de Cattell de la función; es decir, el punto o codo en el que se puede ver que los pesos semánticos se aplanan y tienden a volverse asintóticos. De tal manera, es como Reyes-Lagunes aborda técnicamente su crítica a la manera usual de definir el conjunto SAM (seleccionar las diez palabras con mayores pesos semánticos), ya que con esto se desconoce el tamaño real de la red y, por tanto, el de la muestra que lo representa. En efecto, el número arbitrario de diez palabras puede excluir palabras definidoras importantes (podemos añadir que también puede incluir palabras definidoras de poca importancia) y dicha decisión carece de un sustento teórico para la delimitación del conjunto.

Entre las nuevas maneras de definir los conceptos se tiene que el valor J cambia por tamaño de la red (TR), que consiste en el total de palabras definidoras generadas por los sujetos. El valor M se llama ahora peso semántico (PS), que se explicaría como la jerarquía semántica entre las definidoras. El Valor FMG se cambió por distancia semántica cuantitativa (DSC); además, como índice de representatividad toma el nombre de Índice de Consenso Grupal, que puede definirse a partir de los porcentajes. Una posibilidad de dicho índice es el promedio de los porcentajes de las 10 palabras con mayor frecuencia. Finalmente, el conjunto SAM, las 10

palabras más representativas en la red, se pasa a llamar núcleo de la red, tratando que este valor no se calcule de forma arbitraria ni omitiendo información importante o definidoras irrelevantes.

La metodología es muy similar a la anterior; se pide a los participantes que mencionen las definidoras de una palabra estímulo; cada definidora debe ir con la jerarquía en la que el sujeto cree que mejor describe al concepto. Estos dos elementos (definidoras y su jerarquía) se vacían en una tabla (véase tabla 3). Se suma las frecuencias en cada jerarquía de las palabras repetidas, dejando una fila para cada palabra. A continuación, se muestra el proceso de esta técnica, los datos utilizados no pertenecen a ninguna medición y es para un uso exclusivo de ejemplificación.

Tabla 3.

Ejemplo del vaciado de datos para analizar una red semántica natural modificada

Definidora	Jerarquía									
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
Lluvia	4	2	4	6	4	1	2	2	2	1
Sequía	10	3	3	1	5	5	2	4	3	1
Cambio	8	9	0	2	2	4	4	9	9	2
Desastres	9	2	3	6	0	10	2	3	4	5
Enfermedad	5	4	5	0	4	6	4	0	6	5
Paraguas	6	6	0	7	3	0	4	5	0	0
Sol	7	0	8	9	0	2	3	0	5	6
Nube	2	1	2	3	0	4	5	0	5	4
Gota	3	0	2	0	3	4	6	5	7	0
Animales	1	0	8	9	0	3	2	0	0	2

Nota. En la tabla se pueden notar tanto las definidoras como las jerarquías sumadas. Los datos son solo para ejemplificar.

Se ponderan las frecuencias de acuerdo con la jerarquización que hicieron los participantes, se multiplica la primera jerarquía por diez, la segunda por nueve y así sucesivamente, hasta la jerarquía diez, la cual se multiplicará por uno, si hay más de diez jerarquías las siguientes igualmente se multiplican por uno (ver Tabla 4).

Tabla 4.

Ejemplo del proceso para sacar el peso semántico en la técnica redes semánticas naturales modificadas

Definidora	Jerarquía									
	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
Lluvia	40	18	32	42	24	5	8	6	4	1
Sequía	100	27	24	7	30	25	8	12	6	1
Cambio	80	81	0	14	12	20	16	27	18	2
Desastres	90	18	24	42	0	50	8	9	8	5
Enfermedad	50	36	40	0	24	30	16	0	12	5
Paraguas	60	54	0	49	18	0	16	15	0	0
Sol	70	0	64	63	0	10	12	0	10	6
Nube	20	9	16	21	0	20	20	0	10	4
Gota	30	0	16	0	18	20	24	15	14	0
Animales	10	0	64	63	0	15	8	0	0	2

Nota. Cada frecuencia se multiplica por la jerarquía, se multiplica por diez si el concepto está en primer lugar, por nueve en segundo y así sucesivamente, hasta llegar al décimo (o mayores) en ese caso se multiplica por uno.

Para obtener el peso semántico se suman las frecuencias ponderadas de cada definidora y se ordenan del mayor al menor de acuerdo con este valor (ver Tabla 5), Con esto se puede obtener un bosquejo de las palabras más importantes de la red.

Tabla 5.*Peso semántico de las definidoras*

Definidora	Peso semántico
Cambio	270
Desastres	254
Sequía	240
Sol	235
Enfermedad	213
Paraguas	212
Lluvia	180
Animales	162
Gota	137
Nube	120

Nota. El peso semántico esta ordenado de mayor a menor con base en los cálculos antes mencionados.

Teniendo el peso semántico de todas las definidoras es posible hacer una gráfica lineal (ver Figura 5) para identificar el punto de quiebre de la recta, observando así la definidora donde los pesos semánticos se aplanan y tienden a volverse asintóticos, conformando un codo; en el diagrama vemos que las definiciones se empiezan a aplanar en la palabra “moscas”; en ese caso podrían decirse que hasta allí abarca el núcleo de la red.

Figura 5.

Gráfica lineal de pesos semánticos de las definidoras para obtener el núcleo de la red, usando el punto de quiebre de Cattell, de acuerdo con el método propuesto por Reyes-Lagunes.



Nota. El peso semántico está ordenado de mayor a menor con base en los cálculos antes mencionados.

Para establecer la Distancia Semántica Cuantitativa (valor FMG de Figueroa) se multiplica cada peso por cien y se divide por el peso más alto, se calcula un porcentaje que indica la jerarquía de las definidoras conforme a la importancia en la red, donde el cien por ciento es para el peso más grande y de allí bajando (ver Tabla 7).

Tabla 7.

Definidoras con sus distancias semánticas cuantitativas.

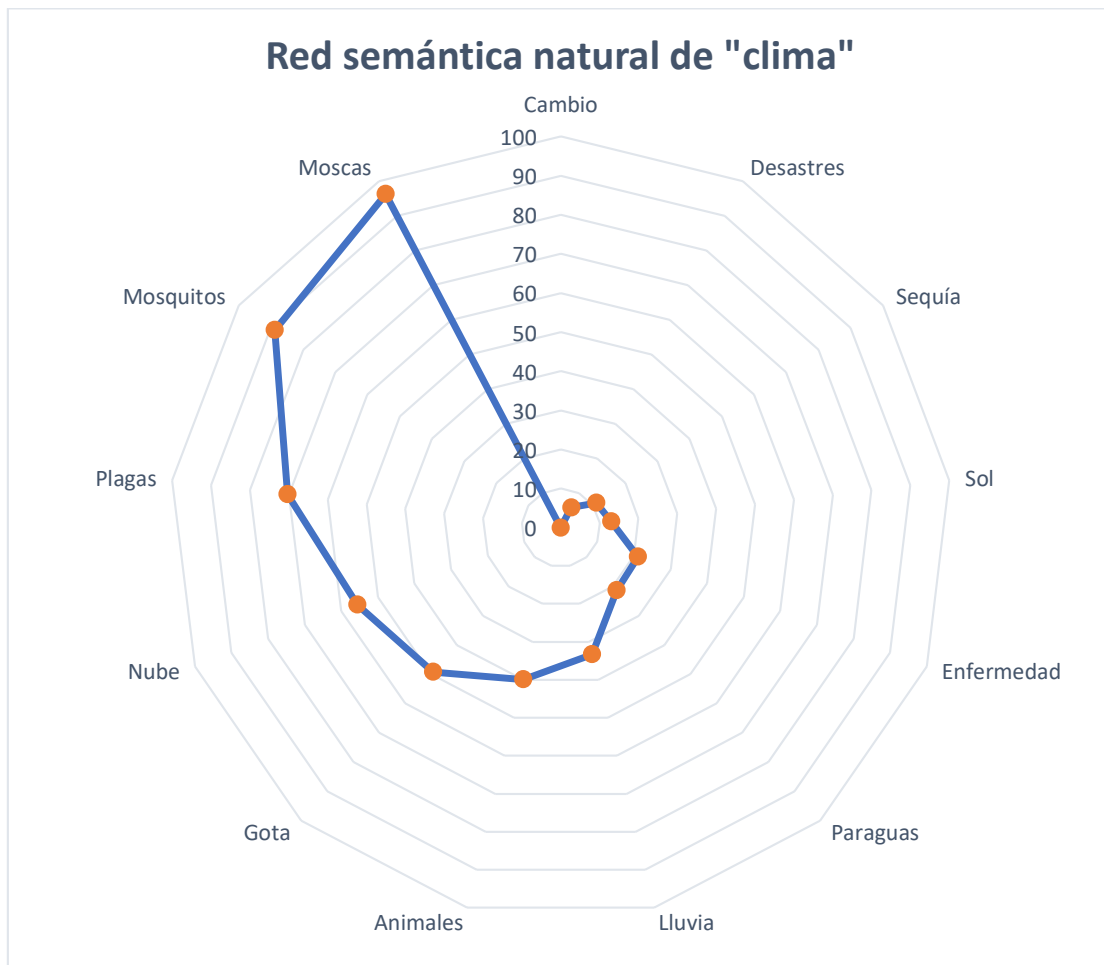
Definidora	Peso semántico (porcentaje)
Cambio	100%
Desastres	94%
Sequía	89%
Sol	87%
Enfermedad	79%
Paraguas	79%
Lluvia	67%
Animales	60%
Gota	51%
Nube	44%
Plagas	33%
Mosquitos	22%
Moscas	10%

Nota. Las distancias semánticas cuantitativas (DSC) peso semántico está ordenado de mayor a menor con base en los cálculos antes mencionados.

Finalmente, los porcentajes de la distancia semántica cuantitativa (DSC) se invierten los valores y los datos se presentan en gráfico de radial, así el grafico la definidora con mayor peso semántico queda en el centro de la misma (ver Figura 6).

Figura 6.

Red semántica natural de la palabra estímulo CLIMA.



Nota. Con el gráfico de radar se tiene una manera visual de cómo las definidoras interactúan con las demás, siendo éste el producto final de las redes semánticas naturales modificadas.

Redes semánticas con escalamiento multidimensional

Mercado, López, & Velasco (Mercado L. E., 2015) se plantean de forma muy directa el problema de estructura en las redes semánticas naturales. ¿Cómo es posible que si el conocimiento semántico tiene una estructura no se pueda apreciar con el modelo de redes semánticas naturales? En palabras de ellos mismos: “Si bien esto proporciona una estructura de relaciones, en la propuesta de Figueroa y colaboradores (1981), son inherentemente lineales. Se puede suponer que existe una estructura de relaciones más rica en una red semántica, que requiere ser rescatada.”

De esta forma, sin hacer grandes cambios al modelo, deciden utilizar el análisis conocido como escalamiento multidimensional, herramienta matemática que permite representar espacialmente, mediante una gráfica, las semejanzas entre los objetos (conceptos), Mercado y cols., señalan que la utilización de esta aproximación refleja las relaciones entre conceptos y sus definidoras, dando una visión acerca de sus propiedades y estructura. Desafortunadamente, esta medida sólo transforma esa medida lineal (las distancias entre pares de definidoras) en una matriz que bidimensiona las medidas, pero no resuelve el problema básico de estructura (Andrade, 2017). Así, por ejemplo, dos palabras o conceptos definidores pueden aparecer muy juntos dentro del análisis de escalamiento multidimensional, pero estar débilmente conectados entre sí. Lo que es peor aún, pueden aparecer separados y pertenecer a un mismo subconcepto, lo que conlleva a un problema de validez frontal, donde el contenido de una escala no cubre de manera adecuada cierto dominio subjetivamente hablando.

El procedimiento es el mismo que en las anteriores metodologías. A los participantes se les pide dar definidoras de una palabra estímulo, estas son jerarquizadas usualmente del 1 al 10,

según sea la definidora que mejor describa al concepto (de acuerdo con el juicio de los participantes). Se suman las frecuencias en cada jerarquía de las definidoras repetidas, se ponderan las frecuencias de acuerdo con la jerarquización, multiplicando la primera jerarquía por diez, la segunda por nueve y así sucesivamente, se suman las frecuencias ponderadas de cada definidora y se obtiene el peso semántico, hasta tener el núcleo de la red (10 principales definidoras).

Posteriormente se hacen combinaciones todos los pares de palabras que se puedan formar mediante el núcleo de la red y las definidoras; para apoyarse en esto es necesario hacer una matriz triangular (ver Tabla 8). Al hacer esto nos permite construir un instrumento donde se les pregunte a los participantes (misma o diferente muestra) que indiquen en una escala del 0 al 10 qué tan relacionados están los pares de palabras, siendo 0 muy relacionado y 10 siendo conceptos muy diferentes.

Tabla 8.

Matriz triangular de la red semántica mueble

Nombres	Mueble	Silla	Mesa	Cama	Grande	Madera
Mueble						
Silla	Silla-Mueble					
Mesa	Mesa-Mueble	Mesa-Silla				
Cama	Cama-Mueble	Cama-Silla	Cama-Mesa			
Grande	Grande-Mueble	Grande-Silla	Grande-Mesa	Grande-Cama		
Madera	Madera-Mueble	Madera-Silla	Madera-Mesa	Madera-Cama	Madera-Grande	

Nota. Ejemplo tomado de Mercado et al (2015), Todos los pares de palabras resultantes en la matriz se les pregunta a los participantes, obteniendo así las distancias de similitud entre ellas.

Al terminar el anterior procedimiento, como resultado mostrará las distancias que los participantes consideraron para cada par de palabras, se vacían los datos en una matriz promedio de las distancias que mencionaron los participantes de forma individual (ver Tabla 9.). Por último, la matriz resultante se somete a un análisis multidimensional que despliega un mapa perceptual de dos dimensiones, en el cual se aprecian las relaciones de cercanía que hay entre los conceptos (ver Figura 7).

Tabla 9.

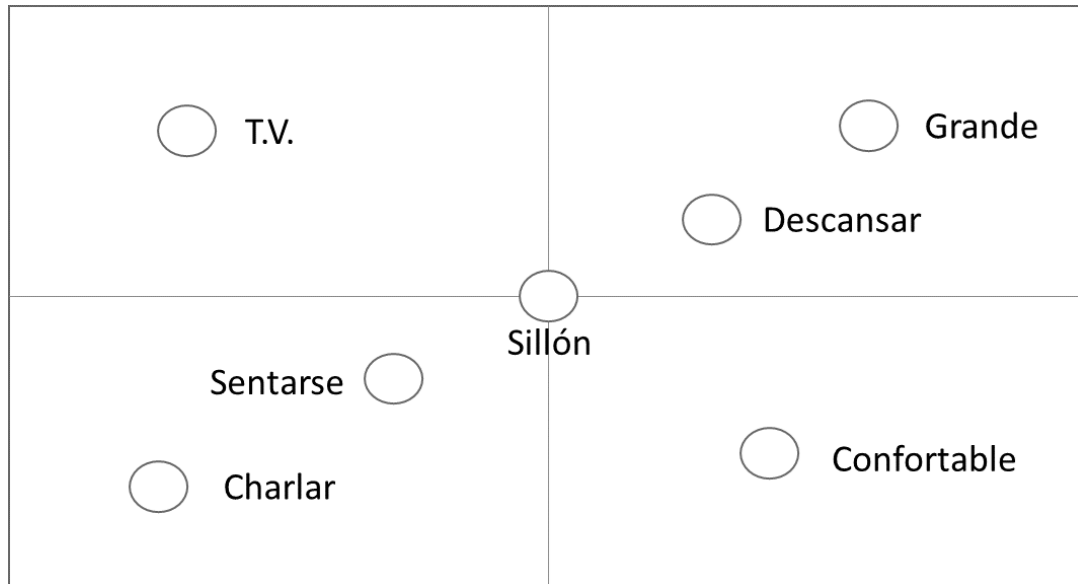
Matriz con las distancias promedio de la red semántica para la palabra estímulo mueble

Nombres	Silla	Mueble	Cama	Grande	Mesa
Silla					
Mueble	5.15625				
Cama	5.90625	5.9375			
Grande	4.40625	5.8125	4.59375		
Mesa	5.9375	5.09375	4.0625	3.65625	

Nota. Ejemplo donde se puede notar el cómo se vacían los datos para así formar la matriz promedio de las distancias. La matriz debe contemplar todas las combinaciones antes hechas (pares de palabras), como este caso solo es para ejemplificar se acortaron.

Figura 7.

Ejemplo de una gráfica de red semántica natural con escalamiento multidimensional



Si se quisiera definir “Sillón” con esta técnica se puede apreciar que las distintas definidoras se separan en diferentes cuadrantes (dimensiones); aquellas que se encuentren más cercanas a la palabra-estímulo que se encuentra justo en el medio del plano cartesiano (centroide) son aquellas que son importantes para su definición. Igualmente, si se desea ver los agrupamientos entre los elementos, se requiere un análisis de apoyo, en este caso un *Cluster Analysis* (también llamado Análisis de grupos), se marca con un círculo en la gráfica las definidoras agrupadas en el análisis.

En pocas palabras, los datos que se utilizan en el análisis son las distancias entre pares de palabras que con el escalamiento multidimensional obtienen una representación en un espacio cartesiano sobre los grados de similitud, pudiendo dar nuevas formas para interpretar las redes semánticas ya que, de acuerdo con Mercado y cols., este enfoque permite aplicar una estructura a los conceptos y las definidoras.

Análisis estructural de redes semánticas naturales

Un grafo es el conjunto de objetos llamados vértices o nodos unidos por enlaces llamados aristas o arcos, que permiten representar relaciones binarias entre los elementos de la red. La teoría de Grafos es una disciplina de las matemáticas discretas, su análisis estructural se utiliza para referirse a propiedades de un grafo, aporta ideas y métodos sobre cómo pueden ser cuantificadas estas propiedades; es decir, ofrece una practicidad al modelo. En el caso de las redes semánticas naturales las aristas representan los conceptos; sin embargo, igualmente se utilizan para referirse a conocimiento, relación entre individuos o proximidad geográfica. Al representar a los individuos (nodos) y sus vínculos (aristas) mediante un grafo, la interpretación e importancia radica en cómo el sendero (red) sigue la influencia entre variables (nodos). El desentrañar la distribución y comunicación de la red se utiliza comúnmente en el análisis de redes sociales e igualmente facilita la optimización de recorridos, procesos, flujos, algoritmos de búsquedas, entre otros, ya que, al conocer la interacción entre sus elementos, se tiene la posibilidad de manipular algunas de sus piezas.

El problema de estructura de las redes semánticas naturales es el de buscar una forma de develar las relaciones entre las palabras que definen a un concepto. Torres y Garza (2014), vislumbran una solución muy cercana a lo que se usa hoy en día, pero sin llegar a echar mano totalmente de la teoría de Grafos. Por otro lado, Echeverría y colaboradores (2008) desarrollaron un software que da con la solución del tan buscado problema estructural y el cual utiliza la teoría de Grafos. El valor de esta iniciativa no es el programa, pues en ese mismo momento se venían desarrollando múltiples plataformas para manipular datos que manejan relaciones (aristas) del tipo de las redes sociales (por ejemplo, NodeXL, Gephi y demás), sino la solución planteada al

problema de representación de las relaciones entre conceptos. Casi paralelamente, hay evidencia de que ya se venía usando comercialmente la misma solución al problema de estructura de las redes semánticas naturales.

La teoría de Grafos y el desarrollo de programas de computadora para análisis de redes sociales dan un nuevo enfoque a las redes semánticas naturales, permitiendo una visualización de la estructura. Con los nuevos avances en las tecnologías de la información y, en particular, del procesamiento del lenguaje natural y sus métodos, las redes semánticas naturales se mimetizan con el análisis de redes de texto; aunque siguen conservando esa aura con la que nacieron. A continuación, una breve descripción de los fundamentos del análisis estructural que permiten conocer cómo es que se utilizan para analizar las redes semánticas naturales.

Métricas del análisis

Freeman (2000), propone diversas mediciones para cuantificar las propiedades de la estructura en un grafo, dependiendo de para que se utilice la red pueden utilizarse diferentes medidas; en este apartado se hablarán de las medidas de tendencia central: grado, intermediación y cercanía.

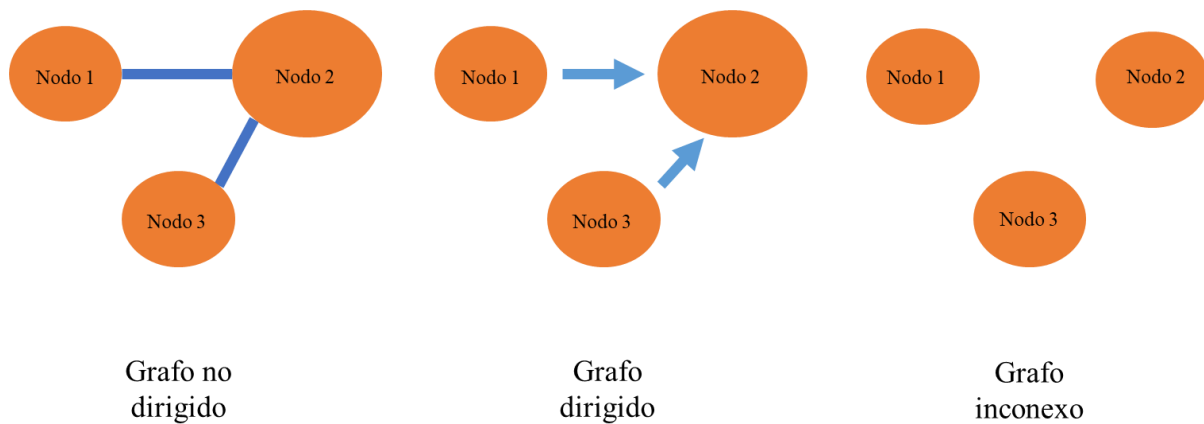
Una red, según Newman (2003), es un conjunto de elementos llamados vértices o nodos con conexiones entre ellos, llamado bordes o aristas. Si no existiera ninguna relación o interacción entre los conceptos se dice que es un grafo inconexo, en el caso de que sea un grafo conexo (los nodos cuentan con vértices) el camino más corto de un nodo a otro se le llama geodésica.

Los nodos pueden ser de diferente tamaño; los que tienen más conexiones son más grandes que los que tienen pocas conexiones, lo cual se deriva del hecho de que los nodos mejor

conectados son más importantes para el desarrollo de la red que los menos conectados. La relación que hay entre dos nodos puede ser recíproca o unilateral; si en la relación la influencia de los nodos es mutua se dice que es no dirigida y la arista simplemente une a los nodos sin indicar dirección; si la relación va en un sentido se dice que es dirigida y la dirección se indica con una flecha (ver Figura 8).

Figura 8.

Ejemplificación de los tipos de grafos



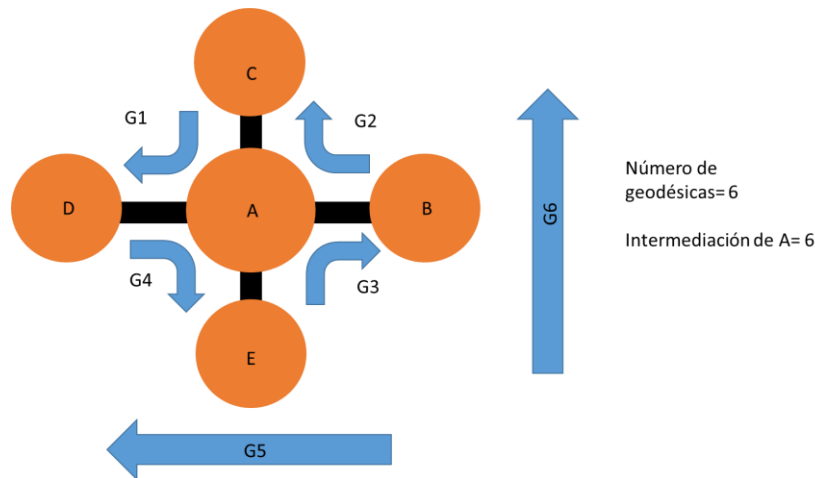
El grado de un nodo se mide con base en la suma de nodos con los que se conecta directamente; es decir, al número de nodos adyacentes a él o de aristas que tiene. En otras palabras, es la frecuencia; sin embargo, debe también equilibrarse con la fuerza del vínculo. Cuántas más elementos estén asociados con un nodo determinado, más fuerte será cualquier asociación. (Berger, 2014).

Cuando un nodo es mediador entre la geodésica de otros dos nodos, el primero se vuelve estratégico para la conexión o relación de ese par de nodos. Por ello es importante que se considere la fuerza con la que lo hace; es decir, el peso que tienen las aristas (número de veces en

las que interaccionan los elementos) que unen a los vértices. El poder de intermediación que tiene un nodo en una red se calcula contando el total de geodésicas que hay entre todos los pares de nodos y el número de veces en que ese nodo se interpone entre ellas. El promedio de las distancias geodésicas de un nodo al resto de los otros nodos del grafo es igual a su cercanía. Un nodo goza de una gran independencia si es muy cercano a todos los demás; en otras palabras, no necesita de otros para establecer una relación o comunicación con cualquier miembro de la red (Figura 9).

Figura 9.

Ejemplificación de las geodésicas y de intermediación en un grafo



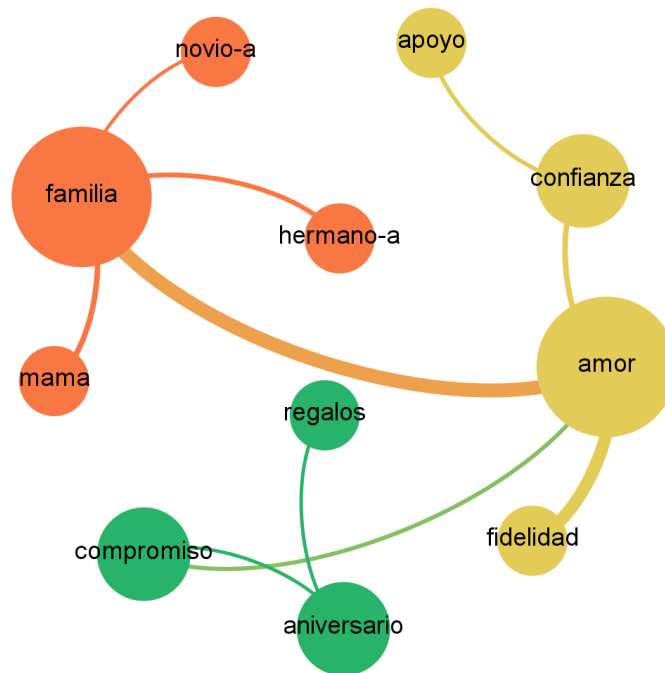
Nota. Las geodésicas están representadas con las flechas, en este ejemplo se muestran seis geodésicas.

Para obtener una estructura en las redes semánticas naturales es necesario construir una matriz de palabras de las definidoras. La metodología es casi la misma que con las redes semánticas naturales con escalamiento dimensional, con la salvedad de que no es necesario que el sujeto proporcione una jerarquización a las definidoras. Al contar con el listado de los pares de palabras (relaciones entre definidoras) se analiza en diferentes programas (como Ghephi,

NodeXL, entre otros) para obtener el grafo. Con el análisis estructural es posible notar las relaciones entre definidoras, el crecimiento, la formación de *clusters* y más mediciones que se necesiten (ver Figura 10).

Figura 10.

Ejemplo de una red semántica natural obtenida con análisis estructural, con la palabra estímulo “pareja”



Nota. El gráfico fue realizado con el programa Gephi. Los colores de cada nodo los proporciona el programa al detectar los grupos entre definidoras.

Análisis comparativos entre aproximaciones

“La escritura y la palabra pueden transformar el sufrimiento” dice Cyrulnik (2020). Las redes semánticas se forman con las palabras recuperadas en la memoria mediante un proceso reconstructivo que no sólo se obtiene por vínculos asociativos, sino también por la naturaleza de los procesos de memoria. El psicólogo, científico, publicista, entre otros, activan su memoria en tan sólo instantes que le permiten tener una mayor profundidad de la información que analiza.

Esta elaboración del pensamiento difícilmente se puede lograr analizando una red semántica natural clásica, debido a lo que bien decían Mercado y colaboradores (2015), no hay estructura en una red donde no se habla de la relación entre sus definidoras. Al construir una matriz simétrica entre los conceptos se puede determinar cómo se definen los conceptos entre ellos, descubriendo así una estructura de conceptos definidores y definidos.

Andrade (2017) menciona que hay pocas diferencias estadísticas significativas entre los diferentes métodos al elaborar las redes semánticas; sin embargo, cualitativamente se pueden notar diferencias que pueden aportar lo mismo en situaciones de investigación académica que en aplicaciones, por ejemplo, de análisis de mercado.

Como se aprecia en la Tabla 10, las diferentes técnicas para analizar una red semántica natural ofrecen casi las mismas características; el investigador puede optar por una u por otra dependiendo de si necesita alguna particularidad, por ejemplo, para hacer alguna división o agrupación entre los datos el escalamiento multidimensional o el análisis estructural puede ayudar. En cuanto al aspecto gráfico, cada uno tiene su forma de exponer los elementos en la red, logrando que sean más llamativas o fáciles de entender. Sin embargo, el análisis estructural ayuda a reconocer la distribución en los enlaces y grupos de subconceptos, proporcionando rutas

a la creatividad, mostrando la relación que hay entre los elementos, permitiendo construir más amplias interpretaciones de éstos. Para ejemplificar se realizó una medición con 35 participantes en la CDMX, utilizando la palabra estímulo “vivienda”, los datos que se muestran a continuación son el resultado de dicho ejercicio, los mismos datos fueron utilizados para las distintas técnicas.

Tabla 10

Diferencias cualitativas entre técnicas de redes semánticas naturales

Tipo de Red	RSN	RSN Modificada	RSN Escalamiento Multidimensional	Análisis Estructural de RSN
Atributos				
Riqueza de la red	Valor J	Tamaño de red (TM)	Tamaño de red (TM)	Numero de conceptos (vértices)
Definidoras principales	Conjunto SAM	Núcleo de la red	Núcleo de la red	Palabras con mayor grado
Peso semántico	Valor M	Peso semántico (PS)	Peso Semántico	Grado y grado ponderado
Distancia semántica	Valor FMG	Distancia semántica cuantitativa (DSC)	Relativo su posición de acuerdo al centroide en la grafica	Distancia Geodésica
Representación gráfica	No	Gráfico de radar	Gráfico Cartesiano	Grafo
Segmentar y dimensionar	No	No	Cluster Analysis (Análisis por aparte, como apoyo)	Modularidad

Los resultados finales de cada una de estas técnicas difieren en su modo de presentación. En las redes semánticas naturales se presentan en forma de tabla con los datos pertinentes, valor J, conjunto SAM, valor M y Valor FMG, tal y como se muestra en la tabla 11. Con esta técnica se pueden notar los elementos que conforman el concepto estímulo “vivienda”; sin embargo, no muestra las interacciones entre las definidoras, el modo de obtener los datos, de manera

“natural”, hace que nos muestre con mayor exactitud la estructura de la memoria semántica de los participantes.

Tabla 11.

Presentación de los datos en las redes semánticas naturales

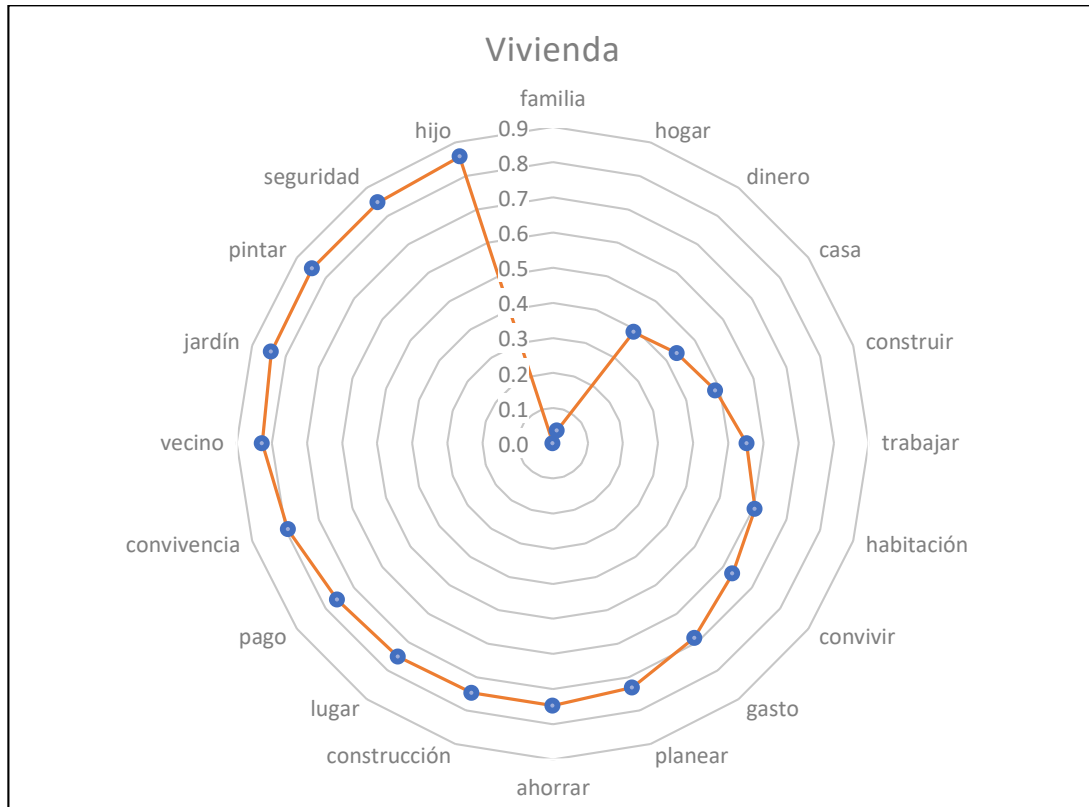
Conjunto SAM	Valor M	Valor FMG	Valor J
			134
familia	261	0	palabras
hogar	251	0.04	
dinero	159	0.39	
casa	147	0.44	
construir	134	0.49	
trabajar	117	0.55	
habitación	103	0.61	
convivir	96	0.63	
gasto	82	0.69	
planear	70	0.73	
ahorrar	66	0.75	
construcción	66	0.75	
lugar	65	0.75	
pago	63	0.76	
convivencia	54	0.79	
vecino	45	0.83	
jardín	41	0.84	
pintar	40	0.85	
seguridad	40	0.85	
hijo	40	0.85	

La técnica de redes semánticas naturales modificadas hace su presentación de una manera gráfica y el renombramiento de los términos ayuda a la comprensión de los elementos (ver en Figura 11.). Al tener el gráfico, las distancias semánticas se localizan de manera intuitiva y con ello se aprecia que el concepto que se acerca al centro del radar es la definidora más cercana a la palabra estímulo. Por ejemplo, en la Figura 11, la palabra familia es la que mejor describe a la

palabra estímulo “vivienda” y, por el contrario, hijo es la definidora con menos similitud de “vivienda”.

Figura 11.

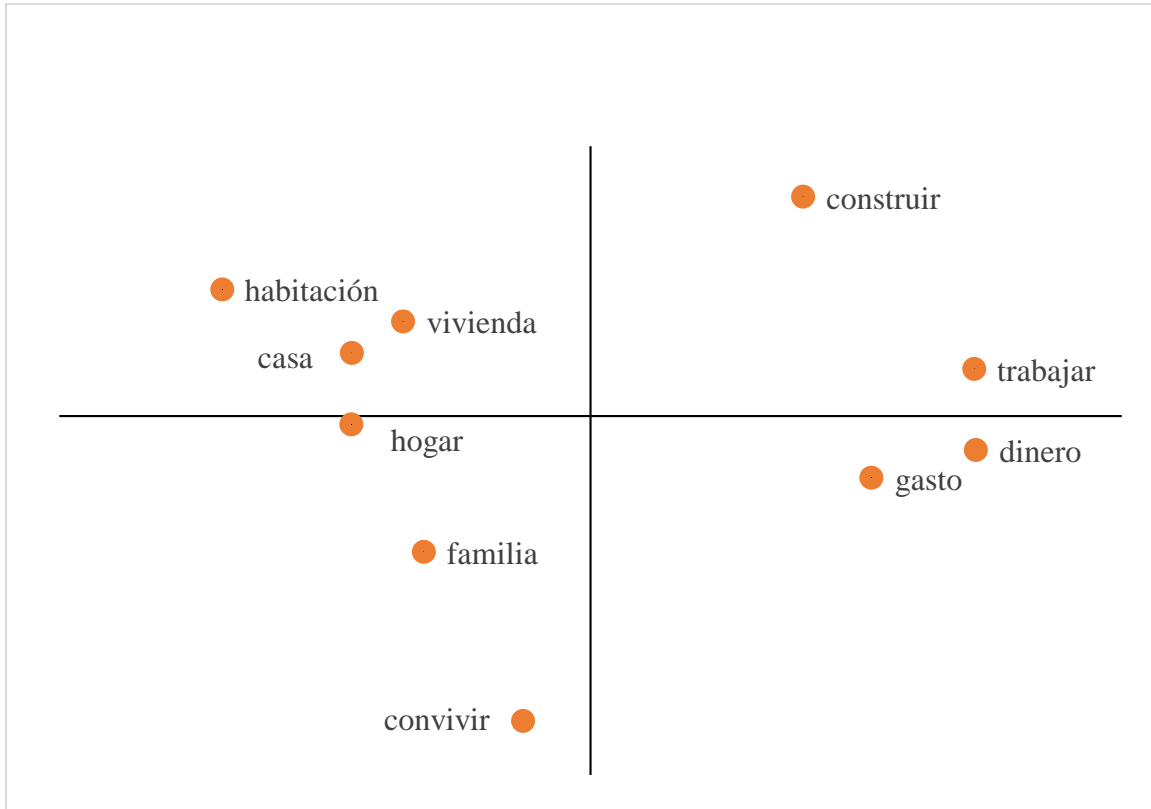
Red semántica natural modificada del concepto “vivienda”



Si se aplica el escalamiento multidimensional, se podrán ver diversas interacciones entre las definidoras; las que se encuentren más cercanas a la palabra estímulo en este caso vivienda (ver Figura 12), son las que mayor grado de similitud presentan con ésta; igualmente, pueden verse de manera sencilla las agrupaciones que se van formando entre ellas; mediante el plano cartesiano se identifican algunas dimensiones (ejes).

Figura 12.

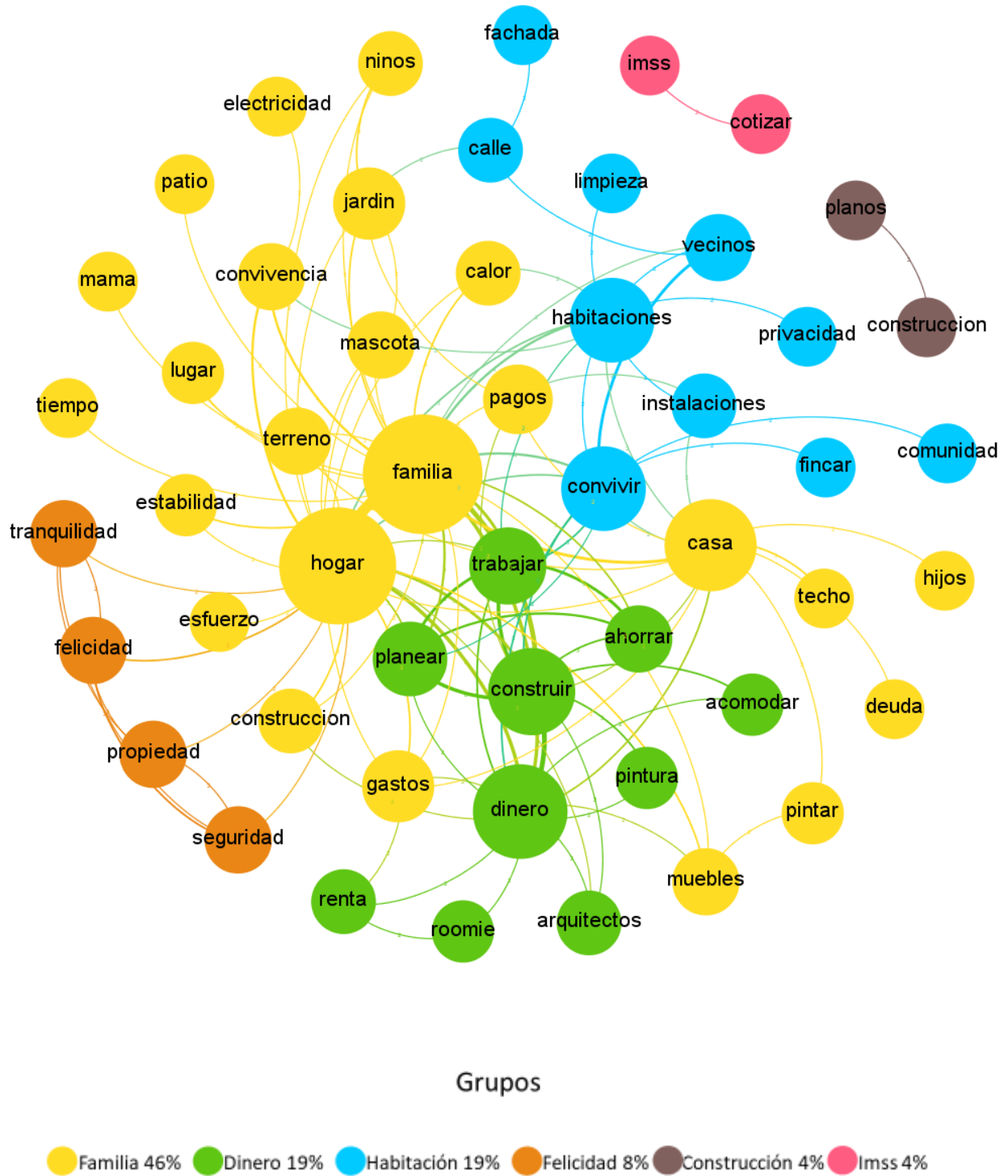
Red semántica natural con escalamiento multidimensional del concepto “vivienda”



Cuando se aplica el análisis estructural a las redes semánticas naturales es posible ver las interacciones que tienen los elementos (definidoras) entre sí. Igualmente, la propia técnica separa por grupos, permitiendo destacar la manera en cómo y con ¿cuáles? interactúan. En la Figura 13. el concepto vivienda está estrechamente relacionado con el grupo “familia”, lo que significa que los participantes conciben como una vivienda a un lugar donde pueda estar su familia, donde se pueda lograr una convivencia entre los integrantes (mascotas, hijos, por ejemplo). En menor medida, la vivienda es un sitio físico y por lo tanto requiere de manutención, algo que se logra con “dinero”, el segundo grupo con mayor relación con la palabra estímulo.

Figura 13.

Análisis Estructural de la red semántica natural del concepto “vivienda”



El grafo da la posibilidad de explorar los diferentes tipos de relaciones entre conceptos y su modularidad, el análisis expresa de una manera sencilla los diferentes grupos que lo conforman; igualmente, se obtiene el porcentaje de cuántos nodos conforman la red. En este ejemplo notamos que conceptos se encuentran mayormente relacionados con el estímulo “vivienda” revelando que están estrechamente enlazados en un contexto de familia, 46% de los nodos pertenecen a este grupo. El crear un hogar requiere esfuerzo, gastos y pagos, ya sea para nuestras mascotas o la construcción, sin embargo, es para llegar a una finalidad, la convivencia ya sea en el jardín o patio. Lo anterior mencionado es una muestra de lo que se puede inferir al someter las partes del grafo, las interacciones entre los nodos crean una historia donde la creatividad del investigador fluye, trayendo información valiosa a los estudios de conceptualizaciones.

Discusión y conclusiones

La memoria es un factor importante en la vida, nos permite saber quiénes somos y como nos relacionamos con otros y nuestro medio, la manera en que se obtiene y maneja la información es un tema de suma importancia que ha sido tratado desde los tiempos de grandes filósofos hasta nuestros días. En la actualidad se aplican diversos modelos para emular el *semantic priming* llevándolo a un sistema computacional, haciendo posible grandes avances en la creación de contenido, pronóstico de tendencias e inteligencia artificial.

Las RSN nos permite conocer el proceso por medio del cual los individuos manejan la información y la formación de conceptos, de acuerdo con sus experiencias y relaciones en su medio. Se trata de una serie de técnicas para el análisis de conceptos, situaciones, objetos o imágenes en específico. Cabe recordar que los conceptos se ven afectados por el contexto y la temporalidad; por aspectos como las condiciones sociales y culturales de cada grupo de estudio, lo que puede implicar la adquisición de cierta información o los cambios sufridos en las relaciones de los elementos de la red, sirviendo como un auxiliar en el estudio de ciertas conductas, llevan a la intervención, prevención y estudio de ciertos fenómenos psicológicos, sociológicos, educativos, mercadológicos, entre otros. Las redes semánticas naturales se han utilizado en el ámbito académico para medir múltiples conceptos y comparar con las ideas y representaciones que las personas construyen acerca de cualquier objeto, tema o situación de su cotidianidad, que son expresados también en términos coloquiales. Su importancia radica en que mediante el lenguaje es posible conocer y comprender muchos de los procesos mentales, puesto que existe una íntima relación del lenguaje, tanto con el pensamiento como con la percepción y el aprendizaje (Chomsky, 1977). De este modo, la técnica de las redes semánticas constituye un recurso empírico de acceso a la organización cognoscitiva de los significados, analizando los

significantes empleados y considerando que el universo cognoscitivo que de ahí resulta es fruto también de aprendizajes sociales.

El presente trabajo muestra las diferentes aproximaciones de un mismo análisis, con esto se puede elegir entre diferentes opciones, dependiendo de la perspectiva en la investigación. Si bien, no hay diferencias perceptibles al examinar las principales definidoras en una red, la forma de presentación y la estructura, fomentan a incluir el análisis estructural, debido a los diversos elementos ofrecidos, cuya utilidad es dado por la creatividad del investigador.

La metodología de las redes semánticas se usa no como método único, sino como acompañante de todo un esquema de herramientas. La técnica brinda una rica posibilidad de comprensión fenomenológica en específico. El análisis estructural es la respuesta directa a tan codiciada estructura y déficit de potencia en el análisis.

Los avances en la ciencia y en las tecnologías permiten ver y analizar diversas cuestiones (por ejemplo, las neurociencias, la inteligencia artificial) que hace unas cuantas décadas eran muy difíciles de conocer. Es muy improbable que los investigadores pioneros de la memoria semántica hubieran imaginado soluciones como las que existen actualmente; y aunque lo hubieran imaginado, se habrían encontrado con el problema de la falta de medios para instrumentar dichas soluciones. El procesamiento de lenguaje natural dispone de un arsenal de herramientas que permite al investigador trabajar no sólo con redes semánticas naturales, sino con múltiples manifestaciones del lenguaje humano. Las soluciones van desde el análisis de redes de texto (pariente cercano del análisis estructural de redes semánticas naturales) hasta el análisis de sentimientos, con la ayuda del aprendizaje de máquina. Igualmente se encuentran múltiples algoritmos de análisis, como el Análisis Semántico Latente (LSA), esbozado aquí,

ideado por Landauer & Dumais (2009). La técnica de redes semánticas naturales es muy importante como herramienta de análisis, pero quizá sea igual o más importante desde el punto de vista psicológico y pedagógico, pues representa un primer atisbo a esa área tan vasta del procesamiento de lenguaje natural y expande los horizontes de una psicología cognoscitiva mediante análisis más sofisticados de manifestaciones y representaciones que nos brinda el lenguaje natural en un contexto específico.

Las redes semánticas naturales son representadas efectivamente usando el análisis estructural, pero seguramente irán evolucionando junto con la teoría de grafos, el avance de la tecnología (sistemas de cómputo) y de los avances en la psicología cognoscitiva sobre los procesos de memoria. Independientemente de lo fascinante que resultan los avances en todos los campos del quehacer científico, se requiere de un esfuerzo extra de los investigadores para actualizarse y aprovechar las nuevas tecnologías, así como los avances teóricos y metodológicos.

En pocas palabras, desde el planteamiento de Quillian y Collins (1969), donde el proceso de jerarquización era artificial y llevaba al “efecto típico”, se ha evolucionado hasta llegar a las redes semánticas naturales, donde cambió la forma de ordenamiento haciendo que fueran los mismos participantes quienes dieran sentido y magnitud a las definidoras, pasando por varias correcciones como dar una interpretación gráfica, combinarlo con técnicas estadísticas como escalamiento multidimensional o análisis estructural, lo que ha llevado a una facilitación en la interpretación de los datos, llevando a la actualización y el refinamiento de las técnicas, creando representaciones del conocimiento que son usadas actualmente para ayudar a los ordenadores a interpretar de mejor manera los datos, por ejemplo inteligencia artificial y procesamiento de lenguaje. Cabe aclarar que el análisis estructural fue posible gracias al avance del cómputo, ya que permite el analizar datos de manera rápida y eficiente, esto ha llevado a mostrar la conexión

entre los conceptos en la memoria humana, la creación de significados psicológicos y ofrece una muestra al proceso del pensamiento. Dependiendo del tratamiento estadístico del Grafo, obtendríamos herramientas de predicción, indicadores para un diagnóstico o herramientas para una intervención. Ya sea mostrando la configuración de conceptos relacionados, modificando o ampliando un significado psicológico, hay una gran ventana de oportunidades al incluir las RSN en diversas investigaciones psicológicas, educativas, mercadológicas, sociales, políticas u otras donde la utilidad es solo marcada por la inventiva de los científicos.

Referencias

- Aguado, L. A. (2001). Aprendizaje y memoria. *Revista de neurología*, 32(4), 373-381.
- Andrade, R. J. (2017). *MARCAS, MEMORIA Y SIGNIFICADO: Análisis de Estructuras Semánticas*. Ciudad de México.
- Bartlett, F. C. (1916). Transformations arising from repeated representation: A contribution towards an experimental study of the process of conventionalisation. . *Fellowship Dissertation*.
- Berger, J. (2014). *Contagio: el poder del boca a boca en la comunicación viral* (1ra ed.). Buenos Aires, Argentina: Temas Grupo Editorial.
- Caramazza, A. &. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of cognitive neuroscience*, 1-34.
- Carrillo-Mora, P. (2010). Sistemas de memoria: reseña histórica, clasificación y conceptos actuales. Primera parte: Historia, taxonomía de la memoria, sistemas de memoria de largo plazo: la memoria semántica. *Salud mental*, 1(33), 85-93.
- Cerdán, L. L. (2011). La memoria en el proceso de enseñanza/aprendizaje. *Pedagogía magna*(11), 311-319.
- Chomsky, N. C. (1977). On wh-movement.
- Collins, A. M. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, 8(2), 240-247.
- Cyrulnik, B. (2020). *Escribí soles de noche: Literatura y resiliencia*. Editorial Gedisa.
- Descartes, R. (1637). *Discurso del método*. Ciudad Autónoma de Buenos Aires: Colihue.
- Echeverría, M. S., R., V., Urzua, P., & Ferreira, R. (2008). DispoGrafo: una nueva herramienta computacional para el análisis de relaciones semánticas en el lexico disponible. *RLA, Revista de Lingüística Teórica y Aplicada*, 81-91.
- Figuroa, J., González, E., & Solís, V. (1976). An Approach to the Problem of Meaning: Semantic Networks. *Journal of Psycholinguistic Research*, 5(2), 107-115.
- Forte, V. (2005). *Mapas conceptuales. La gestión del conocimiento en la didáctica*. México: Alfaomega.

- Freeman, L. (2000). La Centralidad en las Redes Sociales. Clarificación Conceptual. *Política y Sociedad*, 33, 131-148.
- Freeman, L. (2012). *El Desarrollo del Análisis de Redes Sociales: Un Estudio de la Sociología* (1ra ed.). Bloomington, Indiana, United States of America: Palibrio.
- Goldstein, E. B. (2015). *Cognitive Psychology. Connecting Mind, Research and Everyday Experience* (4th edition ed.). Stamford, Connecticut, USA: Wadsworth, Cengage Learning.
- Gutiérrez, M. R. (2005). Análisis Semántico Latente: ¿Teoría psicológica del significado? *Revista Signos*, 303-323.
- Hernández, A. (2005). *Descartes: Discurso del método*. San Vicente, Alicante, España: Editorial Club Universitario.
- Hinojosa, R. G. (2008). El tratamiento estadístico de las redes semánticas naturales. *Revista Internacional de Ciencias Sociales y Humanidades*, 133-154.
- Hume, D. (1748). *Investigación sobre el entendimiento humano*. Madrid: El libro de bolsillos.
- Kant, I. (1781). *Crítica de la razón pura*. Madrid: Alfaguara.
- Kranjec, A. &. (2010). Are temporal concepts embodied? A challenge for cognitive neuroscience. *Frontiers in Psychology*, 1, 240.
- Landauer, T. K. (2009). An introduction to latent semantic analysis. *Discourse Processes*, 259-284.
- Locke, J. (1690). *Ensayo sobre el entendimiento humano*. Oxford: Clarendon Press.
- Lund Kevin, B. C. (1996). Producing high-dimensional semantic spaces. *Behavior Research Methods, Instruments, & Computers*, 203-208.
- MacLeod Colin M., M. P. (2000). Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends in cognitive sciences*, 383-391.
- Manzanero, A. L. (2006). Procesos automáticos y controlados de memoria: Modelo Asociativo (HAM) vs. Sistema de Procesamiento General Abstracto. *Revista de psicología general y aplicada*, 3(59), 373-412.
- Martin, A. (2007). The representation of object concepts in the brain. *The Annual Review of Psychology*(58), 25-45.

- Masson, M. E. (2001). Priming, Cognitive Psychology of. *International Encyclopedia of the Social & Behavioral Sciences*, 12042-12046.
- McNamara, T. P. (2004). *Semantic Priming: Perspectives from Memory and Word Recognition*. Psychology Press.
- Mercado, L. E. (2015). *Manual de redes semánticas estructurales*. México, Coyoacán, México: Universidad Nacional Autónoma de México, Facultad de Psicología.
- Mercado, S., López, E., & Velasco, E. (2015). *Manual de redes semánticas estructurales*. México, Coyoacán, México: Universidad Nacional Autónoma de México, Facultad de Psicología.
- Meyer David E, S. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 227-34.
- Miller, G. A. (1956). *The Magical Number Seven, Plus or Minus Two*. Madrid: Alianza.
- Montealegre, R. (2003). La memoria: operaciones y métodos mnemotécnicos. *Revista Colombiana de Psicología*(12), 99-107.
- Mora, P. C. (2010). Sistemas de memoria: reseña histórica, clasificación y conceptos actuales. Primera parte: Historia, taxonomía de la memoria, sistemas de memoria de largo plazo: la memoria semántica. *Salud Mental*, 33(1), 85-93.
- Newman, M. E. (2003). The Structure and Function of Complex Networks. *SIAM Review*, 167–256.
- Oviedo, G. L. (2004). La definición del concepto de percepción en psicología con base en la teoría Gestalt. *Revista de estudios sociales*(18), 89-96.
- Paivio, A. (1971). *Imagery and verbal processes*. London, Ontario, Canadá: Psychology Press.
- Periñán, P. C. (2012). En defensa del procesamiento del lenguaje. *onomázein*, 13-48.
- Price, A. R. (2015). Semantic memory: cognitive and neuroanatomical perspectives. *Brain Mapping*, 529-536.
- Reyes-Lagunes, I. (1993). Las Redes semánticas naturales, su conceptualización y su utilización. *Revista de Psicología Social y Personalidad*, 81-97.
- Rodríguez, R. J. (2010). Herramientas informáticas para la representación del conocimiento. *Subjetividad y procesos cognitivos*, 232-217.

- Squire, R. L. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory*, 171-177.
- Torres, F., & Garza, S. (Diciembre de 2014). Visualización de Redes Semánticas Naturales a Través de Acomodo de Grafos Dirigido por Fuerzas. *International Journal of Good Consciencie*, 9(3), 68-75.
- Valdez, J. L. (2010). *Las redes semánticas naturales, usos y aplicaciones en psicología social*. (Segunda ed.). Toluca, México: Universidad Autónoma del Estado de México.
- Vásquez, A. C., Vega Huerta, H., & Pa, J. (2009). Procesamiento de lenguaje natural. *Revista de investigación de Sistemas e Informática*, 45-54.
- Vivas, J. R. (2007). Evaluación de las redes semánticas de conceptos académicos en estudiantes universitarios. *PsicoUSF*, 111-119.
- Vivas, J. R. (2008). Distsem. Un Método de captura y graficación de Redes Semánticas. Aplicaciones a Educación y a Neuropsicología. *Revista Mexicana de Psicología*.
- Warrington, E. K. (1984). Category specific semantic impairments. *Brain*, 829-853.