

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

MARCADORES, RASGOS ESPECÍFICOS Y REPRESENTACIÓN DEL CONOCIMIENTO.

Enrique del Teso
Universidad de Oviedo

0. PRELIMINAR.

La semántica se ocupa de emparejar las expresiones lingüísticas con los hechos del mundo. La representación del conocimiento consiste en procedimientos formales que expresan los datos disponibles sobre determinados dominios. La semántica no puede hacer su trabajo literalmente como lo acabamos de decir. No se pueden emparejar con las expresiones lingüísticas los hechos del mundo, sino sustitutos simbólicos suyos expresados en un lenguaje con propiedades adecuadas y controlables. La representación del conocimiento y la tarea ordinaria de la semántica, por tanto, se parecen hasta cierto punto. Aquí hablaremos de ese parecido. En concreto, hablaremos de una cuestión de semántica léxica en la que el problema básico, la relación entre los componentes genéricos y específicos, se plantea de manera análoga a como se plantea en la representación del conocimiento. De todas las formas de representación de conocimiento consideraremos la que está vinculada con las tecnologías de web semántica, las construcciones llamadas ontologías.

1. WEB SEMÁNTICA.

1.1 Representación del conocimiento.

La razón de representar el conocimiento es la misma por la que se codifica o representa cualquier otra cosa. Ideamos señales que representen datos con dos finalidades. Una de ellas es homogeneizar materialmente los datos que manejamos para adaptarlos al medio disponible. Los datos de los que queremos hablar son materialmente muy distintos (unos se comen, otros sólo se ven, otros no son sensibles, ...), pero hablamos siempre envueltos en aire, una masa gaseosa que sólo se deja comprimir y descomprimir. La codificación permite asociar los heterogéneos datos de nuestras conversaciones con homogéneas señales acústicas que se adaptan al canal. La segunda razón es manejar unidades que tengan propiedades adecuadas para algún propósito. En nuestras comunicaciones ordinarias necesitamos manejar unidades que se propaguen y que tengamos siempre disponibles. Si tenemos que hablar de una manzana, sucede que ni la manzana es algo que se propague ni es algo que tengamos siempre a mano para señalarla. El convertir a todos los datos en sonidos nos garantiza que el dato siempre se propagará y siempre tendremos la señal a manos porque nunca estamos en ningún sitio sin nuestro aparato fonador.

El conocimiento es el conjunto de datos que algunos o muchos sujetos tienen sobre determinados dominios. La representación del conocimiento busca expresarlo en

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

señales uniformes que lo doten de las propiedades que se requieren para ser tratado y procesado por las máquinas. En los ordenadores y en los servidores hay depositada una gran cantidad de datos (como los habituales archivos de los tratamientos de texto), pero no hay programas que puedan hacer nada con ellos. Las máquinas son totalmente ciegas a los datos grabados en ellas y se limitan a ser su continente. Las técnicas de representación de conocimiento buscan, pues, hacer tratable computacionalmente el conocimiento que tenemos de las cosas.

1.2 Conocimiento y web.

El tratar computacionalmente el conocimiento debe ser algo que pensemos referido a la web. La web es y tiende a ser el soporte masivo del conocimiento y es además el espacio de acceso universal, tanto para consultar como para introducir los datos. En este momento el código que está detrás de las páginas web que vemos está escrito en HTML y codifica sólo los aspectos formales: los archivos que se juntan en la página; las fotos que aparecen, su tamaño y su localización; las letras, su color y tipo de fuente; ... El conocimiento realmente está en los sujetos humanos que acceden a Internet. El código que expresa lo que vemos en la página es totalmente ciego respecto a su contenido. El sujeto es el que sabe que la página expresa un empeoramiento de la situación en Oriente o un cambio de tendencia en el clima. No podemos esperar que los programas hagan nada basándose en el conocimiento soportado en las páginas web. Así, no se pueden hacer con la web cosas como:

- Formular a un buscador preguntas como “¿qué lingüistas españoles son discípulos directos de Chomsky?”.
- Disponer de páginas web formadas *ad hoc* para el uso concreto que estemos haciendo. Las páginas que podemos ver ahora están diseñadas por alguien y nosotros sólo nos asomamos a ellas. No es posible que se forme una página en caliente tomando datos de diferentes fuentes y juntándolos en una página que antes no existía.
- El manejo de la web no puede iniciar secuencias complejas programadas. El dato de que un vuelo llega con retraso no puede modificar nuestra agenda electrónica y alterar citas sin intervención humana.

1.3 Metadatos.

Parece evidente que para que la web pueda hacer más cosas debe codificar más cosas. Seguirá habiendo código para el color y el tamaño de las letras y para la ubicación de las fotografías. Pero para que un programa sea capaz de juntar en una página todo lo que hay un fin de semana en una ciudad para niños, tendrá que poder “saber” que una película, una representación de teatro, un cuentacuentos o una exposición lúdica tienen carácter infantil y eso sólo podrá ocurrir si en las páginas correspondientes está codificada esa información. Las páginas habrán de tener entonces una serie de “etiquetas” que contengan información relativa a su contenido. Tales etiquetas han de verse como metadatos: las páginas contienen datos y las etiquetas invisibles para el usuario contienen datos sobre los datos, de manera que se puedan escribir programas que trabajen con lo que esas etiquetas dicen acerca del contenido de las páginas web.

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

Las etiquetas se escriben en un lenguaje llamado RDF y constan de secuencias con una estructura constante “sujeto – propiedad – valor”. Con ellas se pueden codificar ideas como: “El curso máster x tiene como alumnos a Elena y Alejandro”, “Alejandro es inspector de trabajo”, “Un inspector de trabajo es un tipo de funcionario”, ... De esta manera una parte del código de las páginas estará expresando contenidos, y no sólo aspectos formales, de esas páginas.

De todas formas, ¿qué es lo que hace que este nuevo código sea semántico? Un vaso no es un objeto semántico y el ponerle etiquetas referidas a lo que es no lo convierte en un objeto semántico. Si el código que nos permite acceder a los datos no codifica realmente el sentido de esos datos, no vamos a obtener sentido sólo por añadir etiquetas sin sentido. Repárese en que la secuencia de caracteres “Alejandro es inspector de trabajo” tiene sentido para un humano, pero no para una máquina. La mejor forma de responder a esto es recordar lo que es para nosotros un objeto semántico, un signo.

En un signo siempre hay una señal que se refiere a un dato merced a una asociación entre ambos que realiza un sujeto. La consecuencia de tal proceso es un determinado estado cognitivo interno del sujeto que, en muchas ocasiones, da lugar a una conducta observable. Aunque pidamos a un vaso etiquetado que se acerque a nosotros, el vaso no se moverá si no hay humano que lo haga, lo que indica que nuestros sonidos no tienen significado para él (de hecho, nada tendrá significado para el vaso). Las reacciones que tienen las personas a nuestras palabras nos dan la prueba externa de que tienen significado para ellas. En el caso de las máquinas debemos prescindir de nada que se parezca a estados cognitivos internos y concentrarnos en las reacciones de conducta. Las conductas de un ordenador consisten en la ejecución de programas. Que para una máquina una etiqueta RDF sea realmente semántica quiere decir que tal etiqueta desencadene la ejecución de un programa o que, en general, intervenga en comportamientos programados de las máquinas, sin intervención humana alguna. Que un ordenador “entienda” un dato quiere decir que el código que expresa ese dato lo hacen funcionar de una manera programada.

Se habla de web semántica precisamente porque estas secuencias RDF que expresan los metadatos son susceptibles de iniciar y entrar en secuencias programadas sin intervención humana. Que la web sea semántica quiere decir esencialmente que es más automática, que se pueden programar más cosas. Cuando las páginas estén etiquetadas con metadatos, habrá una generación de programas que actúe sobre esos metadatos y que permitirán una serie de funciones que hoy no son posibles. El abanico dependerá de la imaginación de los programadores y de la presión de la demanda.

1.4 Ontologías.

Pero aún se necesita algo más que el etiquetado RDFS. Supongamos que tenemos una página del Ministerio de Administraciones Públicas en la que se contienen nuevas leyes sobre la dedicación laboral de los funcionarios, sobre el sistema de jubilación y sobre el disfrute de bajas por nacimiento en el caso de los varones. Imaginemos que esta página está debidamente etiquetada con información semántica y que necesitamos saber cómo va a afectar a las plantillas de la universidad y a sus equipos de investigación. Un programa que deba trabajar en esa cuestión requiere algo más que la información semántica codificada en la página ministerial. Necesita también información sobre la

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

universidad: necesita saber que un profesor es un tipo de funcionario, que en la universidad hay asignaturas que son impartidas por los profesores, que las asignaturas tienen créditos, que la dedicación de un profesor se mide en créditos también, que puede haber equipos de investigación, que hay gente haciendo tesis doctorales, que parte de la dedicación de un profesor puede ser la dirección de tesis y la pertenencia a equipos, ... Junto con la información del Ministerio etiquetada para que haya programas que funcionen con ella, tiene que estar codificado lo que es una universidad. Todas las cosas que sabemos sobre las universidades deben codificarse de manera que una máquina pueda trabajar con esos datos y así es como se pueden escribir programas que aprovechen los datos semánticos de las páginas etiquetadas en dominios específicos. La universidad es un dominio, es decir, un área de la realidad relevante para nuestros propósitos. La ontología es la representación formal de un dominio. Una ontología es un tipo de representación de conocimiento. En una ontología se codifica el conocimiento de manera que pueda ser empleado en tecnologías relacionadas con la web y, eminentemente, la web semántica.

Pensemos un ejemplo más. Supongamos que queremos desarrollar una aplicación que ayude a un visitante que venga a Asturias a aprovechar su tiempo de la mejor manera posible. Necesitamos, evidentemente, una base de datos o similar que contenga información de las cosas que se pueden ver y hacer en Asturias. Pero necesitaremos también un perfil del visitante. Será necesario saber cuántos días estará; si viene con niños o no; si tiene alguna minusvalía; si viene solo, acompañado o es una visita de muchos a la vez porque estén en un congreso; qué cosas le gusta ver y hacer y en qué medida cada una; ... Una ontología puede ser la representación de estos perfiles.

Las ontologías se pueden escribir en RDFS también, pero se tiende más a lenguajes con algo más de capacidad expresiva, como OWL (recomendado por el consorcio W3C). Básicamente OWL aventaja a RDFS en que puede expresar más cosas de las que se necesitan y no puede expresar alguna de las que no deben expresarse; en concreto, RDFS permite expresiones de lógica de segundo orden que la versión más habitual de OWL (OWL DL) no permite, lo que es una ventaja teniendo en cuenta que una ontología con expresiones de segundo orden puede no ser computable.

2. LO GENÉRICO Y LO ESPECÍFICO.

2.1 Análisis componencial y representación formal.

Sin entrar en importantes cuestiones teóricas que aquí no podemos desarrollar, es una práctica habitual la de postular el significado como una entidad compleja. Se discute principalmente la naturaleza de esos componentes y si realmente los procesamos como una acumulación de propiedades sin más o más bien como una acumulación de atributos con un valor activado de entre otros posibles. Esto es, dicho sin más complicaciones, podemos pensar que la idea de ‘plátano’ contiene entre otras las propiedades de ‘fruta’, ‘alargada’ y ‘amarilla’; o podemos pensar que contiene propiedades como ‘forma’ y ‘color’ (a diferencia de la dulzura, por ejemplo), que asumen los valores de ‘alargado’ (una de las formas posibles) y ‘amarillo’ (uno de los colores).

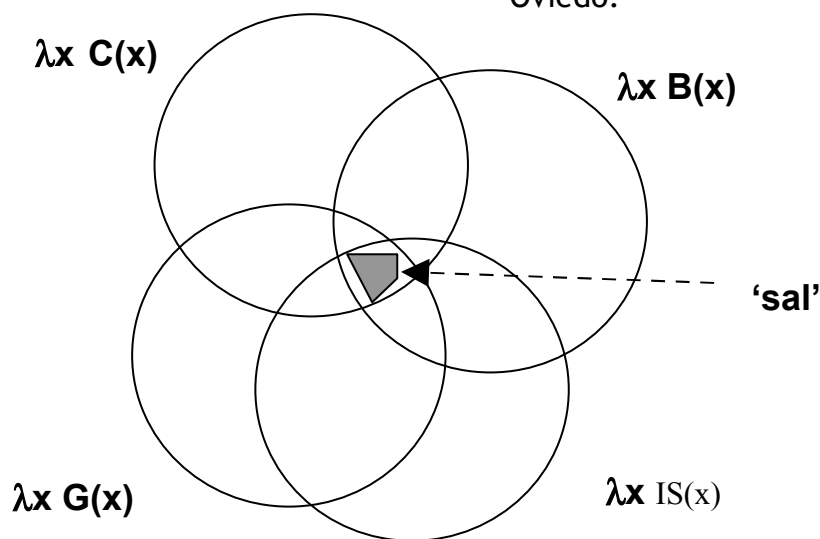
Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

Lo que nos interesa, como decimos, es la naturaleza componencial del significado. Resulta intuitivamente aceptable que el significado tiene componentes por lo naturales que resultan paráfrasis y negaciones que amplían y niegan aspectos parciales del significado. Así, *no será sal, pero es blanco y cristalino* indica que las ideas de ‘blanco’ y ‘cristalino’ son aspectos (componentes) de la idea de ‘sal’; *tendrá sal, pero no sabe a nada* revela que la propiedad de dar sabor es otro aspecto de la idea de ‘sal’.

El significado se representa formalmente. Desde siempre, la representación (gráfica) de los distintos niveles en que se puede analizar una secuencia lingüística se asocia con el final del trabajo del lingüista. Las reflexiones sobre cómo se estructura el plano de la expresión, qué son los fonemas, sus alófonos, etc. sólo parecen culminar cuando vemos una transcripción fonológica completa de un texto real. La transcripción fonológica no es sino la representación *completa* de un texto desde uno de los puntos de vista en que es capaz de aprehenderlo y explicarlo el lingüista. Lo mismo ocurre con la gramática y los conocidos “análisis sintácticos”; no se tiene la sensación de conocer bien un modelo sintáctico hasta que no se tiene destreza suficiente en la representación visual de ese modelo en oraciones concretas. Como en los demás casos, y aunque estemos menos habituados a ello, la culminación de la teoría semántica es la posibilidad de representar semánticamente secuencias idiomáticas. El procedimiento podría parecerse a lo que hacemos con la fonología. La representación semántica de un texto sería algo así como la “transcripción” de su sentido a través de un sistema simbólico convencional. Cuando la necesidad de rigor no es muy grande, como ocurre en los diccionarios con las palabras aisladas, ese sistema simbólico puede ser el propio idioma, utilizado como metalenguaje. Cuando tratamos de expresar el resultado de un análisis más científico, inevitablemente debemos utilizar algún formalismo útil. Naturalmente, la comparación con el caso de la fonología no puede ser llevada muy lejos. La representación que la semántica léxica puede hacer del significado de las palabras tiene el importante condicionamiento de que esa representación ha de ser compatible con el llamado principio de composicionalidad, según el cual el significado de las secuencias complejas es función del significado de sus secuencias simples constituyentes y la estructura gramatical o combinatoria que haya entre ellas. Esto significa que si tenemos debidamente explicitada la estructura gramatical de un enunciado y debidamente representado el significado de sus piezas léxicas mínimas, el significado del enunciado ha de ser en buena medida predecible y explicitable.

Los formalismos que se emplean en semántica varían según los propósitos y los principios teóricos que inspiren el análisis. Pero la mayoría de ellos tienen algo que ver con el lenguaje de predicados, normalmente ampliado con el abstractor lambda. Esto quiere decir que la mayoría de los formalismos tenderán a tratar el significado como un predicado lógico, que es el producto de otros predicados que son sus componentes. Así, si C representa a ‘condimento’, B a ‘blanco’, G a ‘granuloso’ y IS a ‘intensificador de sabor’, ‘sal’ debería ser una conjunción del tipo $\lambda x (C(x) \wedge B(x) \wedge G(x) \wedge IS(x))$. El significado equivale, pues, a la conjunción de sus componentes. Extensionalmente equivaldría a la clase producto de las clases definidas por cada componente:

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
 Oviedo.



(Debemos insistir en que hay muchas discusiones relevantes en las que no podemos entrar aquí; sólo nos detendremos en lo que se requiere para la cuestión concreta que plantearemos).

2.2 Problemas con la conjunción de predicados.

Los formalismos no deben hacernos olvidar la naturaleza sustancial de aquello con lo que estamos tratando. El valor semántico de una palabra o una frase consiste en la representación mental a la que da lugar su recepción. Empleamos formalismos lógicos o similares en la medida en que es lícito que nuestras representaciones cognitivas tienen efectivamente atributos lógicos. Pero pecaríamos de exceso de formalismo si representáramos predicados y, por tanto, clases que razonablemente no se correspondan con representaciones cognitivas reales. Esto es, las palabras no pueden estar expresando algo en lo que ni hablante ni oyente piensa. Algo así acabamos de hacer con la representación anterior. Hicimos que la clase ‘sal’ expresada por la palabra *sal* fuera el producto de las siguientes clases:

- Condimentos: sal, pimienta, laurel, ajo, perejil, azafrán.
- Objetos granulosos: sal, lija, grava, emulsión, arena.
- Objetos blancos: sal, armiño, tiza, cisne, azucena, escayola, estaño.

La primera de estas tres clases sí responde a una representación cognitiva. Hay situaciones en las que pensamos en condimentos. Por ejemplo, no es extraño que haya un sitio en nuestra cocina o en nuestros supermercados para condimentos y, por tanto, que esa categoría de cosas mueva nuestra conducta. Sin embargo, aunque es cierto que los armiños, las tizas y la sal tienen en común el ser blancos, nadie representa mentalmente semejante agrupación (más que como ejercicio académico, como ahora) ni desarrolla conductas movidas por tal categoría. Lo mismo sucede con una clase que englobe a la sal, la lija y las emulsiones. Decir que el significado ‘sal’ tiene como componentes tales propiedades es postular como unidades semánticas categorías sin realidad cognitiva y, por tanto, no podemos estar haciendo semántica y ocupándonos de las ideas que las palabras mueven en la mente de los hablantes.

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

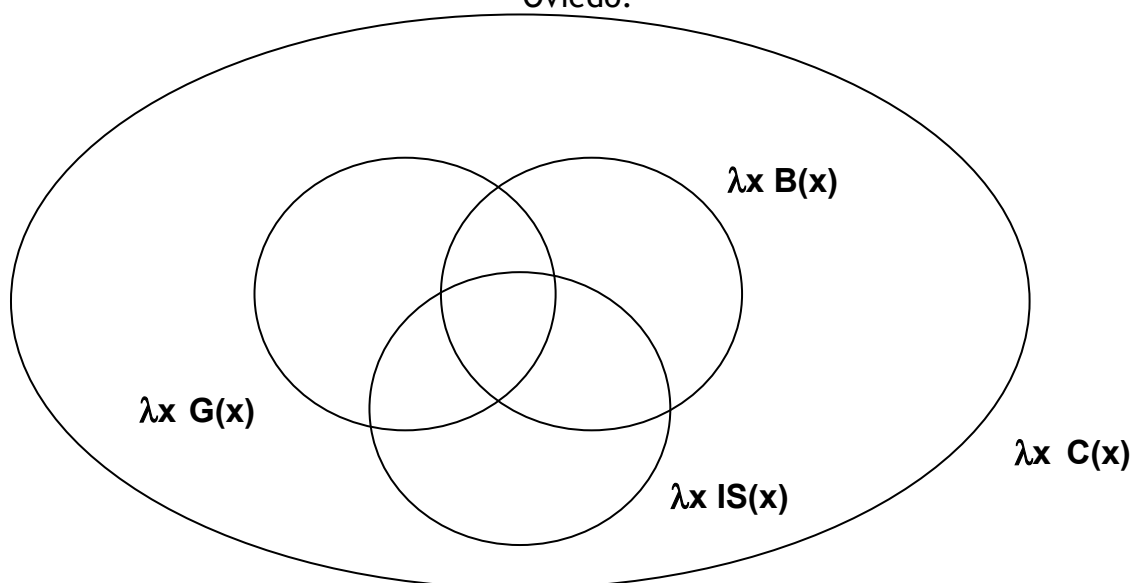
Es difícil establecer qué es lo que hace que ciertas agrupaciones de objetos sean de hecho representadas en nuestra mente y otras no. Sin duda, tiene que ver con la homogeneidad funcional de los objetos que englobamos en una misma clase y con el hecho de que aquello que comparten y define a la clase sea algo relevante para desarrollar conductas adecuadas en alguna situación. Lo cierto es que cuando una clase lógica agrupa objetos que, como tal agrupación, tienen alguna presencia en nuestra mente la clase evoca a sus elementos y los elementos a la clase. La idea de condimento nos trae a la mente cosas como el perejil o el azafrán; e inversamente estos elementos nos hacen pensar en condimentos. Pero en el caso de las clases sin realidad cognitiva, la clase no evoca a sus elementos. Cuando pensamos en blancura, no pensamos en sal y escayolas. Por tanto, el esquema unas veces dicho y otras asumido de un significado con componentes en producto lógico debe ser retocado.

2.3 Genéricos y específicos.

A pesar de lo que llevamos dicho, es evidente que la sal sí es granulosa y de color blanco y que tales características de alguna forma deben ser parte de su significado. Lo que ocurre es que no estábamos planteando la cuestión de una manera formalmente correcta. El error es tratar a todos los componentes como si fueran iguales. Intuitivamente, parece lógico decir que la sal es un condimento blanco y parece raro decir que es un objeto blanco que se usa para condimentar. La única clase cognitivamente real de las que se citan en la descomposición de ‘sal’ es la de ‘condimento’. No hay un producto de predicados del que este sea un factor más. Este componente en realidad define una clase genérica que engloba al producto de todas las demás.

Hace ya mucho tiempo que se viene dando como estrategia más normal de definición la de caracterizar a la clase definida por su género próximo y su diferencia específica. Cuando definimos algo, solemos englobar lo definido en algo más amplio y luego buscar lo que tiene de específico dentro de ese grupo. El género próximo es la clase más pequeña que incluya a la definida. La diferencia específica es el conjunto más pequeño de propiedades que haga a la clase definida disjunta de todas las demás incluidas en el mismo género. La semántica no es tan austera e incluirá entre los componentes del significado cuantos componentes específicos sean habitualmente evocados por la palabra considerada. Pero sí debemos retener la distinción esencial entre componentes genéricos y específicos. La sal es lo que es blanco, granuloso e intensificador de sabor dentro de los condimentos. Extensionalmente tenemos una intersección de clases incluidas en una clase genérica:

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
 Oviedo.



Esto es, en vez de

$$\lambda x [C(x) \wedge B(x) \wedge G(x) \wedge IS(x)]$$

tendríamos

$$\lambda x [C(x) \supset B(x) \wedge G(x) \wedge IS(x)]$$

Así por ejemplo, si tomamos la clase de los vehículos con elementos como bicis, aeronaves, coches, camiones, autobuses y motos, estaremos tratando con una propiedad que puede ser genérica en muchos significados. Si por el contrario tomamos la clase de los objetos pequeños, con elementos como bonsáis, pigmeos, ponis, avionetas, ácaros, botes, amuletos o ampollas, es evidente que la propiedad considerada, 'pequeño', ha de ser específica y englobada en un género mayor. Lo mismo sucederá con la propiedad 'rojo', que define una clase con elementos como amapola, rubí, cereza, petirrojo o sangre. Las palabras establecen relaciones semánticas entre sí cuando coinciden en el sema genérico, aquel que evoca a las clases incluidas en él.

2.4 Estabilidad y carácter funcional de los componentes específicos.

Aunque lo que acabamos de decir parece razonable, hay un problema evidente. Si los componentes específicos definen subclases incluidas en las definidas por los componentes genéricos, entonces componentes específicos incluidos en géneros distintos definirán necesariamente clases disjuntas. Dicho en términos más sencillos, la idea de 'blanco', presente en 'sal' y en 'armiño', serían ideas distintas, como la idea de 'pequeño' en bonsái, poni y pigmeo serían ideas también distintas. Esto es contraintuitivo: lo que hace pequeño a un bonsái es la misma cosa que hace pequeño a un poni; y un armiño es blanco por lo mismo que lo es la sal.

Sin embargo, también parece claro que este tipo de propiedades (deliberadamente ejemplificamos con casos claros, por no introducir complejidades que no caben en este trabajo) pueden definir valores distintos según qué género sea el considerado. un libro es marrón si lo es por fuera, mientras que un periódico sólo es marrón si sus páginas

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

interiores lo son también; una vaca es marrón si ese es su color natural y no deja de serlo porque la pintemos de otro color, mientras que una mesa es marrón si está pintada así, sea cual sea su color natural; una mesa no se hace marrón porque la cubramos con azúcar moreno, pero un pastel sí. Si saltamos al rojo, sucede que una manzana es roja si lo es por fuera, mientras que un pomelo es rojo sólo si es rojo por dentro; y una tela roja no es del mismo color que el pelo rojo. Si pensamos en tamaños como ‘pequeño’, es bien sabido que las dimensiones que hacen a algo pequeño son distintas si hablamos de elefantes o de hormigas. Y hasta las patas que mencionamos cuando decimos que una mesa es un mueble con patas sólo son salientes rígidos si el género incluyente es inanimado.

Todo encaja si pensamos los rasgos específicos, no como un predicado normal, sino como una función que toma por argumento el género incluyente y por valor una clase. De esta manera, si un bonsái es un árbol pequeño, con $P = \text{pequeño}$ y $A = \text{árbol}$, la idea de pequeño no será $\lambda x (Px)$, sino $\lambda x (Px)^{\lambda x (Ax)}$, es decir, pequeño para ser un árbol. La clase designada por $\lambda x (Px)^{\lambda x (Ax)}$ será el conjunto de tamaños o medidas que son pequeñas en función de tratarse de un árbol. La identidad que percibimos en las cosas que son pequeñas (o blancas, o granulosas) se debe a que, como funciones que son, son una correspondencia regular entre dos cosas; la diferencia que se observa se debe a que una función varía inherentemente de valor cuando le cambiamos el argumento.

Lo interesante es que, si observamos la variación conceptual de las palabras, incluidos los usos figurados, los componentes más resistentes son entonces los específicos. Lo que cambia es el componente genérico (llamado muchas veces marcador). Puedo decir que Juan hace de mesa para que rellene un papel, cuando o obviamente no es una mesa. Lo que cambia en el uso de esta palabra es el marcador o género próximo: no se refiere a un mueble. Pero mantiene lo específico, la idea de soporte horizontal con algún tipo de sostén vertical. Al cambiar el componente genérico (el argumento), cambiará en qué consiste ese soporte y ese sostén (el valor, los rasgos específicos).

3. RASGOS GENÉRICOS Y CONTEXTOS.

3.1 Contextos.

La palabra “contexto” se utiliza en muchos ámbitos y con diferentes niveles de precisión teórica. En el caso específico del lenguaje, entendemos por contexto de una señal el conjunto de datos externos a ella que intervienen en su procesamiento correcto. Los enunciados lingüísticos, como cualquier otra señal comunicativa, comportan la presunción de su relevancia. Esto es, se asume siempre que la información de las palabras es relevante para las circunstancias, lo que supone que tal información ha de tener un cierto tipo de interacción con otros datos que están activados o que son susceptibles de ser activados. Ese conjunto de datos forman el contexto y su interacción con la información semántica de las palabras determina el sentido total de estas.

En general, siempre que decimos que algo tiene contexto estamos aceptando que ese algo es variable o incompleto y que se manifiesta distintamente según con qué elementos externos aparezca acompañado. Siendo así que alcanzará valores distintos para acompañantes externos distintos, tendrá sentido decir que las cosas que se

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

contextualizan son en algún sentido funciones que toman por argumentos contextos y por valores interpretaciones o manifestaciones concretas.

3.2 Dimensión contextual de los rasgos genéricos.

Si pensamos otra vez en la palabra “mesa”, sabemos que siempre la referiremos a algo que tenga algo horizontal que haga de soporte, pero lo que estemos dispuestos a conceptualizar como soporte horizontal dependerá del tipo de objeto del que supongamos que nos están hablando. Lo más habitual es que sea un mueble y entonces suponemos que el tablero será una superficie rígida normalmente de madera. Pero podemos tolerar muchos tipos de objetos como referencia de la palabra “mesa”; lo que no dejamos de buscar en el tipo de objeto que sea referido con esta palabra es algo horizontal que sirva de soporte. Los rasgos específicos parecen más resistentes, pero con un valor que se modifica en función de la clase incluyente.

La modificación que la clase incluyente induce en los rasgos específicos es del tipo de la que los contextos inducen en los objetos que se contextualizan. Esto es especialmente visible en el caso de los adjetivos. Los nombres comunes suelen codificar clases que se corresponden con representaciones cognitivas del mundo. No se pone nombre a cosas en las que no se piensa. En cambio los adjetivos típicamente codifican propiedades específicas cuyo género próximo es el indicado por el sustantivo al que se refiere el adjetivo en cuestión. Así, tenemos adjetivos como *rojo* y ya vimos que realmente nuestra mente no representa la clase de todos los objetos rojos y ni siquiera tienen el mismo color todos los objetos a los que aplicaríamos este adjetivo. Este tipo de palabras se emplea asumiendo que ya hay un género introducido por un nombre y lo que indican siempre es una subclase dentro de ese género incluyente. Cuando se usa la palabra *rojo* ya estamos pensando en impermeables, cabellos o pomelos. El género introducido por el sustantivo es a todos los efectos el contexto que determina el valor referencial del adjetivo. Esta relación que se da entre sustantivo y adjetivo estira en la cadena del habla la que se da colapsada en una palabra entre el marcador semántico y los rasgos específicos.

4. VUELTA A LAS ONTOLOGÍAS.

Como ya se dijo, las ontologías tienen su punto de similitud con la semántica. Una y otra actividad supone construcciones formales que modelan hechos. El objetivo de la semántica es emparejar de manera adecuada las expresiones lingüísticas con representaciones formales del mundo que expresan. En el caso de las ontologías, por un lado, las construcciones formales no intentan ser el sentido de ninguna expresión dicha en ningún lenguaje. Y por otro lado, al realizarse en un contexto tecnológico, las ontologías buscan sobre todo que la aplicación informática que sea funcione eficientemente.

Pero en muchos puntos hay una continuidad de métodos bien visible. En concreto, el tema semántico que elegimos, la relación entre los marcadores genéricos y los rasgos específicos, apunta a una cuestión también conocida en el ámbito de la representación del conocimiento, que es la relación entre lo genérico y lo específico. La manera más eficiente de formalizar dominios es la de acumular microontologías, más que hacer

Teso Martín, E. del (2006), “Marcadores semánticos, rasgos específicos y representación del conocimiento”,
I Congreso Internacional de Filología Hispánica: Jóvenes Investigadores,
Oviedo.

ontologías de gran tamaño y complejidad. Es inevitable que un mismo concepto pueda y deba aparecer en varias ontologías y que las ontologías modelen ámbitos más y menos generales. Cada ontología atrapa en un red de relaciones una serie de conceptos e individuos que determinan el comportamiento de un programa en una operación. Cuando la ontología activada es una en particular, es esa red y no otras la que determina el resultado de la operación, lo que no quiere decir que todos los conceptos e individuos representados en ella resulten activados y afecten al comportamiento del programa en esa operación. Cada ontología tiene un papel similar al que tienen los estados cognitivos en las tareas de contextualización: es el conjunto de lo disponible para la operación que se esté haciendo. Lo que efectivamente intervenga en la operación será un subconjunto suyo. El que un concepto esté en distintas ontologías supone que un concepto puede intervenir en una operación concreta de algún programa distintamente contextualizado, si entendemos por contexto el conjunto de clases, relaciones e individuos que se activan en la misma operación que él.

Lo que dijimos antes respecto de los marcadores se plantea de la misma manera en la representación de conocimiento. Algunas ontologías son genéricas respecto de otras. Y algunas son tan genéricas que no corresponden a ningún dominio, sino que modelan categorías recurrentes en todas y sólo sirven para ayudar a modelar ontologías más específicas. Los datos de las ontologías más englobantes pueden ser contextos que determinen el valor que pueden tener los conceptos más específicos. Así por ejemplo, si contextualizamos una iglesia dentro del marco genérico de los objetos geográficos, tal objeto puede ser integrado en una ontología como un recurso turístico, igual que un parque. Si el marco genérico es el de los objetos sociales no agentivos (una de las categorías de la ontología genérica DOLCE), entonces esa iglesia podrá ser tratada como componente de la institución eclesiástica, como una universidad o un ministerio. La manera en que los conceptos de las ontologías genéricas insertados en ontologías de dominio pueden modificar el valor de los mismos elementos sustanciales es una manifestación más del hecho general de la contextualización y se maneja de manera similar a como debemos manejarlo en semántica.