

Organizadores lingüístico-conceptuales del conocimiento del mundo

Herminia Peraita
Uned

Resumen

Pretendemos con este trabajo describir una serie de representaciones lingüístico-conceptuales obtenidas empíricamente en experimentos de Psicología Cognitiva mediante tareas de producción libre de atributos o características, sin más restricción que la temporal y, en el caso concreto de los datos que presentamos, procedentes de una muestra de sujetos bilingües italo-castellanos de 1º y 2º de BUP del Liceo Italiano de Madrid.

El análisis detenido de los conjuntos de atributos producidos y el proceso de "filtraje" a que fué sometida dicha información para ser estructurada y organizada, nos permitió comprobar como, a pesar de la diversidad de sujetos con los que se había trabajado, dichos atributos podían siempre clasificarse en unos bloques tras los cuales subyacían unas determinadas relaciones, diferentes en función del campo semántico conceptual estudiado -animales, plantas, objetos, etc.-

Se decidió diseñar un esquema básico que recogiese dichas relaciones semánticas a partir de una estructura dual. Esto quiere decir que los componentes relacionales básicos se consideraron como configuradores de un "espacio o trama de relaciones" pretendidamente universal, especie de lista finita o conjunto de primitivos que nos permiten organizar el conocimiento del mundo. Por otra parte los contenidos concretos que actualizan o activan dichas relaciones, se consideró pertinente que constituyesen otra parte diferente del esquema ya que no pueden ser finitos, ni universales, ni mucho menos fijos e inamovibles.

Además y a la vista de los resultados, se sugiere la posibilidad de establecer una correlación entre dicho espacio de relaciones conceptuales y las expresiones lingüísticas mediante las cuales se expresan en el lenguaje natural, basándonos en una interpretación de la gramática de casos de Fillmore. Por otra parte los componentes relacionales pueden ser considerados "slots", estableciendo una analogía con el lenguaje de los frames, y los contenidos concretos, sus valores.

ORGANIZADORES LINGÜÍSTICO-CONCEPTUALES DEL CONOCIMIENTO DEL MUNDO

INTRODUCCION.

Pretendemos con este trabajo describir una serie de representaciones lingüístico-conceptuales obtenidas empíricamente en experimentos de Psicología Cognitiva mediante tareas de producción libre de atributos o características, sin más restricción que la temporal y, en el caso concreto de los datos que presentamos, procedentes de una muestra de sujetos bilingües italo-castellanos de 12 y 22 de BUP del Liceo Italiano de Madrid.

Inicialmente este trabajo se generó en el marco teórico de la categorización y la formación de conceptos naturales, tal y como había sido descrito por Rosch (1978), con la finalidad de comprobar ciertas hipótesis sobre el nivel básico de categorización: en concreto sobre el nº y tipo de atributos con que los sujetos caracterizaban a este, frente a los niveles supraordenado y subordinado. (Rosch, 1976)

El análisis detenido de los conjuntos de atributos producidos y el proceso de "filtraje" a que fue sometida dicha información para ser estructurada y organizada, nos permitió comprobar como, a pesar de la diversidad de sujetos con los que se había trabajado, dichos atributos podían siempre clasificarse en unos bloques tras los cuales subyacían unas determinadas relaciones, diferentes en función del campo semántico conceptual estudiado: animales, plantas, objetos, etc. - (Peraita, 1988)

Se decidió diseñar un esquema básico que recogiese dichas relaciones semánticas a partir de una estructura dual. Esto quiere decir que los componentes relacionales básicos se consideraron como configuradores de un "espacio o trama de relaciones" pretendidamente universal, especie de lista finita o conjunto de primitivos que nos permiten organizar el conocimiento del mundo. Por otra parte los contenidos concretos que actualizan o activan dichas relaciones, se consideró pertinente que constituyesen otra parte diferente del esquema ya que no pueden ser finitos, ni universales, ni mucho menos fijos e inamovibles.

Además, y a la vista de los resultados, se sugiere la posibilidad de establecer una correlación entre dicho espacio de relaciones conceptuales y las expresiones lingüísticas mediante las cuales se expresan en el lenguaje natural, basándonos en una interpretación de la gramática de casos de Fillmore (1968). Por otra parte los componentes relacionales pueden ser considerados "slots", estableciendo una analogía con el lenguaje de los frames, y los contenidos concretos, sus valores. (Winograd, 1983)

No pretendemos entrar aquí en la problemática de tipo teórico que subyace al planteamiento de la existencia de universales de significado, o de listas finitas de primitivos, o de la base conceptual y cognitiva de la gramática, pues sobrepasa con mucho nuestros objetivos, pero si nos gustaría apuntar dos aspectos.

IMPLICACIONES

Desde el punto de vista de las posibles implicaciones de

tipo aplicado o práctico del esquema que proponemos, queremos resaltar la relevancia que un planteamiento como este puede tener en el campo de la traducción automática. Para ello revisaremos brevemente un proyecto concreto: el Atlas II de Fujitsu (Soler, 1988).

Desde el punto de las implicaciones teóricas expondremos el punto de vista de Brachman (1983, 1985), cuya posición ante este tema nos parece relevante a la hora de interpretar nuestros datos.

Dentro del conjunto de sistemas de traducción automática actualmente existentes, bien sea en funcionamiento o en fase experimental, algunos de ellos han comprendido que, comprender el lenguaje es, al fin y al cabo, comprender el mundo y por tanto un proceso de traducción que no se limite a textos científicos y técnicos muy neutros, debe trasladar o trasvasar no sólo, o no primeramente estructuras sintácticas, sino más bien conceptuales. En otras palabras debe trasvasar el conocimiento del mundo implícito en el texto de la lengua origen al de la lengua destino. Para ello hay que asumir en algún grado o medida, una representación universal del conocimiento. En concreto este es el marco teórico en el que trabaja el proyecto Atlas II de Fujitsu.

Algunos de sus presupuestos teóricos básicos son:

12- Hay que partir de la representación conceptual que subyace a la representación de la frase, y superar los niveles de traducción guiados por la estrategia "palabra por palabra" e incluso los que trabajan a nivel de cláusula u oración (estrategia sintáctica).

22- Muchas palabras pueden llegar a tener listas interminables de significados, por lo cual requieren reglas complejas para ser abordadas, y además no suele coincidir el significado de las palabras en la lengua origen y en la destino.

32- Existe una heurística, o conjunto de heurísticos, para comprender el mundo, supuestamente universal.

42- No bastan reglas sintácticas y semánticas, sino que para poder transferir conocimiento, que al fin y al cabo es lo que se transfiere de una lengua a la otra, hay que definir reglas de estructura conceptual que permitan dicha transferencia. El papel de la representación sintáctica en este sistema es secundario.

52- En el proceso de traducción, tras el análisis morfológico, y sintáctico-semántico del input, se genera una estructura conceptual, y dado que puede no ser compatible con la de la otra u otras lenguas, se establecen unas reglas de transferencia conceptual, con la finalidad de llegar a una serie de primitivos-interlingua-a los cuales todas las lenguas podrían referirse (agente, objeto, instrumento, etc. con sus restricciones correspondiente). Ello implicaría elaborar frames de todas las palabras de una lengua. No olvidemos que la estructura de casos se puede asignar no sólo a verbos, sino a grupos nominales o nominalizados.

Por lo arriba indicado se puede comprobar como, para este sistema al menos, una representación conceptual, a modo de lista finita de conceptos que nos permita representar el conocimiento del mundo, es absolutamente imprescindible para una comprensión automática del lenguaje natural (Minsky, 1975; Schank, 1977)

Brachman(1983)plantea un problema básico que existe en los lenguajes basados en marcos o frames,ya que estos presentan una cierta ambigüedad al poder ser interpretados de una manera definicional o bien factual.El sistema de representación Krypton diseñado por ellos distingue entre un lenguaje descriptivo,sin implicaciones asercionales,que delimita básicamente una estructura y otro asercional que hace afirmaciones sobre el mundo.Es una forma de evitar el interpretar un lazo como aserción y el siguiente como formando parte del significado de un término... Sugiere la necesidad de integrar un acercamiento basado en la lógica tradicional y el lenguaje centrado en el objeto,lo cual permitiría diferenciar lo que en una representación esquemática de tipo red se refiere al concepto en sí,y lo que se refiere a hechos o afirmaciones sobre hechos-componenete terminológico y asercional respectivamente.En cualquier caso la espina dorsal de la estructura representacional,en el lenguaje de las redes declarativas,son las frases del tipo"ES UN" y lo principal poder especificar con claridad qué es lo implicado en el componente asercional,por cada tipo de estructura en el terminológico.Para ello analiza toda la problemática subyacente al lazo ES UN ,relación aparentemente taxonómica,pero tras la cual se ocultan muchos tipos de relaciones que giran en torno al eje genérico/genérico y genérico/individual.Insiste en que el verdadero problema que plantean tanto las redes smánticas como los lenguajes basados en frames,es cómo integrar representaciones definicionales con afirmaciones sobre la tipicidad de los objetos del mundo.(Brachman,1983)

Los conjuntos de producciones lingüísticas (no en el sentido
de reglas de producción, sino de proceso de lenguaje) de que hemos
partido en nuestro análisis han sido consideradas como textos más
que como oraciones yuxtapuestas o inconexas, expresados bajo un
título o tema común-los nombres de las categorías- que han sido
sometidos a un análisis textual. Los esquemas que proponemos
permiten, por tanto, convertir un texto en una estructura
normalizada; formatear el conocimiento de un campo; constituir una
base de datos para diseñar un sistema de question-answering en
lenguaje natural y programar preguntas para la búsqueda y
recuperación de la información. (Sager, 1972, 1978).

PROPUESTA

Lo que proponemos en último lugar, es una forma de
sistematización de la información conceptual para describir
objetos, situaciones, personas, que puede modelarse tanto en forma
de red como de esquema, y, cuyas relaciones se conceptualizan de
manera semejante a casos, que se generan en la estructura
profunda (conceptual) pero que están sometidas a transformaciones
en la superficial sintáctica. El conjunto de casos para cada
palabra equivale a una entrada de diccionario -con una serie de
restricciones semánticas y sintácticas asociadas- que puede ser
consultada en cualquier momento por el analizador. O lo que es lo
mismo: cada palabra lleva asociada una trama de casos o
relaciones. El diccionario, obviamente está organizado según
criterios semánticos pues se asume el poder estructurante del
léxico.

Los objetivos más globales y a más largo plazo de este trabajo
son:

1º-La construcción empírica de esquemas para determinadas
categorías del lenguaje natural y, de distintos niveles de
generalidad, incorporando datos estadísticos, con lo cual se puede
ver cuáles son las entradas más probables de los esquemas para
distintas muestras de sujetos. (Representaciones basadas en la
frecuencia)

2º-Describir la adquisición y evolución de esquemas para
intentar modelizar el proceso y no sólo el producto.

3º-Definir reglas de producción que expliquen la
actualización de los esquemas.

4º-Pasar del nivel descriptivo al nivel de interpretación
teórica, intentando diseñar un modelo de representación del
conocimiento del mundo útil en el procesamiento del lenguaje.

5º-Intentar demostrar la existencia de un sistema continuo
de organización que va desde lo puramente perceptual (un polo del
continuo) hasta los sistemas de valores y creencias en el otro
polo.

Algunos de los problemas que se nos plantean a la vista de
los datos son:

1º¿Por qué los sujetos a ciertas edades no enuncian y por
tanto no parecen tener disponibles ciertos lazos o
relaciones, por ejemplo "ES UN"? ¿Poseen la representación
mental-conceptual- subyacente que implica una representación
jerárquica, o esta se va construyendo gradualmente? Si este lazo
marca la identidad fundamental de ciertas entidades tiene que ser
conocido necesariamente, pero ¿desde qué edad?

2º¿Por qué las jerarquías parecen estar tan ramificadas
pero tienen tan poca profundidad? (La trucha es un pez o es un
animal, pero no "es un ser viviente", "una entidad física", "un ser
animado", etc)

3º-¿Qué representan los slots de los frames-o

ranuras? conjuntos de condiciones cuantificadas existencialmente, implícitamente universales? Condiciones necesarias para los miembros de una clase?

42-Por último, no hemos podido entrar a analizar el tema de la verdad o falsedad de las aseveraciones, ni la modalidad, (necesario versus contingente), ni el complejo tema de la cuantificación, aspectos todos ellos que requieren un estudio más detenido.

DESCRIPCION DE LOS TIPOS DE RELACIONES BASICAS QUE CONFIGURAN EL ESQUEMA PROPUESTO.

GENERICOS PARA TODAS LAS CATEGORIAS CONCEPTUALES

ES UN/A. Relación taxonómica o de sobreordenación. De subconjunto. Indica inclusión en clases de mayor generalidad. (Responde a la capacidad, o habilidad, lógico-ontológico-cognitiva de clasificar todo aquello que nos rodea en clases más amplias e incluyentes.) (Expresiones alternativas: ES UN TIPO DE/ES UNA ESPECIE DE/ES UNA CLASE DE).

ES . Relación cualitativa, predicativa o atributiva. Indica atribución de distintos tipos de propiedades: color, tamaño, forma, textura, sabor, y toda una gama de atribuciones no tan ligadas a lo perceptual. El rango va desde propiedades "objetivas", muy centradas en el objeto, hasta atribuciones "subjetivas", más o menos interpretativas y evaluativas, centradas en el sujeto, muy mediadas por pautas culturales y contextuales. (Expresiones alternativas: ES (DE +SUST)).

ES EL/LA. Relación de identidad básica.

SIRVE PARA. Relación de utilidad, funcionalidad, finalidad. (Expresiones alternativas: ES PARA+infinitivo; SIRVE (DE+sust) PARA; SIRVE (COMO+sust)).

TIENE. Relación partonímica. Indicación de las partes o elementos que constituyen un objeto, situación o persona. Puede hacer referencia a partes visibles y observables desde el exterior o internas- por ejemplo en seres vivos los órganos-.

PUEDE SER. Relación que indica los diversos tipos o clases o modalidades.

INFERENCIAS asociadas con el conocimiento de un contexto específico y con el conocimiento del mundo.

ESPECIFICAS DEL CAMPO SEMANTICO ANIMAL

VIVE EN. Indica el habitat y sus características.

SE ALIMENTA DE. Indica los hábitos nutricionales.

SE REPRODUCE POR. Sistema de reproducción.

SE DEFIENDE POR/POR MEDIO DE. Sistema de defensa y protección

RESPIRA POR. Sistema de respiración.

ES ALIMENTO PARA. Indica su usuario o destinatario final.

(HACE). Acción o actividad comportamental o expresiva característica. (Nada, vuela, canta)

(INFERENCIAS) asociadas con el conocimiento del contexto de una clase de animales.

ESPECIFICAS DEL CAMPO SEMANTICO VEGETAL.

SE OBTIENE DE/SE COGE DE.Lugar o zona donde se encuentra o produce o genera una determinada planta o fruta.

VIVE EN/SE ENCUENTRA EN/CRECE EN.Lugar.Hábitat.

PRODUCE/GENERA/DA.

ESPECIFICAS DEL CAMPO SEMANTICO ARTEFACTOS/OBJETOS.

ES DE.Indica el material que se ha empleado en la construcción de algo.(Expresión alternativa:ESTA HECHO DE).

FUNCIONA CON/POR.Sistema de alimentación o funcionamiento.

ES PARA.Indica el usuario o destinatario.

Referencias.

BRACHMAN,J.R.(1983) What ISA is and isn't:an analysis of taxonomic links in semantic network.IEEE Computer.September Special Issue.

BRACHMAN,R.J.,FIKES,R.E. and LEVESQUE,H.J.(1983)KRYPTON:A functional approach to knowledge representation.IEEE Computer,16(10)67-73.

FILLMORE,C.(1968) The case for case.In Bach,E. and Harms,R.(Eds.)Universals in linguistic theory.Holt .N.Y.1-88.

MYNSKY,M.(1985)A framework for representing knowledge.In Brachman and Levesque(Eds.)Readings in knowledge representation.Morgan-Kaufmann.

PERAITA,H.(1988).La representación del mundo en el niño de E.G.B.MEC.CIDE.Madrid.

ROSCH,E.,MERVIS,C.B.,GRAY,W.,JOHNSON,D.BOYES-BRAEM,P.(1976)Basic objects in natural categories.Cognitive Psychology8,382-439.

ROSCH,E.(1978).Principles of categorization.In Rosch,E. and Lloyd,B.(Eds.)"Cognition and Categorization". L E A.

SAGER,N.(1972)Syntactic formatting of scientific information.Proc.1972 Fall Joint Computer conf.

SAGER,N.(1981).Natural language information processing.Addison Wesley.

SCHANK,R.C. and ABELSON,R.P.(1977)Scripts,plans and knowledge.Proc.4th.IJCAI.1151-1157.

WINGRAD,T.(1983)Language as a cognitive processes.Volume 1:Syntax.Addison-Wesley.

