

La representación semántica de los eventos y las entidades en FunGramKB

Marcia Barrera, Daniela Ibarra, Pilar Oplustil, Josué Pino,
Camila Gallegos, Sanndy Infante, Valeria Mora, Fredy Núñez
Pontificia Universidad Católica de Chile (Chile)

Índice

- 1 Introducción
- 2 Los eventos en FunGramKB
 - 2.1 Metodología
 - 2.2. Análisis de Casos
 - 2.2.1 *Exclamar*
 - 2.2.2 *Recite*
- 3 Las entidades en FunGramKB
 - 3.1 Trabajo con diccionarios
 - 3.2 Predicaciones estrictas y rebatibles
 - 3.3 Postulado de significado en COREL
- 4 Conclusiones
- Referencias

Resumen

El presente artículo presenta dos metodologías plausibles respecto a dos áreas específicas del componente ontológico de la suite de FunGramKB: eventos y entidades. En el caso de los eventos, se presenta brevemente las propiedades que los caracterizan como unidades conceptuales dentro de FunGramKB. Posteriormente, se describe la metodología aplicada, a partir de ejemplos que involucran unidades conceptuales derivadas del dominio cognitivo de los eventos, #COMMUNICATION. El objetivo es indicar, de manera práctica, cuáles son las decisiones y consideraciones con las que el ingeniero del conocimiento puede encontrarse al trabajar con las unidades conceptuales de los eventos en la suite de FunGramKB. En el caso de las entidades, se pretende establecer criterios de análisis para resolver los problemas derivados de su representación conceptual. Para lograr este objetivo, primero se sistematiza la definición de entidad según el modelo, luego se establecen los criterios para el trabajo con diccionario y, finalmente, el modo de proceder para la formalización en COREL.

Palabras clave: FunGramKB, postulado de significado, evento, entidad.

1 Introducción

Una de las tareas de los ingenieros del conocimiento es poblar la ontología de la base de conocimiento de FunGramKB. Para llevar a cabo esta tarea es necesario realizar una serie de procedimientos que permitan representar y gestionar el conocimiento conceptual. Sin embargo, dependiendo del área ontológica que se esté interviniendo, la gestión de esta información no es homogénea. Por ello, es necesario precisar cuáles son las propiedades que diferencian a eventos y entidades, y cómo a partir de la naturaleza de cada tipo de concepto se desarrolla una metodología adecuada para cada caso. En primer lugar, se aborda el ámbito ontológico de los eventos: cómo se definen éstos desde la representación computacional de FunGramKB, qué particularidades tienen como conceptos, los pasos que constituyen la metodología de

trabajo que proponemos para este ámbito, y luego, la aplicación de esta metodología a ciertos casos ejemplares. En segundo lugar, se abordan las entidades: cómo se representan en la base ontológica, cómo se realiza la edición conceptual de éstas, principalmente a través del trabajo con diccionarios, para finalizar con el resultado de aquella edición. Por último, se presentan las conclusiones derivadas del trabajo, con respecto a cada dimensión conceptual y sus respectivas metodologías.

2 Los eventos en FunGramKB

Los eventos se definen, en la base de FunGramKB, como acciones, realizaciones o estados del mundo. La subontología correspondiente a los eventos está desplegada en cuatro grandes metaconceptos, o dimensiones cognitivas: #COMMUNICATION, #MATERIAL, #PSYCHOLOGICAL y #STATIVE. Bajo estos metaconceptos, se despliegan los conceptos básicos y los conceptos terminales. La labor con que se ejemplificará la edición conceptual corresponde al concepto hijo del metaconcepto #COMMUNICATION, +SAY_00, desde el cual se despliegan otros conceptos básicos y terminales que refieren al ámbito de los eventos de la comunicación.

Las unidades conceptuales de los eventos poseen una formalización estándar dentro de la base ontológica que debe considerarse en el momento de proceder con el trabajo. Los eventos, a diferencia de las entidades y las cualidades, están configurados por un Marco Temático (MT), el cual indica cuáles son los participantes prototípicos del evento del caso, heredados del dominio conceptual en el que se enmarca. Además, tal como las otras unidades que conforman la base ontológica, incorporan un Postulado de Significado (PS), en el cual se describen cuáles son los rasgos que caracterizan a dicha unidad conceptual a través de proposiciones. En el caso de los eventos, el PS se modela a partir de un solo rasgo donde debe estar contenido todo aquello que describe a la unidad conceptual. Así como en las entidades y en las cualidades, estos elementos están formalizados a través del lenguaje COREL (*Conceptual Representation Language*). Además, poseen una descripción, que indica lo que está contenido en el PS, pero expresado en lenguaje natural (en el caso de FunGramKB Suite, en inglés). También contienen, como todas las subontologías, la información respectiva a la herencia de la unidad conceptual, es decir, cuál o cuáles son sus superordinados. Estos son, a grandes rasgos, los elementos que componen la formalización de los eventos en la suite de FunGramKB. Como tales, deben ser considerados tanto en el momento en que se diseña una metodología para el desarrollo del trabajo, como en el instante en que se modelan estas unidades.

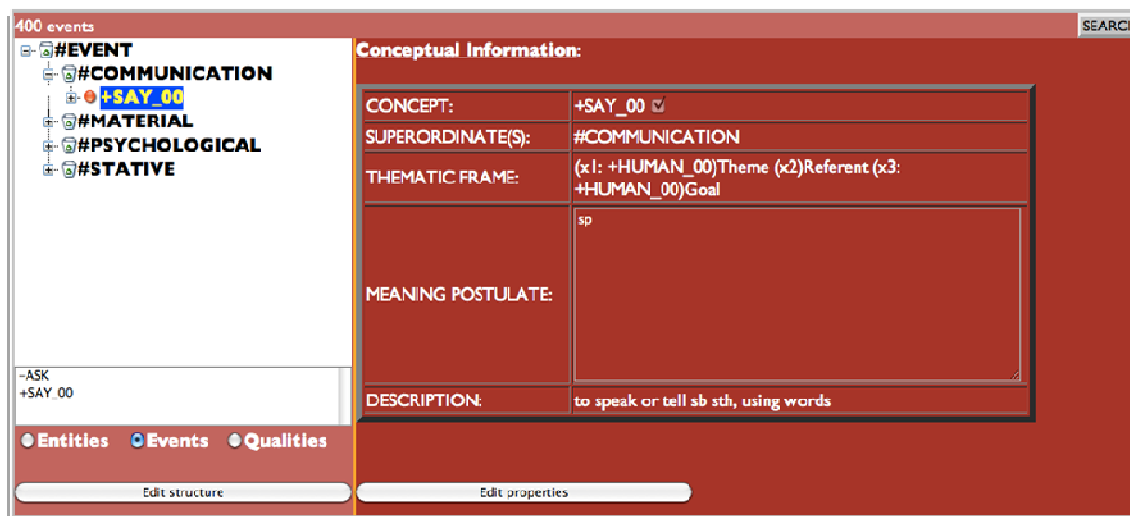
2.1 Metodología

La siguiente metodología se divide en dos grandes áreas: primero se explica el método para trabajar en la subontología de los eventos y posteriormente se realiza un análisis de caso para los eventos \$EXCLAIM_00 y \$RECITE_00. Algunos de los procedimientos aquí planteados se basan en las propuestas de Jiménez y Luzondo (2011).

A continuación, se presentan los pasos a seguir para determinar si un concepto debe ser agregado a la ontología de FunGramKB como un concepto terminal.

En primer lugar, se consideran las unidades conceptuales derivadas del metaconcepto #COMMUNICATION. En particular, este trabajo aborda el análisis del concepto básico +SAY_00.

Figura 1. Metaconceptos de los eventos y entrada conceptual de +SAY_00.



El primer paso consiste en elaborar una lista de ítems léxicos relacionados, ya sea como sinónimos, hiperónimos o hipónimos, para el concepto básico +SAY_00. Con el fin de realizar esta tarea, se recurre a una variedad de diccionarios de inglés y español disponibles tanto en línea como impresos¹. El resultado es la siguiente lista:

arengar	conversar	enunciar	opinar
articular	cuchichear	exclamar	perorar
balbucir	declamar	gritar	proferir
cascar	departir	hablar	pronunciar
charlar	dialogar	murmurar	recitar
comentar	discursear	musitar	susurrar

Tabla 1. Ítems léxicos relacionados con +SAY_00.

Una vez obtenida la lista, se seleccionan los conceptos cuya entrada en la ontología de FunGramKB se encuentra incompleta o no existe y se buscan las definiciones de cada uno en los diccionarios mencionados.

Esta actividad permite determinar el *differentia*², entre cada concepto seleccionado y +SAY_00, para completar así un concepto o para proponer la entrada de un nuevo concepto.

Son analizados aquellos ítems que presentan un *differentia* claramente distinguible del concepto +SAY_00, con el fin de determinar si debe clasificarse como un concepto terminal o como un subconcepto. En otras palabras, si la diferencia se manifiesta solo en las preferencias de selección de alguno de los argumentos del MT, entonces el

¹Los diccionarios consultados fueron: las versiones online de Cambridge, DRAE online, Espasa-Calpe, Longman, MerriamWebster, Multiword Net, Oxford, Wordreference. Y los diccionarios impresos de Manuel Seco y Diccionario de la RAE.

² En el contexto de FunGramKB, se utiliza el término *differentia* para referir a los rasgos que diferencian este concepto de su superordinado y que hacen que sea pertinente la incorporación de este nuevo concepto a la ontología.

ítem se considera “subconcepto” de +SAY_00; en cambio, si la diferencia se presenta con cualquier otra característica, el ítem se considera “concepto terminal” (Periñan y Mairal, 2010)

El trabajo, en este caso, se limita a registrar entradas para aquellos ítems clasificados como conceptos terminales. Para cada uno de ellos, se sigue el siguiente procedimiento:

- Determinar aquellas características expresadas en las definiciones del concepto que son comunes entre los distintos diccionarios consultados.
- Identificar las propiedades esenciales que caracterizan al concepto (*differentiae*).
- Proponer una descripción a partir de las definiciones.
- Establecer los papeles temáticos según el concepto padre.
- Formular proposiciones en lengua natural, las que luego formarán parte del PS del concepto.
- Codificar las propiedades en COREL.
- Ingresar el registro en la suite de FunGramKB.

2.2. Análisis de Casos

Se presentan, a continuación, los ejemplos de \$EXCLAIM_00 y \$RECITE_00 para mostrar cómo fue aplicada la metodología en el análisis de dos conceptos relacionados con +SAY_00.

2.2.1 Exclamar

Al ingresar un nuevo evento a la base de conocimiento es necesario detectar una característica que lo diferencie de su padre, o *differentia*, para que su incorporación sea justificada. Como ya hemos señalado, si se trata de una diferencia que se encuentra en su PS se está frente un concepto terminal; en cambio, si la especificación se encuentra en el MT, se postula un nuevo subconcepto.

Para decidir con qué concepto trabajar es necesario realizar una búsqueda de sinónimos y términos relacionados, y luego establecer las diferencias que tienen éstos con el concepto padre. Se ha escogido, para mostrar la metodología, el verbo “exclamar” que puede ser codificado como el concepto \$EXCLAIM_00.

Después de determinar el concepto con el cual se va a trabajar es necesario realizar una búsqueda en diferentes recursos capaces de proveer las definiciones del término, para luego crear una descripción que contenga lo más relevante y que permita crear un PS. Para esto es necesario contrastar definiciones de diferentes diccionarios y tomar en consideración las características que más se repiten.

Las definiciones de “exclamar” son las siguientes:

Recurso	Definición
Oxford	Cry out suddenly, especially in surprise, anger, or pain
Multiword Net	Utter aloud; often with surprise, horror, or joy.
MerriamWebster	To cry out or speak in strong or sudden emotion.
Cambridge	To say or shout something suddenly because of surprise, fear, pleasure, etc.
Longman	To say something suddenly and loudly because you are surprised, angry, or

	excited.
RAE	Emitir palabras con fuerza o vehemencia para expresar la viveza de un afecto o para dar vigor y eficacia a lo que se dice.
Espasa Calpe	Emitir palabras con fuerza o vehemencia para dar intensidad o eficacia a lo que se dice.

Tabla 2. Definiciones de “exclamar”.

En el caso de “exclamar”, después de haber realizado la búsqueda, es posible afirmar que la diferencia con su padre +SAY_00 se encuentra en su PS. Por lo tanto se puede establecer como un concepto terminal, ya que las principales diferencias con +SAY_00 se encuentran en la manera (Manner) y la razón (Reason) del evento y no en sus participantes.

Es posible agrupar las características que diferencian a \$EXCLAIM_00 de +SAY_00 en dos grandes grupos: aquellas que presentan la causa, que en este caso, sería una emoción y aquellas que dan cuenta de la manera en que el evento se lleva a cabo.

Loudly	Feeling
Vigor	surprised, angry, or excited.
Aloud	surprise, anger, or pain.
Shout	because of surprise, fear, pleasure, etc.
Fuerza	surprise, horror, or joy.
Intensidad	in strong or sudden emotion
Loudly	
Eficacia	

Tabla 3. Clasificación de los componentes semánticos de *exclaim*.

A partir de las características encontradas en los diccionarios y que se presentan como constitutivas de \$EXCLAIM_00, es posible formular una descripción que recoja los elementos más importantes que entregan estos recursos, en este caso: “Say something loudly because of a feeling.”

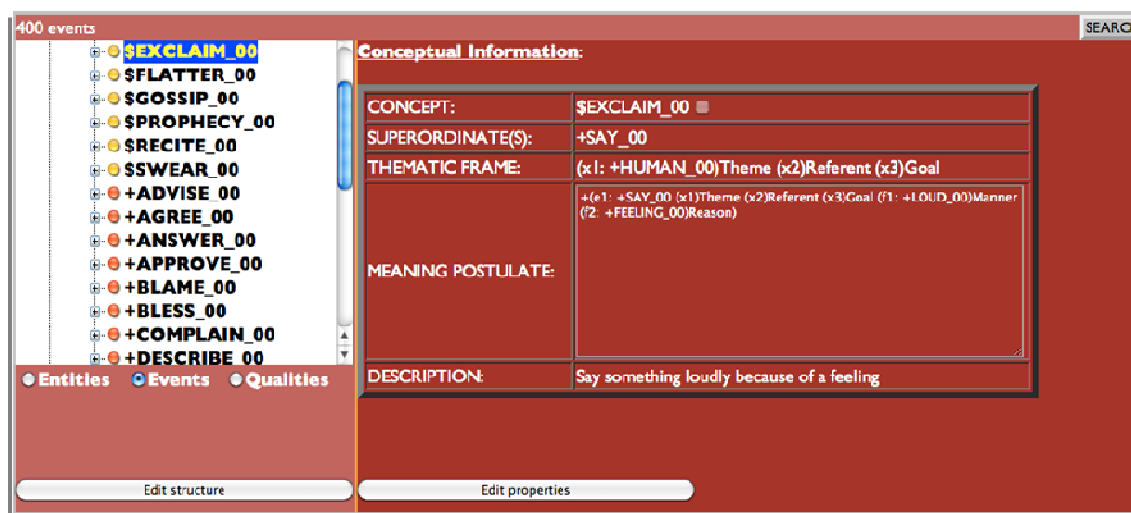
Luego de tener la descripción del evento en lenguaje natural se puede desarrollar el PS, teniendo en consideración el MT heredado de su padre +SAY_00. De este modo se establece la predicación que contiene los participantes y las especificaciones que diferencian este concepto de su padre. A partir de este procedimiento, se puede formular, en primera instancia, la siguiente propuesta:

Alguien (x1 Theme) dice algo (x2 Referent) a otro (x3 Goal) de manera fuerte (+LOUD_00 Manner) debido a una emoción (+FEELING_00 Reason)

Considerando este enunciado como estructura básica se crea el PS del evento \$EXCLAIM_00 en lenguaje COREL.

+ (e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: +LOUD_00)Manner (f2: +FEELING_00)Reason)
--

Figura 2. Entrada conceptual de \$EXCLAIM_00.



2.2.2 Recite

A continuación, se ejemplifica el concepto “recitar”, siguiendo los pasos propuestos en la metodología, para mostrar el trabajo requerido para llegar al registro de \$RECITE_00 en la suite de FunGramKB.

Para el concepto “recitar”, las definiciones son las siguientes:

Recurso	Definición
Oxford	Repeat aloud or declaim (a poem or passage) from memory before an audience.
Multiword Net	Repeat aloud from memory.
MerriamWebster	To repeat from memory or read aloud publicly.
Cambridge	To say a piece of writing aloud from memory.
Longman	To say a poem, piece of literature etc that you have learned, for people to listen to.
RAE	Referir, contar o decir en voz alta un discurso u oración.
Espasa Calpe	Decir algo en voz alta, especialmente versos.

Tabla 4. Definiciones de “recitar”.

A partir de estas definiciones, se determinan las características comunes que expresan las definiciones acerca del evento:

- Repetir, decir.
- En voz alta.
- De memoria.
- En público; para ser escuchado.
- Lo dicho es típicamente un poema o una obra literaria.

Luego, discutimos las características que definen su *differentia* respecto de +SAY_00:

- En voz alta.
- De memoria.
- Lo dicho es un poema.
- Ante una o más personas.

Es en este punto en que la discusión grupal produce las consideraciones más interesantes y relevantes en torno al significado y el uso del concepto “recitar”. Cada integrante del equipo de ingenieros del conocimiento presenta sus ideas, que luego de un análisis en conjunto sirven para determinar un *differentia* que explicita claramente las distinciones entre “decir” y “recitar”. Algunas consideraciones relevantes:

- Recitar necesariamente es un evento que se ejecuta a viva voz. Si así no fuera, podría considerarse que la acción realizada se acerca más a “pensar”.
- Para recitar, lo dicho debe ser algo que ya se ha aprendido, algo que está almacenado en nuestra memoria, de otro modo, “recitar” podría considerarse un sinónimo de “improvisar”, pero no hay evidencia en las definiciones encontradas que nos lleve a aceptar esta relación.
- Al recitar, típicamente lo que espera el sujeto que recita es ser escuchado. Si ya aceptamos que “recitar” es un evento que se ejecuta en voz alta, entonces existe, al menos, una persona que escucha lo recitado, a decir, la persona que recita.

Todas estas diferencias sobrepasan la especificación del marco conceptual del concepto padre +SAY_00. Por lo tanto, en este paso podemos establecer que “recitar” corresponde a un concepto terminal. Codificado en notación COREL se obtiene el siguiente resultado: \$RECITE_00.

El siguiente paso es proponer una descripción para el nuevo concepto terminal. La descripción propuesta a partir de las definiciones es la siguiente:

“Say something from memory, aloud, before one or more person”

A continuación, se establecen las preferencias de selección del concepto. Para realizar esta tarea, hay que guiarse por el MT del concepto padre, +SAY_00, y construir el de \$RECITE_00 a partir de éste. En este caso, se especifica que tanto el primer participante (Theme) como el tercero (Goal) son típicamente humanos, mientras que el segundo participante (Referent) es típicamente un poema:

(x1: +HUMAN_00)Theme (x2: +POEM_00)Referent (x3: +HUMAN_00)Goal

Posteriormente, se determinan las proposiciones inherentes al evento. El análisis nos permitió postular las siguientes:

- Alguien le dice algo a una o más personas, de modo que es escuchado.
- Lo dicho se encuentra almacenado en la memoria.

El paso siguiente es la codificación en COREL. En el caso de los eventos, debe codificarse sólo una predicación que encierre todas las proposiciones del PS. Para lograr expresar asertivamente la propuesta del PS, fue necesario el uso de los satélites que COREL posee como lenguaje, y el uso de la coindización.

Los satélites cumplen la función de indicar aquellas especificaciones que no están incluidas en el MT, como por ejemplo, las causas, los propósitos, situaciones típicas, etc. En este caso, se utilizó el satélite “Condition”, el cual indica una condición necesaria para que la predicación se lleve a cabo.

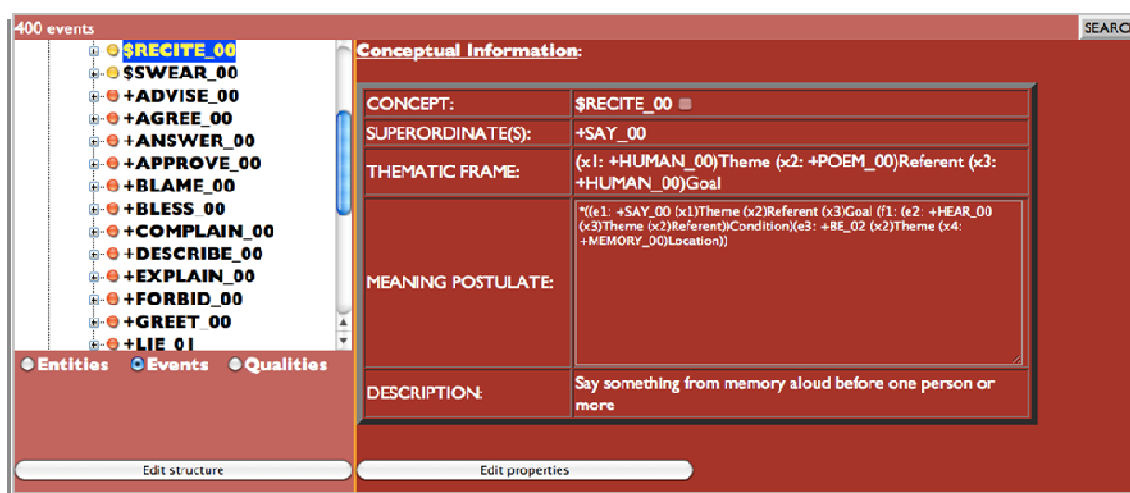
La coindización implica la recuperación de las variables dentro del PS. De esta forma, un mismo argumento puede desempeñar diferentes papeles en diferentes partes del PS. Por ejemplo, el argumento (x2) que participa en el comienzo del PS como “Referent”, reaparece más tarde como el “Theme” en la tercera predicación. Es decir, lo que en un principio era lo dicho, es también aquello que está almacenado en la memoria. De esta manera, a través de la coindización, se establecen las relaciones entre las diferentes partes del PS.

Para el ejemplo, el resultado es el siguiente:

```
*((e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: (e2: +HEAR_00 (x3)Theme (x2)Referent))Condition)(e3: +BE_02 (x2)Theme (x4: +MEMORY_00)Location))
```

Finalmente, se ingresa el registro en la suite de FunGramKB. Una vez definidos los parámetros que la suite exige para cada evento, podemos registrar la entrada. El resultado se muestra en la siguiente imagen:

Figura 3. Entrada conceptual de \$RECITE_00.



A partir de estos ejemplos se puede presentar una posible metodología, de carácter estructurado, con la cual es posible realizar el trabajo conceptual en las unidades de los eventos. Como se ha visto, la metodología expuesta apunta a trabajar con los eventos y con sus particularidades dentro de la base ontológica; el modelado de estas unidades puede requerir una gran abstracción por parte del ingeniero del conocimiento, ya que debe tener en cuenta procesos que para los hablantes se dan por hechos, pero que deben ser especificados para la base de conocimiento. Para ello, se recurre a las diferentes definiciones de los diccionarios, los cuales además son consultados de manera preferente en diferentes idiomas para aminorar dentro de lo posible las variaciones que cada lengua genera en sus piezas léxicas. A su vez, la inclusión de la intuición del hablante modela el impacto que los diccionarios podrían tener dentro de la formalización que realizan del conocimiento. De igual modo es importante tener en cuenta cuál es la estructura que la base ontológica le da a los eventos, la representación de los participantes preferentes en el MT, la conformación de un PS que presenta una sola proposición, las características del lenguaje COREL y las demás piezas conceptuales con que cuenta la base para conformar el PS.

3 Las entidades en FunGramKB

Para los fines de nuestro trabajo, una entidad corresponde a la representación

conceptual de una “cosa” u “objeto” del mundo real. Su inclusión en la base de conocimiento responde, principalmente, al principio de motivación léxica: el concepto que se desee ingresar debe estar lexicalizado en alguna de las lenguas trabajadas en FunGramKB. El conocimiento sobre esta entidad es representado por medio de un PS que agrupa predicaciones estrictas –rasgos necesarios de la unidad conceptual— y rebatibles –rasgos ejemplares—. Los conceptos que no responden a este principio (la minoría) se debe a que son organizadores de la jerarquía, pues actuarían como conceptos “paraguas” (Periñán y Mairal, 2010).

Existen dos tipos de entidades según su lugar en la jerarquía conceptual: las entidades básicas y las terminales. Las básicas corresponden a una lista discreta, trabajada por los autores del modelo, en base al vocabulario definitorio del *Longman Dictionary of Contemporary English*; se representan por medio del símbolo “+”. Los conceptos terminales representan los nodos finales de la estructuración jerárquica conceptual y están representadas por el símbolo “\$”.

Para la ejemplificación de esta edición conceptual se trabajará con la edición conceptual de una entidad básica (+): + BAG_00

A continuación, se proponen una serie de pasos para la edición conceptual de entidades. Estos pasos comprenden la elección de un concepto, búsqueda y delimitación de atributos e ingreso a la base de conocimiento mediante la formalización en COREL.

3.1 Trabajo con diccionarios

Para definir una entidad, considerar la intuición del ingeniero del conocimiento es esencial. Sin embargo, en algunas ocasiones es posible que al intentar definir un concepto surjan ciertas dudas específicas, como el color o la forma prototípica de una u otra entidad.

Para ayudar y guiar la creación del PS, un paso necesario es el trabajo con diccionarios. Dentro de FunGramKB Suite se recomiendan una serie de diccionarios, tanto de inglés como de español, que facilitan el trabajo del ingeniero del conocimiento. Algunos de ellos son el *DRAE: Diccionario de la lengua española (Real Academia)*, *Longman Dictionary of Contemporary English*, *CLAVE* o *OneLookDictionary*.

En este caso, se analizará, a modo de ejemplo, el concepto +BAG_00. En los diferentes diccionarios se han buscado las definiciones de *bolsa* en español, y *bag* en inglés. A continuación, se citan tres de siete definiciones con las que se construyó el PS final:

- “Saquillo de cuero o de otra cosa en que se echa dinero, y que se ata o cierra para que este no se salga.” (*DRAE*)
- “Saco hecho de un material flexible, con o sin asas, que se usa para llevar o guardar algo.” (*CLAVE*)
- “A container made of paper, cloth, or thin plastic which usually opens at the top.” (*Longman Dictionary of Contemporary English*)

Cada uno de los atributos que aparecen en las definiciones se han organizado en una tabla. Esta se ha ordenado de modo tal que en la fila superior se indican los nombres de los diccionarios y, en la primera columna (izquierda), los distintos atributos mencionados por los diccionarios. La tabla resultante se muestra a continuación:

+BAG_00	MANUEL SECO	RAE	WORD MAGIC	WORD REFERENCE	CLAVE	LONGMAN	CAMBRIDGE
MATERIAL FLEXIBLE	X	X	X	X	X	X	X
APERTURA SUPERIOR	X	X				X	
PARCIALMENTE CERRADO		X	X	X			
TRANSPORTE DE COSAS		X	X		X		X
GUARDADO DE COSAS			X	X	X		
ASAS	X				X*		X

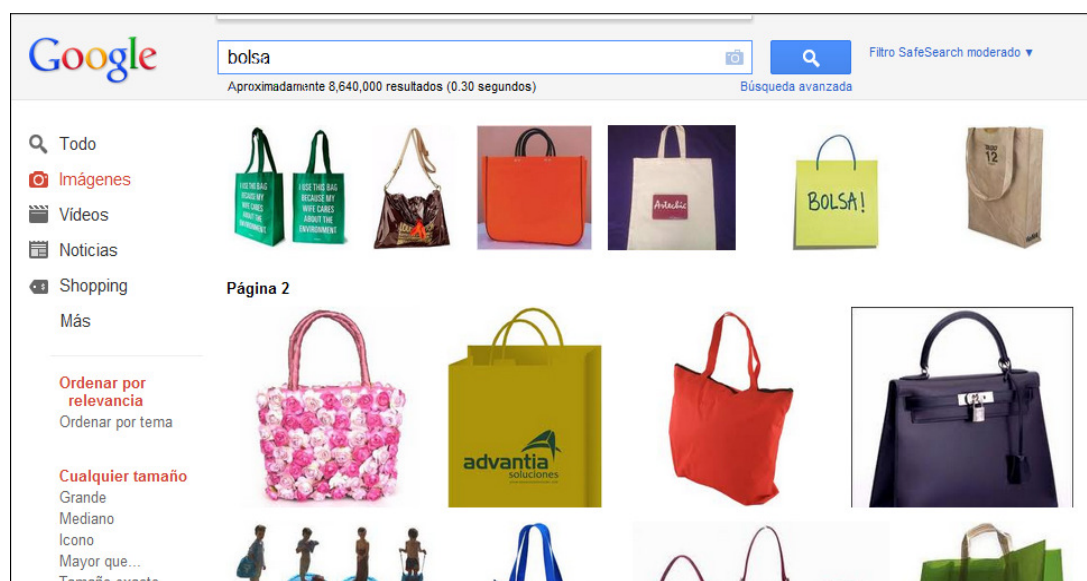
Tabla 5. Componentes semánticos de +BAG_00.

Como se observa en la tabla, las X permiten observar fácilmente qué atributos son los más frecuentes y, por tanto, tendrán una mayor relevancia a la hora de construir el PS.

Puede suceder que el ingeniero del conocimiento tenga dudas respecto a algunos atributos que los diccionarios capturan en sus definiciones, ya sea por su antigüedad o construcción sintáctica. Por ejemplo, una de las acepciones para *bolsa* en el *DRAE* es “Saco pequeño de cuero en que se echa dinero, y que se ata o cierra”. Sin embargo, parece ser que esta definición no es la más adecuada para lo que el analista entiende como *bolsa*, debido a que reúne atributos más bien anticuados, tal como explicita la definición en el diccionario *CLAVE*: “Saco pequeño que antiguamente se usaba para llevar el dinero (...)”. Además, el que se explicita la función de contenedor al comienzo o al final de la definición no significa que ésta sea más o menos importante, pues las definiciones lexicográficas, generalmente, no ordenan su construcción sintáctica de modo jerárquico.

Frente a esto es importante advertir que las definiciones lexicográficas no son de naturaleza conceptual, por lo que la intuición del ingeniero será siempre lo que prime sobre estos recursos. Herramientas útiles para sortear la duda pueden ser las búsquedas en corpus, como el *CREA* o el *Corpus del español* por Mark Davies, entre otros. Otra herramienta menos convencional, pero bastante útil en el trabajo con entidades, resulta ser *Google Images*. Su utilidad, en el caso de +BAG_00 se demuestra en nuestro caso, por ejemplo, cuando, al dudar sobre la prototipicidad del atributo “presencia de asas”, el buscador de Google evidencia la presencia de asas en la mayoría de las imágenes que en la red han sido etiquetadas como *bolsa*:

Figura 4. Imágenes para *bolsa* a partir de los resultados en *Google images*.



3.2 Predicaciones estrictas y rebatibles

Tras obtener las características del concepto, se procede a diferenciarlas entre rasgos nucleares y ejemplares. Con los primeros se hace referencia a los que cumplen una función categorial, es decir, “rasgos obligatoriamente necesarios en el mundo real” para una entidad (Periñán y Mairal, 2010). En el caso de +BAG_00 se puede decir que, para toda bolsa de este mundo, su naturaleza es ser un contenedor. Luego, ésta será una predicación *estricta* y estará precedida por el operador “+” en el PS.

En contraste, los rasgos ejemplares son aquellos rasgos que pueden ser rebatidos, incluso si se postulan situaciones singulares. Nuevamente para +BAG_00, si bien la mayoría de las fuentes consultadas confirman que las bolsas son hechas con materiales flexibles —como cuero, tela o papel—, podría concebirse que en alguna parte del mundo se elabore una bolsa con un nuevo material que no posea un grado tan alto de flexibilidad, como podría ser el caso de fibras fuertemente trenzadas. En base a esta razón, el analista debería considerar el carácter *rebatible* de esta predicación. Luego, su formalización estará precedida por el operador “*”.

Es importante destacar que, tal como apuntan Periñán y Mairal (2010), no es necesario que en todos los PS deba existir por lo menos una predicación estricta; de hecho, los PS pueden estar compuestos sólo por un tipo de rasgo —sea nuclear o ejemplar—. Lo que sí debe ser considerado es que toda predicación estricta implica que cada vez que el concepto aparezca, el rasgo será atribuido a su referente de modo obligatorio.

Respecto al número, el tipo y la relevancia de los rasgos que deben ser incorporados al PS, la suite de FunGramKB proporciona una lista de descriptores que, por una parte, funciona como guía semántica que amplía la información rescatada de los recursos lexicográficos y, por otra, permite trabajar con un criterio que haga posibles la homogeneidad y la consistencia.

La lista de descriptores para las entidades está basada en el modelo del Lexicón Generativo postulado por Pustejovsky (1995). El Lexicón Generativo es un modelo de representación para la Semántica Léxica. Uno de sus niveles de análisis, i.e. la estructura de *qualia*, intenta explicar las propiedades semánticas de los lexemas que contribuyen a la interpretación al interactuar con otros lexemas en construcciones complejas. Se define *qualia*, en tanto atributo subjetivo, como “un conjunto de

propiedades o eventos asociados a un ítem léxico que mejor explican qué quiere significar una palabra” (Pustejovsky, 1995).

La estructura de qualia define estas propiedades mediante cuatro dimensiones o papeles:

- a. Qualia formal: Es la categoría básica. Distingue a una entidad en relación a un dominio más amplio. Se refiere a la magnitud, forma, dimensión, color, posición, etc.
- b. Qualia constitutivo: Expresa la relación entre la entidad y sus partes constituyentes. Se refiere al material, peso y otros componentes de la entidad.
- c. Qualia télico: Expresa la función o propósito de la entidad, si es que lo tiene. Puede, por ejemplo, dar cuenta del propósito de un agente al crear una entidad, o bien de la meta que la ejecución de cierta actividad involucra.
- d. Qualia agentivo: Expresa los factores involucrados en el origen del objeto. Puede dar cuenta del creador, artefacto y/o cadena causal del objeto.

Estos descriptores están vinculados a las dimensiones conceptuales, por lo que en el caso de +BAG_00, al estar subordinado al metaconcepto #SELF_CONECTED_OBJECT, la lista que se ofrece es la siguiente:

- i. Las PARTES de las que está compuesta la entidad
- ii. EL COLOR u otros ATRIBUTOS VISUALES de la entidad
- iii. EL TAMAÑO/EXTENSIÓN de la entidad
- iv. Las CUALIDADES TÁCTILES de la entidad
- v. EL GUSTO de la entidad
- vi. EL PESO de la entidad
- vii. La FORMA de la entidad
- viii. La TEMPERATURA de la entidad
- ix. EL OLOR de la entidad
- x. EL VALOR de la entidad
- xi. EL MATERIAL del que está hecho la entidad
- xii. EL LUGAR donde se encuentra la entidad
- xiii. EL PROPÓSITO/FUNCIÓN para el que se usa la entidad
- xiv. ACTIVIDADES o ACCIONES relacionadas con la entidad
- xv. EL LUGAR DESDE DONDE SE OBTIENE la entidad
- xvi. La ENTIDAD QUE PRODUCE la entidad
- xvii. La FORMA en que la entidad SE PRODUCE/SE OBTIENE

De esta lista, a la hora de identificar los rasgos definitorios para +BAG_00, los puntos que parecen más relevantes fueron i, xi y xiii. De este modo, los rasgos para *bolsa* en tanto objeto artificial serían:

- i) una bolsa suele tener dos asas
- xi) una bolsa está hecha de un material flexible
- xii) una bolsa es utilizada para guardar o trasladar algo

Observamos que cada uno de los rasgos refieren a la entidad que se está definiendo: una “bolsa”. Esto se debe a que todo PS de una entidad debe responder al “principio de la entidad omnirreferencial”, pues “cada una de las predicaciones que configuran un PS debe necesariamente hacer referencia a su *definiendum*” (Periñán y Mairal, 2001). Tal acción puede realizarse explícitamente –como al enunciar “una bolsa es un objeto artificial”– o por coindización.

3.3 Postulado de significado en COREL

Para formar las predicaciones traducidas a lenguaje de COREL, se utilizan exclusivamente conceptos básicos. De este modo, para traducir el rasgo “una bolsa es un contenedor”, utilizamos el concepto básico +BE_00, el cual, acorde a sus papeles temáticos, señala que existe una entidad (*Theme*) que pertenece a una clase (*Referent*); por lo que la primera predicación será:

+(e1: +BE_00 (x1:+BAG_00)Theme (x2: +CONTAINER_00)Referent)

El segundo rasgo a definir es que una bolsa está hecha de un material flexible. Para tal predicación utilizaremos el concepto +BE_01, que vincula dos entidades señalando que una (*Attribute*) es atributo de la otra (*Theme*). Respecto a “material flexible”, si bien no existe un concepto básico o terminal que aluda directamente a esa categoría, lo que proponemos es que el ingeniero del conocimiento exprese este rasgo prototípico en base a la lista de conceptos que la ontología le ofrece. Es así como se acordó identificar “materiales flexibles” con el plástico, papel, cuero y/o tela. La posibilidad de escoger entre uno de estos materiales o la combinación de éstos, se formaliza por medio del conector lógico de disyunción “|”. Si quisiéramos marcar la posibilidad de utilizar todos los materiales juntos, el conector lógico sería el de conjunción “&”; en cambio, si obligatoriamente tuviéramos que escoger uno de los materiales, el conector de exclusión “^” sería el adecuado.

La predicación resultante es la siguiente:

*(e2: +BE_01 (x1)Theme (x3: +PLASTIC_00 | +PAPER_00 | +LEATHER_00 | +CLOTHING_00)Attribute)

En “una bolsa es utilizada para guardar o trasladar algo”, existen tres predicaciones. Primero, para que una bolsa pueda “guardar” o “trasladar” algo, debe necesariamente “contener” el objeto a guardar o trasladar. Es decir, para que el objeto pueda ser guardado o movido, necesariamente debe estar primero dentro de la bolsa—contenido—. Es así como en el PS se expresaría primero la necesidad de que el objeto deba ser contenido por la bolsa. Para eso utilizamos el concepto +CONTAIN_00:

+(e3:+CONTAIN_00 (x4)Theme (x1)Location)

Para expresar la posibilidad de que la bolsa “mueva” o “traslade” algo de un lugar a otro, se utilizará el concepto +MOVE_00. Ya que ser utilizado para mover un objeto, es solo una posibilidad y no una necesidad para la entidad “bolsa”, esta predicación tendrá un carácter rebatible (*):

*(e4:+MOVE_00 (x5)Agent (x4)Theme (x6)Location (x7)Origin (x8)Goal (f1:x1)Instrument)

De modo similar, el hecho de que la bolsa pueda “guardar” un objeto por un período prolongado, se expresará con el concepto básico +STORE_00. Al igual que la predicación anterior, “guardar” o “almacenar” tiene un carácter rebatible:

*(e5:+STORE_00 (x1)Theme (x4)Referent)

El hecho de que la entidad que se contiene sea la misma que puede ser trasladada o guardada, se expresa por medio de la repetición del argumento x4 —el objeto

contenido, guardado y trasladado—. De este modo, por medio de la coindización de x4, es posible señalar que un mismo argumento puede participar en más de una predicación con distintas funciones.

Finalmente, para expresar que “una bolsa suele tener dos asas, se utilizará el concepto +COMPRISE_00, el cual es un evento que señala que una entidad (*Theme*) consta de otra (*Referent*); en este caso, la bolsa podría constar de dos asas:

```
*(e6:+COMPRISE_00 (x1)Theme (x9: 2 +HANDLE_00)Referent)
```

De este modo es cómo se obtiene el PS definitivo para la entidad +BAG_00, la cual es propuesta como un objeto artificial, hecho posiblemente de plástico, papel, cuero o tela, que contendrá objetos en su interior, los cuales pueden ser trasladados y/o guardados. Además, se incluye el hecho de que una bolsa puede tener dos asas.

```
+(e1:+BE_00 (x1:+BAG_00)Theme (x2:+ARTIFICIAL_OBJECT_00)Referent)
*(e2:+BE_01 (x1)Theme (x3: +PLASTIC_00 | +PAPER_00 | +LEATHER_00 |
+CLOTHING_00)Attribute)
+(e3:+CONTAIN_00 (x4)Theme (x1)Location)
*(e4:+MOVE_00 (x5)Agent (x4)Theme (x6)Location (x7)Origin (x8)Goal
(f1:x1)Instrument)
*(e5:+STORE_00 (x1)Theme (x4)Referent)
*(e6:+COMPRISE_00 (x1)Theme (x9: 2 +HANDLE_00)Referent)
```

4 Conclusiones

Formular una metodología estructurada ofrece al ingeniero del conocimiento una base sólida sobre la que trabajar al momento de modelar el conocimiento, resumir los criterios fundamentales, las consideraciones que no deben ser olvidadas y, a la vez, encausar las discusiones que son necesarias para concluir la definición de una unidad conceptual.

Los procedimientos descritos en este artículo son aquellos que se nos develaron durante el trabajo mismo realizado con eventos y entidades. Lo que comenzó como un proceso de “prueba y error” para hacer coincidir nuestras discusiones acerca de un concepto dado, con la jerarquía establecida en la ontología de la suite, poco a poco fue dando forma a un protocolo bien definido que nos permitió avanzar en la producción de nuevas entradas de la base de conocimiento y en la corrección de algunas ya existentes.

Sin embargo, es importante destacar que las metodologías presentadas configuran sólo una de las posibles secuencias procedimentales a seguir, ya que los mismos criterios que nos llevaron a formular nuestro protocolo, se pueden distribuir o encausar a través de otras estrategias.

Este punto es particularmente relevante, porque nos indica claramente que nuestra propuesta no es, en absoluto, un estado final en la definición de metodologías de trabajo con la ontología de FunGramKB, sino que deja abierta la investigación a la exploración de nuevos procesos que permitan generar entradas válidas a la base de conocimiento en menos tiempo o de manera más refinada. También podemos aventurar que, si bien esta metodología se ha probado utilizando las definiciones en inglés y en español de los conceptos analizados, la selección de este par de idiomas es meramente casual y no la limita en ningún sentido, por lo que es posible aplicar el protocolo a cualquier otra lengua que cuente con recursos lexicográficos fiables.

Referencias

- Cambridge Dictionaries Online*, [<http://dictionary.cambridge.org>, fecha de consulta 22 de julio de 2011]
- CLAVE, [<http://clave.librosvivos.net/>, fecha de consulta 22 de julio de 2011]
- Corpus de referencia del español actual (CREA)*, [<http://corpus.rae.es/creanet.html>, fecha de consulta 22 de julio de 2011]
- Davies, Mark: Corpus del Español, [<http://corpus.rae.es/creanet.html>, fecha de consulta 22 de julio de 2011]
- Longman Dictionary of Contemporary English: *Longman English Dictionary Online (LDOCE)*, [<http://www.ldoceonline.com/>, fecha de consulta 22 de julio de 2011]
- Merriam-Webster Dictionary and Thesaurus on line*, [<http://www.merriam-webster.com/>, fecha de consulta 22 de julio de 2011]
- Multilingual Lexical Database-MultiWordNet*, [<http://multiwordnet.fbk.eu/english/home.php>, fecha de consulta 22 de julio de 2011]
- Jiménez Briones, Rocío y Luzondo Oyón, Alba (2011): "Building ontological meaning in a lexico-conceptual knowledge base". *Onomázein* 23, 11-40.
- Periñán Pascual, Carlos y Ricardo Mairal Usón (2010): "La gramática de COREL: un lenguaje de representación conceptual". *Onomázein* 21, 11-45.
- Pustejovsky, James (1995): *The Generative Lexicon*. Cambridge: the MIT Press.
- Real Academia Española (1992): *Diccionario de la Lengua Española*. Madrid: Espasa.
- Seco, Manuel, Olimpia Andrés y Gabino Ramos (1999): *Diccionario del Español Actual*. Madrid: Santillana.
- Oxford Dictionaries*, [<http://www.oxforddictionaries.com>, fecha de consulta 22 de julio de 2011]
- Word Magic Dictionary*, [<http://www.wordmagicsoft.com/diccionario/en-es>, fecha de consulta 22 de julio de 2011]
- WordReference Dictionaries*, [<http://www.wordreference.com>, fecha de consulta 22 de julio de 2011]