

# MODELO COGNITIVO DE LAS CUALIDADES PARA EL PROCESAMIENTO DEL LENGUAJE NATURAL

CARLOS PERIÑÁN PASCUAL  
ALFREDO SANCHO CARAVACA  
*Universidad Católica San Antonio de Murcia*

RESUMEN. *FunGramKB es una base de conocimiento léxico-conceptual multipropósito para su implementación en un sistema del procesamiento del lenguaje natural. FunGramKB comprende dos niveles de conocimiento independientes aunque claramente interconectados: el léxico y el cognitivo. A diferencia de otros sistemas del procesamiento del lenguaje natural, no pensamos que el componente más rico deba ser el módulo léxico, sino más bien el cognitivo, por lo cual en lugar de hablar de una postura lexicista, decimos que FunGramKB adopta en realidad una estrategia conceptualista. El propósito de esta ponencia es describir cómo este enfoque conceptualista afecta al tratamiento de las cualidades en nuestra base de conocimiento.*

PALABRAS CLAVE: *Procesamiento del lenguaje natural, base de conocimiento, ontología, postulado de significado, cualidad.*

ABSTRACT. *FunGramKB is a multipurpose lexico-conceptual knowledge base to be implemented in a natural language processing system. FunGramKB comprises two independent but interconnected levels of knowledge: lexical and cognitive. Unlike other natural language processing systems, our cognitive module is a richer component than the lexical one, so instead of saying that a lexicalist approach is taken in FunGramKB, we prefer to state that a conceptualist strategy is followed indeed. The aim of this paper is to describe how this conceptualist approach affects the treatment of qualities in our knowledge base.*

KEY WORDS: *Natural language processing, knowledge base, ontology, meaning postulate, quality.*

## 1. FUNGRAMKB

*FunGramKB (Functional Grammar Knowledge Base)* es una base de conocimiento léxico-conceptual multipropósito para su aplicación en sistemas del procesamiento del lenguaje natural (PLN). Cada uno de los componentes en los niveles léxico y cognitivo de *FunGramKB* se construye de forma independiente, aunque claramente interconectado con otros componentes del mismo o diferente nivel. Por ejemplo, y a diferencia de otros sistemas del PLN, el lexicón no es el componente central de nuestra base de conocimiento, sino más bien la ontología, por lo cual en *FunGramKB* no adoptamos un enfoque lexicista sino conceptualista.

Además, *FunGramKB* se caracteriza por ser una base de conocimiento multipropósito, en el sentido de que es multilingüista y multifuncional (Nirenburg *et al.* 1996). En otras palabras, no sólo está diseñada para poder trabajar con diversas lenguas naturales<sup>1</sup>, sino que además posibilita su reutilización en diversas tareas del PLN (p.ej. traducción automática, recuperación y extracción de la información, sistemas de diálogo persona-máquina, etc.).

## 2. EL MODELO ONTOLÓGICO

El modelo ontológico de *FunGramKB* se presenta como una taxonomía conceptual universal, entendiendo por *universal* que cada concepto que podamos imaginar tenga un lugar apropiado en la ontología. Además, aunque sea lingüísticamente independiente, nuestra ontología posee una motivación lingüística; en otras palabras, la ontología está ligada a la semántica de las unidades léxicas, aunque su conocimiento no es específico de una lengua

determinada. Nuestro modelo ontológico distingue tres niveles conceptuales diferentes, cada uno de los cuales presenta unidades conceptuales de diferente rango: metaconceptos, conceptos básicos y conceptos terminales. La figura 1 sirve para ilustrar nuestra estructuración conceptual<sup>2</sup>:

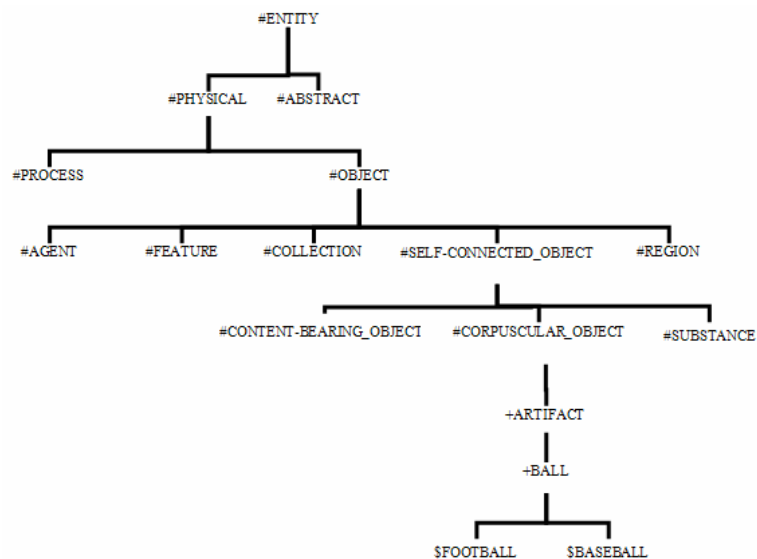


Figura 1. Ejemplo de la estructuración ontológica en *FunGramKB*.

Los metaconceptos, precedidos por el símbolo #, construyen el nivel superior de la taxonomía. El estudio de las principales ontologías lingüísticas de nivel superior ha dado como resultado un modelo metaconceptual cuyo diseño facilita la integración o el intercambio de información con modelos ontológicos diferentes. Los metaconceptos raíz son #ENTITY, #EVENT y #QUALITY, los cuales corresponden a las categorías lingüísticas de nombre, verbo y adjetivo respectivamente.

Los conceptos básicos, precedidos por el símbolo +, se utilizan en *FunGramKB* como unidades conceptuales definitorias que permiten construir los postulados de significado de los conceptos básicos y terminales, además de intervenir como preferencias selectivas en los marcos temáticos y predicativos.

Finalmente, en el nivel inferior de la ontología aparecen los conceptos terminales, los cuales van precedidos por el símbolo \$. La línea divisoria entre tratar un concepto como básico o terminal es decidir si ese concepto posee suficiente capacidad definitoria como para formar parte de algunos postulados de significado.

Los conceptos básicos y terminales comparten el mismo tipo de propiedades, entre las que destacamos la presencia de un postulado de significado y un marco temático, este último sólo atribuible a los conceptos de tipo evento. A modo de ejemplo, ilustramos estas propiedades con el concepto +EXPLAIN:

**Marco temático:**

+EXPLAIN\_00 (x1: +HUMAN\_00)Theme (x2)Referent (x3: +HUMAN\_00)Goal

**Postulado de significado:**

+(e1: +SAY\_00 (x1)Theme (x2)Referent (x3)Goal (f1: +CLEAR\_00)Manner  
(f2: (e2: +UNDERSTAND\_00 (x3)Theme (x2)Referent))Purpose)

### 3. LOS TIPOS COGNITIVOS DE LAS CUALIDADES

#### 3.1. *El análisis semántico de los adjetivos*

Tradicionalmente la convergencia de los parámetros del dinamismo, la gradación y la polaridad han servido para analizar semánticamente a los adjetivos. Sabemos que un adjetivo posee un determinado valor en cada uno de estos rasgos semánticos porque solemos aplicar tests léxico-sintácticos de sustitución a modo de diagnóstico (Cruse 1986). Por ejemplo, suponiendo que  $\alpha$  y  $\beta$  son las realizaciones léxicas de un nombre y un adjetivo respectivamente, podemos reconocer que un adjetivo es dinámico o no aplicando el siguiente test:

*$\alpha$  está siendo  $\beta$*

En otras palabras, un adjetivo dinámico (p.ej. *clever*) puede expresarse a nivel léxico-sintáctico junto a una estructura verbal en presente progresivo, ya que el referente nominal que posee la cualidad descrita por el adjetivo debe hacer un esfuerzo por mantenerla. Por el contrario, el adjetivo estático (p.ej. *tall, German*) expresa una cualidad que siempre se mantiene independientemente de los procesos que sufra el referente nominal.

Con respecto a la gradación, el referente del nombre calificado por un adjetivo graduable (p.ej. *good, expensive*) puede presentar la cualidad expresada por el adjetivo en diferentes grados. En cambio, para un adjetivo no graduable (p.ej. *dead, male*) sólo existen dos posibles realidades: el referente nominal posee o no esa cualidad. Los tests de sustitución para detectar la gradación suelen ser los siguientes, cuya validación nos indica que el adjetivo en cuestión es graduable:

*$\alpha$  es muy  $\beta$*   
 *$\alpha_1$  es más/menos  $\beta$  que  $\alpha_2$*

Finalmente, el tipo semántico de la polaridad permite establecer una oposición de significado con otros adjetivos, dando como resultado dos grandes grupos de adjetivos: aquéllos que establecen relaciones dicotómicas de antonimia (i.e. adjetivos polares, p.ej. *true-false*) y aquéllos que forman parte de series no binarias (i.e. adjetivos seriales, p.ej. *hot-warm-cool-cold, huge-big-small-tiny*). En otras palabras, si un adjetivo comparte con al menos otros dos adjetivos un mismo rasgo semántico, i.e. pertenecen a una misma dimensión semántica, entonces hablamos de adjetivos seriales. En cambio, el adjetivo es polar cuando la dimensión semántica está formada por sólo dos adjetivos, donde si uno de ellos es aplicable, entonces el otro no puede serlo, y viceversa. El test de sustitución para la polaridad tendría la siguiente forma, donde  $\beta_1$  y  $\beta_2$  son adjetivos:

*Si  $\alpha$  no es/está  $\beta_1$ , entonces  $\alpha$  es/está  $\beta_2$*

La validación de este test implica que tratamos con adjetivos polares; en caso contrario, los adjetivos implicados pertenecen a una serie.

#### 3.2. *La descripción tipológica de las cualidades*

En el modelo ontológico de *FunGramKB* aplicamos estos mismos parámetros (i.e. dinamismo, gradación y polaridad) en la descripción tipológica de las cualidades básicas y terminales, pero descartamos rotundamente el uso de los tests de validación por pertenecer

éstos al nivel de descripción léxica. En su lugar, tomamos el propio significado de las cualidades para determinar los valores de sus tipos cognitivos. Por tanto, los conceptos de dinamismo, gradación y polaridad se definen ahora desde un enfoque exclusivamente cognitivo, razón por la cual ya no hablamos de adjetivos y nombres sino de cualidades y entidades respectivamente.

De esta forma, una cualidad es dinámica siempre que la cualidad atribuida a la entidad pueda ser modificada en un momento determinado del tiempo gracias a la acción ejercida por la propia entidad a la que se le atribuye dicha cualidad o por una segunda entidad, la cual puede ser utilizada como instrumento por parte de la primera entidad, o bien esta segunda entidad puede cambiar la cualidad sin mediación de la primera entidad. Por ejemplo, el concepto +CLEVER describe una cualidad que puede ser poseída por una entidad si ésta realiza el esfuerzo debido. En cambio, la cualidad \$GERMAN—entendida como “que nació en Alemania”—no puede ser adquirida voluntariamente por una entidad, ni tampoco la entidad puede renunciar a esa cualidad una vez que la posee, ni tampoco ninguna entidad puede despojar a otra entidad de esa cualidad, por lo que hablamos de una cualidad estática.

Por otra parte, una cualidad es graduable (p.ej. +EXPENSIVE) sólo cuando puede presentarse en una misma entidad o en otras entidades en diferentes grados de intensidad.

Finalmente, *FunGramKB* describe las oposiciones de significado dentro del marco de una dimensión cognitiva, para lo cual determinamos un concepto focal positivo y un concepto focal negativo para cada dimensión. Los calificativos de *positivo* y *negativo* no denotan una zona semántica de connotación positiva y negativa respectivamente, sino la ausencia o no del operador de negación *n* al definir sus significados con respecto al otro. En otras palabras, definimos uno de los conceptos como negación de su opuesto. Por tanto, el focal positivo es aquel concepto que se utiliza para describir el focal negativo. Además, si existe gradación conceptual en una dimensión cognitiva, combinamos los conceptos focales positivo y negativo con los operadores *m* (muy) y *p* (poco), los cuales se introdujeron inicialmente como cuantificadores relativos de las entidades (Periñán y Arcas 2004), pero que con las cualidades sirven para indicar su grado de intensidad. Con estos tres operadores formales (i.e. *n*, *m* y *p*), podemos crear un máximo de siete zonas semánticas de gradación dentro de una misma dimensión cognitiva, donde una zona intermedia o *zona neutra* se origina cuando la definimos como la negación tanto del concepto focal positivo como del negativo (fig. 2):

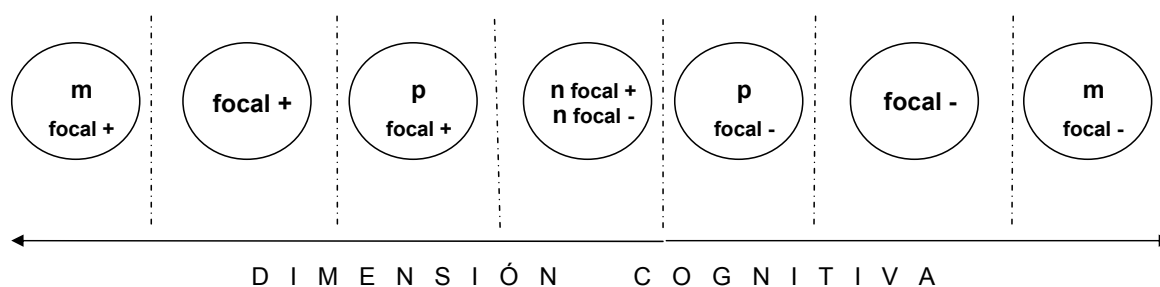


Figura 2. La estructuración canónica de una oposición de significado graduable.

En definitiva, toda oposición de significado implica la construcción de una dimensión cognitiva, con un mínimo de dos y un máximo de siete zonas semánticas cognitivamente factibles. La serie graduable de la temperatura sirve para ilustrar la estructuración canónica de la oposición de significado (fig. 3):

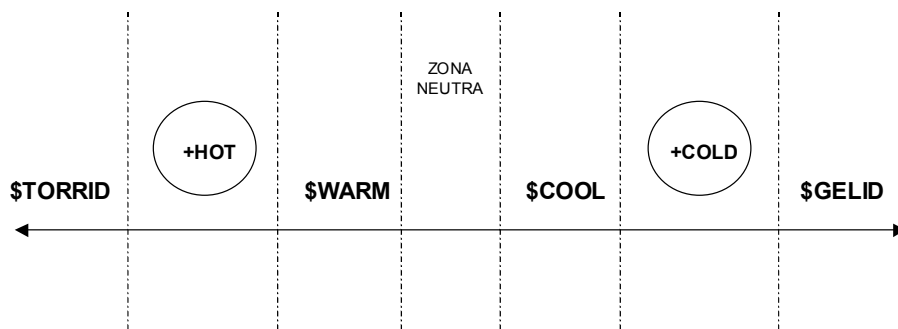


Figura 3. La dimensión cognitiva de la temperatura.

Recordemos que las zonas semánticas de una dimensión cognitiva se describen a través de unidades conceptuales interlingüísticas, cada una de las cuales puede estar más o menos lexicalizada en las diversas lenguas que cubre nuestra base de conocimiento. En *FunGramKB*, resulta muy útil diferenciar entre *factibilidad cognitiva* y *lexicalización* de una zona semántica cuando describimos una oposición de significado. Mientras la lexicalización implica la materialización de un concepto en forma de unidades léxicas, entendemos por factibilidad cognitiva el hecho de que una determinada zona semántica pueda ser susceptible de ser lexicalizada en una lengua, lo cual no implica que esa zona esté realmente lexicalizada. En la dimensión de la temperatura (fig. 3), la zona neutra no se lexicaliza en inglés o español, pero esto no significa que en cualquier otra lengua no pueda existir una unidad léxica cuyo significado corresponda a esta zona semántica, por lo cual decimos que esta zona semántica es cognitivamente factible.

Si nos guiamos por la semántica léxica tradicional, la polaridad implica que una determinada dimensión semántica esté formada únicamente por dos adjetivos, los cuales no pueden ser aplicados simultáneamente a un mismo nombre en un determinado punto en el tiempo (p.ej. *alive-dead*). En cambio, si esa misma dimensión está formada por más de dos cualidades, hablamos de una serie. A nivel cognitivo, el problema de aplicar esta definición tradicional de polaridad es que parte de la premisa de que las cualidades polares no pueden ser graduables, ya que si lo fueran, implicaría la existencia de otras franjas de intensidad en forma de zonas semánticas, lo cual constituiría una serie. En cambio, desde un punto de vista estrictamente cognitivo, oposiciones de significado como DRY-WET deben ser tratadas como polaridad graduable. En *FunGramKB*, la dimensión cognitiva de la humedad se estructura en diversos grados de intensidad, como mostramos en la siguiente representación gráfica (fig. 4):

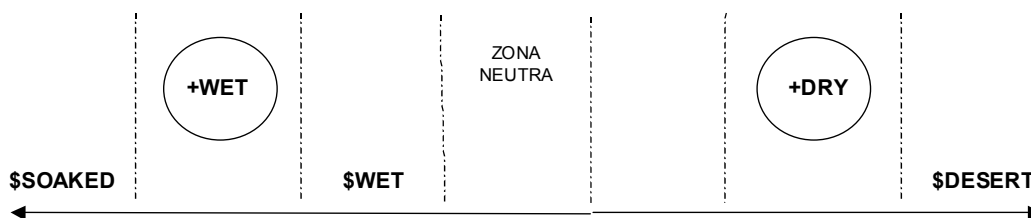


Figura 4. La dimensión cognitiva de la humedad.

A pesar de que las dimensiones de la temperatura (fig. 3) y la humedad (fig. 4) son muy parecidas estructuralmente, sólo en esta última existe una polaridad. Por tanto, en *FunGramKB* definimos la polaridad como la oposición de significado donde la negación de una cualidad en la dimensión cognitiva implica la afirmación de una cualidad ubicada en cualquier otra zona semántica en el lado opuesto de la misma dimensión. Por ejemplo, en el caso de la humedad, decimos que una entidad está húmeda o no si está o no impregnada de

líquido, independientemente de la cantidad de líquido absorbido. Por tanto, si la zona semántica neutra de una dimensión cognitiva no es factible cognitivamente, entonces hablamos de polaridad; en caso contrario, hablamos de serie, aunque la zona neutra no esté lexicalizada. Por tanto, la gradación no puede tomarse como un criterio determinante de la polaridad o la serialidad de una oposición de significado. A diferencia de lo que ocurría en la semántica léxica, la polaridad cognitiva deja de ser necesariamente una relación binaria *stricto sensu*.

### 3.3. Una aplicación de los tipos cognitivos de las cualidades

En *FunGramKB*, un postulado de significado es un conjunto de una o más predicaciones ( $e_1, e_2 \dots e_n$ ) conectadas lógicamente, donde definimos el término *predicación* como una estructura cognitiva que expresa algunos de los rasgos genéricos que configuran el *definiendum* conceptual (Periñán y Arcas 2004). A modo de ejemplo, presentamos algunas de las predicaciones del postulado de significado de +BIRD:

+BIRD

+( $e_1$ : +BE ( $x_1$ : +BIRD)<sub>Theme</sub> ( $x_2$ : +VERTEBRATE)<sub>Referent</sub>)

\*( $e_2$ : +HAVE ( $x_1$ )<sub>Theme</sub> ( $x_3$ : m +FEATHER & 2 +LEG & 2 +WING)<sub>Referent</sub>)

\*( $e_3$ : +FLY ( $x_1$ )<sub>Theme</sub>)

Las predicaciones anteriores se traducen al lenguaje natural de la siguiente forma:

Todos los pájaros son vertebrados.

Un pájaro típico tiene muchas plumas, dos patas y dos alas.

Un pájaro típico vuela.

Como muestra el ejemplo de +BIRD, *FunGramKB* expresa el significado de los conceptos básicos y terminales de su ontología por medio de otros conceptos básicos, convirtiendo así el postulado de significado en una representación interlingüística controlada del conocimiento semántico.

En el caso de las cualidades que intervienen en una oposición de significado, los tipos cognitivos (más concretamente, la gradación y la polaridad) no sólo se utilizan para determinar el patrón estructural de las predicaciones que componen el postulado de significado sino también pueden decidir el rango conceptual de la cualidad que definimos. Veamos algunos ejemplos:

i) +TRUE

+( $e_1$ : +BE ( $x_2$ )<sub>Theme</sub> ( $x_1$ : +TRUE)<sub>Attribute</sub>)

\$FALSE

+( $e_1$ : +BE ( $x_2$ )<sub>Theme</sub> ( $x_1$ : \$FALSE)<sub>Attribute</sub>)

+( $e_2$ : n +BE ( $x_2$ )<sub>Theme</sub> ( $x_3$ : +TRUE)<sub>Attribute</sub>)

ii) +BIG

+( $e_1$ : +BE ( $x_2$ )<sub>Theme</sub> ( $x_1$ : +BIG)<sub>Attribute</sub>)

+( $e_2$ : +BE ( $x_1$ )<sub>Theme</sub> ( $x_3$ : +SIZE)<sub>Referent</sub>)

+SMALL

+( $e_1$ : +BE ( $x_2$ )<sub>Theme</sub> ( $x_1$ : +SMALL)<sub>Attribute</sub>)

+(e<sub>2</sub>: +BE (x<sub>1</sub>)<sub>Theme</sub> (x<sub>3</sub>: +SIZE)<sub>Referent</sub>)  
 +(e<sub>3</sub>: n +BE (x<sub>2</sub>)<sub>Theme</sub> (x<sub>4</sub>: +BIG)<sub>Attribute</sub>)

#### \$HUGE

+(e<sub>1</sub>: +BE (x<sub>2</sub>)<sub>Theme</sub> (x<sub>1</sub>: \$HUGE)<sub>Attribute</sub>)  
 +(e<sub>2</sub>: +BE (x<sub>1</sub>)<sub>Theme</sub> (x<sub>3</sub>: +SIZE)<sub>Referent</sub>)  
 +(e<sub>3</sub>: +BE (x<sub>2</sub>)<sub>Theme</sub> (x<sub>4</sub>: m +BIG)<sub>Attribute</sub>)

#### \$TINY

+(e<sub>1</sub>: +BE (x<sub>2</sub>)<sub>Theme</sub> (x<sub>1</sub>: \$TINY)<sub>Attribute</sub>)  
 +(e<sub>2</sub>: +BE (x<sub>1</sub>)<sub>Theme</sub> (x<sub>3</sub>: +SIZE)<sub>Referent</sub>)  
 +(e<sub>3</sub>: +BE (x<sub>2</sub>)<sub>Theme</sub> (x<sub>4</sub>: m +SMALL)<sub>Attribute</sub>)

Los postulados de significado en (i) sirven para ilustrar una polaridad no graduable, donde el concepto focal negativo (\$FALSE) se describe como la negación del concepto focal positivo (+TRUE). Además, mientras el focal positivo se representa como un concepto básico, el focal negativo siempre se presenta aquí como terminal al no existir gradación en la polaridad. Por otra parte, los postulados de significado en (ii) sirven para ilustrar una serie graduable, donde sólo los focales positivo y negativo pueden ser conceptos básicos, ya que éstos se utilizan para describir los otros conceptos de la dimensión cognitiva. De esta forma, minimizamos la redundancia informativa en la ontología, p.ej. evitando la circularidad al no definir el focal positivo como la negación del focal negativo, y mantenemos la consistencia de las descripciones de significado de las cualidades, p.ej. describiendo todas las oposiciones de significado de características similares con postulados de significado estructurados de forma similar.

#### 4. CONCLUSIONES

A diferencia de otras bases de conocimiento para el PLN, *FunGramKB* se diseñó bajo una sólida fundamentación cognitiva. En este sentido, el enfoque conceptualista de los parámetros del dinamismo, la gradación y la polaridad nos permiten construir una base de conocimiento más robusta. Con este fin, hemos desarrollado una serie de herramientas para la gestión de los niveles léxico y cognitivo de *FunGramKB*, las cuales no sólo ayudan a representar visualmente una extensa cantidad de datos complejos, sino también garantizan el desarrollo consistente y uniforme de nuestra base de conocimiento.

#### NOTAS

1. Actualmente, *FunGramKB* da cobertura a unidades léxicas del alemán, español, francés, inglés e italiano.
2. En este artículo, los nombres de los conceptos son representados en mayúsculas, mientras que las unidades léxicas se presentan en minúsculas.

#### BIBLIOGRAFÍA

- Cruse, D. A. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Nirenburg, S., S. Beale, K. Mahesh, B. Onyshkevych, V. Raskin, E. Viegas, Y. Wilks y R. Zajac. 1996. "Lexicons in the MikroKosmos Project". *AISB'96 Workshop on Multilinguality in the Lexicon*. Brighton, 1-2 April 1996.
- Periñán Pascual, C. y F. Arcas Túnez. 2004. "Meaning postulates in a lexico-conceptual knowledge base". *15<sup>th</sup> International Workshop on Databases and Expert Systems Applications, Zaragoza, 30 August-3 September 2004*. Los Alamitos, CA: IEEE. 38-42.