

Generación de definiciones para un diccionario escolar de la sexualidad

Generation of definitions for a school dictionary of sexuality

Esteban Piedra Matamoros.¹

Fecha de recepción: 25- 06- 2019

Fecha de aceptación: 12-4-2021

Resumen

El objetivo de este artículo es realizar un aporte desde la Ingeniería Lingüística por medio de una herramienta que genere definiciones automáticamente a partir de información conceptual, contextos de uso y patrones lingüísticos. La posibilidad de extraer información de un corpus con marcas que indiquen si es información académica, coloquial o de otro orden, permite clasificar sus contextos definitorios y generar automáticamente diferentes tipos de diccionarios y glosarios dirigidos a distintas audiencias y con diferentes propósitos. Este trabajo nace gracias al apoyo durante la creación de una herramienta metodológica que selecciona la información necesaria para la conformación de un corpus textual recopilado de Internet, que tiene como resultado la confección de un diccionario especializado de la sexualidad para el grupo de escolares mexicanos. Con esto, se pretende crear nuevos documentos que motiven un impacto en la reducción de enfermedades de transmisión sexual, embarazos no deseados y violencia sexual en este segmento de la población.

Palabras clave: Ingeniería lingüística, generación de definiciones, corpus textual, diccionario, glosario.

Abstract

The joint work of disciplines such as Computer Science and Linguistics has resulted in the innovation of technological tools for searching, selecting and organizing written information found on the Internet. The aim of this article is to make a contribution from Linguistic Engineering through a tool that generates definitions automatically from conceptual information, contexts of use and linguistic patterns. The possibility of extracting information from a corpus with marks that indicate whether it is academic, colloquial or other information allows us to classify their defining contexts and automatically generate different types of dictionaries and glossaries aimed at different audiences and with different purposes. This work arose thanks to the support provided by the creation of a methodological tool selects the necessary information for the conformation of a textual corpus compiled from the Internet, which has resulted in the preparation of a specialized “dictionary of sexuality” for a group of Mexican schoolchildren. With this dictionary, we intend to create new materials that lead to the reduction of sexually transmitted diseases, unwanted pregnancies and sexual violence in this segment of the population.

Key words: Linguistics engineering, generation of definitions, textual corpus, dictionary, glossary.

¹ Master en lingüística, Universidad de Costa Rica, Sede de Occidente, Costa Rica. Correo electrónico: estepima@gmail.com

I. Planteamiento del problema

Parte del reto al que se enfrentan los diferentes campos del saber en cuanto a la actualización de la información y al mejoramiento de procesos reside en contar con un diccionario especializado como una herramienta, pues es imprescindible no solamente para profesionales y expertos, sino también para intérpretes y traductores del discurso, en cualquier disciplina científica.

La evolución continua en la utilización de los medios de comunicación y en la aplicación de tecnologías ha traído como consecuencia la estandarización de la terminología utilizada en las ciencias. Una manera de compilar términos utilizados por autoridades en cualquier campo, más los términos que utiliza la población general para referirse a los mismos conceptos es por medio de corpus lingüísticos. En los últimos años, la creación de diccionarios de especialidad se ha desarrollado de la mano de la evolución de la terminología, entendida como la aplicación de herramientas informáticas en la gestión y administración de términos (Alcina, 2001).

Aunque el esfuerzo realizado en el mejoramiento de estas herramientas informáticas ha sido vasto, todavía hay muchos problemas que deben ser abordados por acercamientos teóricos posteriores. Los futuros repositorios terminológicos deben contener el conocimiento de manera más estructurada y sistemática, de manera que éste sea comprensible, tanto para humanos como para ordenadores (Alcina, 2001).

Con el fin de favorecer el uso de estos dispositivos, estos diccionarios deben contar con una herramienta que genere definiciones automáticamente a partir de información conceptual, contextos de uso y patrones lingüísticos. Además, de acuerdo con la cantidad monumental de información accesible en Internet, una problemática es establecer una estrategia de selección de corpus, el cual será destacado con el fin de obtener las definiciones de los términos y sus contextos de uso. Las definiciones de un diccionario especializado deben sustentarse y alimentarse de otras definiciones de diccionarios que gocen de prestigio académico y científico.

La problemática que evidencia el presente trabajo radica en la creación de una metodología para extraer conocimiento lexicográfico de un saber especializado como lo es la educación sexual y para una población específica como lo es el grupo de escolares mexicanos, basado en información de textos recopilados de Internet.

II. Justificación

El presente trabajo se justifica desde un marco de apoyo para la creación de una herramienta metodológica que seleccione los datos necesarios para conformar corpus textuales recopilados de Internet, lo cual tendrá como resultado la confección de un diccionario especializado de la sexualidad para el grupo de escolares. La utilidad de este trabajo se evidencia en realizar un aporte al impulso de la educación sexual en el grupo de escolares mexicanos. Con ello, se pretende crear nuevos documentos que motiven un impacto en la reducción de enfermedades de transmisión sexual, embarazos no deseados y violencia sexual en este segmento de la población.

Lo anterior debido a que produce mejores resultados el procesamiento del lenguaje en ámbitos más delimitados como la sexualidad y para una franja etaria específica como lo es el grupo de los escolares. El número y la pertinencia de las acepciones en los diccionarios será menor y más acertado en un diccionario de terminología especializada que cuando toma como base el lenguaje en general.

En el ámbito terminológico sobre la educación sexual en México, se destaca que no existe un consenso en acercamientos sobre esta temática, además de que hay pocos glosarios y vocabularios con la terminología referente a este campo (Reyes-Careaga, Medina Urrea, Sierra Martínez, 2011). Aunado a esto, es posible afirmar que el tema de la sexualidad ha sido tratado a lo largo de los años como tabú por parte de los organismos responsables de la educación y formación de jóvenes en el país.

En el siguiente ejemplo, se define el concepto “pene”, tomado de dos diccionarios diferentes como lo son: 1. Diccionario Escolar de la Real Academia Española (1996)

y 2. Diccionario de la Real Academia Española, en su versión vigésima segunda (2001). Lo anterior con el fin de comparar las definiciones lexicográficas y su pertenencia según el público para el que son dirigidos. El primer ejemplo es tomado del Diccionario Escolar y el segundo del Diccionario de la Real Academia.

1. pene. m. miembro viril. (p. 853).
2. pene. (Del lat. *penis*). m. *Anat.* Órgano masculino del hombre y de algunos animales que sirve para miccionar y copular. (p. 1721).

De lo anterior, se destaca que, por un lado, la definición que se extrae del Diccionario Escolar de la Real Academia Española gana en simplicidad de estructura, aunque posiblemente el escolar necesite una guía y acompañamiento para entender el concepto a cabalidad. Por otro lado, el término del Diccionario de la Real Academia se define de manera más compleja, debido a la generalidad de su público meta. Los ejemplos anteriores contribuyen en destacar la necesidad de formar una estrategia de establecimiento de parámetros para definir términos, según las exigencias específicas de la población en estudio.

El tema de esta investigación es la educación sexual humana para el grupo de escolares de México. Por su importancia en la salud pública y por ser un tema tan complejo, requiere de atención terminológica especializada. Los esfuerzos educativos en el campo de la sexualidad deberían orientarse a que los estudiantes reconozcan su desarrollo sexual como parte de su trayecto en la búsqueda de identidad, de sentido de la vida, y de personalización, reconociéndose como un ser sexuado sin angustias ni miedos.

De acuerdo con la escogencia de textos de Internet de forma sistemática, se explica el proceso lingüístico-científico que se sigue para la extracción de términos y conceptos, con el fin de sentar las bases para construir metodologías en la creación de herramientas que ayuden en la elaboración de diccionarios, extractores conceptuales, redes semánticas, ontologías, etcétera.

La extracción terminológica dentro de la labor lexicográfica es la recopilación de material de trabajo en un principio para luego seleccionar qué material es importante de todo el material recopilado. El trabajo lexicográfico agrupa terminología especializada de cualquier área de especialidad, estructura tal información y agrupa los fragmentos textuales por categorías. Además, delimita y define los usos y particularidades de ciertos tipos de palabras necesarias para la adecuada explotación de recursos lingüísticos disponibles. Son procesos de estructuración semántica bajo criterios de organización léxica especializada.

El trabajo que se presenta debe contemplar diferentes pasos como: estrategia de escogencia de textos y conformación de corpus, extracción terminológica a partir de ese corpus o de la red, reconocimiento de contextos definitorios, aplicación de herramientas de análisis y agrupamiento conceptual de la información por categorías. Esa estructura de trabajo es aplicable a la mayoría de áreas temáticas y proporciona una metodología unificada para la obtención de información relevante a partir de Internet con la ayuda de herramientas computacionales. Actualmente, estas herramientas se encuentran en proceso de desarrollo. Algunos ejemplos son *ECODE* y *DESCRIBE*, las cuales son, según Medina Urrea (2008): “motores de búsqueda conceptual que tienen como fin agrupar diversas definiciones de un término a partir de la información que suministra la red” (p. 8).

Esta investigación toma en cuenta que este tipo de trabajos necesitan una actualización de términos y una adecuación de registros continuas. Es claro entonces, el aporte que ciencias como la Ingeniería Lingüística y Lingüística computacional realizan al desarrollo de los campos del saber, en este caso particular, al trabajo y educación de la sexualidad humana.

III. Objetivo

Es fundamental que el estudio y trabajo en cualquier área del saber, y en este caso particular de la sexualidad humana, cuente con una herramienta actualizada que permita consensos, integración y claridad en las definiciones especializadas con el fin de procurar un

entendimiento común. Al ser esta una temática polémica en países de tradición conservadora, es necesaria la validación de un instrumento riguroso que cuente con una terminología armoniosa y sistematizada. Con este fin, la Organización Mundial de la Salud (OMS) ha buscado opciones en pro de la buena comunicación en materia de salud sexual en México.

El objetivo de este trabajo es realizar un aporte al esfuerzo de sistematización de la terminología de la sexualidad para una población con una franja etaria específica como lo es el grupo de escolares mexicanos, con el fin de facilitar la comunicación sobre este tema. Para ello, es indispensable la obtención de un corpus terminológico de los conceptos referidos a la sexualidad en México, construcción del sistema conceptual correspondiente a la terminología de la sexualidad, construcción de la base terminológica de las sexualidades en México, revisión del contenido de la base de datos desde un punto de vista terminológico y sexológico; y elaboración de un diccionario electrónico accesible al público en general.

Debido a la naturaleza polémica del tema, es necesario hacer sesiones de trabajo para sensibilizar al equipo profesional y llegar a acuerdos unificando criterios y mejorando habilidades en el uso del programa. Para la realización de este proyecto es necesario un equipo humano compuesto por ingenieros lingüísticos que aporten conocimiento en cuanto a la creación de diccionarios electrónicos por medio de búsquedas inteligentes, terminólogos expertos en la elaboración de diccionarios especializados y expertos en el área de sexología que validen y revisen el contenido técnico del diccionario.

IV. Metodología

Con certeza, uno de los usos más importantes que las personas pueden hacer de Internet es realizar consultas sobre lenguaje especializado en cualquier campo del saber. Los usuarios siempre optarán por obtener información que provenga de fuentes prestigiosas y de confianza. Por esa razón, es fundamental considerar los aportes que existen en los diccionarios académicos, artículos científicos, tesis y textos escritos por especialistas en el

campo, además de cualquier otro texto que se considere relevante en el tema.

Para efectos del presente trabajo, se entiende el lenguaje especializado como los subcódigos usados por los hablantes y que escogen según las necesidades expresivas y las características particulares del contexto comunicativo. De acuerdo con lo anterior, se rescata que el hombre necesita una comunicación especializada con otros que compartan un conocimiento similar, para asegurarse la correcta realización de cierto tipo de labores. Para esto, es necesario desarrollar un mecanismo de adquisición de léxico en algún área, definir su conformación en bloques que permitan la utilización de ese léxico en los contextos adecuados e implementar la selección y agrupación de conceptos.

El primer paso del trabajo es realizar un aporte en el diseño de una estrategia de selección de corpus que se encuentra en la red, de acuerdo con estudios previos realizados por expertos en la materia de sexualidad. En este caso particular, esa labor consiste en reunir si no todo, gran parte del material textual disponible en Internet que aborde el tema de la sexualidad en México. Ese corpus debe pasar, en primera instancia, por un proceso de estandarización de formato y de tipo de información antes de someterse a análisis.

Es fundamental que la selección del corpus sea una muestra representativa de las fuentes, ya que esta es una manera práctica y viable de conocer la terminología de un área en particular. La extracción de la información es entendida como el núcleo principal de las actuales tecnologías del lenguaje, de ahí que resulte necesario trabajar con sistemas de cómputo capaces de buscar, localizar y brindar información relevante de cualquier tipo a un usuario (Wilks, citado por Sierra, 2009).

La información recopilada debe estar en formato electrónico y en una codificación uniforme para que puedan ser eficaces para el investigador. Posteriormente, esta información será clasificada en bloques de conformidad con sus características, desde textos altamente especializados como diccionarios, tesis académicas, artículos especializados y libros de

investigación científica hasta textos provenientes de foros médicos o foros de difusión con acceso a toda la población. Además, se contemplaría cualquier otro tipo de información que la estrategia de selección permita recopilar.

La extracción automática de definiciones mediante el reconocimiento y análisis de patrones lingüísticos a partir de textos de especialidad se ha convertido en un punto clave en el desarrollo de disciplinas encargadas del Procesamiento de Lenguaje Natural. La posibilidad de extraer información especializada con datos que reflejen si son confiables académicamente, o sin son de carácter coloquial, o de otro tipo, permite clasificar los contextos definitorios según estos rasgos y generar semiautomáticamente diversos tipos de glosarios y diccionarios para diferentes audiencias y fines. Este es el principal objetivo de las llamadas ingenierías del conocimiento (Reyes-Careaga, Medina Urrea, Sierra Martínez, 2011).

Posteriormente, y de acuerdo con lo planteado por el doctor Medina Urrea (2008) en acercamientos teóricos anteriores, se etiquetan los textos en *XML (Extensible Markup Language)* con la herramienta *TOK* previamente desarrollada por un grupo profesional. Esta etiqueta permite codificar un texto para que cualquier persona pueda consultarlo desde su ordenador. El etiquetador tiene la libertad de utilizar las marcas que más se acoplen a su forma de trabajo o a la teoría que maneje. Por otro lado, *TOK* es un pequeño programa que permite crear etiquetas vacías que *XML* puede leer y modificar. El producto esperado es una metodología para la generación automática de definiciones con el fin de obtener un diccionario onomasiológico de la sexualidad para mexicanos escolares. Para realizar un diccionario de esta naturaleza, es necesario contar con una base de conocimientos léxica lo suficientemente rica que contenga una diversidad de definiciones para cada uno de los términos que se está buscando.

La organización terminológica consiste en aplicar análisis lingüísticos a los elementos del corpus para determinar si son o no del área de estudio. Para tal fin, se usa la herramienta *Wordsmith*, que es un programa de

extracción y organización que obtiene las palabras más frecuentes, su concordancia y su posición en el texto. Eso dará como resultado una lista de términos del corpus. Una vez que se tiene la lista, se ordenan por frecuencia y por peso. Por frecuencia se entiende la aparición de la palabra y sus derivados; y por peso la representatividad del elemento en la constitución general del corpus no solo por su frecuencia de aparición, sino por su posición y su relación con otros términos. Posteriormente, se separan los vocablos de acuerdo con la subárea a la que pertenecen. La intención es que cada una de las subáreas posea un número parecido de integrantes, con el fin de hacer de la metodología un procedimiento más confiable.

En resumen, la extracción terminológica debe contemplar diferentes etapas. Primero, se realiza una lista de palabras plenas; para esto se eliminan palabras vacías como conjunciones, determinantes y preposiciones. Estos componentes se extraen del corpus por medios estadísticos automáticos con el uso de herramientas como *Wordsmith*. Posteriormente, se realiza una lista de palabras plenas por frecuencia de aparición con el fin de establecer candidatos a términos. Este paso determinará el peso y *ranking* del concepto.

Búsqueda de definiciones:

En el caso del sistema *DESCRIBE*, este debe tener como entrada el término que se quiere consultar y conocer con más profundidad y exactitud. La plataforma funciona de la siguiente manera:

- El usuario escribe una palabra o frase en la barra de búsqueda.
- El sistema despliega un resumen de los resultados obtenidos.
- El usuario escoge alguna de las tres opciones desplegadas, según su necesidad.
- El sistema muestra una lista de definiciones encontradas del término, no solamente de diccionarios, sino de textos publicados en internet con sus respectivas referencias URL.

- Si el usuario se interesa por alguna definición específica puede consultar el sitio web de donde se obtuvo la información.
- El usuario podrá consultar los documentos que haya elegido. De esa forma la búsqueda concluye satisfactoriamente.

DESCRIBE permite obtener y ampliar el conocimiento de los usuarios en relación con terminología específica. Es una herramienta para especialistas en lengua interesados en Internet como generador de corpus, y permite que Internet tenga otra herramienta de búsqueda más específica y especializada. Es necesario recordar que la plataforma *DESCRIBE* se encuentra en pleno desarrollo y continua mejora. Todavía no ha sido posible ubicar y analizar la totalidad de textos en la red. En la parte del etiquetado, sería necesario destacar la importancia de contar con etiquetas adicionales para trabajar conceptos más especializados. En el caso de las áreas temáticas, ampliar el repertorio para distribuir los textos que se encuentran etiquetados como pendientes. Asimismo, se deben explorar métodos para asignar documentos a varias áreas temáticas, pues es posible que un documento haga referencia a más de un tema (Reyes-Careaga, Medina Urrea, Sierra Martínez, 2011).

En conclusión, la posibilidad de extraer información de un corpus con marcas que indiquen si es información académica, coloquial o de otro orden, permite clasificar sus contextos definitorios y generar automáticamente diferentes tipos de diccionarios y glosarios dirigidos a diferentes audiencias y con diferentes propósitos. Es importante recalcar que estas herramientas no son todavía productos terminados. Al contrario, existe un grupo de personas profesionales en diferentes campos del saber realizando esfuerzos con el fin de mejorar el funcionamiento de las plataformas *DESCRIBE* y *ECODE*.

V. Marco teórico

La base teórica de la presente investigación se apoya en conocimientos provenientes de una rama incipiente de la lingüística como lo es la lingüística computacional o ingeniería lingüística, o desde un paradigma más moderno, ingeniería del conocimiento. De acuerdo con

lo anterior, *terminótica* es la aplicación de herramientas informáticas en la gestión terminológica (Alcina, 2001). Realizar un aporte en la creación de la norma y ser fuente de otras obras son ejemplos de la importancia de contar con definiciones de calidad en cualquier ámbito de la ciencia. En los últimos años, la elaboración de diccionarios de especialidad ha ido de la mano con la revolución tecnológica, pero la calidad de los programas de análisis textual y gestores de bases terminológicas y de conocimiento lamentablemente no han aumentado al mismo ritmo que los avances de la ciencia y la tecnología.

El incipiente campo de la ingeniería lingüística ha obtenido muy buenos resultados en cuanto a la extracción terminológica, no así en cuanto a la extracción de conceptos, proceso que representa un reto más complejo, como consecuencia del paradigma entre una visión normativa sostenida por modelos de autoridades académicas en la elaboración de diccionarios y una visión que contemple aspectos comunicativos y cognitivos subyacentes en la elaboración de conceptos. Se entiende el término como una unidad de significación especializada que cuenta con rasgos particulares como categoría morfológica, capacidad referencial y normativa, así como un significado especializado. En contraste, el concepto puede ser visto como una unidad de conocimiento abstracto con una serie de rasgos propios de un objeto, evento o relación con el fin de situarlo en el mundo (Sierra, 2009). Esta unidad es representada por una definición, la cual se constituye, según el concepto aristotélico de dos elementos básicos: un género próximo y una diferencia específica. El *genus* es un descriptor que hace referencia a la clase a la cual pertenece un objeto y la *differentia* es la serie de rasgos propios que distinguen a dicho objeto o evento respecto de otros agrupados en esta misma clase.

Para los lexicógrafos, el acto de definir términos involucra dos aspectos complementarios: optar por una concepción de significado y explicarlo por medio de una formulación lingüística. En ese sentido, una teoría semántica puede ofrecer información importante que contribuya en ambos procesos. En lo referente al acceso del contenido semántico de las unidades léxicas, dicha teoría semántica determinaría bajo cuál perspectiva se debe indagar el problema. En lo que respecta al proceso de expresar el

contenido mediante una formulación lingüística, la teoría semántica contribuye a seleccionar cuáles informaciones se actualizarán en la definición (Bugueño y Sita, 2013).

Se asume la hipótesis que toda definición está constituida por una dimensión intensional y otra extensional. En toda definición debe haber una dimensión que se refiera al contenido de una unidad léxica, al mismo tiempo que esta se proyecta sobre la realidad infinita en la medida que conviene o no a determinadas entidades del mundo. De acuerdo con tres teorías semánticas analizadas, se destaca que la distinción que hace la semántica de las condiciones de verdad sobre los conceptos extensión-intensión, se reconoce el valor real de la extensión en la formulación de las definiciones; por otro lado, evidencia las dificultades de la categorización. Como lo señalan Buñuego y Sita (2013):

El modelo formal es capaz, por ejemplo, de describir de forma satisfactoria el concepto *ave* desde el punto de vista intensional, mediante un conjunto de condiciones necesarias y suficientes: «ser un animal», «tener pico», «tener el cuerpo cubierto de plumas», «tener alas», «poner huevos». Sin embargo, por medio de estos rasgos, no logra determinar la categoría de la palabra *ave*, dado que no se aplican de forma homogénea al conjunto de todas las aves existentes (p. 31).

La aplicación del análisis componencial ha sido catalogada por la semántica estructural como poco eficaz y limitado. Este análisis solo se aplica a las palabras con significado lexical (lexemas) y aun así, deja por fuera el estudio de algunas de ellas, como por ejemplo, las especies naturales y los colores. La semántica prototípica incluye el componente extralingüístico o enciclopédico en el significado, aspecto considerado como una contribución al estudio lexicográfico. Sin embargo, no es posible evaluar la eficacia de las consideraciones lingüísticas en la definición, ni tampoco cuándo habría que usarlas.

En ese sentido, la definición terminológica es una fórmula lingüística que tiene como propósito realizar una detallada descripción del concepto que una denominación representa. Se diferencian tres tipos de definición: lingüística, ontológica y terminológica. Esta

última describe el concepto en función de un campo de especialidad, a diferencia de la definición lingüística que parte en relación al sistema lingüístico en general. Por otro lado, la ontológica incluye todos los aspectos característicos de un concepto.

En ese mismo orden, Felber (1984) distingue dos tipos de definiciones terminológicas; las *intensionales*, las cuales enumeran las características propias del concepto, exponiendo el hiperónimo o superordinado en el sistema conceptual y enumerando las características que lo distinguen del resto de términos; y las *extensionales*, que consisten en la enumeración exhaustiva de todas las especies que se encuentran al mismo nivel de abstracción o de todos los objetos individuales que pertenecen al concepto definido (p. 163). De acuerdo con lo anteriormente expuesto, se destaca que la mayoría de las definiciones que se encuentran en los diccionarios son intensionales, debido a la dificultad de delimitar completamente las jerarquías conceptuales, de manera que se incluyan todas las especies de un grupo.

La definición terminológica debe contemplar un acercamiento que haga referencia al concepto genérico y además, debe proporcionar las características que lo diferencian con el resto de conceptos del mismo grupo. Ndi-Kimbi, citado por Alcina (2001), manifiesta que:

En el año 1916, Josephs estableció la estructura de la definición terminológica bajo la siguiente fórmula: *Definiendum* = *genus* + *differentiae*. *Definiendum* es el concepto que queremos definir, *genus* es el concepto genérico o descriptor del que el *definiendum* hereda las características y *differentiae* son las características que los distinguen con el resto de conceptos englobados en la extensión del descriptor (p. 2).

De acuerdo con lo anteriormente expuesto, las definiciones deben cumplir con las siguientes condiciones:

- Expresión adecuada.
- Principios lexicográficos.
- Adecuación general.
- Adecuación específica.

Un patrón de definiciones está formado por el hiperónimo de un grupo de conceptos y las características esenciales que describen a los conceptos de ese grupo. Estos patrones se podrían usar para generar las definiciones de una manera semiautomática (Alcina, 2001). El resultado de esta aplicación de patrones será un sistema que recoja las dimensiones pertinentes dentro de cada concepto dentro del campo de la especialidad que permita ver el sistema conceptual subyacente y en el que los términos se empleen de manera unificada. El patrón tendrá los siguientes elementos:

1. Descriptor: concepto más amplio que el que estamos definiendo. Es el hiperónimo. Procede de la estructuración jerárquica previa del campo de especialidad. Ejemplo: el descriptor *defecto* se usa en todos los términos englobados en la rama “defectos del producto acabado”.
2. Nombres de las características: cada concepto se define mediante un grupo de características determinadas a partir de un análisis conceptual previo.
3. Valores de las características: en algunos casos, los conceptos tienen una lista de valores identificados, en otros casos no.
4. Expresión lingüística.

VI. Conclusión

La tarea de elaborar definiciones terminológicas es compleja. Requiere de un análisis formal y sistemático del concepto con el fin de encontrar relaciones con otros elementos e identificar sus características esenciales. Como solución a dicho problema, se propone una metodología de elaboración de definiciones de búsqueda automática basada en el reconocimiento de patrones de definición, el cual es un modelo que seguirán todas las definiciones de un mismo grupo conceptual. Dicho patrón servirá de plantilla al terminólogo y garantizará un sistema de definiciones basado en principios ontológicos con una base coherente, compleja y conforme a los principios establecidos para tal fin.

Se considera relevante el aporte que se pueda realizar en cuanto a la confección de un diccionario para un grupo de escolares, especialmente si está enfocado en temas clave para el desarrollo de valores en nuestras sociedades. En este caso la sexualidad, al ser históricamente un tema caracterizado por la desinformación, ha traído de forma colateral consecuencias negativas en nuestras sociedades latinoamericanas, problemas asociados a la violencia y la ignorancia. El acceso al conocimiento en temas de sexualidad saludable por parte de los jóvenes garantiza una sociedad más sana, capaz de enfrentar la educación de las nuevas generaciones. Este esfuerzo es importante de reconocer con el fin de ser replicado en Costa Rica. Lo anterior para seguir por la senda de la información y educación escolar.

VII. Referencias bibliográficas

- Alcina, A. Valero, E. (2001). *Análisis de las definiciones del diccionario cerámico científico-práctico. Sugerencias para la elaboración de patrones de definición*. Universidad de Jaume I.
- Bugueño Miranda, F., Sita Farias, V. (2013). Los subsidios de tres teorías semánticas para la generación de definiciones. En *Revista de Filología Universidad de Laguna*. Tenerife, España. Vol. 31, p.p. 19-34.
- Felber, H. y Heribert Picht. (1984). *Métodos de terminografía y principios de investigación terminológica*. Madrid: Instituto Miguel de Cervantes-CSIC.
- Medina Urrea, A. (2008). *Extracción de conocimiento lexicográfico a partir de textos en internet*. México: Universidad Autónoma de México.
- Miranda, M. (1990). *Guía para la educación sexual escolar*. Chile: Alborada, S.A.
- Real Academia Española. *Diccionario de la Lengua Española*. (2001). Vigésimo segunda edición. Madrid: Espasa Libros, S.L.U.

Real Academia Española. *Diccionario Escolar*. (1996).
Primera edición. Madrid: Espasa Calpe, S.A.

Reyes-Careaga, T. A.; Medina Urrea, A.; Sierra Martínez,
G. E. Un corpus para la investigación en la
extracción de términos y conceptos definitorios:
hacia un diccionario de las sexualidades desde
México. *Debate terminológico*. No. 07, abril 2011.;
pp. 24-35.

Sierra, G. (2009). Extracción de contextos definitorios en
textos de especialidad a partir del reconocimiento
de patrones lingüísticos. En *Linguamática*. (2):
19-37.