

Análisis de la confluencia entre término y descriptor en la elaboración de tesauros

CATALINA NAUMIS PEÑA

Centro Universitario de Investigaciones Bibliotecológicas
de la UNAM, 04510, México D.F., Tel: (525)623-03-29

E-Mail: naumis@servidor.unam.mx

RESUMEN

Se confronta la función que desempeña el término para la comunicación científica, y el descriptor para la indización y obtención de información, así como las características de uno y otro, y las relaciones entre ellos. El término se observa en el marco de la terminología y el descriptor en el proceso de organización del conocimiento, el término es utilizado para otorgarle al descriptor, claridad y precisión en la comunicación cuando se fija su contenido en la elaboración de un tesauro; el descriptor adquiere una articulación flexible, al combinarse con otros descriptores para la recuperación de información.

Palabras Clave: Terminología, Descriptor, Tesauros

AN ANALYSIS OF THE CONFLUENCE BETWEEN TERM AND DESCRIPTOR IN THE ELABORATION OF THESAURUSES

CATALINA NAUMIS-PEÑA

ABSTRACT

The paper examines the role of the term used in scientific communication and the descriptor employed for indexing and reference, as well as the characteristics of each and their interrelationship. The term is observed within the framework of Terminology, while the descriptor is observed as lying within the process of organization of knowledge. The term, when its content is fixed in the elaboration of a thesaurus, is used to provide the descriptor clarity and accuracy in communication. The descriptor, on the other hand, acquires flexibility as it works in combination with other descriptors to aid in the collection of information.

Key Word: Terminology, Descriptors, Thesaurus

Trabajo recibido el
24 de noviembre
de 2000

*

Trabajo
aceptado el
10 de enero de 2001

La terminología es una disciplina que se ocupa del estudio de los términos especializados en un área del conocimiento para alcanzar un mejor entendimiento entre quienes la practican. En el marco de la bibliotecología se desarrollan vocabularios controlados o documentales, compuestos por descriptores, con el objeto de lograr una mejor organización del conocimiento registrado y de facilitar su posterior recuperación. Entre estas dos disciplinas existe una confluencia sobre la que

vale la pena reflexionar, el descriptor, unidad lingüística que permite indizar los documentos.

Este trabajo parte del supuesto de que el término y el descriptor son unidades lingüísticas con particularidades que los distinguen entre sí y que son propias de la función cumplida por ellos en el contexto de los lenguajes especializados.

TERMINOLOGÍA Y TÉRMINOS

En la actualidad, la terminología como disciplina está en un proceso de revisión o ampliación de los principios que rigen la Teoría General de la Terminología, cuya finalidad es la normalización conceptual y denominativa de los términos y el estudio de las unidades terminológicas en el marco de la comunicación especializada. En este sentido, expresa Teresa Cabré,

[...]Las unidades terminológicas son en sí mismas lenguaje y conocimiento, y su finalidad primera es la de representar y transferir este conocimiento; constituyen por tanto la materia misma de los términos.¹

La Teoría General de la Terminología concibe al lenguaje especializado como un subconjunto del lenguaje natural. El término es la unidad terminológica de un subconjunto de unidades terminológicas que reflejan los rasgos de un lenguaje especializado, a través de la limitación de esas unidades terminológicas en sus aspectos pragmáticos y semánticos, los cuales le otorgan validez en un ámbito específico del conocimiento y su desincorporación de las palabras del lenguaje general. Al respecto, Sager expresa:

[...] Los elementos caracterizados por una referencia especial dentro de una disciplina son los *términos* de esa disciplina, y colectivamente forman su *terminología*, aquellos que funcionan como referentes generales dentro de una variedad de sublenguajes simplemente se llaman *palabras* y la totalidad forma el *vocabulario*.²

La finalidad de esa limitación en el ámbito de un lenguaje especializado y sus términos es evitar la superposición de conceptos y buscar la disminución de la flexibilidad que le otorga el lenguaje natural a la unidad terminológica. En el lenguaje natural la polisemia ayuda a la memoria porque a veces una palabra favorece la economía y la flexibilidad del lenguaje, y es el contexto el que ayudará a ubicar el uso de la palabra.³ Sin embargo, antes de continuar con las explicaciones acerca de la teoría general de la terminología es necesario establecer algunas precisiones sobre el uso de los términos *concepto* y *contexto* en el marco de la disciplina.

El concepto en terminología no se estudia tomando en cuenta los aspectos psicológicos o filosóficos sino su expresión a través del lenguaje. Sager explica acerca de la formación de conceptos: La formación de conceptos es un proceso que consiste en

1 M. Teresa Cabré (1999) "Terminología y documentación". p. 233

2 Juan C. Sager (1993) *Curso práctico sobre el procesamiento de la terminología*. p. 43

3 Ángel Raimundo Fernández González, Salvador Hervás y Valerio Báez (1989) *Introducción a la semántica*. p. 90.

la agrupación y ordenación, dentro de categorías abstractas, de los objetos materiales e inmateriales que sentimos, percibimos o imaginamos”.⁴ Los conceptos pueden representar no sólo seres sino cosas en la mente y, en un sentido amplio, cualidades, acciones, situaciones, relaciones, funciones, etcétera. Por otra parte los conceptos nunca aparecen aislados sino que siempre tienen relaciones directas o indirectas con otros conceptos con los que comparten un cierto número de características iguales o similares.⁵

Cuando se observa un objeto se busca identificar en él los rasgos comunes que tiene con otros objetos conocidos y se abstraen algunos de los elementos que lo componen para identificar similitudes. Un vaso puede estar construido con un material diferente al vidrio y tener una forma diferente a la de los vasos que se acostumbra usar en un ámbito determinado; sin embargo extrayendo una serie de particularidades del recipiente se obtiene el concepto que pertenece a la categoría de vaso.

En cuanto al contexto éste se puede comprender a través de la explicación de Coseriu “Constituye contexto del hablar toda la realidad que rodea un signo, un acto verbal o un discurso, como presencia física, como saber de los interlocutores y como actividad”.⁶ Es claro que además del contexto idiomático y verbal es importante el contexto físico, histórico, cultural y ocasional, tal como clasifica Coseriu, a los elementos del contexto.

La comunicación en el ámbito especializado exige claridad en las unidades terminológicas y, en lo posible, un término para cada concepto dado. Las unidades terminológicas se utilizan como términos si el usuario del lenguaje ubica el término en la estructura de ese conocimiento. Sager expone... “un término adquirido sin tener conciencia de la configuración convencional de conocimiento a la que se refiere es inútil comunicativamente”.⁷

Una unidad léxica puede tener elementos que le otorguen una diferente función en el lenguaje natural y en los lenguajes especializados o incluso puede ser una unidad que tiene diferentes significados en distintos lenguajes especializados. La palabra “ondulación”, por ejemplo, puede significar un tipo de peinado o el serpenteo de una bandera en el mástil o el oleaje del mar, por lo que será el contexto en que se utilice el que aclarará la acepción correspondiente del término.

El ejemplo anterior muestra significados por semejanza y esa variación de significado nos obliga a contextualizarlo, pero no se está aseverando que las palabras pierdan significado fuera de un contexto, sino que: “las palabras aisladas tienen significados más o menos permanentes; de otro modo sería imposible la comunicación. El contexto sólo modifica el sentido de las palabras dentro de ciertos límites”.⁸

4 *Op. cit.* p.43

5 Peter Weissenhofer (1995) *Conceptology in terminology theory, semantics and word formation.* p. 2.

6 *Ibid.* p. 91.

7 *Op. cit.* p.43.

8 Ángel R. Fernández González (1989) *Introducción a la semántica.* p. 51

En un lenguaje especializado la connotación actúa mediante una limitación previa de los atributos que posee para un sistema de conocimientos. La referencia de un término en un lenguaje especializado, en el ámbito terminológico, se fija a través de una definición donde cada término queda enmarcado en el sistema conceptual de la disciplina. Desde el punto de vista de la teoría terminológica la definición es la referencia de un término a un concepto por medios lingüísticos. Los términos definidos son a su vez integrantes de una estructura que está agrupada por principios cognitivos.

La terminología desarrolla los vocabularios que construye a partir de un término que agrupa a otros términos asociados entre sí por pertenecer al grupo, y continúa hacia niveles más específicos de profundidad para formar la estructura de conocimiento de disciplina. Pero en lugar de agrupar los términos en forma alfabética, como se hace en el ámbito de la bibliotecología, se ordenan de acuerdo con su nivel de especificidad. Otra de las características del término es su ubicación en un contexto comunicativo oral o escrito en el realiza relaciones de significado.

Todo esfuerzo de aprendizaje implica ubicar el concepto dentro de un sistema de conocimientos que lo delimita y le otorga una definición que lo fija dentro de la estructura. La llamada función ideacional del lenguaje consiste en la formulación, a través del lenguaje, de pensamientos a los que se refiere quien habla tanto como quien recibe, y en consecuencia sirve para poder comunicarse.

El terminólogo propone la incorporación de locuciones adverbiales o verbos o cualquier otra unidad sintagmática que en el contexto del lenguaje especializado al que pertenecen, evocan o remiten a una unidad conceptual que sólo tiene validez en el campo específico de ese saber humano y no de otro.

La terminología estudia además el comportamiento, los cambios, las evoluciones y/o las adaptaciones que van sufriendo los términos de una especialidad. También analiza los comportamientos contrastivos que ocurren entre las variantes dialectales desde el punto de vista morfo-sintáctico para proponer y sugerir el uso adecuado de ciertos términos.

BIBLIOTECOLOGÍA Y DESCRIPTORES

Las disquisiciones anteriores acerca de los lenguajes de especialidad, la terminología y los términos están relacionadas con lo que ocurre en el análisis y sistematización de la información para su posterior recuperación, temas muy importantes de la bibliotecología.

La recuperación de un documento a partir de sus aspectos temáticos es una tarea que ha preocupado al hombre desde que se comenzaron a generar documentos escritos. La búsqueda de mayor claridad en la tarea de almacenar y recuperar información tiene una historia, pero para no entrar en detalles históricos se enmarcará la organización del conocimiento desde el punto de vista temático en un proceso global del que forma parte.

De acuerdo con el estudio realizado por Vizcaya Alonso⁹ a partir de Andrés Cruz Paz, V. M. García Suárez, Mijailov, Chernii y Guiliariievskii, el proceso de la información es uno de los objetos de estudio de la bibliotecología o ciencia de la información, que incluye el proceso de presentación, recolección, procesamiento (organización de la información), conservación, búsqueda, diseminación y percepción de la información, y uno de los temas de estudio lo constituye la estructura en que se representa dicha información.

La organización de la información surge de la necesidad de acceder a los documentos a través de cualquiera de los elementos que los componen sin tener que revisar el documento completo. Para lograr el acceso se toman en cuenta los diferentes intereses de usuarios y el grado de profundidad de esas necesidades de información. El resultado de esa organización implica la obtención de un producto que representa a un documento, a través de diferentes elementos que conforman un todo mucho más claro, conciso y fácil que intenta dar una idea rápida del documento completo.

Otra característica de esa organización es que cada una de las representaciones de un documento estará armonizada con las representaciones de otros documentos ligados por una consistencia en su presentación, lo que a su vez conformará un todo nuevo que puede ser un sistema de información, una base de datos o una bibliografía. En concreto las actividades que engloba la organización de la información son: registro bibliográfico, clasificación y/o indización y resumen.

Estas actividades conforman dos categorías: forma y contenido. La forma se representa a través de la descripción que se hace en el registro bibliográfico, y el contenido por medio del resumen, la clasificación y/o la indización. Estos elementos son el resultado de dos procesos que se realizan ante un documento: análisis y síntesis.

El análisis, como proceso, se define como la separación de las partes de un todo hasta llegar a conocer sus principios o elementos. En tanto que la síntesis sería la composición de un todo mediante la reunión de sus partes. Estos procesos de análisis y síntesis son fundamentales para representar un documento, tanto desde el punto de vista formal como del contenido. Sin embargo no son procesos independientes que se aplican a la representación de documentos. No se puede perder de vista el objetivo del sistema de información ni el contexto en el cual se encuentra inserto.

Para obtener ese nuevo documento consistente y eficaz cuya finalidad es la obtención de cada uno de los elementos iniciales que conformaron cada uno de los registros bibliográficos, se debe contar con llaves o claves que permitan acceder a cada uno de los elementos o a conjuntos de ellos.

Uno de los elementos importantes para disponer de llaves que nos permitan acceder a los documentos a través de su contenido temático es el lenguaje documental, el cual deberá ser insertado en el sistema de información a medida que se vayan integrando cada uno de los registros bibliográficos.

9 Dolores Vizcaya Alonso (1997) *Información: procesamiento de contenido*, p. 29.

Para ser eficaz, el lenguaje documental debe responder a las necesidades del sistema de información una vez que éste ha sido conceptualizado; es decir una vez que se ha establecido la perspectiva del sistema de información y se han definido sus usuarios, será necesario pensar en la herramienta que sea capaz de introducir nuevamente al documento primario a través de los términos que representen los contenidos de impresos, imágenes, argumento, sonido, música y/o voz, dependiendo del soporte en el cual se encuentre la información que se quiere organizar.

En definitiva el contenido documental se expresa a través de una clasificación, una indización y un resumen, los cuales buscan fijar los significados de un documento y expresar en la clasificación temática, la clasificación topográfica y luego el resumen como síntesis del documento completo.

De estas tres tareas que conforman una parte de una de las etapas del proceso de organización documental, sólo retomaremos una, la indización o clasificación temática o facetada, para lo cual utilizaremos lenguajes documentales que normalizan los términos escogidos tanto en el momento de hacer la indización, como en el de recuperar la información. Por tanto es importante tener presente todo el proceso en el que interviene esta fijación de contenidos a través de una clasificación temática que ayude a entender la diferencia entre el descriptor y el término, porque la función del descriptor está relacionada con la búsqueda y recuperación de la información—etapa que ya se mencionó como fundamental en el proceso de la organización documental—, a diferencia de la función del término, que está destinado a la comunicación científica general.

La diferencia entre la agrupación temática o facetada es un aspecto que merece una amplia discusión, sin embargo para los fines del presente trabajo se puede partir de la definición expresada por Chaumier “[...]La reagrupación por temas se hace según el dominio (campo semántico) la reagrupación por facetas se hace según la naturaleza de los conceptos”.¹⁰

La fijación producida por el análisis de contenido para indizar o clasificar supone la existencia previa de sistemas clasificatorios que establezcan la pertenencia de ese documento a una temática o una faceta en especial, lo cual es imprescindible en el proceso de organización de la información. Con respecto a la tarea de indización Cabré opina: “La indización es en el fondo una operación terminológica, es una actividad de detección de la terminología más indicativa del contenido documental, por la que se atribuye un valor de contenido a determinadas unidades y expresiones del texto, que son las que servirán de anclaje en la recuperación de información”.¹¹ Como se puede observar, la terminología le atribuye aquí una función muy específica al descriptor al definirlo como el término más indicativo del contenido documental e incluso como el anclaje para hacer la recuperación de información.

10 Jacques Chaumier (1986) *Análisis y lenguajes documentales*. p. 90.

11 *Op cit* p. 237.

Para normalizar e indizar la organización de los documentos que luego serán recuperados para usarse, una de las clasificaciones más utilizadas en las bases de datos es la temática o facetada, la cual permite agrupar en estructuras jerarquizadas el conocimiento que se tiene sobre una especialidad con los documentos que se relacionan entre sí. La clasificación temática a diferencia de la sistemática –más parecida en el agrupamiento a las terminologías– se representa a través de términos normalizados; es decir términos que son como unidades transmisoras de conocimiento especializado y se estructuran en un lenguaje de búsqueda y recuperación de información documental en orden alfabético de entrada, por lo general de tipo descriptor. Estos términos funcionan dotando a esas unidades lingüísticas de características propias adecuadas a las necesidades del sistema de información del que forman parte, tal como se explicará más adelante.

Existe sin duda una función pragmática en la clasificación porque persigue la agrupación del conocimiento para ser recuperado como información, pero está sometida a elementos teóricos y normativos que proceden de disciplinas del conocimiento que le brindan su apoyo. Estas reflexiones parten de la relación con la terminología, fundamentada en la teoría terminológica, asumiendo que los términos son unidades lingüísticas a la vez que unidades de conocimiento especializado y de comunicación especializada.

La terminología nutre a través de sus elementos teóricos y la práctica terminológica a la bibliotecología o documentación para construir los fundamentos esenciales que garanticen el diseño de los sistemas de organización y representación del conocimiento. Los instrumentos clasificatorios de la bibliotecología o documentación usan el trabajo descriptivo y normativo de la terminología para fijar los contenidos documentales.

Las clasificaciones temáticas

Las clasificaciones bibliográficas que surgieron en el siglo XIX están basadas en principios empíricos o prácticos que se aplican a la organización de los documentos en las bibliotecas y establecen límites de competencias temáticas para contrastar los conocimientos compartidos con las comunidades que utilizarán los documentos. A este respecto Eric de Grolier, citado por San Segundo expresa:

[...] cabe afirmar que la clasificación es un artefacto cultural que depende no sólo de los parámetros culturales, sino también de las formas políticas, económicas, de las condiciones sociales y otras.¹²

Estas diferencias culturales y coyunturales determinan la necesidad de crear clasificaciones para los diferentes entornos informativos que deben ser representados: en la biblioteca se utilizan los llamados encabezamientos de materia y las clasificaciones sistemáticas para el ordenamiento de los libros en la estantería, mientras que

12 Rosa San Segundo Manuel (1996) *Sistemas de organización del conocimiento: la organización del conocimiento en las bibliotecas españolas*. p. 32.

en los sistemas de información automatizados si bien a veces se usa una clasificación sistemática, se emplean más los vocabularios controlados o tesauros. El tesoro, a diferencia del vocabulario controlado, está constituido por elementos estructurales de las disciplinas en las que se construye: “Un tesoro combina características estructurales de los esquemas de clasificación y de las listas de encabezamientos de materia”,¹³ pero esto está dirigido hacia un área especializada.

Desde finales del siglo pasado y hasta la aparición de las bases de datos, las colecciones de fondos bibliográficos –además de su organización basada en clasificaciones numéricas o alfanuméricas que precisaban la ubicación topográfica de los documentos para su acomodo en la estantería– se indizaban con vocabularios controlados representados por encabezamientos de materia normalizados, lo que permitía luego su localización a través de los catálogos manuales.

Las clasificaciones numéricas o alfanuméricas se explican claramente en estas palabras:

El objetivo prioritario de la organización del conocimiento documental o clasificación bibliotecario-bibliográfica es la ordenación por grupos temáticos de fondos bibliográficos, y la elaboración de catálogos y bibliografías sistemáticos que posibiliten y faciliten a los usuarios o lectores el acceso por materias a los libros o documentos. De esta forma, tras la aplicación de un determinado sistema clasificatorio se pueden reunir y agrupar los fondos documentales según la materia sobre la que versan. Así los libros, bibliografías o catálogos estarán regidos por un orden correlativo interrelacionado.¹⁴

Las listas de encabezamientos de materia son colecciones de términos que incluyen las expresiones más usuales de los conceptos más otras expresiones equivalentes a las anteriores pero que no deben ser usadas para indizar y por tanto tampoco para recuperar información en los catálogos. Estas listas incorporan además una guía de términos que son más generales o más específicos. Los encabezamientos de materia se utilizan para recuperar información de los catálogos que responden a la colección de una biblioteca más bien de tipo general; es decir que abarca varias ramas del conocimiento.

Para evitar un crecimiento desmesurado de los catálogos, estas expresiones ligan conceptos que logran la condensación necesaria para hacer una rápida recuperación de información. Esta condensación de información, denominada pre-coordinación, era mayor en los catálogos manuales, pero con la aparición de los catálogos automatizados se ha producido un fenómeno que se conoce como la tesaurización de los encabezamientos de materia, porque la tendencia es a post-coordinar mayor cantidad de términos; es decir hacerlos más cortos y menos condensados de información. En el ejemplo que sigue se presenta la entrada de una lista de encabezamientos

13 María Texia Iglesias Maturana (1992) *Análisis y recuperación de información documental*. p. 47.

14 Rosa San Segundo Manuel (1996) *Sistemas de organización del conocimiento*. p.70.

de materia que tienen mayor nivel de pre-coordinación y otra que incluye un nivel mayor de post-coordinación:

- ❖ EDUCACIÓN ESCOLAR PARA MINUSVÁLIDOS EN MÉXICO
Un solo término pre-coordinado para la recuperación
- ❖ EDUCACIÓN ESCOLAR, MINUSVÁLIDOS, MÉXICO
Tres términos post-coordinados para la recuperación

Tesauros

Con la aparición de la computadora se cuestionó el uso de los lenguajes documentales precoordinados y controlados para indizar y recuperar información, ya que parecía muy fácil recuperar usando cualquier término incluido en el título o asignado por el indizador sin someterlo a una normalización previa. En la actualidad la polisemia de los lenguajes especializados hace imprescindible la fijación de contenidos y una representación más precisa que le permita a los sistemas automatizados de información hacer postcoordinaciones a partir de cualquier término o unidad lingüística. El propósito de esto es agrupar información sobre una determinada temática en el sistema de información al que está destinado.

El proceso de agrupar información legitima al descriptor como elemento fundamental en la estructura de un tesauro y ha sido favorecido por la llegada de la informática documental, lo cual facilita las combinaciones de los elementos del lenguaje para recuperar información.

Como los tesauros deben servir a sistemas de información que aspiran a recuperar información a partir de la utilización de inteligencia artificial, deben estar estructurados como mapas conceptuales designados para que la información acumule un mayor volumen de contenidos semánticos e incluya una clasificación jerárquica, con un nivel cada vez menor de volúmenes de contenidos semánticos que a la vez pondere las diferentes temáticas o facetas que abarca la disciplina o sistema de información.

No siempre los sistemas de información comprenden información sobre una disciplina en particular, algunas veces el ámbito es mayor que el de una sola disciplina y el lenguaje documental que se utilice deberá responder a ese ámbito de cobertura del sistema, lo que implicará la consecuente necesidad de hacer tesauros acordes con los sistemas de información a los cuáles están destinados.

Si por ejemplo se hace un tesauro para una base de datos de contenidos educativos, el descriptor AGUA puede admitir varios hiperónimos: es un elemento desde el punto de vista químico; un líquido desde el punto de vista físico, y en geografía cubre buena parte del suelo de la Tierra y es un elemento jurídico ya que existen aguas territoriales y aguas internacionales. Aquí el AGUA no es una noción monosémica, pero para el sistema del tesauro se debe estructurar en un área del conocimiento. Y su ubicación será determinada por la necesidad del sistema considerando únicamente aspectos prácticos.

La opción léxica del tesoro exige desambiguar lo que puede ser polisémico para elaborar la red nocional que exige el instrumento. El AGUA quedará asentada en un área del conocimiento, pero estará relacionada por afinidad con los demás descriptores que puedan ser utilizados para asociar en la indización. Es decir, para indizar un documento que trate sobre AGUA, se elegirá el descriptor AGUA, pero si el hiperónimo que le corresponde en el tesoro no es el mismo que le corresponde al hiperónimo del AGUA al que se refiere el documento que se está indizando, no se pierde claridad, porque en la red de términos asociados al descriptor AGUA, aparecen también los otros hiperónimos.

Un tesoro se define como **una estructura lingüística de clasificación temática o facetada, cuya organización básica está conformada por una relación de descriptores que representan o describen autoridades o contenidos semánticos representados por unidades lingüísticas extraídas del lenguaje de una disciplina o área específica del conocimiento**. Estos términos controlados denominados descriptores son normalizados y usados para indizar contenidos que posteriormente pueden ser buscados y recuperados en un sistema de información. Es importante aclarar que el descriptor representa conceptos pero no como elementos aislados, sino que cada concepto individual es parte integrante de un sistema conceptual que a la vez refleja la ordenación y la estructura de algún campo del conocimiento. Un tesoro bien elaborado brinda precisión, claridad, confiabilidad e integridad a los elementos que lo componen porque opera como una estructura de base para organizar la información.

La palabra tesoro, en su origen *thesaurus* o *thesauros*, era el nombre dado en las casas de la época griega a la cámara donde se guardaba el tesoro, pero posteriormente se usó para designar a ciertos diccionarios, índices o repertorios de términos¹⁵. El primero que usó la palabra con este sentido fue el florentino Brunetto Latini en su enciclopedia *Livre du Trésor*, compuesta en francés entre los años 1262 y 1268. El término se continuó usando pero se popularizó a partir del *Thesaurus of English Words and Phrases*, un diccionario ideológico recopilado por Peter Mark Roget en 1852 y reeditado unas noventa veces, desde ese año. En la introducción de este diccionario se podía leer:

Un conjunto... clasificado, no en el orden alfabético como lo están (las palabras) en un diccionario, sino según las ideas que expresan... El objetivo consiste, una vez dada la idea, en encontrar la palabra o las palabras a través de las cuales puede expresarse conveniente y justamente esta idea.¹⁶

En la actualidad la connotación de tesoro es diferente, en especial en el ámbito de la bibliotecología y la información, ya que aún aparece en algunos diccionarios y enciclopedias usado en el sentido de diccionario, catálogo, antología.¹⁷

15 José Martínez de Souza (1993) *Diccionario de bibliología y ciencias afines*: p 776.

16 Jacques Chaumier (1986) *Análisis y lenguajes documentales*. p. 12.

17 *El pequeño Larousse ilustrado 1996 en color*. México: Ediciones Larousse. 1792 p.

Las diferencias entre el tesoro y el diccionario, además de las derivadas de su función, son: la representación de la estructura conceptual de un mismo campo temático en un sistema de información; la aclaración sobre el sentido único en el cual se debe utilizar la unidad lingüística para ese sistema de información; la eliminación de los equívocos mediante el envío a términos normalizados o descriptores de equivalencia univalente, y las referencias recíprocas entre descriptores relacionados. Es decir, el tesoro nos muestra el campo de aplicación de una unidad lingüística y le asigna una posición en la disciplina.

Con el uso masivo de los sistemas computarizados se ha generalizado el criterio de que la utilización del lenguaje natural o la búsqueda en texto libre, a través de palabras claves, es suficiente para hacer búsquedas y recuperaciones en los sistemas de información, porque por una parte se ahorra tiempo y se simplifica el ingreso de información. y por la otra, se protegen las expresiones del autor y es posible revelar diferencias culturales, psicológicas o lingüísticas en el texto de un documento.

Pero también existen muchos argumentos en pro de la indización que utiliza un vocabulario controlado. Debido a la riqueza de las expresiones, los sistemas con lenguaje natural se tornan confusos cuando se los consulta en busca de conceptos más que de palabras. El proceso se hace largo y complejo cuando el usuario no localiza el tema y debe buscar todas las posibles formas en que un mismo concepto puede haber sido expresado.

Buscando con exhaustividad en toda la colección de documentos que compone una base de datos se encontraría la información solicitada pero luego de un desmesurado gasto de tiempo y esfuerzo. Para que la inversión de tiempo en dicha tarea se minimizara, la información se indizaría con objeto de señalar la existencia o no de documentos relevantes acerca de distintas materias, y se almacenaría siguiendo ciertos criterios de clasificación.

Para no seguir abundando en las funciones del tesoro en especial, se presenta como resumen la clasificación de las funciones del tesoro que hace Chaumier.¹⁸

Entrada. Descripción del documento.

- Clasificar los conceptos
- Proporcionar el vocabulario que traduzca los conceptos
- Presentar un vocabulario unívoco
- Favorecer la coherencia del análisis documental
- Presentar el entorno semántico de los descriptores escogidos para traducir los conceptos

Tratamiento. Creación de los ficheros.

- Facilitar el almacenamiento de la información
- Transformar la información en datos manipulables

18 Jacques Chaumier.(1986) *Análisis y lenguajes documentales*. p. 17.

- Favorecer el control y la validación de los datos
- Salida. Formulación de la pregunta
- Ordenar los conceptos de la demanda
 - Proporcionar el vocabulario de los conceptos de la demanda
 - Facilitar el diálogo hombre/sistema
 - Focalizar o ampliar la pregunta

LOS DESCRIPTORES Y LOS TÉRMINOS

Para organizar la información se utilizan palabras claves, descriptores y encabezamientos de materia que ya fueron definidos acotando su campo de aplicación. Las palabras claves son palabras significativas extraídas directamente de los textos y usadas para la indización en los sistemas automatizados de búsqueda.

Los descriptores son términos validados para emplearse en temas definidos previamente y que se usarán para representar un documento al indizar su contenido. Esto se hace mediante una doble búsqueda en los equivalentes de las palabras usadas por el autor entre los descriptores validados en el tesauro y en la definición de la estrategia de búsqueda para localizar el documento cuando es solicitado.

- ❖ Una palabra clave tiene un peso significativo en una frase y es una palabra o grupo de palabras escogidas a partir del título o texto de un resumen o documento para representar el contenido y facilitar su recuperación.
- ❖ “Un descriptor puede definirse como un término o un símbolo autorizado y formalizado que figura en un tesauro y que se usa para representar, sin ambigüedad, los conceptos contenidos en los documentos y en las preguntas.”¹⁹ Agregan las normas ISO 2788-1986: “[...] la representación de un concepto, preferentemente en forma de sustantivo o frase nominal.”²⁰ El descriptor puede ser una palabra simple, compuesta o una frase que describe un contenido documental.

Los descriptores constituyen las unidades básicas que conforman los tesauros y unidades lingüísticas normalizadas para la indización, y han sido extraídos por un proceso de análisis-síntesis a partir del lenguaje especializado expresado en los documentos que se analizan para un sistema y validados mediante vocabularios especializados.

En el campo de la organización de la información el descriptor, al igual que el término, es una unidad lingüística cuya función primordial es describir un concepto en forma unívoca en el contexto de un sistema de información. El conjunto de unidades lingüísticas denominadas descriptores y sus relaciones correspondientes conforman el Tesauro, herramienta que se utiliza para indizar de manera normalizada el conoci-

19 María Texia Iglesias Maturana (1992) *Análisis y recuperación de información documental*. p. 101.

20 Catalina Naumis (1998) “Dinámica en la construcción de tesauros y el acceso a la información”. p. 83.

miento que conllevan los documentos y para evitar la dispersión provocada por el uso en la búsqueda de términos que no han sido utilizados en la indización previa.

El descriptor sirve como puente entre el lenguaje usado por los autores y el que suele utilizar el usuario, e implica la reunión de varios términos de la especialidad del sistema de información que contextualizan un descriptor como único término normalizado que se podrá utilizar para representar un concepto. En este sentido tanto el descriptor como el término tienen en común una función clarificadora sobre el uso de los términos especializados en un área del conocimiento, pero mientras el término se usa intercalado con el lenguaje general, un descriptor agrupa otros descriptores y los relaciona.

El uso de un lenguaje artificial nos permite realizar una indización de conceptos, más que una indización de términos. Esto significa que establecemos una definición standard para cada concepto y luego utilizamos esa descripción cada vez que nos encontramos con el concepto, aunque el autor lo haya expresado con diferentes palabras. Así, por ejemplo, si decidimos que la definición standard es “educación para la salud”, utilizaremos siempre esta expresión para representar este tema, aún cuando un autor en su documento hable de “educación sanitaria”.²¹

La normalización de las unidades lingüísticas en el ambiente documental se realiza a nivel del sistema de información mediante un acuerdo entre especialistas del conocimiento, quienes transmiten un sistema del cual participan los documentalistas. Esto también implica la garantía literaria, es decir el uso que hacen los especialistas más reconocidos de un área del conocimiento, de esos términos en el contexto de los documentos: la normalización de los descriptores que constituyen los tesauros conlleva la priorización de una denominación sobre sus alternativas, el control de la sinonimia y la regularización de su forma de presentación”²². En cambio la normalización del término se establece a través de organismos autorizados que revisan el uso del término tanto en el contexto documental como en el de la lengua hablada.

Sin embargo la documentación utiliza los términos normalizados por la terminología como base para sus descriptores, porque siempre busca los usos fijados en los vocabularios de cada especialidad para tomar la decisión final de utilizar los descriptores que sean los más conocidos por la comunidad de destino, sobre todo cuando algo no queda claro en el contexto documental del sistema de información, que siempre es la primera fuente de consulta.

En su calidad de término normalizado el descriptor debe ser único, porque su función es pragmática y se lleva a la praxis mediante la agrupación de información que representa un contenido temático bajo una sola denominación. La función de los descriptores es denominar conceptos, pero no incluye expresamente el uso de los descriptores en el contexto del lenguaje de especialidad (como los términos), sino que los contextualiza en un grupo de descriptores porque entre ellos mantienen entre sí rela-

21 María Texia Iglesias Maturana (1992) *Análisis y recuperación de información documental*. p.- 47.

22 M. Teresa Cabré (1999) “Terminología y documentación”. p. 238

ciones de diferente tipo. El conjunto de estas relaciones entre descriptores constituye la estructura conceptual del tesaurus, para ser más precisos, de la especialidad a que se refiere el sistema de información.

Cuanto mayor sea la delimitación de un descriptor más posibilidades habrá de situar el concepto para la comunicación, ya que se hace explícito un “punto de vista”²³ desde el cual se observa al descriptor, y que la “clase de pertinencia”²⁴ de la que partimos “decide no describir los hechos recogidos a no ser desde un único punto de vista,²⁵ para lo cual han de tomarse en consideración, dentro de la masa heterogénea de estos hechos, sólo los rasgos que afectan a este punto de vista, excluyendo todos los demás (por este motivo tales rasgos se llaman pertinentes”²⁶

El objetivo del descriptor es atenerse a lo nocional estabilizando el contenido para el sistema documental, por eso cada descriptor debe tener términos sinónimos asociados llamados no descriptores y descriptores afines que también pueden ser utilizados para la indización así como notas que expliquen el uso exacto de ese término para el sistema de información en el que se utiliza. También todo descriptor debe incluir relaciones genéricas y partitivas que guíen tanto al indizador como al usuario que busca información hacia el descriptor que representa un núcleo importante de información en el marco de un lenguaje especializado.

El descriptor usa fuentes documentales de la especialidad para describir y normalizar el uso de los términos que serán base de los descriptores, y no descriptores que alimentan el sistema de información. Ese uso de los instrumentos terminológicos asocia un único significado a la red de no descriptores y le confiere validez al descriptor a la hora de hacer la indización.

En la sinonimia la significación tiene una participación importante porque se debe escoger como descriptor al término cuyo “estereotipo” sea el más representativo para la comunidad de usuarios a quienes se destina el sistema. “Un estereotipo no es un conjunto de propiedades fácticas del objeto, obtenidas en alguna forma directamente de él mismo, sino una construcción elaborada por la sociedad en su memoria de experiencias compartidas en relación con el objeto. Se puede decir que es una especie de abstracción, en la cual ciertas características perceptibles del objeto no se toman en cuenta y otras se destacan.”²⁷

Rendón Rojas²⁸ menciona la necesidad de explicitar los sinónimos para la comunicación: “Si hay conciencia de que un mismo objeto es denominado por diferentes términos, no existe un grave problema de comunicación y entendimiento, y sólo es

23 Félix Sagredo Fernández y José Ma. Izquierdo Arroyo (1983) *Concepción lógico-lingüística de la documentación*. p.14.

24 *Ibid* p. 15.

25 Sucede en los tesauros facetados, en los tesauros ordenados por tema a veces se utiliza la polijerarquía para recoger diferentes punto de vista.

26 *Ibid* p. 15 cita a Roland Barthes (1971, p. 99).

27 Luis Fernando Lara (1997) *Teoría del diccionario monolingüe*. p.185.

28 Miguel Ángel Rendón Rojas (1997) *Bases teóricas y filosóficas de bibliotecología*. p. 51.

necesario explicitar esa sinonimia”. El tesoro no sólo expresa la sinonimia sino que además guía el uso hacia el descriptor autorizado.

Los descriptores afines son aquellos que tienen una relación de mayor o menor volumen de contenido semántico, pero cuya inclusión no fue decidida como jerarquizada porque hay otro descriptor cuyo contenido es más conveniente para el sistema de información y/o los descriptores que comparten el mismo término de mayor volumen de contenido semántico y/o los descriptores relacionados con el descriptor de la entrada. Es decir estos descriptores están asociados y conviene recordárselos al indizador y al recuperador de información para ayudarle a evitar confusiones. Por ejemplo si en la entrada alfabética de un tesoro el descriptor es:

DOCUMENTOS ADUANALES

TG: DOCUMENTOS COMERCIALES

TE: CERTIFICADOS DE ORIGEN

CERTIFICADOS FITOSANITARIOS

TR: ADMINISTRACIÓN DE ADUANAS

DOCUMENTOS DE IMPORTACIÓN

En este caso el término genérico, es decir el de mayor volumen de contenido semántico es DOCUMENTOS COMERCIALES, y los términos específicos, es decir los de menor contenido semántico son CERTIFICADOS DE ORIGEN y CERTIFICADOS FITOSANITARIOS. Los términos relacionados son ADMINISTRACIÓN DE ADUANAS, que desde otro punto de vista podría haber sido escogido como el de mayor contenido semántico en lugar de DOCUMENTOS COMERCIALES, en tanto que DOCUMENTOS DE IMPORTACIÓN podría haber sido, desde otro punto de vista, como un descriptor de menor contenido semántico que DOCUMENTOS ADUANALES.

Esta entrada del descriptor permite realizar la ubicación en una estructura acorde con el sistema de información para el cual está estructurado y recordarle tanto al indizador como al usuario que existen otros descriptores que pueden ser revisados y que quizá se adecuan mejor al documento que tiene entre manos.

Las notas de alcance son descripciones sobre el uso del término cuya finalidad es limitar el uso de un descriptor, excluir ciertas características o sugerir algún método de inserción en el sistema de información. En este aspecto la tendencia más moderna es incluir la mayor cantidad de este tipo de aclaraciones que están más ligadas a la definición del descriptor que la tradicional nota de alcance. La definición en este contexto se entiende como: la operación por medio de la cual, se establece el nexo entre una expresión lingüística y el objeto que ella designa²⁹. Se puede decir que la tendencia es a terminologizar los tesauros para parafrasear lo que ocurre con las listas de encabezamientos.

BIBLIOGRAFÍAS

NA: Usar para documentos que en sí son una compilación de referencias bibliográficas. Para documentos que describen y analizan la actividad bibliográfica, usar BIBLIOGRAFÍA.

Las relaciones categoriales implican la reagrupación en campos semánticos o con base en características comunes de cada conjunto de descriptores. Tanto en un tesauro con clasificación temática como uno facetado ocurren situaciones parecidas con las agrupaciones de descriptores, pero existen diferencias en cuanto a las jerarquías a las cuales se somete un descriptor. Las facetas son fruto de una abstracción mayor que las categorías temáticas, porque un descriptor puede representar diferentes puntos de vista y estar en varias categorías temáticas pero en una sola faceta.

La dificultad para ubicar un descriptor en una sola temática es explicable porque la revisión de especialidades muestra que el saber es un continuo y que su segmentación en temáticas lo que busca es dividir con fines prácticos la masa enorme de conocimientos científicos que ha acumulado el hombre. Un descriptor puede aparecer en materias distintas o ser analizado desde distintos puntos de vista, como ocurre con el ABORTO, que aunque debe ser practicado por un médico, tiene aspectos que debe dilucidar la psicología o el derecho. Una INDUSTRIA influye en la economía de un país, pero tiene problemas prácticos que debe resolver la ingeniería o la tecnología.

En este sentido el descriptor se ordena en forma sistemática de acuerdo con su volumen de contenido semántico para formar la estructura de conocimientos del sistema de información al que sirve, y luego aparecer integrado en orden alfabético en cada entrada de descriptor. Disponemos entonces de una entrada por nivel de especificidad y otra alfabética.

El aspecto de la jerarquización se observa en el ejemplo que aparece líneas arriba para explicar que los términos relacionados CERTIFICADOS DE ORIGEN y CERTIFICADOS FITOSANITARIOS son tipos de DOCUMENTOS ADUANALES, y que por esa razón DOCUMENTOS ADUANALES es la clase, y CERTIFICADOS DE ORIGEN y CERTIFICADOS FITOSANITARIOS son la subclase y están subordinados jerárquicamente.

La entrada alfabética del tesauro incluye todo el contexto, la nota de alcance, la categoría a la que corresponde el descriptor y las clases que agrupa. Incluye también los sinónimos y los términos asociados, y además en el orden alfabético general de entradas se alternan los sinónimos con la categoría de los no-descriptores y la referencia al descriptor validado. La entrada por nivel de especificidad incluye únicamente los niveles jerárquicos.

Además de su presentación en papel los tesauros incluyen una presentación en páginas Web que permite otras entradas y relaciones más directas para brindarle apoyo directo al sistema de información para el cual fueran implementados.

Como corolario del presente estudio se puede plantear una nueva definición del descriptor: *un término normalizado y relacionado dentro de un contexto lingüístico, en el marco de un tesauro, que describe y representa un determinado contenido de los documentos para integrarse a un sistema de información.*

En conclusión en el descriptor confluyen varios aspectos: el contenido o representación del conocimiento (semántica), la definición del objeto que el descriptor designa (terminología); la estructuración en una organización a través de las relaciones conceptuales con otros descriptores (lingüística); la estructura morfo-sintáctica del descriptor (lingüística); la función para el almacenamiento en un sistema de información (documental), y la función de recuperador de información para satisfacer la necesidad del usuario (documental).

CONSIDERACIONES FINALES

El término normalizado es una fuente para validar del descriptor.

El descriptor y el término tienen una diferente funcionalidad que se refleja en los instrumentos lingüísticos que se construyen con ellos.

La terminología contribuye a la funcionalidad de la comunicación entre hablantes de una especialidad y sirve a la documentación en la transmisión de conocimientos.

En el proceso de búsqueda y recuperación que se realiza al diseñar una búsqueda informativa los descriptores se pueden asociar unos con otros, con lo cual surgen otros significados en los descriptores conjuntados mediante la postcoordinación. Más aún, la tendencia es asociar los descriptores con el lenguaje natural transmitido a través de expresiones especializadas que han comenzado a validarse en los documentos que integran el sistema de información.

El análisis de la función del descriptor responde a la inquietud planteada en el trabajo de Cabré:

[...]Cabe pues explorar nuevas vías de solución para que la descripción documental pueda ser sistemática, pero su recuperación sea mucho más flexible y adaptada al usuario.³⁰

Es necesario tomar en cuenta que un documento no está representado por un solo descriptor, sino que cada documento incluirá tantos descriptores como temáticas tratadas contenga.

Si este mismo documento, por ejemplo, fuera a indizarse se representaría para un sistema de información especializado en Bibliotecología o Documentación como TERMINOLOGÍA, TESAURO, DESCRIPTOR, TÉRMINO.

Y podría ser recuperado a la hora de la búsqueda en el mismo sistema a través de asociaciones de descriptores como TESAURO Y TERMINOLOGÍA o TESAURO Y DESCRIPTOR, y adquirir frente al usuario que hace la búsqueda informati-

30 *Op cit.* p. 246.

va una utilidad parcial porque su expectativa es diferente, en tanto que busca la significación adecuada a sus intereses y que el sistema de información le permite rescatar un cierto contenido que es dado por el nivel de postcoordinación que permiten los descriptores.

En la riqueza de la búsqueda en sistemas informáticos, además de esta flexibilidad que adquieren los términos que se han usado para la indización, se le agrega otra más que permite asociar estos descriptores normalizados con dos términos científicos que aparecen en los resúmenes producido por el análisis documental realizado por el indizador, y las palabras clave del título captadas por el sistema de información.

En el momento de hacer la búsqueda existe una alternancia y combinación del lenguaje natural con el lenguaje de la disciplina, aspecto que le otorga al descriptor una gran flexibilidad y una actualización que sería imposible lograr si la búsqueda se restringiera al mismo instrumento lingüístico que fue creado para el sistema de información.

El término tiene una función descriptiva y prescriptiva que sirve de base para la construcción del descriptor, pero éste adquiere una función activa y dinámica diferente al término que le sirvió de base, dada su función de índice para la búsqueda de información integrada a un sistema.

OBRAS CONSULTADAS

- CABRÉ, María Teresa. "Terminología y documentación". p. 231-247. En su *La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Barcelona: IULA, 1999.
- CHAUMIER, J. *Análisis y lenguajes documentales: el tratamiento lingüístico de la información documental*. – Barcelona: Mitre, 1986. – 173 p.
- FERNÁNDEZ GONZÁLEZ, Ángel Raimundo, Salvador Hervás y Valero Báez. *Introducción a la semántica*. 5ª. Ed. – Madrid: Cátedra, 1989. 252 p.
- FODOR, Jerry A. *Conceptos: donde la ciencia cognitiva se equivocó* / traducción de Liza Skidelsky. Barcelona: Gedisa, 1999. 236 p. (Ciencias cognitivas: serie CLA-DE-MA)
- GUINCHAT, Claire y Michel Menou. *Introducción general a las ciencias y técnicas de la información y documentación*. – 2ª ed. correg. y aum. / Marie France Blanquet. – Madrid: UNESCO, 1990. – 555 p.: il.
- IGLESIAS MATURANA, María Texia. *Análisis y recuperación de información documental*. Santiago de Chile: Ediciones Instituto Profesional de Santiago, 1992. xv, 237 p.
- LARA, Luis Fernando. *Teoría del diccionario monolingüe*. México: El Colegio de México, Centro de Estudios Lingüísticos y Literarios, 1996. 274 p.

- LERAT, Pierre. *Las lenguas especializadas*. / Traducción de Albert Ribas. Barcelona: Ed. Ariel, 1997. 221 p. (Ariel lingüística)
- MARTÍNEZ DE SOUSA, José. *Diccionario de bibliología y ciencias afines*. 2ª ed. aum. y act. Madrid: Fundación Germán Sánchez Ruipérez. xvii, 961 p. (Biblioteca del Libro)
- MOLHOLT, Pat y Toni Petersen. "The role of the 'Art and Architecture Thesaurus' in communicating about visual art". En: Knowledge Organization. Vol. 20, no. 1 (1993), p. 30-34.
- NAUMIS PEÑA, Catalina. "Dinámica en la construcción de tesauros y el acceso a la información". p. 69-90. En: *La información en el inicio de la era electrónica: organización del conocimiento y sistemas de información* / Centro Universitario de Investigaciones Bibliotecológicas. México D.F.: UNAM, El Centro, 1998. Vol 1. (Monografías; 25)
- RENDÓN ROJAS, Miguel Ángel. *Bases teóricas y filosóficas de la bibliotecología*. México: UNAM. Centro Universitario de Investigaciones Bibliotecológicas, 1997. xi, 132 p. (Monografías; 24)
- SAGER, Juan C. *Curso práctico sobre el procesamiento de la terminología* / con un capítulo adicional de Joan Torruella y Gloria Clavería; bibliografía de Blaise Nkwenti-Azeh; traducción del inglés Laura Chumillas Moya. — Madrid: Fundación Germán Sánchez Ruipérez, 1993. 448 p. — (Biblioteca del libro; 57)
- SAGREDO FERNÁNDEZ, Félix y José Ma. Izquierdo Arroyo. *Concepción lógico-lingüística de la documentación*. Madrid: Ibercom-Red Comnet de la Unesco, 1983.—xv, 440 p.
- SAN SEGUNDO MANUEL, Rosa *Sistemas de organización del conocimiento: la organización del conocimiento en las bibliotecas españolas*. Madrid: Universidad Carlos III en Madrid, 1996.
- SLYPE, Georges Van. *Los lenguajes de indización : concepción, construcción y utilización en los sistemas documentales*. — Madrid : Fundación Germán Sánchez Ruipérez, 1991. — 200 p.
- THOMAS, Alan R. "Terminological problems : ocurrence, causes, consequences and solutions in guidance, counselling and associated fields". En *International Journal of Information and Library Research*. Vol. 1, no. 1 (1989) : p. 40 - 52.
- VALDIVIA DOUNCE, Lourdes *Palabras y cosas: una semántica cognitiva de los términos singulares*. México DF: UNAM. Coordinación de Humanidades. Centro de Neurobiología, 1998.
- VIZCAYA ALONSO, Dolores *Información: procesamiento de contenido*. Rosario: Parhadigma, 1997. 187 p
- WEISSENHOFER, Peter. *Conceptology in terminology theory, semantics and word-formation: a morpho- conceptually based approach to classification as exemplified by the English baseball terminology*. Vienna: International Network for Terminology, 1995. xi, 270 p. (IITF series; 6)