



La norma argentina de constucción de tesauros IRAM 32057: necesidad de actualización. 1.El vocabulario

Autor:

Martínez, Ana M.

Revista

Información, Cultura y Sociedad.

2009, N°21, pp. 11-30



Homenaje



**LA NORMA ARGENTINA DE CONSTRUCCIÓN DE
TESAURUS IRAM 32057: NECESIDAD DE ACTUALIZACIÓN.
1. EL VOCABULARIO CONTROLADO**

[THE ARGENTINE STANDARD FOR THESAURUS CONSTRUCTION IRAM 32057:
A NEED FOR UPDATING. 1. THE CONTROLLED VOCABULARY]

ANA M. MARTÍNEZ
PAOLA MENDES

Resumen: Se comparó la norma argentina de construcción de tesauros monolingües (IRAM 32057:1983) con las últimas ediciones de la norma internacional (ISO 2788:1986), las normas nacionales de España (UNE 50106:1990), Estados Unidos (Z39.19:2005) y Gran Bretaña (BSI 8723-2:2005b), así como diversos manuales reconocidos. Se analizaron tanto el alcance y los aspectos formales de la norma, como los conceptos referidos a la normativa sobre el vocabulario controlado. El alcance de la norma argentina se limita a la construcción de tesauros monolingües, mientras que las nuevas normas estadounidense y británica cubren también diferentes sistemas de organización del conocimiento. En relación con los aspectos formales, la norma argentina carece de tabla de contenido, índice analítico, glosario, bibliografía y algunas secciones y anexos, contando además con muy pocos ejemplos. En cuanto a los aspectos conceptuales, algunas definiciones son incorrectas o desactualizadas, lo mismo que algunas de las recomendaciones para el control del vocabulario. Se concluye que la norma argentina debería revisarse y actualizarse de conformidad con las normas publicadas recientemente y la norma ISO 25964 aún en preparación.

Palabras clave: Tesauros; Normalización; Norma IRAM 32057.

Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación
Cátedra de Organización del Conocimiento I. Calle 48 entre 5 y 6, 5º piso. 1900 La Plata,
Argentina. Correo electrónico: ammarti@speedy.com.ar.

Artículo recibido: 04-06-09. Aceptado: 12-11-09.

INFORMACIÓN, CULTURA Y SOCIEDAD. No. 21 (2009) p. 11-30

© Universidad de Buenos Aires. Facultad de Filosofía y Letras. Instituto de Investigaciones
Bibliotecológicas (INIBI), ISSN: 1514-8327.

Abstract: The Argentine standard for the construction of monolingual thesaurus (IRAM 32057:1983) was compared with the last editions of the international standard (ISO 2788:1986), and the national standards of Spain (UNE 50106:1990), United States (Z39.19:2005), and Great Britain (BSI 8723-2:2005b), and very well known manuals. The scope and formal issues were analyzed, as well as the conceptual issues for controlled vocabulary. The scope of the Argentine standard is limited to the construction of monolingual thesauri, while the new North-American and British standards also cover different knowledge organization systems. In relation to the formal issues, the Argentine standard has not table of content, analytical index, glossary, bibliography and some sections and annexes, having very few examples. In reference to conceptual issues, some definitions are incorrect or outdated, as well as some recommendations for vocabulary control. We conclude that the Argentine standards should be reviewed and updated in agreement to the recently published standards, and the ISO 25964 standard, still in preparation.

Keywords: Thesauri; Standardization; IRAM Standard 32057.

Introducción

En la década de 1970 se compilaron los primeros documentos normativos para la construcción de tesauros monolingües, entre ellos las directrices de Unisist (UNESCO, 1973), la norma internacional ISO 2788 (International Organization for Standardization, 1974), las normas nacionales de Estados Unidos (American National Standards Institute, 1974), Gran Bretaña (BSI Group, 1979), Chile (Instituto Nacional de Normalización de Chile, 1978), Colombia (Instituto Colombiano de Normas Técnicas, 1980) y otros (Aitchison y Dextre Clarke, 2004), así como los conocidos manuales de Aitchison & Gilchrist (1972) y Lancaster (1972).

En 1983, el entonces Instituto Argentino de Racionalización de Materiales (IRAM) publicó la correspondiente norma argentina IRAM 32057, que se basó en las primeras ediciones de las directrices de Unisist, la norma ISO 2788 de 1974, las normas nacionales de Colombia, Chile y Estados Unidos, así como en un documento preparado por un grupo de trabajo del Centro Argentino de Información Científica y Técnica (Instituto Argentino de Normalización, 1983).

Apenas tres años después de publicada la norma argentina, apareció la segunda edición de la norma ISO 2788 (International Organization for Standardization, 1986), que fue adoptada como norma nacional en varios países

de Europa, incluyendo España, donde la Asociación Española de Normalización (AENOR), luego de traducirla al español, la aprobó como norma UNE 50106 (Asociación Española de Normalización, 1990). También se publicaron nuevas ediciones de los manuales de Aitchison et al. (2000) y Lancaster (1986), éste último también traducido a nuestra lengua (Lancaster, 1995).

Ya en el siglo veintiuno, precisamente entre 2005 y 2007, se publicaron nuevas ediciones de las normas estadounidense (National Information Standards Organization, 2005) y británica (BSI Group, 2005a,b-2007a,b). Estas nuevas ediciones presentan dos novedades principales: a) no solamente normalizan los tesauros, sino también otros sistemas de organización del conocimiento como los sistemas de clasificación, las taxonomías, ontologías, etc. y b) hacen hincapié en el formato en línea y no en el impreso como sucedía en ediciones anteriores. Actualmente, la International Organization for Standardization (ISO) está desarrollando una nueva norma, la ISO 25964, basada en las nuevas normas estadounidense y británica, que sustituirá a las actuales normas ISO 2788 e ISO 5964 para tesauros monolingües y multilingües, respectivamente. La primera parte de esta nueva norma internacional (ISO 25964-1), que se encuentra aún en etapa de proyecto, se titula *Information and documentation: thesauri and interoperability with other vocabularies: Part 1. Thesauri for information retrieval* (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=53657).

En consecuencia, el propósito de este trabajo es comparar la norma argentina IRAM 32057 con las ediciones de las normas, directrices y manuales más recientes, con el fin de identificar aquellos aspectos que requieren la actualización de la norma que se encuentra aún vigente en nuestro país. Debido a la extensión del trabajo, se ha decidido dividirlo en varias partes: 1) el vocabulario controlado, 2) las relaciones entre los términos y 3) la aplicación de la tecnología a la gestión del tesoro.

Metodología



Se comparó la norma argentina IRAM 32057 de 1983, con las siguientes fuentes:

- Norma internacional ISO 2788 de 1986, para tesauros monolingües. Algunos conceptos de la norma argentina coinciden con la primera edición de la ISO 2788 de 1974, pero aquí se comparan con la segunda edición, que es la vigente.
- Normas nacionales de España UNE 50106 (1990), Estados Unidos Z39.19 (2005) y Gran Bretaña BS 8723-2 (2005).
- Manuales de construcción de tesauros de Aitchison et al. (2000), Lancaster (1995) en su versión española y Gil Urdiciain (2004).

Los aspectos analizados fueron:

- Alcance de la norma.
- Aspectos formales de la norma: tabla de contenido, secciones de la norma, ordenamiento de las secciones, bibliografía, índice analítico, glosario, etc.
- Aspectos conceptuales referidos al vocabulario del tesoro: términos empleados en la norma, definiciones y normativa.

Debido a la influencia que la Terminología ha tenido sobre las normas de construcción de tesauros, consideramos importante tener en cuenta los siguientes conceptos establecidos por las normas ISO 704 e ISO 1087-1 sobre principios, métodos y trabajo terminológicos (International Organization for Standardization 2000a,b):

- Un *objeto* es cualquier cosa que pueda percibirse o concebirse. Puede ser concreto como una silla, abstracto como la belleza o imaginario como una sirena.
- Un *concepto* es la representación mental de un objeto. Se define como una unidad de pensamiento y, dentro del campo de la Terminología, como una unidad de conocimiento creada por una combinación única de características, dentro de un dominio o campo de especialidad determinado. Un concepto se expresa o comunica por medio de una definición o una designación. Esta última puede ser verbal (un apelativo o un término) o un símbolo. Hay dos tipos de conceptos: concepto individual y concepto general.
- Un *concepto individual* es aquel que corresponde a un objeto único y se representa por medio de un apelativo o nombre propio o bien por un símbolo, por ejemplo *World Wide Web* o , *América* o . Por lo tanto, un *apelativo* o nombre propio es la designación verbal de un concepto individual.
- Un *concepto general* es aquel que corresponde a dos o más objetos, que forman un grupo en razón de propiedades comunes. Se representa por un término (*marca registrada*, *pesos*) o un símbolo (@, \$). Un *término* se define entonces como la designación verbal de un concepto general en un dominio específico. Aunque puede contener números, letras griegas, etc. (α -fetoproteína, carbono 14), básicamente es una designación verbal.

Algunos conceptos referidos al vocabulario de los tesauros se pueden comprender y explicar mejor cuando se asocian con los conceptos terminológicos antes mencionados, particularmente la relación concepto general/descriptor y concepto individual/nombre propio.

Resultados

1. Alcance de la norma

La norma argentina, así como la internacional y la española, limitan su alcance a la construcción de tesauros monolingües. En cambio, las nuevas normas estadounidense y británica han ampliado su alcance para incluir diversos sistemas de organización de conocimiento, como se indica a continuación:

- Norma ANSI/NISO Z19.39: abarca la construcción de tesauros, listas de términos, anillos de sinónimos y taxonomías.
- Norma BS 8723: abarca la construcción de tesauros, listas de epígrafes, sistemas de clasificación, taxonomías y ontologías.

2. Aspectos formales

En la Tabla 1 (p. 16) se muestran las distintas secciones en que se ha dividido el texto de las normas analizadas.

Como se puede observar, la norma argentina carece de algunas secciones importantes, por ejemplo:

- a. Tabla de contenido e índice analítico: resulta innecesario señalar que son recursos imprescindibles para una recuperación efectiva de la información en cualquier documento.
- b. Definiciones: es una sección importante porque en ella se definen los términos empleados en la norma, con lo cual se evitan confusiones. La norma argentina proporciona las definiciones pertinentes a lo largo del texto, pero no en una sección específica que se pueda consultar sin pérdida de tiempo. Las nuevas normas estadounidense y británica incluyen, en adición a esta sección, un glosario más amplio que en nuestra opinión resulta útil y recomendable.
- c. Abreviaturas y símbolos: en esta sección se indican las abreviaturas y símbolos utilizados para indicar las referencias correspondientes a las relaciones, notas, etc. Las abreviaturas son USE, UP (usado por), TG (término genérico), TE (término específico) y TR (término relacionado), así como aquellas referidas al término tope (TT), las relaciones jerárquicas de tipo género/especie (TGG/TEG), parte/todo (TGP/TEP) y la nota de alcance (NA). En la norma argentina, la mayoría de estas abreviaturas se encuentran dispersas a lo largo del texto, pero no en una sección específica o en un anexo como sucede con las otras normas. Por otro lado, la norma argentina no incluye los símbolos correspondientes a varias de las referencias anteriores, que son '(USE), = (UP), < (TG), <- (TGP), > (TE), >- (TEP) y - (TR) (Martínez, 2007).
- d. Control del vocabulario: es una sección introductoria en la que se explican las formas en que se logra el control del vocabulario en un tesauro, principalmente el control de la sinonimia, homonimia y polisemia, que

IRAM 32057	ISO 2788	UNE 50106	ANSI/NISO Z39.19	BS 8723-2
--	Contents	Contenido	Contents	Contents
--	--	--	Foreword	Foreword
1. Introducción	0. Introduction	0. Introducción	1. Introduction	Introduction
2. Objeto y campo de aplicación	1. Scope and field of application	1. Objeto y campo de aplicación	2. Scope	1. Scope
Antecedentes	2. References	2. Referencias	3. Referenced standards	2. Normative referents
3. Descripción, clasificación y relación con otros lenguajes documentarios	--	--	--	--
--	3. Definitions	3. Definiciones	4.1 Definitions	3 Definitions
--	4. Abbreviations and symbols	4. Abreviaturas y símbolos	4.2 Abbreviations and symbols	4. Symbols, abbreviations, and other conventions
--	5. Vocabulary control	5. Control del vocabulario	5. Controlled vocabularies	5. Vocabulary control in a thesaurus
4. Estructura del tesoro	6. Indexing terms	6. Los términos de indización	6. Term choice, scope, and form	6. Thesaurus terms scope, form and selection
4.2.2 Representación de conceptos mediante la asociación de varios descriptores	7. Compound terms	7. Términos compuestos	7. Compound terms	7. Complex concepts
4.4 Relaciones entre descriptores	8. Basic relationships in a thesaurus	8. Relaciones básicas en el tesoro	8. Relationships	8. Relationships between term
--	--	--	--	9. Facet analysis
5. Presentación y disposición del tesoro	9. Display of terms and their relationships	9. Presentación de los términos y sus relaciones	9. Displaying controlled vocabularies	10. Presentation and layout
--	--	--	10. Interoperability	Toda la Parte 4 de la norma BS 8723
--	--	--	--	11. Thesaurus functions in electronic systems
6. Metodología para la elaboración del tesoro	10. Management aspects of thesaurus	10. Aspectos organizativos de la construcción de tesoros	11. Construction, testing, maintenance, and management systems	12. Management aspects of thesaurus construction
7. Actualización	--	--	--	13. Updating
Anexo: Instituciones que informan sobre lenguajes de indización existentes	--	--	--	--
--	Anexo: Symbolization of thesaural relationships :	Anexo: Símbolos utilizados par alas relaciones en un tesoro	--	En Parte 1 de la norma BS 8723
--	--	--	Appendix A through F*	--
--	--	--	Bibliography	Bibliography
--	--	--	Index	Index
--	--	--	Glossary	En Parte 1 de la norma BS 8723

Tabla 1. Comparación de las secciones en que se divide el texto de las normas de construcción de tesoros

luego se describen con mayor detalle en las siguientes secciones. La norma argentina es la única de las normas analizadas que carece de esta sección introductoria.

- e. Bibliografía: en esta sección se aportan las referencias bibliográficas a otras normas, así como a manuales, tesauros, artículos de publicaciones periódicas, etc. Si bien solo las nuevas normas estadounidense y británica incluyen esta sección, sin duda puede resultar muy útil para quien decida construir un tesoro.

La norma argentina es la única norma que incluye la sección 3. *Descripción, clasificación y relación con otros lenguajes documentarios*, que se limita a una diferenciación muy pobre entre tesoro, lista de encabezamientos de materia y sistema de clasificación. No debe confundirse esta sección con la modalidad de las nuevas normas estadounidense y británica que aportan la normativa para la construcción de distintos sistemas de organización del conocimiento, como se indicó en el apartado anterior.

Por otro lado, la norma argentina, en su extensa sección 4. *Estructura del tesoro*, incluye la información referida a la forma, selección y desambiguación de los términos, el tratamiento de los términos compuestos y las relaciones entre términos, temas que en las restantes normas han sido organizados en tres secciones diferentes, menos extensas pero más fáciles de consultar.

También se observa en la Tabla 1 que el orden de las secciones en la norma argentina difiere del que presentan las restantes normas, que mantienen entre sí cierta similitud en el ordenamiento.

Con respecto a la tipografía que debe emplearse en el tesoro, las normas argentina (apartado 4.2.9.1), internacional y española (apartado 4.4 en ambos casos) recomiendan utilizar letras mayúsculas compactas para los descriptores y las referencias, así como letras minúsculas para los no descriptores, por ejemplo: patata USE PAPA. En cambio, las nuevas normas estadounidense y británica recomiendan limitar las mayúsculas a las iniciales de los nombres propios, las siglas y acrónimos y otros casos establecidos por las reglas gramaticales de cada idioma. La norma británica recomienda además la letra regular para los descriptores y cursiva o itálica para los no descriptores, con las referencias en mayúscula compacta y negrita, por ejemplo: *patata* **USE** papa. La norma estadounidense, por su parte, recomienda usar una letra regular o itálica para los no descriptores, negrita para los descriptores y mayúscula compacta regular para las referencias, por ejemplo: patata USE **papa**. Como se observa, no existe acuerdo respecto a la tipografía, excepto en la necesidad de distinguir los descriptores de los no descriptores.

En cuanto a los signos de puntuación, en general todas las normas coinciden en que deben evitarse, a menos que sean imprescindibles para

mantener el significado del descriptor, por ejemplo en *p-benzoquinona*. Los paréntesis deben evitarse porque están recomendados para los calificadores. Otros signos como el punto, la coma, los dos puntos, el punto y coma, el ampersand, etc. no deben utilizarse. Respecto al punto, la norma argentina indica que puede utilizarse “cuando una palabra debe cortarse porque la longitud del descriptor es limitada” (apartado 4.2.9.2); esta indicación es incorrecta, porque los descriptores no deben *cortarse*, pues se perdería precisión en la recuperación. Las siglas y acrónimos deben escribirse sin puntos intermedios, formando una palabra, por ejemplo: OEA, AFIP, SIDA. Las normas también coinciden en que los signos diacríticos pueden producir problemas en los sistemas informáticos, pero no se deben eliminar.

La norma argentina, la internacional y la española sugieren una notación de CDU para lograr una clasificación normalizada, pero no coinciden en las notaciones asignadas:

- Norma argentina: la notación es 001.815:025.3, donde 001.815 es *Elaboración de índices. Índices de autores. Índices de materia* y 025.3 es *Catalogación. Técnica de catalogación*. Es una notación incorrecta, pues el tesoro no es un índice y tampoco se usa en la catalogación, sino en la indización que se representa en CDU como 025.4.
- Normas internacional y española: la notación es 025.48 *Combinaciones de tesauros y clasificaciones. Tesauros facetados*. También podría tenerse en cuenta la notación de CDU 025.43 *Listas de términos controlados. Tesauros. Listas de descriptores*.

Las restantes normas no sugieren ninguna notación para su clasificación.

Una observación adicional en cuanto a los aspectos formales es la falta de ejemplificación en la norma argentina. En el apartado 4.1.3 no se ofrece ningún ejemplo de los conceptos que pueden ser descriptores, cuando las restantes normas ofrecen entre 18 y 35 ejemplos. De igual modo, en el mismo apartado la norma argentina carece de ejemplos para los distintos tipos de nombres propios, mientras que las restantes normas incluyen entre 2 y 12 ejemplos. Otras secciones de la norma argentina, si bien ofrecen ejemplos, lo hacen en una cantidad por lo general muy limitada.

La norma argentina no indica ISBN ni derechos de autor.

3. Aspectos conceptuales

3.1. Tesoro

La norma argentina proporciona dos definiciones de tesoro, tomadas de la primera edición de la norma internacional:

- a. Desde el punto de vista de su función: “es una herramienta de control terminológico utilizada para traducir a un lenguaje más restringido

- (lenguaje documentario, lenguaje de información) el lenguaje natural usado en los documentos y utilizado por los indizadores y los usuarios”.
- b. Desde el punto de vista de su estructura: “es un vocabulario controlado y dinámico de términos que guardan, entre sí, relaciones semánticas, el cual se aplica a un campo específico del conocimiento”.

Las restantes normas y manuales consultados, incluyendo la segunda edición de la propia norma internacional, no hacen esta diferencia por puntos de vista, siendo las distintas definiciones aceptadas coincidentes con la definición (b). En todos los documentos hay acuerdo en considerar al tesauro como un vocabulario controlado, es decir un vocabulario unívoco o monosémico. El término *lenguaje documentario* se usa poco hoy en día, teniendo más difusión su equivalente *lenguaje documental* (Barité, 2009; Asociación Española de Normalización, 1997; Gil Urdiciain, 2004). Actualmente otros autores consideran que el tesauro es un *sistema de organización del conocimiento* (Hodge, 2000; Tudhope, 2005).

En el apartado 3.2, la norma argentina distingue dos tipos de tesauros, también de acuerdo con la primera edición de la norma internacional:

- a. Tesauros que hacen explícitas las relaciones entre descriptores, ya sean de equivalencia (USE/UP), jerárquicas (TG/TE) o asociativas (TR/TR).
- b. Tesauros que no hacen explícitas las relaciones y que también se conocen como *listas de términos*.

Las normas y manuales consultados no aceptan más esta tipología, siendo considerados tesauros solamente los mencionados en el inciso (a), es decir los que hacen explícitas las relaciones entre términos. Más aún, la nueva norma estadounidense hace una clara distinción entre la lista de términos y el tesauro, como dos vocabularios controlados distintos.

3.2. Descriptores

La norma argentina usa preferentemente el término *descriptor*. Las otras normas, directrices y manuales han optado por los términos *preferred term* o *término preferente*, aunque hacen explícita la sinonimia con *descriptor*. Incluso la norma estadounidense, que en todas sus ediciones anteriores prefirió *descriptor*, en la última edición de 2005 ha optado por *preferred term*. La norma española usa *término preferente*.

En algunos apartados (por ejemplo 3.3.3 y 4.1.4), la norma argentina menciona *término preferencial* y en otros (5.3.1 y 5.3.2) *término preferido*, pero no hace una mención explícita de la sinonimia entre estos términos y *descriptor*. El *Diccionario de la lengua española* (Real Academia Española, 2001) reconoce que *preferente* y *preferencial* son sinónimos (lo que tiene

preferencia o superioridad sobre algo), pero no incluye *preferido* como adjetivo, sino como participio del verbo *preferir*.

Puede decirse que en nuestro medio *descriptor* está más difundido que *término preferente* y tal vez sea conveniente mantener su uso, pero la norma argentina debería hacer explícita la sinonimia entre ambos términos.

Por otro lado, la norma argentina, en su apartado 4.1.1, define *descriptor* como “un término o símbolo autorizado y formalizado, que figura en un tesauruso y se usa para representar, sin ambigüedad, los conceptos contenidos en los documentos y en los pedidos de recuperación de información”. De nuevo, esta definición coincide con la primera edición de la norma internacional.

Las restantes normas y los manuales consultados establecen que el descriptor o término preferente siempre es un término, es decir una designación verbal, de modo que no se contempla que pueda ser un símbolo como afirma la norma argentina. Por ejemplo, el concepto “acción y efecto de informar, dar noticia de algo” debe representarse con el término *información* y no con el símbolo ④.

En el apartado 4.1.3, la norma argentina establece que pueden considerarse como descriptores “los términos que representan conceptos o combinaciones de éstos”. Sin embargo, dado que un descriptor es un término, debe representar un solo concepto, no una combinación de ellos. La segunda edición de la norma internacional y la norma española definen *término preferente* como “un término usado coherentemente cuando se indiza para representar un concepto dado”, mientras que la norma británica lo define como “término usado para representar un concepto”.

En el apartado 4.1.2 la norma argentina distingue dos tipos de descriptores:

- “Descriptores generales, que podrán ser objeto de acuerdos multinacionales después de su traducción a diversas lenguas”.
- “Descriptores específicos que podrán ser objeto de acuerdos institucionales en los campos especializados”.

Esta distinción no figura en ninguna de las normas vigentes y manuales consultados. En nuestra opinión es una diferenciación innecesaria, porque los acuerdos a los que se hace referencia no modifican en nada la normativa aplicable a los descriptores o términos preferentes y, por otra parte, pueden llevar a confusión con el nivel de especificidad en una relación jerárquica (términos genéricos o términos específicos).

Respecto a las formas gramaticales de los descriptores, la norma argentina indica, en su apartado 4.2.1, que “una palabra o varias palabras pueden conformar un descriptor” y ofrece al respecto un solo ejemplo que es *teoría del conocimiento*. En las restantes normas y manuales se establece que un descriptor o término preferente puede ser una palabra o una *frase*, que no es lo mismo que *varias palabras*.

Cuando el descriptor está conformado por una sola palabra, ésta debe ser un sustantivo y se conoce como un descriptor o término preferente simple. Por ejemplo: edificios, procesos, vuelos.

Cuando el descriptor está conformado por una frase, ésta debe ser nominal, es decir el núcleo de la frase debe ser un sustantivo. El modificador de la frase, que califica al núcleo, puede ser:

- Directo: el modificador es un adjetivo y constituye el tipo de frase más recomendado por las normas y manuales. Por ejemplo: *edificios públicos, procesos técnicos, vuelos espaciales*.
- Indirecto: el modificador es una frase preposicional. Las normas y manuales recomiendan evitar en lo posible este tipo de frase, pero la terminología propia de cada dominio hace necesario utilizarlas con más frecuencia de la deseada. Por ejemplo: *edificios de oficinas, procesos de fabricación, vuelos de cabotaje*.

En el apartado 4.2.2, la norma argentina hace referencia al uso de adjetivos como descriptores y afirma que “existe, evidentemente, un cierto número de casos donde solamente pueden usarse adjetivos u otras formas no sustantivas”. En las restantes normas y en los manuales consultados no resulta tan evidente el uso de adjetivos como única palabra del descriptor. Es más, se presenta como una excepción. En lenguas como la nuestra, en las que el adjetivo se pospone al sustantivo, incluir los adjetivos como descriptores sólo sirve para que se puedan construir descriptores compuestos en el momento de la indización. Sin embargo, tanto la norma internacional como la española y la británica afirman que esta modalidad debe evitarse, pues tiende a producir incoherencias entre los indizadores.

Las normas y manuales coinciden en prohibir el uso de verbos y adverbios como descriptores. La norma argentina hace referencia a los verbos en el apartado 4.2.5, pero no menciona el uso de adverbios, que incluso pueden ser objeto de excepción.

En cuanto a la elección de la forma singular o plural para los descriptores o términos preferentes, la norma argentina en su sección 4.2.6 establece que “se realiza de acuerdo con los usos del idioma en el cual se construye el tesoro” y que “es necesario seguir las costumbres lexicográficas de la lengua”. Esto no es así exactamente, ya que en muchos casos podría elegirse el singular o el plural indistintamente, sin transgredir los usos y costumbres de una lengua. Por ejemplo: *árbol* o *árboles*, *brazo* o *brazos*, *reacción química* o *reacciones químicas*. Las normas y manuales consultados establecen reglas específicas para esta elección:

- Conceptos concretos: si responden a la pregunta ¿cuántos? el descriptor se indica en plural, por ejemplo *aves*, *bibliotecas*, *cámaras fotográficas*, *hospitales de día*. Si responde a la pregunta ¿cuánto? el descriptor debe

ir en singular, por ejemplo *acero, calcio, fósforo, sangre*. Un caso especial es el de las partes del cuerpo: si hay solo una en un mismo organismo va en singular y si hay dos o más en plural, por ejemplo: *boca, cabeza, nariz, pero arterias, ojos, piernas*.

- Conceptos abstractos: generalmente los descriptores que representan conceptos abstractos como los fenómenos, propiedades, sistemas de creencias, actividades o disciplinas se redactan en singular, por ejemplo: *verano, oscuridad, budismo, digestión, psicología*. No obstante se escriben en plural los descriptores que representan conceptos abstractos que abarcan distintos tipos o clases, por ejemplo *reacciones químicas, ciencias biológicas, conjuntos*.
- Cuando la forma singular y la plural representan conceptos distintos, las normas recomiendan una aclaración, ya sea por medio de un calificador o de una nota de alcance.

Por último, las normas vigentes y los manuales consultados incluyen numerosos ejemplos de descriptores, ordenados por las siguientes categorías: a) conceptos concretos: cosas y sus partes, materiales, b) conceptos abstractos: actividades o procesos, eventos u ocurrencias, propiedades, disciplinas o campos temáticos, unidades de medición, tipos de personas y organizaciones. La norma argentina carece de esta categorización y sus respectivos ejemplos, que deberían figurar en el apartado 4.1.3.

3.3 Desambiguación de descriptores

Hay dos formas de evitar la ambigüedad: los calificadores y las notas. La norma argentina trata estos temas en el apartado 4.3.

- Calificador: la norma argentina utiliza *término calificativo* en lugar de *calificador*, como usan otras fuentes en español, por ejemplo la norma española de construcción de tesauros, la norma española UNE 50113/6 sobre terminología de los lenguajes documentales (Asociación Española de Normalización, 1997) y los manuales de Gil Urdiciain (2004) y Lancaster (1995). Las normas y manuales consultados que se encuentran en inglés utilizan el término *qualifier* que coincide con *calificador*. El calificador es un término que se agrega entre paréntesis al descriptor y forma parte de éste. Su propósito es distinguir entre distintos homónimos o términos polisémicos que se encuentren en el tesauro, por ejemplo *tanques (contenedores)* y *tanques (vehículos)*. La norma argentina indica en el apartado 4.3.1 que el calificador puede ser un término o símbolo, pero las restantes normas no contemplan que pueda ser un símbolo.
- Nota de alcance: la norma argentina utiliza el término *nota de alcance*, pero también sugiere el sinónimo *nota de uso*, que no está contemplado

en las restantes normas ni en los manuales consultados. Otras fuentes en español como la norma española de construcción de tesauros, la norma UNE 50114/6 sobre terminología de los lenguajes documentales (Asociación Española de Normalización, 1997) y los manuales de Gil Urdiciain (2004) y Lancaster (1995) prefieren los términos *nota de aplicación* o *nota aclaratoria*. Las normas y manuales en inglés coinciden en el término *scope note*. El diccionario español/inglés Merriam Webster Online (2009) traduce *scope* como *alcance*, de modo que la norma argentina ha optado por un término correcto. La referencia para indicar la nota de alcance es NA que, si bien está incluida en los ejemplos de la norma argentina, no se hace explícita en el texto, como sucede con otras referencias.

La norma argentina, la internacional y la española, así como los manuales consultados, reconocen básicamente cuatro tipos de notas de alcance: a) definición del descriptor, b) limitación del alcance del descriptor, que puede ser positiva (este término incluye...) o negativa (este término excluye ...), c) instrucción para el indizador y d) historia del descriptor (fecha de incorporación, modificación o eliminación del término preferente). Todos estos tipos de notas se indican con la referencia NA. En cambio, las nuevas normas estadounidense y británica establecen tres tipos de notas: 1) *scope note*, cuya referencia es SN, equivalente en nuestra lengua a nota de alcance (NA) 2) *definition* o definición, cuya referencia es DEF y 3) *historical note*, cuya referencia es HN y podría traducirse al español como nota histórica o NH. El orden recomendado de estas notas es:

- Descriptor
- CC Notación o código de clasificación
- NA nota de alcance
- UP usado por
- TT término tope
- TG término genérico
- TE término específico
- TR término relacionado
- DEF definición
- NH nota histórica

Las mismas normas estadounidense y británica contemplan la posibilidad de establecer notas de alcance recíprocas, que remitan por ejemplo del descriptor A al descriptor B y viceversa. La norma estadounidense establece la referencia SN para remitir de A a B y la referencia X SN para la reciprocidad de B a A, lo cual constituye un verdadero disparate. En cambio, para la norma británica la referencia es SN en ambos casos, que es lo correcto. Un ejemplo de nota de alcance recíproca sería el siguiente:

teatro

NA Úsese para el género literario. Para el local donde se representa la obra teatral use *teatros*.

teatros

NA Úsese para el local donde se representa la obra teatral. Para el género literario use *teatro*.

3.4 Términos compuestos

La norma argentina, al igual que las normas internacional y española, parece confundir *precoordinación* con el tratamiento de los términos compuestos. De acuerdo con las nuevas normas estadounidense y británica, así como otros autores reconocidos (Foskett, 1996:97; Miller & Teitelbaum, 2002:87-93), es necesario recordar que el tesauro siempre es poscoordinado, mientras que la *precoordinación* es propia de otros sistemas de organización del conocimiento como las listas de encabezamientos de materia o los sistemas de clasificación.

Para explicar mejor esta diferencia, en la Tabla 2 se presentan 18 ejemplos de epígrafes tomados de la tercera edición impresa de la *Lista de encabezamientos de materia para bibliotecas*. Los tres primeros ejemplos son epígrafes simples (una palabra) y los siguientes seis son epígrafes compuestos (una frase). Ahora bien, cada uno de estos nueve ejemplos representa un solo concepto, motivo por el cual pueden ser considerados términos y, en consecuencia, serían aceptables como descriptores. No existe en ellos *precoordinación*.

En cambio, los nueve últimos ejemplos de la Tabla 2 son epígrafes compuestos que representan dos o más conceptos. Por lo tanto no pueden ser considerados términos y no serían aceptables como descriptores. En estos últimos nueve ejemplos es donde existe *precoordinación*.

Por otro lado, las normas de construcción de tesauros incluyen diversas reglas y recomendaciones para el tratamiento de los términos compuestos, que se orientan principalmente a decidir si un término compuesto se mantiene o se fracciona en dos o más términos.

La norma internacional y la española reconocen dos tipos de fraccionamiento: a) el fraccionamiento semántico (*división morfológica* en la norma argentina, apartado 4.2.2.1) y b) el fraccionamiento sintáctico (*división semántica* en la norma argentina, apartado 4.2.2.1), aunque sólo el segundo está aceptado por las normas y manuales. Las nuevas normas estadounidense y británica ni siquiera mencionan el fraccionamiento semántico. Si bien las reglas para el fraccionamiento son diversas y complejas, la nueva norma británica considera que, en la práctica, el fraccionamiento de términos compuestos se puede resumir en los siguientes casos:

Tipo de epígrafes	Epígrafe	Conceptos representados		¿Precoordinación?
		Cantidad	Concepto	
Simple (una palabra)	Ornitólogos	1	Ornitólogos	No
	Psicoterapia	1	Psicoterapia	No
	Sencillez	1	Sencillez	No
Compuesto (una frase nominal con modificador directo)	Radiación solar	1	Radiación solar	No
	Retiros espirituales	1	Retiros espirituales	No
	Salud pública	1	Salud pública	No
Compuesto (una frase nominal con modificador indirecto)	Instrumentos de medición	1	Instrumentos de medición	No
	Libertad de conciencia	1	Libertad de conciencia	No
	Tiro al blanco	1	Tiro al blanco	No
Frases que representan más de un concepto	Hidratos de carbono en el organismo	2	Hidratos de carbono Organismo	Si
	Inyección de agua en campos petroleros	3	Agua Campos petroleros Inyección	Si
	Levadura como alimento para animales	3	Alimento Animales Levadura	Si
Frase coordinadas con la conjunción Y (o E si corresponde)	Cine y televisión	2	Cine Televisión	Si
	Libros y lectura	2	Lectura Libros	Si
	Limpieza y mantenimiento de hoteles	3	Hoteles Limpieza Mantenimiento	Si
Cadena precoordinada (epígrafe + subdivisiones)	Egipto-historia-insurrección-1919	4	1919 Egipto Historia Insurrección	Si
	Música-historia y crítica-siglo XX	4	Crítica Historia Música Siglo XX	Si
	Sociología-investigaciones-Argentina-Bibliografía-CDROM	5	Argentina Bibliografía CDROM Investigaciones Sociología	Si

Tabla 2. Ejemplos de epígrafes precoordinados y no precoordinados, tomados de la tercera edición impresa de la *Lista de encabezamientos de materia para bibliotecas (LEMB)*

- a. Cuando el núcleo está calificado por dos o más adjetivos, por ejemplo *escuelas primarias públicas*, debería fraccionarse en dos descriptores: *escuelas primarias* y *escuelas públicas*. En el momento de la recuperación, el usuario tendría que poscoordinar estos dos descriptores con el operador AND: *escuelas primarias AND escuelas públicas*.
- b. Cuando el núcleo representa una propiedad, parte o componente del modificador, por ejemplo *motores de avión* puede fraccionarse en los descriptores *motores* y *aviones*, que deberán ser asignados conjuntamente por el bibliotecario en el momento de la indización. Cuando la poscoordinación es de este tipo, se indica en el tesauro con un signo de adición (+) de la siguiente manera:

aviones
UP+ motores de avión
motores
UP+ motores de avión
motores de avión
USE motores + aviones

La norma argentina recomienda otra referencia para esta última relación, que es UPC (usado por combinación), que no se usa en las restantes normas y manuales consultados.

El fraccionamiento debe evitarse en los siguientes casos:

- Cuando el término compuesto es aceptado por la terminología del dominio del tesauro, por ejemplo en Química el término *cromatografía líquida de alta resolución*.
- Cuando el modificador posee un significado distinto o es una metáfora: *mesa redonda, estructura de árbol, silla de tijera*.
- Cuando el núcleo pierde su significado en ausencia del modificador. Por ejemplo *flores artificiales* o *huevos de Pascua*, que no son flores ni huevos, respectivamente.

3.5 Identificadores o nombres propios

En el apartado 4.1.3, la norma argentina establece, al igual que las restantes normas, que pueden considerarse como descriptores “los términos que representan entidades individuales; estos términos se llaman también nombres propios o identificadores”. Si se respetan las normas de trabajo terminológico (International Organization for Standardization, 2000a,b), debe considerarse que los nombres propios o identificadores no son *términos* sino *apelativos*: los primeros sólo representan conceptos generales, mientras que los segundos representan conceptos individuales, por lo tanto un nombre propio no puede ser descriptor.

La norma internacional y la norma española de construcción de tesauros no reconocen el término *identificador* para este tipo de designación. En cambio, la norma estadounidense y la norma española UNE 50113/6 sobre terminología de lenguajes documentales (Asociación Española de Normalización, 1997) y los manuales de Aitchison et al. (2000) y Lancaster (1995) sí reconocen la sinonimia entre *nombre propio* e *identificador*. El problema más serio se presenta con la nueva edición de la norma británica, que le ha dado un significado distinto al término *identificador*, definido como “un conjunto de símbolos, por lo general alfanuméricos, que designan un concepto o un término u otra entidad, con propósitos de identificación única dentro de un contexto determinado de recursos, en un sistema de computación o red [...] Algunas veces, una notación puede usarse como identificador”. En otras palabras, se trata de una ID única para cada signo del tesoro e incluso para los términos candidatos aún no autorizados, que facilita la gestión de dichos signos en la base de datos del tesoro. Si este último fuera el significado que prevaleciera en el futuro, sería necesario eliminar el significado anterior de *identificador* y utilizar solamente el término *nombre propio* para referirse al signo del tesoro que consiste en un apelativo o designación verbal de un concepto individual.

3.6 No descriptores

La norma argentina utiliza el término *no descriptor*, que está aceptado en las restantes normas y manuales. Sin embargo, las normas y manuales prefieren usar el sinónimo *término no preferente* o *non-preferred term*. Sería conveniente que la norma argentina hiciera explícita la sinonimia entre estos términos.

Por otro lado, la norma argentina expresa en el apartado 4.1.4, que “los términos que representan conceptos no establecidos para la indización, deben considerarse como términos no autorizados. Se llaman no descriptores”. Es una definición incorrecta, porque los no descriptores sí representan conceptos establecidos para la indización, aunque lo hagan con términos distintos al descriptor. En otras palabras, son sinónimos del descriptor (coinciden en el significado o concepto, pero no en el significante o término).

Conclusiones

La conclusión general es que la norma argentina IRAM 32057:1983 debería revisarse y actualizarse. Existen en nuestra opinión distintas opciones para ello:

Una opción sería adoptar la norma internacional ISO 25964, una vez que sea publicada y, por supuesto, traducida al español. Es una modalidad practicada desde hace tiempo por la Asociación Española de Normalización y tiene la ventaja de lograr una conformidad total entre la norma internacional y la nacional.

Si por algún motivo es necesario crear una norma nacional distinta a la norma ISO 25964, entonces deberían tenerse en cuenta los aspectos analizados aquí u otros que pudieran surgir en el futuro, para su discusión. En particular:

- Alcance de la norma: habrá que decidir si la norma argentina seguirá limitada al tesoro monolingüe o se ampliará a otros sistemas de organización del conocimiento.
- Aspectos formales: es necesario incorporar en la norma argentina una tabla de contenido, un índice analítico, un glosario, una bibliografía y otras secciones y anexos como sucede en las restantes normas analizadas, así como lograr una mejor ejemplificación.
- Aspectos conceptuales: se requiere revisar y actualizar los términos usados en la norma, sus sinónimos, las definiciones de dichos términos y las recomendaciones de la norma, muchas de las cuales provienen de la primera edición de la norma ISO 2788, publicada hace 35 años.

Por último recordemos que en este trabajo solo hemos tomado en cuenta los aspectos conceptuales del control del vocabulario y que más adelante analizaremos las recomendaciones de la norma argentina para las relaciones entre los términos del tesoro, así como la utilización de la tecnología informática en la construcción y uso de los tesauros.

Referencias bibliográficas

- Aitchison, Jean y Stella Dextre Clarke. 2004. The thesaurus: a historical viewpoint, with a look to the future. En Roe, Sandra K. y Alan R. Thomas, ed. The thesaurus: review, renaissance, and revision. New York: Haworth Information Press. p. 23-34.
- Aitchison, Jean y Alan Gilchrist. 1972. Thesaurus construction and use: a practical manual. London: Aslib. 95 p.
- Aitchison, Jean; Alan Gilchrist y David Bawden. 2000. Thesaurus construction and use: a practical manual. 4th ed. London: Aslib. 218 p.
- American National Standards Institute. 1974. Guidelines for thesaurus structure, construction, and use. New York: The Institute. 20 p. (ANSI Z39.19-1974).

- Asociación Española de Normalización. 1990. Directrices para el establecimiento y desarrollo de tesauros monolingües. Madrid: La Asociación. 47 p. (UNE 50106:1990).
- Asociación Española de Normalización. 1997. Documentación e información: vocabulario. Parte 6: lenguajes documentales. Madrid: La Asociación. 32 p. (UNE 50113/6).
- Barité, Mario. 2009. Diccionario de organización del conocimiento: clasificación, indización, terminología. Montevideo: El Autor. <<http://www.eubca.edu.uy/barite/espaniol/diccionario.php>>. [Consulta: 3 marzo 2009].
- BSI Group. 1979. Guide for the establishment and development of monolingual thesauri. London: The Group. 15 p. (BS 5723:1979).
- BSI Group. 2005a. Structured vocabularies for information retrieval: guide. Part 1: definitions, symbols and abbreviations. London: The Group. 14 p. (BS 8723-1:2005).
- BSI Group. 2005b. Structured vocabularies for information retrieval: guide. Part 2: thesauri. London: The Group. 64 p. (BS 8723-2:2005).
- BSI Group. 2007a. Structured vocabularies for information retrieval: guide. Part 3: vocabularies other than thesauri. London: The Group. 52 p. (BS 8723-3:2007).
- BSI Group. 2007b. Structured vocabularies for information retrieval: guide. Part 4: Interoperation between multiple vocabularies. London: The Group. 62 p. (BS 8723-4:2007).
- Foskett, Antony C. 1996. Subject approach to information. 5ª ed. London: Library Association. 456 p.
- Gil Urdiciain, Blanca. 2004. Manual de lenguajes documentales. 2ª ed. Gijón: Trea. 280 p.
- Hodge, Gail. 2000. Systems of knowledge organization for digital libraries. Bethesda M: Council of Library Resources. <<http://www.clir.org/pubs/reports/pub91/contents.html>> [Consulta: 9 marzo 2009].
- Instituto Argentino de Normalización. 1983. Tesauros monolingües para la recuperación de la información: estructuración y desarrollo. Buenos Aires: El Instituto. 32 p. (IRAM 32057:1983).
- Instituto Colombiano de Normas Técnicas. 1980. Documentación: directrices para el establecimiento y desarrollo de tesauros monolingües. Bogotá: El Instituto. 71 p. (ICONTEC 1476:1980).
- Instituto Nacional de Normalización. 1978. Estructuración y desarrollo de tesauros monolingües destinados a la recuperación de información. Santiago de Chile: El Instituto. 35 p. (NCH 1278:1978).

- International Organization for Standardization. 1974. Documentation: guidelines for the establishment and development of monolingual thesauri. Geneva: The Organization. 14 p. (ISO 2788:1974).
- International Organization for Standardization. 1986. Documentation: guidelines for the establishment and development of monolingual thesauri. 2nd ed. Geneva: The Organization. 32 p. (ISO 2788:1986).
- International Organization for Standardization. 2000a. Terminology work: principles and methods. 2nd ed. Geneva: The Organization. 38 p. (ISO 704:2000).
- International Organization for Standardization. 2000b. Terminology work: vocabulary. Part 1: theory and application = Travaux terminologiques: vocabulaire. Partie 1: théorie et application. 2nd ed. Geneva: The Organization. 41 p. (ISO 1087-1:2000).
- Lancaster, Frederic W. 1972. Vocabulary control for information retrieval. Washington DC: Information Resources Press. 233 p.
- Lancaster, Frederic W. 1986. Vocabulary control for information retrieval. 2nd ed. Arlington: Information Resources Press. 270 p.
- Lancaster, Frederic W. 1995. Control del vocabulario para la recuperación de información. Valencia: Universitat de Valencia. 286 p.
- Martínez, Ana M. 2007. Abbreviations and mark symbols recommended by thesaurus standards and authority guidelines: a need for uniform criteria. En *International Cataloging and Bibliographic Control*. Vol. 26, no. 4, 83-86.
- Merriam Webster online dictionary. Springfield MA: Merriam Webster Inc. <<http://www.merriam-webster.com>> [Consulta: 9 marzo 2009].
- Miller, Uri y Ruth Teitelbaum. 2002. Pre-coordination and post-coordination: past and future. En *Knowledge Organization*. Vol. 29, no. 2, 87-93.
- National Information Standards Organization. 2005. Guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda MD: NISO Press. 184 p. (ANSI/NISO Z39.19:2005). <<http://www.niso.org/standards/resources/Z39-19-2005.pdf>> [Consulta: 28 diciembre 2008].
- Real Academia Española. 2001. Diccionario de la lengua española. 22ª ed. Madrid: La Academia. <<http://www.rae.es>> [Consulta: 9 marzo 2009].
- Tudhope, Doug. 2005. Report on the Fourth European Networked Knowledge Organization Systems/Services (NKOS) Workshop. D-Magazine. Vol. 11, no. 11. <<http://www.dlib.org/dlib/november05/tudhope/11tudhope.html>> [Consulta: 9 marzo 2009].
- UNESCO. 1973. UNISIST Guidelines for the establishment and development of monolingual thesauri. París: Unesco. 37 p.