

Listas de encabezamientos de materia y Thesauri en perspectiva comparada

Por José María IZQUIERDO ARROYO y
Luis Miguel MORENO FERNÁNDEZ
Universidad de Murcia

RESUMEN

Se efectúa un análisis comparativo sistemático de las Listas de Encabezamiento de Materias y los Thesauri, para poner de relieve su semejanza y diferencias, así como las posibilidades de utilización que ofrecen. El procedimiento comparativo comienza con una breve incursión histórica en las circunstancias en que surgieron esos lenguajes, a fin de conocer las necesidades que vinieron a satisfacer y las causas de su posterior evolución en el decurso del tiempo; y sigue con el análisis de sus principios y características, estructura y costes de construcción y utilización. Además de las notables diferencias existentes entre ambos lenguajes documentales, las listas de encabezamientos parecen haber iniciado un proceso de thesaurificación creciente e irreversible en los últimos años, lo que nos lleva a plantearnos la interrogante acerca de si los cambios introducidos en su filosofía y estructura generarán un lenguaje híbrido o darán al traste definitivamente con los lenguajes más tradicionales de la indización por materias.

Palabras clave: Listas de encabezamientos de materia / Tesoros / Análisis comparativo / Lenguajes documentales / Indización por materias.

PRINCIPIOS Y CARACTERÍSTICAS GENERALES

Génesis de las LEM.

Aunque la práctica o ciertos aspectos prácticos de la **indización por materias** se conocieran desde la antigüedad (1), parece comúnmente admitido

que fue un contemporáneo de M. Dewey, Ch. A. Cutter, el primero en tratar de reglar el procedimiento y en enunciar sugerencias acerca de la forma que debían tener los encabezamientos de materias en su «Rules for a Dictionary Catalog», publicado por vez primera en 1876 (2-3). El catálogo alfabético por materias había de permitir al usuario encontrar un libro sobre una materia conocida, o bien facilitarle cuanta información tuviese una biblioteca sobre una materia dada. Objetivos que —se suponía— sólo serían alcanzables si se utilizaba un vocabulario controlado y normalizado (4).

En realidad, los encabezamientos de materias no eran sino el complemento de las clasificaciones sistemáticas para recuperar la información. Quizá sea buena prueba de ello el hecho de que, por ejemplo, los subencabezamientos coincidan con las subdivisiones de ciertos lenguajes clasificatorios como la CDU: Punto de vista, lugar, tiempo, forma. Ambos constituyeron el producto de una época en la cual la preocupación por clasificar convenientemente los documentos llevaba a organizar con eficacia la búsqueda en favor de los usuarios (5-6).

En vida de Cutter y hasta principios del siglo xx el indizador debía suplir los encabezamientos de una inexistente lista con los que él mismo preparaba. Confeccionar el **catálogo alfabético** con la sola ayuda de la experiencia personal y de una serie de reglas y observaciones resultaba un trabajo creativo, pero también arduo y plagado de riesgos. Para evitarlos, la American Library Association procedió a elaborar una lista con los encabezamientos utilizados en las mayores bibliotecas del país. La edición final de esta lista vio la luz en 1911, unos cuantos años después de la primera edición de la LCSH, tan impregnada de los postulados de Cutter (2).

Filosofía y rasgos de las LEM

TABLA 1
Marco general en los lenguajes documentales

Características generales de los lenguajes documentales		LEM	THESAUR I
1.1	Lenguaje controlado	+	+
1.2	Lenguaje "universal"	+	-
1.3	Lenguaje especializado	-	+

De esta manera nacieron las LEM. Estas son listas de términos controlados, igual que los thesauri, dispuestos en orden alfabético, donde se muestran las relaciones que han contraído los conceptos representados por tales términos, susceptibles de utilizarse en un índice, catálogo o base de datos para la descripción característica de las materias (7).

Si bien el criterio clasificatorio del control aproxima a ambos lenguajes documentales, dado que pretenden eliminar o paliar los riesgos derivados de la ambigüedad inherente al lenguaje natural, su filosofía o principios les diferencian sensiblemente en una doble faceta:

- a) Campo de aplicación;
- b) Precoordinación y modo de representar el contenido del documento.

De estos dos aspectos el de mayor peso en la diferenciación de los dos lenguajes documentales lo aporta el principio de la precoordinación, como intentaremos demostrar.

a) Campo de aplicación.

Las diferencias tocantes al campo de aplicación son a nuestro juicio más accidentales o circunstanciales, si se quiere, que de fondo.

El catálogo alfabético de materias y las LEM constituyen los instrumentos por excelencia de la indización y búsqueda documental de las bibliotecas nacionales, municipales y universitarias. Se utilizan con preferencia en las bibliotecas de tipo enciclopédico, que no limitan sus colecciones a un dominio particular. En cambio, los thesauri suelen centrarse en un campo concreto del conocimiento y se nutren de un lenguaje más especializado y técnico que las listas.

No es fácil averiguar hasta qué extremo esto viene impuesto por la tradición y los convencionalismos o por la naturaleza del lenguaje. Pero N. Richter advierte con firmeza que la indización alfabética mediante LEM es susceptible de aplicarse a «dossiers» documentales, documentos iconográficos y cartográficos, y, en general, a todos los soportes de información (6).

Mejores perspectivas para el contraste de estos dos tipos de lenguajes documentales tienen los aspectos relacionados con la **coordinación** y el modo de representar la información.

b) Precoordinación y modo de representar el contenido del documento.

Según Van Dijk y Van Slype la **indización coordinada** «es el principio de base de todos los lenguajes de indización» (5); y una indización es coordinada en la medida que los conceptos y las palabras o cualesquiera otros signos alfanuméricos empleados para representarlos son susceptibles de combinarse entre sí para caracterizar lo que convenimos en llamar contenido noético de los documentos. Al describir éste, la coordinación o combinación, tal y como se ha expuesto, no diferencia entre sí sustancialmente a los lenguajes documentales; es más, puesto que hablamos de coordinar o combinar términos o códigos alfanuméricos para representar el conteni-

do neotético de los documentos, cabe inferir que la coordinación es una regla o facultad común a todos los lenguajes documentales en el momento de indizar los contenidos.

Ahora bien, las **construcciones gramaticales** de los encabezamientos de materias de las listas poseen una elevada tasa de precoordinación, mientras que los descriptores de los thesauri responden al principio de la postcoordinación. Los lenguajes documentales **precoordinados** reciben este nombre porque proporcionan al indizador ya vinculados los diferentes-conceptos que forman un tema, de suerte que se utilizan juntos —precoordinados— en el momento de la indización del documento, componiendo la expresión que utilizaremos en su momento para describirlo (8). Esto elimina la posibilidad de falsas combinaciones y da precisión a la indización. Courier refiere al respecto que los sistemas precoordinados prevén por lo general el empleo de un solo encabezamiento por documento (8-9).

Sin embargo, no hay razones científicas que determinen que los lenguajes precoordinados utilicen por regla general sólo un encabezamiento de materia para indizar un documento. Otra cosa es que la práctica propicie el uso de un encabezamiento. Salvada, pues, esta circunstancia, puede decirse que un documento admite varias unidades de indización si versa sobre varias materias de manera sucesiva o si es susceptible de enfocarse bajo varios aspectos (10-11). Pero los planteamientos acerca del número de encabezamientos que se considera más conveniente para indizar un tema han evolucionado a lo largo del tiempo; por ejemplo, la especialista en la LCSH Lois M. Chan manifiesta que el promedio de encabezamientos asignados con la LCSH es inferior a dos, aunque esto depende del tipo de documento, y, entre otros factores, de la política impuesta por la biblioteca (12), mientras que el manual de indización REBUS señala que se pueden emplear hasta nueve encabezamientos (10). (TABLA 2)

Esto nos lleva a pensar que las LEM, al igual que otros lenguajes documentales precoordinados, tienden a efectuar una representación del contenido de los documentos más **sintética**, o menos **analítica** si se quiere, que los lenguajes postcoordinados como los thesauri. En nuestra opinión —como ya dijimos en un trabajo anterior—, el binomio analítico/sintético no debe plantearse en términos radicales, pues al igual que sucede con otros aspectos de los lenguajes documentales esa dualidad no desemboca en un problema de absolutos, sino de grados. Los extremos, en todo caso, se toman como punto de referencia para facilitar la comparación. Tal vez sea por eso por lo que en un reciente trabajo se ha dicho que las listas son instrumentos para la clasificación «analítico-sintética» (13).

Además, cuando se habla de lenguaje documental sintético/analítico debe tenerse en cuenta una cosa importante: Que el término «sintético», y, por lo tanto, su correlato, denota no sólo en número de términos utilizables en la descripción del contenido noético de los documentos —lo que se ha dado en denominar exhaustividad de la indización—. También expresa

TABLA 2
Principios y codificación del contenido

PRINCIPIOS DE UTILIZACIÓN Y REPERCUSIÓN EN EL TRATAMIENTO DOCUMENTAL DE LAS / LOS		LEM	THESAURI
2.1	Representación analítica del contenido		+
2.2	Codificación de la información mediante expresiones extraídas de la lengua natural	+	+
2.3 <i>PRECOORDINACIÓN</i>	2.3.1 No trasluce la multidimensionalidad de las relaciones entre los términos	+	-
	2.3.2 Los términos se disponen en un orden prefijado	+	-
	2.3.3 No se combinan los términos en el momento de la búsqueda	+	-
	2.3.4 Potencia semántica	-	+
2.4 <i>POSTCOORDINACIÓN</i>	2.4.1 Multidimensionalidad de las relaciones entre los términos	-	+
	2.4.2 Los términos tienen el mismo peso específico	-	+
	2.4.3 Se combinan los términos en el momento de la búsqueda	-	+
	2.4.4 Potencia semántica	-	+

que la indización mediante LEM busca la captación «global» del tópico o tema del que trata el documento. Así, la caracterización temática lleva implícita la «síntesis» o, lo que los anglosajones denominan «sumarización» del contenido del documento. Dicho con otras palabras: el ideal es que un

solo encabezamiento represente el contenido del documento en su totalidad, en lugar de realizar una selección más detallada de aquellos conceptos que sirvan para expresárnoslo.

Lancaster ha señalado que la precoordinación en los lenguajes documentales tiene estas consecuencias:

a) Es difícil describir la **multidimensionalidad** de las relaciones entre los términos;

b) Los términos se disponen siguiendo una secuencia determinada (A, B, C, D, E), que implica que el primer término es más importante que los otros;

c) No es fácil, si bien no completamente imposible, **combinar** los términos en el momento de la búsqueda (14). Es decir, cuando un índice por materias aparece impreso o bajo la forma de un catálogo de fichas convencional, se sacrifica la verdadera multidimensionalidad. Es posible construir unas entradas en las que retengamos todos los términos de indización, pero esos términos deben ordenarse siguiendo una secuencia lineal, de manera que el acceso al documento puede efectuarse solamente a través del primer término de la cadena (15).

Génesis de los THESAURI

Los lenguajes clasificatorios y las LEM aparecieron en una época en la que predominaba la preocupación por clasificar los documentos física y temáticamente para localizarlos con rapidez. Pero la «explosión de la información», su diversificación y especialización, desbordaron la capacidad de aquéllos para representar y recuperar el contenido de los documentos.

Las investigaciones desarrolladas para superar esos inconvenientes vinieron a coincidir en que la solución pasaba por **profundizar** en la indización y **flexibilizar** las búsquedas. Lo demuestra la aparición del sistema unitérmino de *Taube* —1951-53—, la elaboración del índice permutado KWIC —*Luhn*, 1959— y la indización por «conceptos» —«descriptores»— de *Moövers* en 1960. Estas tres personalidades, además de revolucionar las técnicas de indización (10), pusieron los jalones teóricos y metodológicos que permitirían la creación y uso de los thesauri en la década de los 60. El nacimiento del fichero invertido —*Taube*—, gracias al cual es posible acceder a la información de manera directa y no secuencial, la combinación de términos en la búsqueda, la indización por conceptos y el empleo de índices permutados, son otros tantos pilares sobre los que se orientan los modernos sistemas de documentación y los thesauri. En cambio, hay autores que subrayan la nula incidencia práctica de las LEM en la creación del thesaurus (15).

Filosofía y rasgos generales de los THESAURI

Estos lenguajes controlados tienen por objeto representar el contenido de los documentos y de las consultas de forma analítica (16). Su utiliza-

ción responde al principio de la **postcoordinación**; esto significa que los conceptos y palabras manejados en la indización podrán combinarse entre sí para formular las consultas que permitan recuperar los documentos que los contengan. Es decir; los lenguajes postcoordinados permiten yuxtaponer o coordinar los conceptos en la fase del análisis —indización de documentos— para poder combinarlos después del almacenamiento en la búsqueda documental (8). Esto les confiere una **potencia semántica** superior a los que tienen como norte la precoordinación. La **potencia semántica** es el número de expresiones no sinónimas correctamente construidas, e incide en la eficacia de los lenguajes de indización (17).

Eso no impide que un lenguaje documental posea ciertas características consideradas propias de otro; los thesauri presentan una determinada tasa de precoordinación que oscila entre 1,50 y 2,00, dependiendo del idioma en el que han sido hechos (16-18). Todo ello obedece a que no existen los lenguajes documentales «puros», y, como ya se indicó antes, el contraste de algunos de sus rasgos y el desarrollo de otros es cuestión de grado.

La oposición precoordinación/postcoordinación tiene amplias repercusiones en el manejo de los lenguajes documentales, como se verá más adelante. Pero Lancaster ha detectado estas **características en los sistemas postcoordinados**:

- a) Conservan la multidimensionalidad de las relaciones entre los términos;
- b) Todos los términos asignados al documento tienen igual peso específico, o sea, uno no es más importante que otro, aunque posteriormente, si se procede a ponderar la indización, se rompa este equilibrio (14).

ESTRUCTURA

Recientemente se ha subrayado que una de las diferencias entre las LEM y los thesauri radican en que éstos contienen términos más específicos que aquéllas(7-13). Pero esto no es más que una consecuencia del distinto ámbito al que se aplican los lenguajes documentales. Hay, a nuestro juicio, aspectos relativos a las unidades léxicas que brindan mejores posibilidades para el estudio comparado de estos lenguajes documentales. Nosotros destacamos los siguientes: (TABLA 3)

Elementos constituyentes

- a) Las LEM poseen mayor **riqueza terminológica** que los thesauri.

Tal vez porque aquéllas se extienden, al igual que las clasificaciones a las que suelen acompañar, por los dominios enciclopédicos del conocimiento. Al contrario que las LEM, los thesauri no suelen incluir como tér-

TABLA 3
Comparación de la estructura lingüística general

ESTRUCTURA DE LAS / LOS		LEM	THESAUR I	
3.1 UNIDADES LÉXICAS <i>Elementos</i>	3.1.1	Términos específicos	-	+
	3.1.2	Riqueza terminológica	+	-
	3.1.3	Volumen	+	-
	3.1.4	Descriptorios (= Términos preferidos)	+	+
	3.1.5	No descriptorios	+	+
	3.1.6	Microdisciplinas (= Microthesauri)	-	+
3.2 GRAMÁTICA		Sintaxis variada	-	+
3.3 RED SEMÁNTICA <i>Relaciones semánticas entre las unidades léxicas</i>	3.3.1	Relaciones de inclusión (# pertenencia) entre términos y clases (microthesauri)	+	+
	3.3.2	Relaciones de equivalencia semántica intralingüística	+	+
	3.3.3	Relaciones de equivalencia semántica interlingüística	?	+
	3.3.4	Relaciones jerárquicas	+	+
	3.3.5	Asociación entre términos	+	+
3.4 ESTRUCTURA SINDÉTICA COMPLEJA			-	+

minos de indización nombres de ciudades, comunidades, montañas, ríos o desiertos, ni nombres de personas, de organizaciones o proyectos; lo cual no obsta para que haya excepciones, como es el caso del *Tesoro de Defensa* (19). Esos elementos pueden figurar en una lista de identificadores como, por ejemplo, sucede en el caso del thesaurus de la UNESCO (20).

Esta riqueza terminológica guarda correlación con el **tamaño** de los lenguajes documentales. Las LEM son, en promedio, más voluminosas que los thesauri. Lo que implica que hay algunas menos voluminosas que un thesaurus; por ejemplo, la del Ministerio de Cultura es menos voluminosa que el thesaurus de la UNESCO (21).

Claro está que en el tamaño de las LEM influyen otros factores:

- El origen y formación de la lista, así como la naturaleza del centro o centros de documentación a los que sirve.
- La precoordinación del lenguaje documental.

Al elaborar un lenguaje precoordinado se requieren más términos para representar el mismo campo documental que si se pretende abarcarlo con un lenguaje postcoordinado (14). Esto supo verlo M. Taube cuando decidió fraccionar los **subencabezamientos** (o subdivisiones) de la Biblioteca del Congreso en elementos unitarios que llamó unitérminos, demostrando que unos 50.000 encabezamientos podían reducirse a 3.000 unitérminos.

b) Las LEM carecen de las agrupaciones de términos que originan las **microdisciplinas**.

En efecto, LEM y thesauri contienen esencialmente dos tipos de términos: los preferentes, empleados en la indización de documentos; y los equivalentes no preferentes, punto de partida en la indización de documentos y de cuestiones. Su misión no deja lugar a dudas; incardinados en la trama semántica del lenguaje documental ayudan a limitar y controlar la ambigüedad inherente al lenguaje humano. Pero los thesauri, a fin de facilitar al usuario la aprehensión de su estructura y de las relaciones semánticas que sus unidades léxicas mantienen, proceden a agrupar éstas bien por temas o campos semánticos, bien por facetas, constituyendo agrupaciones de términos denominadas generalmente **microthesauri** o microdisciplinas. Y ello independientemente de que en el interior de cada una de esas grandes categorías temáticas se adopte la decisión de organizar la jerarquía sobre **temas** o **facetats** (16-22). Pues bien, la sistematización del contenido del lenguaje documental mediante estas microdisciplinas se halla del todo ausente de las LEM. Es por eso por lo que ha podido llegarse a decir del catálogo de la Library of Congress que es una gran herramienta bibliográfica donde la información se pierde en el orden alfabético (2).

c) La **forma gramatical** que adoptan las unidades léxicas de las LEM es más compleja y menos natural que la de los thesauri. Éstos rechazan la in-

versión de términos, y, además, sus términos de indización no se subdividen. En los thesauri, las situaciones donde las subdivisiones de los términos de indización podrían tener alguna utilidad, han sido solventadas mediante la coordinación de los términos de indización durante la búsqueda. Pese a ello se advierte una tendencia en las LEM a simplificar el vocabulario y a rechazar la inversión en beneficio de la expresión en lenguaje natural (23).

Según Van Slype el tipo de enlaces sintácticos entre los descriptores asignados por indización humana depende del sistema de almacenamiento y recuperación documental; y es independiente del tipo de lenguaje de indización; excepto en aquellos sistemas, muy pocos, en los que el uso de la sintaxis exige una división previa de los términos en categorías predefinidas como, por ejemplo, el sistema PRECIS.

Esto significa que los elementos separadores o signos de puntuación en uno y otro caso son los mismos, pero con algunos matices diferenciadores que conviene resaltar porque están en consonancia con las respectivas naturalezas y funciones de estos lenguajes documentales. Así, tanto si se maneja una LEM como si se indiza con un thesaurus los términos pueden vincularse o separarse entre sí a base de elementos tradicionales, tan conocidos como :

- los dos puntos (:),
- la coma (,),
- el paréntesis (),
- el punto y coma (;),
- la barra (/) o la doble barra oblicua (//),
- los espacios en blanco, el guión.

Sin embargo, en la indización efectuada con un thesaurus pueden expresarse relaciones entre términos mediante **vínculos** o **roles**, recurriendo a símbolos alfanuméricos. Elementos ambos que no están disponibles en la indización con LEM, para señalar los lazos entre los términos escogidos para indizar.

Relaciones semánticas.

En cuanto se refiere a las relaciones semánticas entre los términos, los thesauri presentan hasta un máximo de cinco tipos, mientras que en las LEM sólo encontramos un máximo de tres o cuatro.

En los thesauri localizamos las siguientes:

a) Inclusión propia o subsunción.

Por virtud de ella los distintos términos adscritos a temas o facetas enlazan con un grupo o campo semántico o microdisciplina a la que pertenecen. Por cuanto las LEM carecen de términos que constituyan microdisciplinas o microthesauri, las relaciones de inclusión propia o subsunción se hallan ausentes de ellas.

b) Equivalencia semántica interlingüística.

Se trata de relaciones biunívocas y simétricas que facilitan el enlace entre términos de dos o más lenguas diferentes que representan el mismo concepto o un concepto similar. La norma francesa Z 47-103 de abril de 1980 recomienda intercalar entre los descriptores en cuestión y su equivalencia lingüística un símbolo de la lengua del equivalente, seguido del signo dos puntos (:) y de un espacio en blanco.

MER

E: SEA

Estas relaciones se prodigan con mucha frecuencia en los thesauri, por lo que están muy estudiadas y reguladas. De hecho, la ISO ha emitido directrices para el establecimiento y desarrollo de thesauri multilingües. Sin embargo, las LEM apenas se han aprovechado de su utilidad. Hay LEM como la de la LCSH que ha sido ampliamente traducida y adaptada pero, con la excepción del BILINDEX, no conocemos listas bilingües ni multilingües (24).

c) «Equivalencia» semántica intralingüística.

Relación asimétrica entre un descriptor y un no descriptor, cuando existen varios términos que pueden considerarse representativos de un mismo concepto para necesidades de indización. Esta referencia o reenvío de exclusión remite desde el término descartado o rechazado al término escogido para la indización. También pueden establecerse reenvíos inversos, que nos previenen sobre los términos aptos para indizar y los recusados. (TABLA 4)

Los dos lenguajes, pues, tienen de común denominador las relaciones de equivalencia semántica intralingüística. Ahora bien, en los thesauri están más estandarizadas. En las LEM, la estandarización tradicional en la representación de las mismas se ha cuarteado en la medida que unas listas han decidido escoger la simbología y terminología de los thesauri, a la par que otras han preferido seguir con los conocidos «Véase» y «X».

d) Relaciones jerárquicas.

Relación antisimétrica entre dos términos en los que uno es superior al otro por carácter normativo. Las normas ISO distinguen tres tipos: la relación genérica; la relación partitiva o todo-parte; y la relación de **instancia** (o «ejemplificación») (25). En principio parece evidente que thesauri y LEM cuentan en su haber con elementos vinculados por relaciones de jerarquía. Pero incluso en este ámbito cabe hallar divergencias nada despreciables entre los dos lenguajes documentales: (TABLA 5)

Los thesauri las tienen perfectamente *reconocidas y normalizadas*, con sus cauces de expresión reglados. En cambio, las LEM *se limitan por regla general a despacharlas con la expresión «Véase además», que no articula*

TABLA 4
Referencias entre términos

TIPO DE REFERENCIA		IDIOMA	Simbolismo - Referidores	
			THESAURI	LEM
4.1 Directa <i>De</i> <i>no descriptor</i> <i>a</i> <i>descriptor</i>	4.1.1	Español	USE	VEASE (V.)
	4.1.2	Inglés	USE	SEE
	4.1.3	Francés	EMPLOYER (E)	VOIR
4.2 Inversa <i>De descriptor</i> <i>a</i> <i>no descriptor</i>	4.2.1	Español	USADO POR (U.P.)	UP (X)
	4.2.2	Inglés	USE FOR (U.f.)	UF (X)
	4.2.3	Francés	EMPLOYER POUR (EP)	EP (X)

mucho que digamos la gama de relaciones lógicas que pueden localizarse en el seno de estos lenguajes documentales. Además, la práctica de referencias inversas, o sea, el reenvío desde un descriptor específico hacia el descriptor situado a nivel jerárquico inmediatamente superior es incluso poco practicada en las bibliotecas.

Esto es tanto más chocante cuanto que teóricos de la indización alfabética como Kaiser y Coates han insistido en el valor de los reenvíos jerárquicos inversos para las búsquedas globales (10).

Incluso algún manual y el método de indización alfabética de Gleyze admiten el reenvío desde el término específico al término genérico. Por el contrario, la norma francesa NF Z 44-070 se opone a esta práctica, aunque la NF Z 47-200 recomienda expresar en las LEM las relaciones jerárquicas entre descriptores por las notaciones TG y TE (26). En esta dirección se ha movido la LCSH, que ha incorporado recientemente las indicaciones TG y TE, no sin que esto haya desatado una fuerte polémica (27-28-29).

En España coexisten dos tendencias: La LEM del MC y la adaptación al catalán de la lista de la Biblioteca del Congreso, hecha por el Instituto

TABLA 5
Representación de relaciones jerárquicas

IDIOMA		Simbolismo - Relatores jerárquicos	
		THESAURI	LEM
5.1	Español	TG , TE ; TEG ; TGG , TEP, TGP	V.a. ; TG , TE (XX)
5.2	Inglés	BT , NT ; NTG ; BTG , NTP, BTP	See also (XX)
5.3	Francés	TG , TS ; TSG ; TGG , TSP, TGP	Voir aussi (XX)

Catalán de Bibliografía, siguen el tradicional sistema de «Véase además». En la lista del MC, la referencia complementaria de «V.a.» se dirige tanto al indizador como al usuario, por lo que junto a la de «Véase» y «Usase por» deberá estar siempre presente en el catálogo. Por medio de esta referencia de orientación no se remite a temas más específicos, sino a temas más genéricos. En consecuencia se reservan las indicaciones «R.e.» para informar nada más que al indizador de la posibilidad de usar encabezamientos más específicos dentro de una materia. Esta información se ve ampliada por los avisos de «Pueden usarse además»...

La representante de la otra tendencia es la «LEM en la base de datos CIRBIC», que en su tercera edición ha preferido distinguir «ex profeso» los TG de los TE y éstos de los TR, siguiendo así a la LCSH y apeándose de su postura anterior: «R.e.» para el TG; «V.a.» para el TE; y asimismo «V.a.» para el TR.

De aquí se infieren dos importantes **puntos de diferencia** entre las LEM y los thesauri:

1) La referencia «V.a.» es utilizada como una especie de tapiz que uniformiza o encubre las distintas relaciones lógicas de los elementos integrantes de la LEM, puesto que no se especifican aquéllas por regla general con claridad. Lancaster ha subrayado que las LEM difieren del thesaurus en que

incorporan una estructura jerárquica imperfecta y fracasan en distinguir con nitidez entre las relaciones jerárquicas y asociativas (14).

2) En cuanto se refiere a la utilización y significado de las siglas en las distintas LEM existe menor uniformidad normativa o, al menos práctica, que en el ámbito de los thesauri, donde se aprecian perfectamente y sin peculiaridades equívocas los diversos niveles de jerarquización incorporados.

e) Relaciones asociativas.

Es una relación simétrica entre dos descriptores que designan conceptos susceptibles de evocarse mutuamente por asociación de ideas.

No existen relaciones asociativas entre no descriptores y descriptores, ni entre descriptores ya ligados por una relación jerárquica.

En los thesauri y en las LEM se expresan, respectivamente, mediante los relatores que señalamos en la:

TABLA 6
Representación de relaciones asociativas

IDIOMA		Simbolismo - Relatores asociativos	
		THESAURI	LEM
6.1	Español	TR	V.a., TR
6.2	Inglés	RT	See also
6.3	Francés	TĀ	Voír aussi

Según Gleyze, en las LEM los **reenvíos de orientación** desempeñan un importante papel; en el catálogo alfabético de materias invitan a los usuarios a diversificar sus modos de búsqueda de la información, y comportan además un paliativo a la dispersión producida por la clasificación alfabética (26). Gleyze lleva razón. Pero el inconveniente estriba en que las LEM tradicionales no distinguen entre las relaciones jerárquicas y las asociativas. De nuevo el peso de la información recae sobre el socorrido «V.a.», que desempeña una función polivalente; cuando en los thesauri hay un término destinado en exclusiva a esa indicación —TR—. Bien es verdad que

listas como la LCSH y la de la base de datos CIRBIC han preterido su secular «V.a.», sustituyéndolo también por relaciones «thesaurales», incorporando el TR.

En suma, por lo que incumbe a las relaciones semánticas entre los términos, cabe decir que las de los thesauri son más ricas y están más reguladas, lo que se refleja no sólo en una amplia normativa internacional, sino en la homogeneidad que en ese aspecto revelan estos lenguajes documentales. Esto tiene su contrapartida en la menor complejidad de la estructura sintáctica de las LEM y en la tendencia a la sustitución de los tradicionales reenvíos de éstas por los de los thesauri, lo que ha desencadenado protestas.

Presentación

La **presentación** de las LEM y la de los thesauri difiere bastante, aunque también tengan ciertos puntos en común, como intentaremos reseñar. (TABLA 7)

La estructura semántica de los thesauri les permite adoptar un gran número de presentaciones:

- Lista alfabética estructurada completa o por grupos, permutada o no;
- Índice alfabético, completo o por grupos;
- Lista jerárquica o sistemática;
- Presentación gráfica, en forma de flechas de grupo y de terminograma de grupo (16-18).

Las LEM se presentan sólo bajo la modalidad de lista alfabética completa, donde hay dos tipos de entradas alfabéticas: términos de indización; y términos recusados, indicándose en cada entrada toda la información semántica con ella relacionada, incluidas las notas de alcance. Esto no obstante, el impacto de la informática ha propiciado que surjan dos posturas en torno a la manera en la que debería organizarse el catálogo «on line» (OPAC) para el usuario:

a) El «enfoque estructurado» (structured approach), que podríamos traducir por sistemático, defendido en *The Library of Congress Filing Rules* (1980), e inclinado por una ordenación de los encabezamientos basada en las categorías a las que pertenecen; por ejemplo, los encabezamientos de historia se ordenan cronológicamente antes que alfabéticamente. Para hacer esto los encabezamientos se dividen en más de un elemento de clasificación, efectuándose la ordenación sólo por el primer elemento.

b) El «enfoque alfabético estricto» (strict alphabetical approach) tiene como valedor a la *ALA Filing Rules* (1980), aunque acepte también una presentación sistematizada de las subdivisiones cronológicas. La mayoría de los catálogos «on line» de acceso al público siguen la ordenación alfabética incluso con más rigor que el recomendado por las reglas de la ALA (30).

En los encabezamientos de materia puede asimismo practicarse la rotación de términos o permutación. Lo que ocurre es que la permutación no

TABLA 7
Presentación de estos lenguajes documentales

PRESENTACIÓN DE LOS LENGUAJES DOCUMENTALES				
	<i>Tipo de representación</i>	<i>Características</i>	LE M	THESAUR I
7.1	Lista alfabética estructurada	* Completa	+	+
7.2		* Por grupos	-	+
7.3		* Permutada	-	+
7.4	Índice alfabético	* Completo	-	+
7.5		* Por grupos	-	+
7.6	Lista jerárquica		-	+
7.7	Presentaciones gráficas	* Diagrama de flechas	-	+
7.8		* Terminograma	-	+

se efectúa de una manera sistemática, sino que sigue ciertas pautas. Neet afirma que la rotación debe limitarse sólo a aquellos encabezamientos de materias que comportan un término en posición de subencabezamiento, que designa por sí mismo una noción bastante específica y autónoma y que podrá, de hecho, figurar como tema principal. La rotación de los términos contribuye a incrementar los ficheros; y puede llegar a ser muy compleja si la precoordinación es exhaustiva. De ahí que haya LEM que prefieran la fragmentación de los temas en materias independientes — LCSH —, lo que no obsta —repetimos— para que otras propugnen la rotación de los tér-

minos precoordinados, con la excepción de las fechas y los subencabezamientos de forma (10).

UTILIZACION Y CONSTRUCCION

Para el profesional y el usuario también hay diferencias a la hora de manejar estos lenguajes documentales.

Ventajas e inconvenientes de la utilización de estos lenguajes documentales.

Indizar con LEM requiere menos tiempo, y, generalmente, un menor número de términos de indización que operar con thesaurus. Como contrapartida, el indizador se ve constreñido a disponer los términos en un orden determinado; es decir, no se pueden combinar libremente la mayoría de los términos de indización utilizados. Este sistema de indización impide a seleccionar el término principal con un/os auxiliar/es limitados, que desempeñan una función limitada (31).

Por otro lado, la precoordinación permite describir con precisión las unidades de información extraídas de los documentos por medio de términos constituidos por la combinación de ellos; y ofrece la ventaja de evitar la posibilidad de falsas combinaciones. Sin embargo, la otra cara de la moneda nos muestra que en el momento de la búsqueda no es posible utilizar más que aquellos términos de primer orden —los encabezamientos de materias—. A los términos consecutivos no se accede directamente, salvo en los casos donde se han practicado permutaciones (5-26). (TABLA 8)

Eso ocurre porque las LEM han sido concebidas para búsquedas secuenciales, ora en ficheros manuales, ora en ficheros informatizados. Para Neet, el encabezamiento de materia procede a una precoordinación de los términos en función de las **búsquedas secuenciales** en los ficheros alfabéticos de materias. (TABLA 9)

Los thesauri han sido concebidos para que los temas de los documentos puedan expresarse independientemente unos de otros, con la ayuda de descriptores aislados, porque se ha pensado en una **búsqueda combinatoria**. Así, el indizador o el usuario proceden a coordinar —postcoordinar— los términos en virtud de las necesidades que impone la búsqueda. (TABLA 10)

La **búsqueda secuencial** por encabezamientos de materias precoordinados impone una estructura compleja al catálogo alfabéticos de materias:

— El orden natural de las palabras que forman las expresiones es abandonado para realzar el término focal; por ejemplo, en lugar de «Sistema nervioso» tendríamos NERVIOSO, SISTEMA.

TABLA 8
Construcción y uso de estos lenguajes

CONSTRUCCIÓN Y UTILIZACIÓN DE LAS / LOS			LEM	THESAURI
8.1.1	8.1 Indización	Tiempo	-	+
8.1.2		Núm. de términos	-	+
8.2.1	8.2 Búsquedas	Secuenciales	+	-
8.2.2		Combinatorias	-	+
8.3	Coherencia de la indización		-	+
8.4.1	8.4 Costes	de construcción	+	+
8.4.2		de indización	-	+
8.4.3		de formulación de consultas	+	-

— Los encabezamientos de materias que representan temas muy frecuentes en la colección de documentos, deben subdividirse a fin de guiar mejor al documentalista o al usuario en la manipulación de numerosas fichas o registros informatizados.

— A su vez, la subdivisión de los encabezamientos exige que se especifiquen cuáles son las nociones relegadas a la posición de subencabezamiento; y eso hace introducir numerosos reenvíos o proceder a la rotación de los términos.

TABLA 9
Precoordinación con subencabezamientos en una LEM

<i>Ejemplo de un encabezamiento de materia con precoordinación, mediante signos de puntuación</i>
ENERGÍA NUCLEAR - CONSUMO - ESPAÑA - 1960-1994

TABLA 10
Descriptorios aislados coordinables en un THESAURUS

<i>Ejemplo de indización coordinada mediante descriptorios aislados</i>
ENERGÍA NUCLEAR CONSUMO ESPAÑA 1960-1994

Ejemplo:

TABLA 11
Estructura compleja del Catálogo Alfabético de Materias

DOCUMENTALISTAS - FORMACIÓN PROFESIONAL
FORMACIÓN PROFESIONAL Usase sólo como subencabezamiento bajo los nombres de las distintas profesiones

En este sentido, la indización **postcoordinada** se configura como un método más simple; favorece las búsquedas puntuales y variadas. Gracias a la ausencia de subdivisiones preestablecidas, los descriptores se prestan a la multiplicación de combinaciones durante las búsquedas; aunque con el riesgo incluido, por supuesto, del ruido y las falsas llamadas.

No cabe duda de que las LEM o los encabezamientos de materias expresan los temas de un documento de una manera clara y precisa; pero el usuario se ve obligado a aceptar una indización que le ha sido preparada de antemano, pensada en función de búsquedas hipotéticas, mientras que la postcoordinación permite escoger los términos en virtud de las necesidades reales de información.

Los centros de documentación que se sirven de LEM y desean introducir la postcoordinación con vistas a efectuar búsquedas booleanas han de proceder a descomponer los encabezamientos de materias subdivididos, lo que entraña una reestructuración de la lista, es decir, una reformulación de los términos susceptibles de emplearse en la indización.

Por lo que incumbe a la **coherencia de la indización** cabe decir que no resulta sencilla la comparación, ni demasiado rigurosa. Porque los estudios realizados al respecto (14) se han centrado en el cotejo de la indización efectuada con un mismo tipo de lenguaje documental, por ejemplo, con un determinado thesaurus o con una determinada LEM. Además, mientras que hay autores que claman contra la incoherencia en la asignación de encabezamientos, que achacan a la propia estructura del lenguaje documental —la LCSH en este caso—, y propugnan para subsanarla la conveniencia de utilizar un thesaurus simplificado que recoja todos los términos de indización en inglés, con las referencias de «Véase» y «Véase además» (32), otros cuestionan la objetividad metodológica de tales trabajos, manteniendo la hipótesis según la cual hay aciertos y errores en la asignación de encabezamientos con la LCSH, pero que la asignación de un encabezamiento es correcta si se apoya en el documento textual y se construyen de acuerdo con el manual de la LCSH (4).

Esto no obstante, creemos pueden apuntarse algunos datos interesantes en espera de que trabajos cuantitativos proporcionen más luz sobre la cuestión. Mientras tanto, diremos que *parece* que hay indicios que auguran mayor coherencia para la indización con thesaurus que con LEM.

Si aceptamos la tesis según la cual el punto álgido de acuerdo entre dos indizadores empieza a descender a partir de los dos términos, aun cuando vuelva a recuperarse después de asignarse al documento un número elevado de términos, convalidaremos que en principio la coherencia en la indización mediante LEM y thesauri es muy semejante a la hora de describir el tópico principal del documento, aunque quizá haya un ligero margen favorable a los lenguajes documentales postcoordinados. Por una razón. Porque muchas veces los indizadores no están de acuerdo acerca de cuál es el término principal que encabezará la descripción del documento; aspecto

este que carece de importancia en la indización con thesauri, pero que se revela vital en la indización con lenguajes precoordinaados. Si consideramos también los subencabezamientos parece que la diferencia se acrecienta. En 1968, en su evaluación de la coherencia de la indización con la **MeSH**, Lancaster descubrió que ésta era del 46,10 por ciento en la asignación de los encabezamientos, descendiendo al 34,40 por ciento cuando los indizadores se servían de los subencabezamientos (33).

La posibilidad de aumentar mediante thesaurus el número de términos de indización, mejora la **precisión** y la **coherencia** de la indización, factor al que no se puede recurrir cuando se usa una LEM. Esto se debe a que, por regla general, cuantos más términos de indización se asigne a un documento, hay más probabilidades de recoger los principales conceptos que lo representan. Esto significa que con los métodos de indización tradicionales se pierden fácilmente ciertos aspectos del contenido de los documentos durante el proceso de indización, y, por lo tanto, no son accesibles en el momento de la recuperación. La posibilidad de aumentar mediante thesaurus el número de descriptores para caracterizar un documento, es decir, siendo más exhaustiva la indización, mejora no sólo la calidad de la representación temática —precisión—, sino también la coherencia y las vías para una recuperación más existosa (31). En cambio, la indización mediante encabezamientos de materia en los catálogos on line de acceso al público, persigue el propósito de la sumarización o representación global del documento.

Comparando cifras absolutas referentes a la coherencia de la indización con LEM y thesauri tampoco el resultado favorece a las primeras. Van Slype cifra la coherencia de la indización con thesauri en un umbral situado entre el 50 y el 80 por ciento. Ya se ha dicho que Lancaster calculó la de la **MeSH** en el 46,10 por ciento. Claro está que para que la comparación fuese más rigurosa, habría que indizar —en nuestra opinión— con estos dos tipos de lenguajes el mismo campo documental. Por otro lado, parece demostrado que la **calidad de la indización** mejora conforme se acrecienta la formalización de las relaciones semánticas en los lenguajes documentales (34). De donde se infiere, si estamos de acuerdo en que la red de relaciones semánticas es más elaborada y se encuentra mejor explicitada en los thesauri que en las LEM, que «a priori» resulta más factible obtener una indización coherente o de mayor calidad con aquéllos, y, en definitiva, una recuperación más eficiente.

Costes de construcción y utilización.

Comparar los **costes de construcción y utilización** también resulta útil. El coste de elaboración de un thesaurus y de una lista de encabezamientos es elevado. Y debe correr muy parejo, porque en la creación de los dos lenguajes documentales, como ha señalado J. Rowley, se sigue un proceso si-

milar, aunque diverjan en aspectos importantes como la especificidad que queremos imprimir al sistema, la exhaustividad, el nivel de precoordinación, las relaciones semánticas y su especificación.

Mayor interés ofrece la comparación de los **costes de indización** de los documentos y formulación de consultas. La indización mediante thesaurus tiene un coste más elevado porque requiere un promedio de términos superior e incluso la ocasional utilización enlaces sintácticos complejos entre los términos. Por el contrario, Van Slype no duda en absoluto cuando afirma que en la formulación de las consultas el thesaurus tiene un «coste menos elevado» y una «eficacia mucho mayor» que «cualquier otro lenguaje», gracias a su red semántica. Otros estudios también demuestran que en la búsqueda on-line el thesaurus ofrece un mayor rendimiento que las LEM (35-36-37).

Conclusiones.

Por último, a modo de conclusión, avanzamos una síntesis interpretativa de las ideas que hemos ido exponiendo a lo largo de este trabajo. Dichas ideas-clave son las siguientes:

a) Los encabezamientos tienden a ser expresiones muy genéricas, y no representan con precisión materias complejas o específicas.

b) Los encabezamientos no están contruidos ni seleccionados sistemáticamente. La inversión o la entrada directa no responden muchas veces a un criterio uniforme, lo que hace difícil al indizador insertar un nuevo encabezamiento de uso local, ya que es difícil conocer los principios que en un momento determinado inspiraron la adopción de una forma de encabezamiento.

c) La estructura semántica y las referencias están menos articuladas que en los thesauri, y, además, en opinión de algunos especialistas, tampoco están contruidas sistemáticamente.

d) La indización efectuada mediante LEM es de menos calidad.

e) La construcción de un lista de encabezamientos no es más barata ni sencilla que la de un thesaurus al parecer; y, sin embargo, es menos eficaz en la búsqueda de información que este último, fundamentalmente en una base de datos.

f) A título de mera hipótesis de trabajo: Es probable que las LEM se encuentren en una fase de transición hacia una conversión en thesaurus; el primer paso ha sido ya dado al reajustar su estructura sindética de acuerdo con la de aquéllos.

BIBLIOGRAFIA REFERIDA

- (1) WITTY, F.J.: «The beginnings of indexing and abstracting». *The Indexer*, 8,4, 1973, 193-198.

- (2) DAILY, J.E.: «Subject Headings». En DYM, E.D., ed.: *Subject and Information analysis*. New York: Marcel Dekker, 1985, pp. 256-269.
- (3) DAILY, J.E.: «Subject Headings». En KENT, A., LANCOUR, H., eds.: *Encyclopedia of Library and Information Science*, vol. 29, New York: Marcel Dekker, 1969, pp. 178-191.
- (4) SVENONIUS, E., MCGARRY, D.: «Objectivity in Evaluating Subject Heading Assignment». *Cataloging & Classification Quarterly*, 16, 2, 1993, 5-41; 7.
- (5) VAN DIJK, M., VAN SLYPE, G.: *El servicio de documentación frente a la explosión de la información*. Buenos Aires: Consejo Nacional de Investigaciones Científicas, 1972 p. 53.
- (6) RICHTER, N.: *Grammaire de l'indexation alphabétique*. 5 ed., Le Mans: Bibliothèque de l'Université du Maine, 1988, pp.4-5.
- (7) ROWLEY, J. E.: *Organizing Knowledge: An Introduction to Information Retrieval*. Aldershot: Gower, 1988, p. 265; 285; 285.
- (8) COURRIER, Y.: «Analyse et langage documentaires». *Documentaliste*, 13, 5-6, 1976, 178-189; 183; 183-184.
- (9) *Tesaurus SPINES*, vol. 1, ICYT, UNESCO, Madrid, 1988, p. XII.
- (10) NEET, H.E.: *A la recherche du mot clé. Analyse documentaire et indexation alphabétique*. Genève, Suisse: Institut d'Etudes Sociales, 1989, pp. 19-21; 113 y ss; 118; 21-23; 166; 38-39; 39; 53-56; 75; 149-153; 168.
- (11) STUDWELL, W.E.: «Subject suggestions: some concerns relating to quantity of subjects». *Cataloging & Classification Quarterly*, 10, 4, 1990, 99-104.
- (12) CHAN, L.M.: *Library of Congress Subject Headings. Principles and Application*, 2ª ed., Littleton, Colorado: Libraries Unlimited, Inc., 1986, p. 182.
- (13) BAILAC I PUIGDELLIVOL, A.: «Les llistes d'encapçalaments de materia i la biblioteca pública». *Revista de Biblioteconomia y Documentació ITEM*, 2-3, 1988, 95-103; 97: 96.
- (14) LANCASTER, F.W.: *Indexing and Abstracting in Theory and Practice*. London: The Library Association, 1991, p. 41; 30-32; 179-180; 14; 60 y ss.
- (15) LANCASTER, F.W.: *Vocabulary Control for Information Retrieval*, 2ª ed., Arlington, Virginia: Information Resources Press, 1986, pp. 10-11; 31.
- (16) VAN SLYPE, G.: *Los lenguajes de indización: Concepción, construcción y utilización en los sistemas documentales*. Madrid, Salamanca: Fundación Germán Sánchez Ruipérez, Pirámide, 1991, p. 21; 18; 94; 38-40; 119 y ss.; 60 ss.; 35.
- (17) RIVIER, A.: «Construction des langages d'indexation. Aspects théoriques». *Documentaliste*, 27, 6, 1990, 263-274; 270.
- (18) CHAUMIER, J.: *Le traitement linguistique de l'information documentaire: Les langages documentaires*. 3ed., Paris: Entreprise Moderne d'Édition, 1988, p. 85; 87 ss.
- (19) *Tesaurus de Defensa*. Madrid: Secretaría General Técnica, Ministerio de Defensa, 1991, p. 21.

- (20) AITCHISON, J.: *Tesaurus de la UNESCO*, vol. I, París: UNESCO, 1984, p. X.
- (21) *Lista de Encabezamientos de materia para las bibliotecas públicas*, 2 ed., Madrid: Dirección General del Libro y Bibliotecas, Ministerio de Cultura, 1987, p. 23; 23-24
- (22) FOSKETT, D.: «Thesaurus». En DYM, E.D., ed.: *Subject and Information Analysis*. New York: Marcel Dekker, 1985, p. 270-316; 275-276.
- (23) *Lista de Encabezamientos de materia en la base de datos CIRBIC*. 3 ed., Madrid: UCB-CSIC, 1991, prólogo, s.p. observaciones de uso, s.p.
- (24) *BILINDEX: A Bilingual Spanish-English Subject Heading List Spanish Equivalents to Library of Congress Subject Headings*. Oakland, California: California Spanish Language Data Base, 1984.
- (25) ISO 2788 — 1986 (F) : «Documentation — Principes directeurs pour l'établissement et le développement de thésaurus monolingues». En *Récueil de Normes ISO I Documentation et Information*. 3 ed., Genève, Suisse: ISO, 1988, p. 539.
- (26) GLEIZE, A.: *Pour une méthode d'indexation alphabétique de matières*. Villeurbanne: ENSB, 1983, p. 81; 84.
- (27) DYKSTRA, M.: «LC Subject Headings Disguised as a Thesaurus». *Library Journal*, 113, 4, 1988. 42-46.
- (28) DYKSTRA, M.: «Can Subject Headings Be Saved?». *Library Journal*, 15, 1988, 55-58.
- (29) DYKSTRA, M.: «Subject Analysis and thesauri: a background». *Art Documentation*, 8, 4, 1989. 173-174.
- (30) *Headings for Tomorrow: Public Access Display or Subject Headings*. Chicago and London: ALA, 1992, p. ix.
- (31) ALURI, R., ALASDIR KEMP, D., BOLL, J.J.: *Subject analysis in online catalogs*. Englewood, Colorado: Libraries Unlimited, inc., 1991. pp. 41-42.
- (32) BLOOMFIELD, M.: «A look at Subject Headings: A Plea for Standardization». *Cataloging & Classification Quarterly*, 16, 1, 1993, 119-123.
- (33) LANCASTER, F.W.: *Evaluation of Medlars Demand Search Service*. National Library of Medicine. Bethesda, 1968 / [44].
- (34) SLAMECKA, V., JACOBY, J.: *Effect of Indexing Aids on the Reliability of Indexers*. Final Technical Note. Bethesda, Documentation Inc., 1963 / [45 = p. 30].
- (35) HOSONO, K.: «Demerits of the Library of Congress Subject Headings as search Key at the mechanized post-coordinate information retrieval systems». *Library and Information Science*, 15, 1977, 181-188.
- (36) SLADE, R., KELLY, A.M.: «Sources of popular literature online: New York Times Information Bank and the Magazine Index». *Database*, 2, 1, 1979, 70-83.
- (37) MARKEY, K., VIZINE-GOETZ, D.: «Sources in the machine-readable Library of Congress subject headings». *Library Resources and Technical Services*, 33, 1, 1989, 37-53.