

Problemas de terminología metalingüística en los lenguajes documentales de estructura combinatoria

José-María Izquierdo Arroyo

Luis Miguel Moreno Fernández

Departamento de Información y Documentación
Universidad de Murcia

0.1. Resumen

Este avance de un estudio más detenido sobre «terminología metalingüística» en los lenguajes documentales intenta precisar el alcance de vocablos tales como «combinatorio», «coordinación», «pre-» y «postcoordinación». Para ello, se compara el significado, con frecuencia discordante, que se otorga a estas palabras en cierta bibliografía y normas internacionales. La falta de precisión con que muchas veces se manejan, además de contribuir a sembrar el confusiónismo terminológico en la ciencia de la información, induce a errores conceptuales, como incluir a las listas de encabezamientos de materias entre los lenguajes documentales «combinatorios» o excluirlas, sin justificación, de la poco precisa clasificación «lenguajes documentales clasificatorios-lenguajes documentales combinatorios».

Palabras clave: Lenguajes documentales postcoordinados. Terminología científica.

0.2. Abstract

Advance of a more detailed paper on the metalinguistic terminology applied to indexing languages, that tries to precise the meaning of such terms as “combinatory”, “coordinated”, “pre-” and “postcoordination”. Therefore, the meaning attributed to this words in different international standards and papers is compared, and it is found frequently divergent. This lack of precision drives to terminological confusion and serious conceptual errors, as including subject headings lists into “combinatory” indexing languages, or excluding them from the not very precise dimension “classificatory-combinatory languages”.

Keywords: Postcoordinated indexing languages. Scientific terminology.

1. Introducción

En nuestro Seminario de *Lingüística Documental* —y dentro de una de las líneas de investigación del *Grupo SEMIODOC*—, nos hemos propuesto aplicar los contenidos comúnmente englobadas bajo ‘Lingüística documental’ (LD), en el marco de una sistematización de la disciplina designada por esa expresión. Ello nos obliga a proceder con rigor y coherencia en la descripción y presentación de tales contenidos. Mas rigor y coherencia requieren ante todo —previamente al sistematismo— la constante actualización de un *control terminológico* que haga transparentes los conceptos. En consecuencia, venimos trabajando en la revisión crítica de algunas expresiones habitualmente utilizadas cuando: a) se describen los lenguajes documentales, b) se habla sobre la teoría-metodología-práctica del resumen o de la indización, o c) se mencionan tipos de objetos, actividades o procesos de algún modo implicados en el *Tratamiento Documental del Contenido* de documentos cualesquiera.

Porque no se trata, por ejemplo, de la terminología obrante en un determinado lenguaje documental (ld), sino más bien de aquella de que nos servimos para hablar sobre él. Por lo mismo, se impone distinguir entre: a) problemas de “terminología lingüística” (planteados en la construcción, mantenimiento y uso de un ld), y b) problemas de “terminología *metalingüística*” (surgidos con el uso discorde del lenguaje que habla sobre aquéllos otros problemas); siendo justamente la resolución de estos últimos lo que ahora nos ocupa. Con todo, en la presente *comunicación* restringiremos esos objetivos al esbozo en avance de algunos problemas representativos, de ese tipo, que detectamos en nuestra comunidad científica cuando en ésta se habla sobre una de las categorías de ld: los lenguajes documentales de estructura combinatoria. No porque los consideremos más útiles o efectivos que los dotados de una estructura sintáctica más sólida, sino en razón de estar históricamente más posados y consolidados.

Sabido es que tras los términos están los conceptos, y que sólo fijando inequívocamente aquéllos podremos habérnoslas científicamente con éstos. En tal contexto, estamos en la creencia de que los problemas metalingüísticos requieren un tratamiento de urgencia.

Desde hace tiempo una parte de la bibliografía que se ocupa de los lenguajes documentales adopta como uno de los criterios para clasificarlos oponer a los llamados lenguajes de “estructura jerárquica” los de “estructura combinatoria” —estructura jerárquica vs. estructura combinatoria—. Frente a los lenguajes clasificatorios, identificados con los jerárquicos, se alzan los combinatorios, entre los cuales suelen citarse las “listas de materias”, los “unitérminos” y los “thesauri”. Sin embargo, al desarrollar las características de estos lenguajes documentales esa bibliografía se centra en los unitérminos y thesauri, olvidándose de las listas de

encabezamientos de materias, quizá porque es francamente difícil justificar su inclusión en este grupo. La cuestión terminológica y conceptual se complica si se considera que a los “lenguajes combinatorios” o de “estructura combinatoria” — la expresión la han empleado autores franceses y españoles con este sentido— (1) se les ha llamado asimismo “lenguajes de indización”, “lenguajes analíticos”, “lenguajes de indizados” —“langages d’indexats”— y de “términos”. Precisamente bajo el epígrafe de “lenguajes de indizados”, de “términos”, G. Brunel y S. Lafont engloban aquellos lenguajes cuyo léxico de base está constituido por “palabras claves”, “descriptores” y “encabezamientos de materias” (Brunel y Lafont, 1987, p. 110). Algunos autores españoles siguen esta tendencia, aunque introduciendo determinados retoques; por ejemplo, Amat Noguera distingue entre estructura jerárquica y “estructura asociativa” (Amat, 1989, p. 192-202); y García Gutiérrez entre jerárquica, combinatoria y sintáctica (García Gutiérrez, 1984; p. 170-72; 1989, p. 313-346). Finalmente hay quien, como R. Roy (1987, p. 18, 50-63), contrapone a los lenguajes jerárquicos los analíticos. Dentro de esta clase quedan comprendidos los unitérminos, thesauri y listas de “autoridades de materias” o de encabezamientos de materias, pero él estima que los lenguajes combinatorios son un “caso particular” de “lenguajes analíticos”.

Sistematizando estas posturas, aun a costa de simplificarlas, podemos decir que se enfrentan a los lenguajes jerárquicos los combinatorios, con la reserva de incluir o no, a veces, en este apartado a las listas de encabezamientos de materias (LEM); pero siempre sin justificar una u otra decisión.

Cabe preguntarse dónde radica la diferencia de opiniones o, dicho de otra manera, por qué en unos casos las LEM aparecen caracterizadas como lenguajes combinatorios y en otros no. La respuesta que en principio puede aventurarse es que a menudo manejamos los términos sin cuidarnos de precisar su alcance, desatendiendo la precisión lingüística que tanto solemos preconizar en este área de conocimiento. En tales circunstancias, palabras como “combinatorio”, “coordinación”, “pre” y “postcoordinación”, están tan manidas que a veces se utilizan sin el más mínimo cuidado respecto de su significado y lo que queremos decir con ellas cuando las empleamos mecánicamente en un contexto determinado. Quizá porque con frecuencia, inexorablemente, la práctica cotidiana y el desarrollo de la profesión han ido modificando su alcance original sin que nos hayamos percatado de ello.

El determinante *combinatorio* viene a ser, en francés y, en consecuencia, en español, el término cuasi-equivalente de coordinado y hasta de postcoordinado. Cosa bien distinta es que este ejercicio de sinonimia sea o no afortunado.

Históricamente, la expresión *indización coordinada* se ha vinculado a M. Taube, pese a que otros autores se sirvieron de este método de indización antes

que él —Batten en Inglaterra y Cordonnier en Francia—. Con una acepción restringida pero clara, S. Artandi (1985, p. 181-184) explica la coordinación como la combinación de dos o más términos para expresar un tema distinto del que en principio representaban esos mismos términos por separado o mediante alguna otra combinación. Cuando S. Artandi explica su concepción de la indización coordinada está refiriéndose a la verificada con unitérminos y thesauri. Esto no obstante, partiendo incluso de su propio concepto, éste podría ampliarse si aceptamos que una indización es coordinada en la medida que los conceptos y las palabras o cualesquiera signos alfanuméricos empleados para representarlos son susceptibles de combinarse entre sí para caracterizar lo que convenimos en llamar contenido noético de los documentos. En la fase de la descripción, la coordinación o combinación tal y como la hemos descrito 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 contenido 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* (Van Dijk y Van Slype, 1972, p. 48). Tal vez para evitar la asimilación del concepto de coordinación con ciertos lenguajes, la ISO prefiere la expresión “relación coordinada”, que define como «relación formal que yuxtapone términos o índices» (ISO 5127/6, 1983). La sintaxis en la CDU, por ejemplo, nos sirve para vincular los números principales, tanto los auxiliares comunes como los especiales; y facilita la formación de temas nuevos o complejos con el símbolo “:” —relación— y la coordinación de una secuencia de clases adyacentes con la “.../...” —barra—. Los encabezamientos de materias complementan la información de la que son portadores gracias a una sucesión de subencabezamientos de diversa índole. Los unitérminos y los descriptores simplemente se yuxtaponen.

2. Postcoordinación

Ahora bien, las cosas cambian al desplazarnos desde la etapa de la descripción y almacenamiento de la información hasta la de la búsqueda. Porque es aquí donde expresiones tales como coordinado o combinatorio deberían mostrarse relevantes para clasificar los lenguajes documentales. Coordinación o combinación de términos en el momento de la recuperación de la información significa —o no significa nada— *la posibilidad de asociarlos entre sí mediante operadores booleanos* —factor que suele omitirse, acaso porque resulta obvio— para seleccionar un número determinado de documentos comprendidos bajo esos términos. El empleo de ordenadores ha potenciado la explotación de la búsqueda combinatoria. Mientras que las fichas “uniterms” y el método “Peek-a-boo” limitan en la práctica la combinación al empleo del conectivo lógico “Y”, los siste-

mas informáticos amplían la disponibilidad de conectivos booleanos al “O” y al “NO” (Artandi, 1985, p. 182). Desde esta perspectiva, la coordinación se conoce generalmente como postcoordinación. Ésta ha sido definida por algunos autores y organismos como la facultad de establecer las combinaciones entre los descriptores en el curso de la búsqueda documental (*Vid. Vocabulaire ...*, 1987). Sin embargo, la ISO se revela muy ambigua a este respecto, al considerar la “indización postcoordinada” —así la llama— como «indización sin orden preestablecido de los términos o de los índices» (ISO 5127/3 A, 1981). La falta de precisión de esta norma obedece a dos razones:

1ª. Si el término “indización postcoordinada” lo tomamos como referente del uso del lenguaje documental para describir los documentos y almacenar los datos correspondientes, se *asemeja* bastante al concepto que hemos barajado de “indización coordinada”, pero con dos salvedades: a) que sólo los unitérminos encajarían plenamente en este esquema, y b) que la inclusión de los thesauri, por no citar otros lenguajes documentales, queda mal parada, dado que incluyen términos con un orden prefijado; y si se ha previsto esta última eventualidad, no se refleja con claridad en la redacción de la norma.

2ª. Si el término “indización postcoordinada”, por el contrario, nos sugiere la idea de una indización de cuestiones, es decir, nos remite a la facultad de combinar términos en la fase de la búsqueda documental, nos encontramos ante el concepto de postcoordinación expuesto más arriba, aunque expresado con poco rigor.

La postcoordinación ha evolucionado desde la época de M. Taube, aunque obviamente sigue manteniendo su carta de naturaleza. El cambio principal que ha experimentado reside en que ha dejado atrás la combinación de términos simples, utilizándose ya expresiones compuestas de dos o más palabras. Éstas pueden considerarse términos precoordinados que son postcoordinados en el momento de la búsqueda. El prefijo *post* pone de relieve, además de las diferencias expuestas entre ambas coordinaciones, la existencia de un tercer nivel de coordinación en un plano distinto del que nos ocuparemos más adelante: la *precoordinación*.

Como expusiera Lancaster, el sistema postcoordinado auspicia la “multidimensionalidad” de las relaciones entre los términos, de modo que al barajarse todos durante la recuperación de la información se alcanza el nivel más alto de coordinación y, por ende, de precisión (Lancaster, 1991, p. 30 ss.). Pero las LEM no fueron concebidas para realizar búsquedas combinatorias, sino secuenciales, ora en ficheros manuales, ora en ficheros automatizados. En los índices impresos, o en las entradas del catálogo de materias de las bibliotecas, los términos combinados por el indizador para describir el contenido de los documentos u otros aspectos del mismo constituyen una suerte de cadena “lineal”, por lo que el indi-

zador se ve forzado a servirse de esos términos y del orden que se les ha asignado dentro de esa cadena para recuperar la información (Aithchison y Gilchrist, 1987, p. 62-63). De aquí se deducen dos cosas:

1ª. Que las LEM no responden a las características inherentes a los lenguajes combinatorios. En este nivel, términos como lenguaje “asociativo”, “combinatorio” o “coordinado” funcionan prácticamente como sinónimos de “lenguajes postcoordinados” y las LEM no se basan en el principio de la postcoordinación. Y si aceptamos que “combinatorio” es sinónimo de “coordinación”, LEM, uni-términos y thesauri no son los únicos lenguajes documentales cuyos términos se coordinan para describir documentos.

2ª. Corolario de la anterior: Que la clasificación de los lenguajes documentales en función de la estructura jerárquica vs. combinatoria no resulta adecuada, puesto que margina a las LEM. Éstas no son lenguajes combinatorios, entendiéndose por tales a aquellos lenguajes documentales que permiten la asociación o coordinación de términos durante la *indización de cuestiones* con el fin de representar las preguntas relativas al contenido de los documentos. Por lo tanto, pre-coordinación y postcoordinación no son sino la combinación de los términos o cualesquiera otros signos alfanuméricos de los que integran el cuerpo léxico de los lenguajes documentales en dos fases bien diferenciadas del tratamiento documental de contenido; respectivamente, la indización de documentos y la indización de cuestiones.

Existen además otras razones que ponen en tela de juicio una clasificación de los lenguajes documentales fundada en la contraposición jerárquico-combinatorio. La total identificación de los lenguajes clasificatorios con los jerárquicos, al margen de que en el seno de los combinatorios incluyamos o no las LEM, tampoco constituye un gran acierto terminológico. Se supone que las clasificaciones son los lenguajes documentales de estructura jerárquica por excelencia; pero es que se olvida que las relaciones jerárquicas, que muestran los niveles de superordenación y subordinación de unos términos o índices respecto de otros, no son un rasgo exclusivo y básico sólo de los lenguajes documentales clasificatorios. Forman parte también de la naturaleza del thesaurus, hasta el punto que se ha dicho —y con toda razón— que constituyen una “característica básica” del thesaurus (Aithchison y Gilchrist, 1987, p. 39) (2) en sus distintas facetas: relaciones genéricas, jerárquicas, todo-parte y de instancia o ejemplificación. Luego tomar la jerarquía como elemento diferenciador de los lenguajes documentales conduce a situaciones equívocas, por su falta de rigor.

3. Precoordinación

El término precoordinación goza también de un amplio uso, aunque a veces

esto no halle correlación en la unidad de criterio con que se baraja; pues hay matices en el mismo. Hace ya años que Y. Courrier (1976, p. 183-184), clasificando los lenguajes documentales en virtud de la coordinación, expuso que los lenguajes precoordinados se denominan así porque coordinan antes del almacenamiento los diferentes conceptos que forman un tema, de modo que el indizador los utiliza juntos, ya coordinados, en el momento de la indización de documentos. Es por eso por lo que las entradas de los índices basados en la precoordinación son tan complejas como requiere la descripción del tema tratado (Rowley, 1988, p. 104, 125).

La ISO y, posteriormente, la AFNOR, en sus respectivas formulaciones acerca de la precoordinación, se ajustan bastante a la idea que de ella tiene el autor francés, pero la formulan de una manera más precisa. La ISO no se refiere textualmente a lenguajes precoordinados, sino al tipo de indización derivado de su uso. La norma 5127/3 a)-1981 (E/F) habla de indización precoordinada, definiéndola como «indización de documentos donde el orden de los términos o de los índices ha sido establecido según un lenguaje documental o sistema de indización» (*Recueil ...*, 1988, p. 81). Por su parte, la norma 2788 vuelve a ocuparse de la precoordinación, al explicitar lo que entiende por términos compuestos en un thesaurus, igualmente conocidos como *términos precoordinados*. Así, refiere que es una regla general que los términos de un thesaurus deben representar conceptos simples o unitarios tanto como sea posible, y que los términos compuestos deben fraccionarse en elementos simples, excepto cuando esto afecte a la comprensión del usuario (*ibidem*, p. 526, 533). La norma francesa parece que intenta abarcar los lenguajes documentales y sus usos, enunciando la precoordinación como “principio”, aunque su definición la remite a la composición de los términos en el seno del lenguaje documental. Por lo tanto la AFNOR la define como «principio según el cual las combinaciones entre los términos de un lenguaje documental se efectúan en el curso de su elaboración, por ejemplo la creación de términos compuestos en un thesaurus» (*Vocabulaire ...*, 1987, p. 95). Esto nos permite hacer dos puntualizaciones:

1ª. Ya hablemos de lenguajes documentales precoordinados, ya de sistemas de indización precoordinados, por precoordinación se entiende —o debe entenderse— la combinación de términos o índices dada por un lenguaje documental al indizador, para que éste se sirva de ellos al caracterizar el contenido de los documentos.

2ª. Este concepto de precoordinación complementa el que hemos expuesto de postcoordinación, y no se confunde —o no debe confundirse— con la mera coordinación; que es lo que ocurre algunas veces cuando se expresa que en la indización precoordinada se combinan los términos «en el momento de la indización» (Foskett, 1990, p. 86) (3) o, de forma más ambigua, que los «términos son com-

binados por el indizador...» (Aitchison y Gilchrist, 1987, p. 62).

El concepto de precoordinación origina asimismo determinados problemas cuando se utiliza como variable para evaluar lenguajes documentales. Si bien el comportamiento de un lenguaje documental sólo puede evaluarse *a posteriori*, analizando sus resultados en un sistema de información, también es factible realizar una aproximación al mismo *a priori*, dirigida a esbozar la calidad con que ha sido construido. Uno de estos elementos es la llamada *tasa de precoordinación*. Para J. Chaumier, la tasa de precoordinación es el promedio de palabras significativas por descriptor (Chaumier, 1988a, p. 85). Criterio compartido, por ejemplo, por los autores del Tesouro Spines, al considerar el “nivel de precoordinación” como el «número de palabras por descriptor» (*Tesouro Spines*, 1988, vol. I, p. XII y XVI). Aunque no la formulen detalladamente, puede deducirse que se refieren al cociente fruto de dividir el *número de palabras significativas que constituyen los descriptores* entre el *número total de descriptores*. Van Slype enuncia la tasa de precoordinación de esta forma: «el cociente entre el número de ocurrencias de las palabras significativas constitutivas de los descriptores del thesaurus y el número de descriptores» (Van Slype, 1987, p. 123). Desgraciadamente, la traducción española reduce sus palabras en este punto a un absurdo galimatías (Van Slype, 1991, p. 94).

Estos procedimientos nos conducen a medir el *promedio de palabras por descriptor*, un resultado que en nuestra opinión no responde satisfactoriamente a la filosofía dimanada de la norma ISO 2788. En efecto, según esta norma, un descriptor compuesto es el que está formado por varias palabras; si cada una de ellas figura en el thesaurus como descriptor simple se le conoce como descriptor precoordinado. Estos se incluyen en el thesaurus cuando se usan con frecuencia y/o siempre que los descriptores que los constituyan se utilicen solos con mucha frecuencia; o bien tengan un significado diferente según que se empleen solos o combinados (*Recueil ...*, 1988, p. 534). Por lo tanto, si convenimos que los descriptores deben: a) tener un significado unívoco, b) representar en la medida de lo posible conceptos simples o unitarios, y c) poder fraccionarse en elementos simples, la tasa de precoordinación ha de orientarse más a medir el *número de conceptos nuevos representados por los términos compuestos, incluyendo el número de palabras que los integran*, que el promedio de palabras que forman todos los descriptores; habida cuenta de que los términos que figuran en las expresiones complejas aparecen por separado en el lenguaje documental, a fin de que puedan combinarse y utilizarse libremente como claves de búsqueda en un sistema postcoordinado. Dicho de otra manera: pretendemos conocer el *promedio de palabras que integran los descriptores que expresan conceptos distintos o nuevos*; para lo cual habremos de seleccionar palabras no repetidas en los descriptores compuestos en relación con el número total de descriptores.

4. Conclusión

Hay bastantes más binomios de términos, aparte los revisados, con que juegan las clasificaciones de la indización —más que de los lenguajes documentales mismos-, de cuyo control nos estamos ocupando. Aquellos y éstos habrán de integrarse en un *sistema meta-terminológico* que normalice *constelaciones* de términos tales como: *a*) «libre» / «controlado» (: vocabulario), *a'*) «libre» / «completo» (: texto) (4); *b*) «automática» / «humana» (: indización y/o resumen) (5); *b'*) «extracción» / «asignación» (: indización por) (6); *c*) «clasificación» / «indización» (7); *d*) «convencional» / «por lazos de citación» (: búsqueda (e indización)) (8); *e*) «monojerarquía» / «polijerarquía» (: relaciones jerárquicas) (9), *e'*) «temática» / «facetada» (: clasificación ; indización ; thesaurus) (10), *e''*) «jerárquicas» / «asociativas» (: relaciones entre descriptores) (11); ... etc.

Especialmente preocupante es esa última *constelación* de términos, que gira sobre el eje conceptual *Polijerarquía-Faceta-Relación asociativa*. Ese subconjunto ternario de binomios de términos, aparentemente inconexo, plantea problemas por doquier; y en nuestra investigación vamos entreviendo que el confusio-nismo por ellos creado obedece a una carencia de percepción de la trama conceptual que liga los vocablos constituyentes. Pues sólo en un sistema polijerárquico tiene sentido hablar de «facetas», y sólo desde las facetas cobran sentido esas relaciones (no-monojerárquicas) llamadas «asociativas». Pensar de otro modo conduce a situaciones aporéticas tales como no poder definir las relaciones «asociativas», dentro de un sistema (por ej., en la *red de relaciones semánticas* de un thesaurus), más que *negativamente*; reservando para ellas cuantas relaciones entre términos preferidos (*descriptores*) no sean lógicamente formalizables por la igualdad (interlingüística), ni por la inclusión (especie-género o parte-todo), ni por la circunclusión (género-especie o todo-parte) (Menne, 1976, p. 146). Atisbos —mas mucho discurso implícito— hay ya, sin salir de la bibliografía referida, sobre la íntima conexión tanto entre asociatividad y polijerarquía (12), como entre faceta y polijerarquía (13). La *polijerarquía* es, pues, el concepto clave que organiza la *constelación*.

5. Notas

- (1) Hace más de veinte años que J. Chaumier (1971, p. 49 ss., 62 y ss.; 88, p. 60) expuso esta clasificación, que han seguido otros autores. A su vez, el esquema ha pasado a España con algún retoque a través de ciertos manuales. Aunque en el último trabajo mencionado (Chaumier, 1988, p. 60) identifica los lenguajes combinatorios con los thesauri tan sólo, sin mencionar los otros.
- (2) Las LEM son, obviamente lenguajes pre-coordinados: «Un modo de lograr cierta pre-coordinación (*pre-coordination*), sin mucho incremento del tamaño de un vocabulario controlado (*controlled vocabulary*), es a través del uso de subencabeza-

- mientos (*subheadings*).» (Lancaster, 1991, p. 179).
- (3) Lancaster (1991, p.59) distingue *dos* tipos de pre-coordinación relacionados con la indización de documentos, frente al concepto de post-coordinación, relacionada con la indización de consultas: «Se da, pues, una cierta *coordinación* (*coordination*) (de conceptos o de términos que los representan) en las características del vocabulario, y puede haber otra (clase de) coordinación en el momento de la indización (*time of indexing*). Estas pueden considerarse como formas de *pre-coordinación* en que la coordinación se establece dentro de los registros que constituyen la entrada en una base de datos. El último nivel de coordinación es el que se obtiene mediante la manipulación de los términos en la formulación de la búsqueda (i.e., *post-coordinación*).».
 - (4) Por un lado, «texto libre» se hace sinónimo de «lenguaje natural», y se opone a «vocabulario controlado» (“*controlled vocabulary*”). (Vid. Lancaster, 1991, p. 193; Cfr. Dubois, 1987, p. 243-253; Knapp, 1982, p. 143-153; Pérez, 1982, p. 183-192). Por otra parte, el lenguaje natural o «libre» está claro que puede o no ser «completo»; por lo que «es importante reconocer la distinción entre los términos *texto libre* (*free text*) y *texto completo* (*full text*)» (Vid. Lancaster, 1991, p. 213). Con todo, se constata la existencia de sistemas híbridos de indización (*Hybrid Systems*), que incluyen tanto algunos términos *controlados* como algún tipo de *texto libre* (Vid. Lancaster, 1991, p. 214-215). El problema terminológico se complica más cuando se dan cita los prefijos *pre-* / *post-* con el binomio «libre» / «controlado». Tal es el caso de la expresión «Vocabulario Post-Controlado» (*The Post-Controlled Vocabulary*), en tanto que opuesta a la expresión «vocabulario controlado convencional» (*conventional controlled vocabulary*) (Vid. Lancaster, 1991, p. 215-218).
 - (5) Como puntualiza Lancaster, «los más de los sistemas (llamados) de indización automática no son realmente «automáticos», en el sentido de que sustituyan a los humanos por ordenadores, sino que pretenden *asistir* (*to assist*) al indizador humano. El mejor término para esto es «asistido por ordenador» (“*machine-aided*”).» (Lancaster, 1991, p. 228).
 - (6) En lo que atañe a la indización, queda justificada la «distinción entre indización por *asignación* (*assignment indexing*) e indización por *extracción* (*indexing by extraction*)», entendiéndose que «la mayor parte de la *indización humana* (*human indexing*) es indización por asignación». (Lancaster, 1991, p. 1; cfr. *ibidem*, p. 225). En lo que afecta al resumen, la programación de los ordenadores para seleccionar enunciados a partir de los documentos, «es la base de lo que generalmente se suele denominar «*resumen automático*» (“*automatic abstracting*”), aunque es más adecuadamente referible como «*extracción automática*» (“*automatic extracting*”)». El producto de tal proceso es el llamado «auto-resumen» (“*auto-abstract*”); bastante distante, por cierto, del humanamente elaborado. Y no parece que el uso de bases con armazones de conocimiento (*frame knowledge-base*) gestionadas con sistemas expertos puedan acortar mucho esas distancias. (cfr., *ibidem*, p. 236-237).
 - (7) La indización es sólo un tipo de clasificación, entre otros posibles: «la asignación (*assignment*) de un término a un ítem, situando ese ítem en una clase (*class*) al lado de otras (distintas de aquella) a la que ese término ha sido asignado»; mas «son posi-

- bles otros tipos de clasificación (*classification*).» (*Ibidem*, p. 229).
- (8) Abundan los «estudios que comparan las clases formadas mediante la indización *convencional* por materias (*conventional subject indexing*), con las formadas a partir de los *lazos de citación* (*citation linkages*).» (*ibidem*, p. 235).
 - (9) De acuerdo con J. Chaumier, «las relaciones genéricas (léase «jerárquicas»; tanto las genéricas como las partitivas) pueden ser de carácter bien *monojerárquico* (un término no puede tener sino un solo padre) bien *polijerárquico* (un término puede tener varios padres)» (Chaumier, 1986, p. 86). Ateniéndonos a su ejemplificación y diagramas, en la polijerarquía, un mismo término (más específico, o bien designador de parte) caería en la jerarquía bajo *al menos dos* términos (más genéricos, o designadores de totalidades). Mas, si así fuera, la polijerarquía se opondría a la monojerarquía tal como se opone la división lógicamente *mal* hecha con la *bien* efectuada. En efecto, o bien el término general (o todo) se divide *uniformemente* conforme a un mismo criterio —como dicta la regla—, en cuyo caso todos los *dividentia* quedan divisivamente subordinados a un mismo y *único dividens*; o bien la división no es uniforme, en cuyo caso está mal hecha (*Cfr.* Chaumier, 1978, p. 36-38).
 - (10) Por mantenernos en el mismo autor, tras describir las que denomina «relaciones *categoriales*» —que opone a las «de equivalencia», «jerárquicas», «asociativas» y «de definición»— como aquellas que «reagrupan» los términos del thesaurus en conjuntos que poseen una característica común, distingue entre reagrupación «por temas» y reagrupación «por *facetas*»; con la inocente diferencia de que aquélla se hace «según el *dominio* (campo semántico)», mientras que ésta se realiza atendiendo a «la *naturaleza* de los conceptos» (!). Lo cual no obsta para líneas abajo se sirva de la expresión «*categorías* o *facetas*», describiendo estas últimas. (*Vid.* Chaumier, 1986, p. 90-91). De ello derivará la oposición establecida entre «tesauro temático» y «tesauro *facetado*» (*Vid.* Chaumier, 1986, p. 91; *cfr.* 1978, p. 41). Raíces análogas pueden encontrarse en el “*Thesaurofacet*” de Aitchison y Gilchrist (*Vid.* Aitchison, Gomersall y Ireland, 1969; *cfr.* Aitchison y Gilchrist, 1987). Más recientemente, se ha definido la *faceta* como «grupo de descriptores jerárquicamente subordinados (hierarchically subordinate) a una (misma) categoría (category)» (Fuggmann, 1993, p. X, 16).
 - (11) Si bien el primer miembro de la oposición habría de completarse así: “«jerárquicas» y/o «de equivalencia» y/o «definicionales» y/o «categoriales»”. Las «relaciones *asociativas*» se vienen definiendo como «relaciones de vecindad» (!) que «sirven para indicar las analogías que pueden existir entre dos descriptores (términos asociados)», expresadas con el relator *TA* (o *RT*) (*Vid.* Chaumier, 1988, p. 35 y 38).
 - (12) El propio Chaumier, señala que las «relaciones *asociativas*» se utilizan a menudo «para evitar el empleo de la *polijerarquía*» (Chaumier, 1988a, p. 39; *cfr. ibidem*, p. 22; *cfr.* Chaumier, 1986, p. 89).
 - (13) En este caso, *faceta* y *polijerarquía* quedan indirectamente ligados mediante la oposición establecida entre «monocategorial» (“*monocategorial*”) y «policategorial» (*vid.* Fuggmann, 1993, p. 141, núm. 688); concretamente: «Si estos descriptores «policategoriales» (*polycategorial*) dominan por igual en el vocabulario, pueden destruir la capacidad del vocabulario de ser utilizado fidedignamente. Este es un

problema crucial para un vocabulario, porque el poder de supervivencia del sistema entero depende de la exactitud con que el vocabulario se usa en la práctica." (*ibidem*, p. 141, núm. 687). Mas, retornando a las *LEM*, su conexión con las facetas parece obvia: Justamente porque los subencabezamientos quedan categorizados en facetas, se puede decir que «los mejores candidatos para subencabezamientos son aquellos términos que pueden aplicarse potencialmente a muchos otros términos en el vocabulario.» (Lancaster, 1991, p. 179).

6. Referencias

- Aitchison, J., Gilchrist, A. (1987). *Thesaurus construction*. 2ª ed. London : Aslib, 1987.
- Aitchison, J.; Gomersall, A; Ireland, R. (1969). *Thesaurifacet : a thesaurus and faceted classification for engineering and related subjects*. Whetstone, Leicester : English Electric Co., 1969.
- Amat Noguera, N. (1989). *Documentación científica y nuevas tecnologías de la información*. 3ªed. Madrid : Pirámide, 1989.
- Artandi, S. (1985). *Coordinate indexing*. // Dym, E.D. *Subject and Information Analysis*. New York : Marcel Dekker, Inc., 1985.
- Brunel, G. ; Lafont, S. ; et al. (1987). *Traitement de l'information documentaire*. Paris : PUF, 1987.
- Courrier, Y. (1976). *Analyse et langage documentaires*. // *Documentaliste*. 13 : 5-6 (1976) 178-189.
- Chaumier, J. (1971). *Les Techniques documentaires*. 1ª ed Paris : PUF, 1971.
- Chaumier, J. (1978). *Le traitement linguistique de l'information documentaire : Les langages documentaires*. Paris : Entreprise Moderne d'Édition, 1978.
- Chaumier, J. (1986). *Análisis y lenguajes documentales*. Barcelona : Mitre, 1986.
- Chaumier, J. (1988a). *Le Traitement linguistique de l'information*. 3º ed. Paris : Entreprise Moderne d'Édition, 1988.
- Chaumier, J. (1988b) *Travail et méthodes du/de la documentaliste*. 3ª ed. Paris : Les Editions ESF-Entreprise Moderne d'Édition, 1988.
- Dubois, C. P. R. (1987). *Free text vs. controlled vocabulary : a reassessment*. // *Online Review*. 11 (1987) 243-253.
- Foskett, A.C. (1990). *The subject approach to Information*. 4º ed. London : Clive Bingley & Linnet Books, 1990.
- Fugmann, R. (1993). *Subject Analysis and Indexing: Theoretical Foundation and Practical Advice*, Frankfurt/Main : INDEKS Verlag, 1993. (Textbooks for Knowledge Organization ; 1).
- García Gutiérrez, A.L. (1984). *Lingüística documental: Aplicación a la documentación de la comunicación social*, Mitre, Barcelona, 1984.
- García Gutiérrez, A.L. (1989). *Los lenguajes documentales*. // López Yepes, J. (comp.). *Fundamentos de Información y Documentación*. Madrid : EUDEMA, 1989.
- ISO 5127/3 A)-1981 (E/F) // *Recueil de normes ISO 1 : Documentation et information*. Genève, Suisse : ISO, 1988. p. 81.

- ISO 5127/6-1983 (E/F) // Recueil de normes ISO 1 : Documentation et information. Genève, Suisse : ISO, 1988. p. 98.
- Knapp, S. D. (1982). Free-text searching of online databases. // *Reference Librarian*. 5 : 6 (1982) 143-153.
- Lancaster, F.W. (1991). *Indexing and Abstracting in Theory and Practice*. 1nd ed. London : Library Association, 1991.
- Menne, A. (1976). *Introducción a la Lógica*. 2ª ed. Madrid : Gredos, 1976.
- Pérez, E. (1982) Text enhancement : controlled vocabulary vs. free text. // *Special Libraries*. 73 (1982) 183-192.
- Recueil de normes ISO 1. Documentation et information. Genève, Suisse : ISO, 1988.
- Rowley, J. (1988). *Abstracting and Indexing*. 2nd ed. London : Clive Bingley, 1988.
- Roy, R. (1987). *Classer & Indexer: Introduction a l'indexation documentaire*. 2ª ed. Le Mans : Bibliothèque de l'Université du Maine, 1987.
- Tesauro Spines. UNESCO. Madrid : ICYT, 1988.
- Van Dijk, M., Van Slype, G. *El servicio de documentación frente a la explosión de la información*. Buenos Aires : Cons. Nacional de Investigaciones Científicas, 1972.
- Van Slype, G. (1987). *Les langages d'indexation: conception, construction et utilisation dans les systèmes documentaires*. Paris : Les Éditions d'organisation, 1987.
- Van Slype, G. (1991). *Los lenguajes de indización: concepción, construcción y utilización en los sistemas documentales*. Madrid : Fundación Germán Sánchez Ruipérez ; Pirámide, 1991.
- Vocabulaire de la documentation. 2ª ed. Paris : AFNOR, 1987.