

DEFINICION POR GENERO PROXIMO Y DIFERENCIA ESPECIFICA:
UNA CONTRIBUCION PARA LAS NOTAS DE ALCANCE DE LOS TESAuros

Ana M. Martínez (1), Rosa Z. Pisarello (2) , Amelia Aguado (3) , Cecilia Delorenzi (4)

Cátedras de (1) Clasificación, (2) Referencia y (3) Bibliografía y Selección, Departamento de Bibliotecología y (4) Cátedra de Lingüística, Departamento de Letras. Facultad de Humanidades y Ciencias de la Educación. Universidad Nacional de La Plata. Calle 48 e/6 y 7-5° piso, 1900 La Plata. E-mail: dhubi@huma.fahce.unlp.edu.ar.

RESUMEN

Se analizaron las definiciones utilizadas en diversas fuentes terminológicas, con el propósito de evaluar la utilización de la técnica de definición por género próximo y diferencia específica en la redacción de notas de alcance para todos los descriptores de un tesauro. Se seleccionó una muestra de 78 términos de la norma ISO 5127/3a sobre terminología de la documentación (documentos tradicionales) y su contraparte española la UNE 50113/3a; se obtuvieron 445 definiciones: 156 provenientes de ambas normas, 176 de tres glosarios y 113 de dos diccionarios. Luego se extrajeron los términos equivalentes de dos vocabularios controlados y se analizaron sus respectivas notas de alcance y relaciones jerárquicas. El 77% de las definiciones de las normas, 37% de las de los glosarios y 29% de diccionarios utilizaban la técnica analizada, siendo marcada la preferencia en aquellas fuentes que presentaban un ordenamiento sistemático (característica común al tesauro). En los vocabularios controlados analizados, menos de un tercio de los descriptores contaba con notas de alcance. Se concluye que la inclusión de una definición en cada uno de los descriptores del tesauro es necesaria para los usuarios de sistemas en línea y que la técnica de definición por género próximo y diferencia específica, ampliamente utilizada en normas, glosarios y diccionarios puede adaptarse sin dificultades a los tesauros, debido a que el ordenamiento sistemático de éstos pareciera facilitar su redacción.

INTRODUCCION

Como es sabido, en un vocabulario controlado (sea una lista de epígrafes, una lista de descriptores o un tesauro), la nota de alcance permite clarificar el uso de un término con el fin de evitar ambigüedades durante la indización y la recuperación. Esta nota de alcance puede ser una definición, una indicación de restricciones en el significado o la cobertura temática, una instrucción para los indizadores, una indicación de términos recolectores, la historia del descriptor cuando ha sido dado de baja o modificado, o una nota de reciprocidad (Aitchison et al., 1997; IFLA, 1993; Lancaster, 1986; ISO 2788, 1986).

Hasta ahora, la nota de alcance como definición ha sido una de las menos comunes, a pesar de que numerosos vocabularios controlados están especializados en áreas temáticas en las que el bibliotecario por lo general no es un especialista y, en consecuencia, le serían de gran ayuda. Adicionalmente, con la proliferación de los sistemas de información en línea, los usuarios también necesitan consultar los vocabularios controlados e incluso se han propuesto tesauros especiales para la búsqueda (Bates, 1989; Knapp, 1984; Kristensen, 1993; López Huertas, 1997; Pollard, 1993; Strong y Drott, 1996). Las ventajas de acceder al vocabulario controlado en línea y disponer de adecuadas notas de alcance han sido evaluadas, por ejemplo, en distintas interfaces de las

bibliografías en línea Medline (Schoonbaert, 1996) y LISA (Browne, 1992), así como en diversos catálogos en línea (Connell, 1997).

En una investigación realizada por Raya Fidel (1991) respecto a la elección de los usuarios entre el vocabulario controlado y la lengua natural, en 15 bases de datos de distintas disciplinas, el 75% de los usuarios prefirieron el vocabulario controlado, pero la tercera parte de ellos (25% del total) no encontró un término que coincidiera con su tema de interés. Fidel recomienda insistentemente que debe mejorarse la calidad de los vocabularios controlados y, en ese sentido, podemos plantear que una definición para cada término también beneficiaría a los usuarios, permitiéndoles aclarar sus dudas respecto al significado asignado a un descriptor en un determinado tesoro.

Otra investigación realizada en España por Alvaro Bermejo et al. (1989), sobre 47 tesauros en lengua española, ha permitido detectar importantes deficiencias en estos vocabularios controlados, entre ellos la escasa cantidad de notas de alcance, ya que en dos tercios de los tesauros analizados, menos del 10% de los descriptores contaban con estas notas y en el tercio restante la proporción oscilaba entre 11% y 30% de los descriptores; excepciones a la regla fueron el tesoro de Psicología con 65% de descriptores que tenían notas de alcance (García y Monfasani, 1983; Universidad de Buenos Aires, 1986) y el tesoro de ciencias sociales con 100% (Blasco Génova y Blass Konug, 1978). Las cifras anteriores implican sin duda una alta tasa de ambigüedad.

Aitchison et al. (1997) señalan también que la tendencia actual a eliminar las diferencias entre los tesauros y los bancos terminológicos es un motivo adicional para el incremento del número de definiciones en un tesoro. Debe recordarse además, que las definiciones tienden a ser más necesarias cuando la cobertura temática es interdisciplinaria, o en áreas de las ciencias sociales y las humanidades debido a que su terminología es menos precisa que aquella de las ciencias exactas y naturales o la tecnología. Recientemente, tanto la 4th International Society of Knowledge Organization Conference, celebrada en Washington DC del 15 al 18 de julio de 1996, como la 6th International Study Conference on Classification Research, realizada en el University College de Londres del 16 al 18 de junio de 1997, han puesto de manifiesto la necesidad de agregar una definición para cada término de un vocabulario controlado. En la primera de estas conferencias, Michèle Hudon (1996), de la University of Toronto, centró su atención en un modelo para la preparación de definiciones lógicas construidas especialmente para tesauros, mientras que en la segunda conferencia, Elaine Svenonius (1997), de la University of California at Los Angeles, recomendó la inclusión de definiciones tan amplias como fuera posible.

Arntz y Picht (1995) adhieren a la norma DIN 2342 para la cual "la definición consiste en la determinación de un concepto con medios lingüísticos". Las definiciones pueden analizarse desde un punto de vista lógico o lexicográfico; mientras que la definición lógica debe identificar en forma inequívoca el objeto definido o referente, para lo cual debe caracterizarlo como miembro de la clase más próxima, la definición lexicográfica se limita a enumerar los rasgos semánticos más importantes de la unidad léxica definida (Zgusta, 1971). De algún modo, en la definición de la norma DIN 2342 ya mencionada, se da una posición intermedia que aborda la estructura formal de las definiciones, sin olvidar el medio de expresión de los conceptos. Ambos tipos de definición se asemejan cuando en la definición lexicográfica se indica el hiperónimo y el rasgo semántico que la diferencia, en forma análoga al género próximo y la diferencia específica de la definición lógica. No obstante, no se debe perder de vista que la definición lógica se refiere a conceptos, en tanto que la

definición lexicográfica alude a los rasgos semánticos que aparezcan como relevantes para el hablante de la lengua.

Los términos técnicos presentan dificultades adicionales, porque su definición "debe ser científicamente correcta, describir el objeto correctamente, debe reflejar la noción generalmente aceptada del objeto y resultar inteligible para todos" (Zgusta, 1971). Para satisfacer estos requerimientos, se hace difícil no utilizar elementos enciclopédicos en la definición de los términos.

Las técnicas de la definición pueden agruparse en dos tipos: a) la definición denotativa o por extensión, que caracteriza el término mencionando a los individuos que éste comprende, es decir se enumeran todos los conceptos subordinados que se encuentran en el mismo nivel de clasificación y b) la definición connotativa o por comprensión que señala notas distintivas del concepto, es decir partiendo de un concepto superordinado conocido o ya definido, se indican las características específicas que permiten distinguirlo con respecto a otros conceptos de una misma serie. La definición connotativa o por comprensión puede ser por sinonimia (el término se define por sus sinónimos), operacional (el término se define por una operación ejecutada) y por género próximo y diferencia específica (Arntz y Picht, 1995; Copi, 1980; Haensch et al., 1982).

El género es la clase cuyos miembros se agrupan en subclases; las especies son las diversas subclases en que se divide el género y las características que permiten distinguir una especie de otra constituyen la diferencia específica. En este sentido, consideramos que esta técnica de definición resulta adecuada para un tesoro, ya que se complementa con la estructura jerárquica del mismo. Por ejemplo, tomemos el siguiente término:

PUBLICACIONES PERIODICAS
TG PUBLICACIONES SERIADAS
TE BOLETINES, DIARIOS, REVISTAS

La definición se construiría de la siguiente manera:

- Descriptor: PUBLICACIONES PERIODICAS
- El género próximo coincidiría con el término genérico (TG): PUBLICACIONES SERIADAS.
- La diferencia específica sería la característica que distingue a las publicaciones periódicas de los anuarios, memorias, transacciones, etc., es decir la periodicidad menor al año.
- Las especies serían los términos específicos (TE), que podrían mencionarse como ejemplos, en este caso BOLETINES, DIARIOS, REVISTAS.

Se tendría así la siguiente definición: "Las publicaciones periódicas son publicaciones seriadas con una periodicidad menor de un año, por ejemplo los boletines, diarios y revistas".

Seguendo a Copi (1980), hay cinco reglas para la definición por género próximo y diferencia específica: la definición debe incluir los atributos esenciales de la especie, no debe ser circular, ni ser demasiado amplia o demasiado estrecha, no debe formularse en términos ambiguos, oscuros y figurados, ni ser negativa cuando puede ser afirmativa.

En el presente trabajo analizamos las definiciones utilizadas en diversas fuentes terminológicas, con el propósito de evaluar la posibilidad de usar la técnica de definición por género próximo y diferencia específica en la redacción de notas de alcance de definición para todos los descriptores de un tesoro. Si bien las fuentes analizadas pertenecen al área de la Bibliotecología y la Documentación (porque nos permite trabajar con la terminología de una disciplina que conocemos y porque este trabajo se enmarca en un proyecto más amplio), consideramos que esta técnica de definición puede aplicarse a un tesoro sobre cualquier disciplina o campo del conocimiento.

METODOLOGIA

Se seleccionó una muestra de términos, bajo las siguientes consideraciones:

- Terminología de referencia: se tomó como parámetro de referencia la norma ISO 5127/3a (1983) sobre documentos tradicionales, por entender que se trata de una terminología normalizada y aceptada internacionalmente, con sus respectivas definiciones. Como complemento, se utilizó también la norma española UNE 50113/3a (1991), que constituye prácticamente una traducción al español de las norma ISO ya mencionada. Es importante señalar que ambas normas incluyen los términos y sus definiciones en un ordenamiento sistemático, de acuerdo con un esquema de clasificación propio.
- Glosarios: se eligieron los de Wersig y Neveling (1976), Harrod (1990) y Buonocore (1980). El primero de ellos también presenta los términos y definiciones en un ordenamiento sistemático, mientras que los otros dos utilizan un ordenamiento alfabético.
- Diccionarios de la lengua: se utilizaron tanto el Diccionario de la lengua española (Real Academia Española, 1992) como el de María Moliner (1996), que por supuesto poseen un ordenamiento alfabético.
- Vocabularios controlados: "Vocabulario controlado de bibliotecología y ciencias de la información" (VCBCI) (Peniche Sánchez McGregor, 1992) y "ASIS thesaurus of information science and librarianship" (ASIS) (Milstead, 1994). El primero es una lista de descriptores que no incluye relaciones jerárquicas, mientras que el segundo sí es un tesoro, en el que todos sus descriptores están organizados en un ordenamiento sistemático.

Una vez reunidas las fuentes, se completaron los siguientes pasos:

- 1) Se extrajeron 78 términos de la norma ISO 5127/3a (1983), así como 445 definiciones de todas las fuentes consultadas: 156 de las dos normas, 176 de los tres glosarios y 113 de los dos diccionarios.
- 2) Se analizaron todas las definiciones para establecer cuáles de ellas utilizaban la técnica de definición por género próximo y diferencia específica.
- 3) Se extrajeron los términos equivalentes de los dos vocabularios controlados seleccionados y se analizaron sus correspondientes notas de alcance y relaciones jerárquicas, para comparar su correspondencia con la técnica de definición elegida.

RESULTADOS Y DISCUSION

De las 156 definiciones extraídas de las dos normas, 121 (77%) seguían la técnica de género próximo y diferencia específica, mientras que 65 de las 176

definiciones extraídas de los tres glosarios (37%) y 33 de las 113 obtenidas de los diccionarios (29%) seguían la misma técnica. Resulta evidente, entonces, que las normas muestran una preferencia mucho mayor por la técnica de género próximo y diferencia específica en comparación con los glosarios y diccionarios; sin embargo, la proporción que presentan estos dos últimos (aproximadamente un tercio en ambos casos) es suficientemente importante para confirmar la validez y el uso de este tipo de definiciones.

Por otro lado, es importante señalar que el 51% de las definiciones incluidas en Wersig y Neveling (1976), 38% de las de Harrod (1990) y 32% de las de Buonocore (1980) son por género próximo y diferencia específica. De manera que, si se tiene en cuenta que tanto las normas como el glosario de Wersig y Neveling poseen un ordenamiento sistemático (característico también de los tesauros), nuestros resultados apoyarían la opinión expresada en la Introducción de que este tipo de ordenamiento favorecería la construcción de definiciones por género próximo y diferencia específica.

Por otra parte, del total de 78 términos extraídos de la norma ISO 5127/3a (1983), sólo 26 (36%) estaban incluidos en los dos vocabularios controlados, mientras que 23 términos (29%) se encontraban solamente en uno. Otros 29 términos (37%) no se incluían en ninguno de los dos vocabularios.

De los 45 términos incluidos en VCBCI, 14 (31%) incluían una nota de alcance y 13 de ellas eran definiciones, aunque sólo 3 por la técnica de género próximo y diferencia específica; como ya se explicó, ninguno de los descriptores contaba con relaciones jerárquicas. En contraste, de los 30 términos incluidos en ASIS, sólo uno incluía una nota de alcance con instrucciones para el indizador y no como definición; sin embargo, todos los descriptores mostraban términos genéricos y 11 de ellos (37%) agregaban términos específicos.

Estos resultados coinciden con los de Alvaro Bermejo et al. (1989) en cuanto a la escasa cantidad de descriptores que incluyen notas de alcance (menos de un tercio). Sin embargo, llama la atención favorablemente la tendencia del VCBCI a incluir definiciones como notas de alcance.

Debe recordarse que el VCBCI no posee una estructura jerárquica, mientras que todos los descriptores de ASIS poseen una indicación de término genérico y/o término específico. La comparación de esta característica con la disponibilidad de notas de alcance, en ambos vocabularios controlados, parece sugerir que la estructura jerárquica haría innecesaria la inclusión de una nota de alcance y vice versa, la falta de una estructura jerárquica requeriría mayor cantidad de estas notas.

Sin embargo, la forma narrativa de la nota de alcance puede facilitar la comprensión por parte de los usuarios, ya que las tradicionales referencias de TG/TE, las sangrías, estructuras arborescentes u otras formas de indicar la relación jerárquica, pueden ser difíciles de comprender para los usuarios o de implementar en sistemas de información con recursos limitados.

CONCLUSIONES

Los tesauros y otros vocabularios controlados, que antes eran propiedad exclusiva del indizador, se han convertido en herramientas de recuperación útiles para cualquier usuario de sistemas de información en línea, a quien se debe proporcionar la mayor cantidad de información posible para facilitar la elección de los términos de búsqueda. Una contribución importante en este

sentido es la inclusión de una definición para cada uno de los descriptores u otros términos de indización.

La técnica de definición por género próximo y diferencia específica es ampliamente utilizada en normas, glosarios y diccionarios y puede adaptarse sin mayores dificultades a los tesauros, debido a que el ordenamiento sistemático de éstos pareciera facilitar su redacción.

No obstante, se requiere continuar con este tipo de investigaciones para establecer la utilidad de otras técnicas de definición que contribuyan a enriquecer la comunicación entre un sistema de información y sus usuarios, eliminando la ambigüedad y la vaguedad de los términos.

AGRADECIMIENTOS

Esta investigación ha sido subsidiada por el Programa de Incentivos del Ministerio de Educación de la Nación.

BIBLIOGRAFIA

Arntz R, Picht H. Introducción a la terminología. Madrid: Fundación Germán Sánchez Ruipérez, 1995.

Aitchison J, Gilchrist A, Bawden D (1997). Thesaurus construction. 3 ed. London: Aslib.

Alvaro Bermejo C, Villagrà Rubio A, Sorli Rojo A (1989) Desarrollo de lenguajes documentales formalizados en lengua española. II. Evaluación de los tesauros disponibles en lengua española. Revista española de documentación científica; 12(3):283-97.

Bates M (1989) Rethinking subject cataloguing in the online environment. Library resources and technical services; 33(4):400-12.

Blasco Génova R, Blass Konug M (1978) Tesauro de epígrafes para la clasificación de bibliografía en diccionario sistemático. Madrid: Idice.

Browne G (1992) Scope notes for LISA subject headings. Online and CDROM review; 16(1):3-16.

Buonocore D (1980) Diccionario de bibliotecología: términos relacionados a la bibliología, bibliografía, bibliofilia, biblioteconomía, archivología, documentología, tipografía y materias afines. 2 ed. Buenos Aires: Marymar.

Connell TH (1997) Use of the LCSH system: realities. Cataloging and classification quarterly 23(1):73-98.

Copi IM (1980) Introducción a la lógica. 21 ed. Buenos Aires: EUDEBA (Manuales).

Fidel R (1991) Searchers' selection of search keys: II. Controlled vocabulary or free-text searching. Journal of the American Society for Information Science 42(7):401-14.

García MB, Monfasani RE (1983) Construcción de un tesoro de psicología. Buenos Aires: UBA, 1983.

Harrod LM (1990) The librarian's glossary of terms used in librarianship, documentation and the book crafts and reference book. 5 ed. Bowker.

Haensch G, Wolf L, Ettinger S, Werner R (1982). La lexicografía: de la lingüística teórica a la lexicografía práctica. Madrid: Gredos (Biblioteca románica hispánica).

Hudon M (1996) Preparing terminological definitions for indexing and retrieval thesauri: a model. En: Knowledge organization and change: Proceedings of the 4th International ISKO Conference, Washington DC, 15-18 July 1996. Advances in knowledge organization; 5:363-9.

IFLA (1993). Guidelines for subject authorities and reference entries. München: KG Saur.

ISO 2788 (1986) Documentation: Guidelines for the establishment and development of monolingual thesauri. 2 ed. Geneva: ISO.

ISO 5127/3a (1983) Documentation and information: Vocabulary: Traditional documents. Geneva: ISO.

Knapp SD (1984) BRS/TERM: a vocabulary database for searchers. Database; 7(4):70-5.

Kristensen J (1993) Expanding end-users' query statements for free-text searching with a search-aid thesaurus. Information processing and management; 29(6):733-44.

Lancaster FW (1986) Vocabulary control and information retrieval. 2 ed. Arlington, Va: Information Resources Press.

López Huertas MJ (1997) Thesaurus structure design: a conceptual approach for improved interaction. Journal of documentation; 53(2):139-77.

Milstead JL (1994) ASIS Thesaurus of information science and librarianship. Silver Spring: ASIS.

Moliner M (1996) Diccionario de uso del español. Edición en CDROM. Madrid: Gredos.

Peniche Sánchez McGregor S (1992) Vocabulario controlado en bibliotecología y ciencia de la información. México: CUIB, UNAM.

Pollard RA (1993) Hypertext-based thesaurus as a subject browsing aid for bibliographic databases. Information processing and management 29(3):345-57.

Real Academia Española (1992) Diccionario de la lengua española, 21 ed. Madrid: Espasa Calpe.

Schoonbaert D (1996) Spirs, Winspirs and Ovid: a comparison of three Medline databases on CDROM interfaces. Bulletin of the Medical Library Association; 84(1):63-70.

Strong GW, Drott MC (1996) A thesaurus for end-user indexing and retrieval. Information processing and management 22(6):487-92.

Svenonius E (1997) Definitional approaches in the design of classification and thesauri and their implications for retrieval and for automatic classification. En: Knowledge organization for information retrieval: Proceedings of the 6th International Study Conference on Classification Research, University College, London, 16-18 June 1997. [En prensa].

UNE 50113/3a (1983): Documentación e información: Vocabulario: Documentos tradicionales. Madrid: AENOR.

Wersig G, Neveling U (1976) Terminology of documentation. 2 ed. Paris: Unesco.

Zgusta L (1971) Manual of lexicography. The Hague, Paris: Mouton.