

CAPÍTULO I

LENGUAJES EPISTEMOGRÁFICOS: Consideraciones Generales

1.1. Objeto y aproximación

Esta investigación se centra en el desarrollo de una metodología que permite la construcción de lenguajes asociativos de tipo epistemográfico, a la vez que se proponen sus fundamentos, estructuras y fines generales para la gestión de un sistema de información en materia de Patrimonio histórico de Andalucía. Para ello, se parte de la observación y descripción de los errores, lagunas y problemas desprendidos del Thesaurus de Patrimonio histórico andaluz -TPHA-, lenguaje elaborado para el Sistema de información de Patrimonio histórico de Andalucía -SIPHA-, con vistas a su transformación en un lenguaje documental que potencie su vocabulario, estructuras y, en consecuencia, sus funciones y utilidades.

Tras un análisis pormenorizado de la normativa nacional e internacional sobre construcción de tesauros, contrastada con los elementos innovadores del TPHA, se llega a la hipótesis de que es posible organizar una epistemografía integrada del Patrimonio histórico andaluz -PHA- (teóricamente de cualquier patrimonio histórico que presente categorías y estructuras semejantes, aun sin coincidencias en lo temático) utilizando elementos de Semántica estructural, Lingüística teórica, Lingüística documental y Análisis del Discurso organizados en una marco de tecnología hipermedia. Es éste, por lo tanto, un proyecto experimental que redundará en la propia extensión y consolidación de la Lingüística Documental -LD-, disciplina en la que se halla inserto.

Una vez formulados oportunamente los problemas léxicos y estructurales, inherentes a todo tesoro, se opera en cuatro niveles:

- definiendo el aparato terminológico y conceptual utilizado indicando, con especial énfasis, aquellas modificaciones sémicas introducidas en elementos conocidos y que necesitan revisión y realizando nuevos encuadres y ajustes epistemológicos. Este ejercicio aporta la mayor carga teórica a la LD y justifica la primera parte del título del trabajo (principios de lenguaje epistemográfico).
- partiendo de las relaciones jerárquicas y asociativas convencionales para acceder al concepto de macroestructura introducido por la relaciones binarias lógico-semánticas y explicitables, elementos que ponen al descubierto la necesidad imperiosa de una actualización del metalenguaje de nuestra disciplina: macrocategorías, escenarios, microestructuras o encadenados son conceptos teóricos que deben formar parte del nuevo sistema de descripción.
- equiparando la estructura del TPHA con otras organizaciones categoriales resultando un cuadro macrocategorial más amplio y acorde con la cobertura y profundidad del nuevo lenguaje.
- construyendo las herramientas necesarias para, junto al sistema de descripción y la metodología, elaborar el producto deseado.
- zanjando el problema de la reducción de los elementos del vocabulario a una sola categoría léxica, mediante la admisión de las "aisladamente" significativas, cualquiera que sea su naturaleza.
- creando el corpus de macrorreglas y reglas que afecta al vocabulario, estructura y uso y a las relaciones entre dichos niveles.

Los cuatro últimos puntos justifican el subtítulo de este trabajo: la representación del conocimiento sobre PHA. La precisión, en la formulación del objeto y su método de abordaje, permite realizar consideraciones de primer orden sobre las que descansarán las conjeturas, las variables que determinan el objeto, las hipótesis generales de trabajo y los objetivos trazados.

Conjeturas y considerandos:

- los lenguajes documentales convencionales -ldc-, en general y los tesauros en particular, no disponen de recursos metodológicos ni instrumentales, debido a las limitaciones emanadas de su propia naturaleza, para solucionar, con suficiencia, los problemas y lagunas de los sistemas de información actuales.

- los ldc no han experimentado evolución sustancial en su concepción teórica, metodológicas de construcción o interfaces desde hace cuarenta años en que irrumpieron los primeros tesauros en la escena documental. No potencian la gestión de sus bases léxicas mediante operadores booleanos u otras gramáticas más complejas ni desarrollan satisfactoriamente el eje sintagmático de sus universos cognitivos, insuficiencias que los aleja de lo lingüístico.

- los ldc no han experimentado una evolución pareja o acorde con las posibilidades de las nuevas tecnologías de la información, esencialmente, con las herramientas de gestión de bases de datos y las potencialidades de transmisión y depósito masivo de datos que revolucionan los sistemas de información desde principio de los noventa. La programación avanzada en IA, los sistemas hipermedia y los protocolos de comu-

nicaiones actuales ofrecen un horizonte sin precedentes para reflexionar e intervenir sobre la organización y circulación del conocimiento.

- los softwares de construcción y gestión de thesauri no han experimentado cambios que acompañen al mencionado desarrollo tecnológico y, en consecuencia, ofrecen pobres perspectivas a la metodología, los documentalistas y los usuarios finales.

Una de las causas, que nos ha movido a afrontar esta investigación, es el extrañamiento de las estructura discursivas respecto a las estructuras lógico-semánticas de los lenguajes documentales: el distanciamiento del registro empírico y la nomenclatura teórica que no logra reproducirlo a menos que sacrifique excesivo significado o se atenga a una teoría científica rechazando las demás. El objeto central de este trabajo radica en la construcción de un esquema que compatibiliza las estructuras textuales-enunciados de usuarios con las estructuras apriorísticas del lenguaje documental: las macrocategorías organizadas en escenarios establecen ese puente.

Los lenguajes documentales absorben la terminología tecno-científica mediante formalizaciones que ignoran el comportamiento enunciativo. El proceso y las condiciones de producción textual son elementos que el documentólogo no puede soslayar. Los lenguajes documentales deben reflejar, no sólo un sistema nocional dado sino también como lo ordenan y manipulan los actantes involucrados en la acción discursiva. En este sentido, el esquema macrocategorial (o arquitectura global del escenario) es un marco suficientemente amplio que permite hacer sus articulaciones sobre las estructuras discursivas.

1.2 Hipótesis y objetivos generales

Las observaciones anteriores reubican la investigación y la aplicación de la Lingüística documental en una nueva plataforma desde la que pueden enunciarse las premisas, los objetivos y las hipótesis de trabajo:

- Es posible construir lenguajes documentales avanzados sobre campos especializados del conocimiento que, a la vez, sustituyan a los convencionales tesauros, mejorando su rendimiento y aprovechando sus léxicos y estructuras y ofreciendo nuevas funciones. Estos "lenguajes epistemográficos" son representaciones cognitivas creadas a partir de construcciones lógico-semánticas y discursivas.

- Si lo anterior es viable, deben introducirse modificaciones sustanciales en las categorías léxicas y estructurales que intervienen en el lenguaje, así como en los operadores, en la morfología, en las reglas y en los sistemas de representación, uso y recuperación.

- Los lenguajes documentales, en general, y los asociativos, en particular, son constructos artificiales con comportamientos discursivos. Por tanto, en su descripción y formación intervienen elementos y categorías provistas por el Análisis del Discurso y disciplinas afines. En la confluencia de los conceptos teóricos del discurso con los programas hipertexto es posible construir epistemografías equivalentes a representaciones pragmáticas de mapas cognitivos.

- Se potencia el nivel sintagmático y la articulación del lenguaje documental: los usuarios construyen el léxico y la estructura al utilizarlo mediante un sistema de retroalimentación basado en la "precoordinación realizada por el usuario" posible gracias a las nuevas categorías léxicas autorizadas como descriptores y a las gramáticas internas.

- La base léxica se aproxima más a la lengua natural superando la reducción, a sustantivos o sintagmas nominales, propia de los tesauros, mediante la incorporación de adjetivos, adverbios, prefijos, desinencias y verbos en gerundio (que expresa movimiento) y participio (que expresa estado), elementos que inciden en el cambio del paisaje léxico y genera modifica-

ciones profundas en la organización y manipulación del lenguaje.

- La jerarquización pierde su rol dominante en la definición de la estructura para prevalecer como método de construcción tan sólo en primera instancia, es decir, para servir de andamiaje del edificio horizontal. Una vez construida la macroestructura, la lógica jerárquica se transforma en una microestructura del mismo nivel de importancia y apariencia de cualquier otra relación habida en el lenguaje e, incluso, pasa a depender de la estructura asociativa.

- Los lenguajes epistemográficos son lenguajes documentales de estructura asociativa y cumplen sus dos funciones básicas y dialógicas: a) normalizan el vocabulario, b) sugieren alternativas. Por lo tanto, mantienen el rol instrumental y mediador de los lenguajes asociativos pero, además, c) sirven de bases de datos de respuesta factual, es decir, agregan el rol de producto final mediante sus constructos epistemográficos y la información que obtiene, de ellos, el usuario.

- Los lenguajes epistemográficos simulan enunciados del discurso mediante la macroestructura de orden lógico-semántico (nivel teórico-virtual) y la microestructura, de orden discursivo y pragmático.

- Cualquier regla o mecanismo lógico-semántico determinado en la construcción o en el uso debe responder al conocimiento y al raciocinio del experto y ha de explicitarse a petición del usuario, lo que convierte a los lenguajes epistemográficos en instrumentos científicos y evaluables.

1.3 Concepto de lenguaje epistemográfico como lenguaje asociativo.

En los actuales marcos tecnológicos, el usuario no recibe asistencia de reglas procedimentales que le indiquen o sugieran, con suficiencia, el camino a seguir, según sus necesidades, o los elementos que podría seleccionar. Esta situación de los lenguajes documentales, que no acompaña el desarrollo informático, se debe al bloqueo en que se encuentran la teoría y la metodología de tesauros las cuales, a pesar de su desfase y rechazo por una buena parte de la comunidad de profesionales y usuarios de la información,

siguen, paradójicamente, respaldadas por la normativa vigente nacional e internacional, por estudios, manuales, docencia y proyectos institucionales. Todo ello provoca una mayor consolidación en lugar de una ruptura tal y como los propios thesauri propiciaron respecto a las clasificaciones enciclopédicas introduciendo la lengua natural y fresca estructural.

El actual análisis de contenido aplicado con fines documentales en textos, imágenes y objetos requiere herramientas de representación más sugerentes, versátiles y perfeccionadas que los tradicionales tesauros, gestionadas por poderosos e intuitivos programas que no precisen instrucciones complicadas y que realicen un verdadero papel asesor del documentalista obligado a inmiscuirse en vericuetos profundos de los textos que lee. Es decir, el conocimiento de los expertos sobre un determinado tema debe “volcarse” en un lenguaje o sistema sígnico con el fin de simular la consulta o, en su caso, la respuesta del experto como asesor del documentalista.

Por otro lado, el tipo de lenguaje que propugnamos debe recoger en sus recursos y estructura el conocimiento de otro experto, el documentalista, para guiar al usuario no especializado en técnicas documentales en su búsqueda de información reproduciendo, mediante un interfaz intuitivo, la presencia y el comportamiento del propio documentalista consultado por el usuario.

Así, estos nuevos lenguajes documentales, que proponemos en denominar **lenguajes epistemográficos** -LE-, pertenecen al género de los asociativos desde el punto de vista estructural y se configuran como sistemas expertos en varias vertientes:

- acumulan el conocimiento y las construcciones según el criterio del experto.
- ofrecen los recursos y conocimientos del documentalista como especialista en análisis y recuperación evitando su presencia en accesos remotos, mediante un asistente (sistema tutorial e interfaz).
- mantiene la función de representación propia de los lenguajes documentales a la par de ser una base de conocimiento.

En consecuencia, un LE tiende a representar epistemologías de los diferentes ámbitos del conocimiento, a ofrecer un nuevo nomenclátor de referencias formales y quiere ser ambas cosas: mapas conceptuales y cognitivos estructurados con fines documentológicos que mantengan equidistancia respecto a la Epistemología, a la Lexicografía y a la Terminología.

La denominación **lenguaje epistemográfico** ha sido cuidadosamente elegida para representar todo ese mundo de confluencias. Por un lado, la adjetivación inspirada en la expresiva “Epistemología práctica” de Gardin¹ refleja la vocación científica de estos productos tanto en su método de construcción como en sus áreas de aplicación. Una epistemografía supone la representación formalizada, en una base de conocimiento con fines documentales (de recuperación de información), de un constructo cognitivo.

De otro lado, se mantiene y potencia la base sustantiva: el lenguaje. En efecto, lenguaje significa sistema sígnico sujeto a convenciones estructurales, funcionales y pragmáticas y tiene una especial significancia en el ámbito de la Documentación y de la disciplina que se ocupa de los mismos, la Lingüística documental. La presencia del lenguaje, en la denominación, implica la configuración que adoptarán los corpus conceptuales y el propio orden, apariencia y conjunto de reglas de raigambre también lingüística.

Un LE se declara e identifica en primera instancia con el lenguaje documental, a saber, se configura como dispositivo léxico construido artificialmente para el análisis y la recuperación de información. Es más, su concepción no parte *ex-nihil*, sino de la propia superación empírica de los más evolucionados tesauros que permite la normativa tradicional. Además de suponer una evolución teórico-metodológica de los thesauri, los LE debieran ser construidos, en el nivel práctico, a partir de lenguajes documentales con cierto grado de sistematización y experiencia como los thesauri. En definitiva, tal es el grado de imbricación, el tesoro ha de ser tomado como referencia teórica para pensar estos nuevos lenguajes y como plataforma empírica a la hora de elaborarlos. Esto no debe suponer dependencia

teórico-práctica de los LE respecto a ninguna concepción, vocabulario o estructuración.

Un LE es un tipo más del género de los lenguajes asociativos. No emerge como una alternativa o especie distinta, paralela o divergente en relación a estos lenguajes documentales. Si la característica de identificación principal de los lenguajes asociativos o combinatorios es la lengua natural de los descriptores sumada a una estructuración horizontal en campos, los LE cumplen y potencian, sobradamente, esos principios. Más adelante se verá el compromiso del LE con la estructuración asociativa del conocimiento y la posición de los mismos respecto a las jerarquizaciones.

En suma, los LE son instrumentos auxiliares de documentalistas y usuarios en sus respectivas tareas de análisis y búsqueda de datos con la características propias de los lenguajes asociativos a lo que añaden, respecto a otros tipos de lenguajes documentales, una decidida vinculación con la epistemología que pretende organizar y representar, con los usuarios a quienes asesora y que, a su vez, participan en la construcción usando el lenguaje y una directa articulación con la tecnología fuera de la cual no son concebibles ni viables.

En el modelo documentológico (lectura-transformación-representación-recuperación) se inscriben sin dificultades y asumen los principios de concepción simultánea de todas las fases, experimentalidad, aplicación y validación lo que permite considerarlos como típico producto documental².

El grado de complejidad y especialización en su construcción es inversamente proporcional al nivel de facilidad en su manejo gracias a la asistencia de la alta tecnología. Sin embargo, el documentalista o el metodólogo no realizan incursiones en esos territorios. Lo que viene a quedar paladinamente claro es que, a partir de ahora, la construcción de lenguajes documentales no es una tarea gremial: a la imprescindible labor de los documentalistas se añade la no menos esencial de epistemólogos, expertos y tecnólogos cuya intervención ha de ser ordenada y dirigida por especialistas de la información.

Finalmente, la expresión **lenguaje epistemográfico** representa más una tipología dentro

de la clasificación de los lenguajes documentales que un producto concreto (clasificación enciclopédica, tesoro, glosario, etc.). Somos conscientes de la generalización, debida a la innovación de la metodología y a las características de un lenguaje propuesto *a priori*, que las aplicaciones se ocuparán de concretar.

1.4 El lenguaje asociativo como discurso

No cabe duda de que todo lenguaje documental se construye siguiendo unas reglas procedimentales, con una finalidad como meta y utilizando códigos de distinta naturaleza (ideológicos, persuasivos o culturales) que, consciente o inconscientemente, forman parte de la estrategia y de las condiciones de producción del lenguaje. En suma, las clasificaciones, los tesauros y también los lenguajes epistemográficos se organizan como discurso. Según sus características, cada uno de ellos atenderá a uno u otro esquema discursivo aunque siempre coinciden en un número mínimo y común de constantes y variables (relaciones, códigos, finalidad) que les hace reconocibles y diferentes respecto a otras entidades próximas.

Esta constatación sitúa a la disciplina "Análisis del Discurso"³ como una fuente prioritaria de recursos teóricos, conceptuales y metodológicos para la comprensión y construcción de los lenguajes asociativos⁴. El Análisis del Discurso es un conjunto de teorías y procedimientos que explican, aliado con otras disciplinas, los usos del lenguaje. En ese aspecto, entenderíamos, por ejemplo, que el esquema común abstracto, subyacente y repetitivo de todo lenguaje documental se identifica con lo que denomina Van Dijk superestructura. Los lenguajes documentales, como la dramaturgia neoclásica, los cuentos rusos, los textos científicos o las noticias se ajustan a unas superestructuras que rigen, desde el mayor nivel de abstracción, los comportamientos del conjunto de términos, reglas y operadores en el nivel de superficie y las interacciones entre esos elementos. Los cánones estructurales permiten reconocer los lenguajes y construirlos. Sobre este particular reincidiremos al tratar las estructuras que intervienen en los LE.

Por ser lenguajes documentales avanzados, los procedimientos de construcción de los LE se acometen bajo la metacognición, es decir, bajo la presencia consciente, constante, vigilante y explícita de la superestructura y de la intervención del operador sobre el sistema de macrorreglas generado lo que da carácter científico al proceso. Son instrumentos que modulan los registros empíricos de un área de conocimiento (uni, multi, inter o intradisciplinar) los acomodan a una "Macroestructura" de categorías, convencionales y asepticas respecto a los temas que articulan, que coadyuvan al montaje de construcciones más fácilmente manipulables y abordables dentro de la complejidad del discurso en cuestión. En definitiva, la primera lógica metodológica en la construcción de un lenguaje epistemográfico la introduce el Análisis del Discurso y, a partir del mismo, puede someterse un lenguaje documental a observación y evaluación.

Nos interesa especialmente el hecho de que, entre otros campos, Van Dijk destaque las interacciones del Análisis del Discurso con la Psicología cognitiva y la Inteligencia artificial: disciplinas que se han venido ocupando de presentar "modelos de la producción del discurso y la comprensión por parte de los usuarios del lenguaje (...) este enfoque cognitivo se formula en términos de estructuras de la memoria y de los procesos implicados en la interpretación, el almacenamiento y la reintegración del discurso y en el rol del conocimiento y las creencias en estos procesos de comprensión"⁵. En general, según Van Dijk, esta disciplina viene a demostrar que las propiedades de los usos del lenguaje tienen una naturaleza sistemática y pueden explicarse mediante reglas⁶. Es palmaria la conexión de estas materias con los objetivos del Análisis documental y, en lo que nos atañe ahora, con los modos de representación del conocimiento mediante lenguajes controlados: si contamos con elementos que permiten la descripción de los discursos, esos mismos elementos pueden ser transformados en herramientas de construcción del discurso en diversas representaciones: una de ellas podría ser el lenguaje epistemográfico.

El Análisis del Discurso, en suma, nos aporta claves esenciales para una base de partida diferente en las estrategias de construcción, mante-

nimiento y utilización de un lenguaje documental: no sólo nos presta instrumentos que permiten la presencia de la metacognición⁷ en la fase de elaboración de esos lenguajes sino que también facilita su comprensión, uso y evaluación mediante conceptos teóricos y métodos de desestructuración. Por tanto, la contribución del Análisis del Discurso sitúa a la Lingüística documental lejos de la improvisación y de la intuición, herramientas usuales en la elaboración de lenguajes documentales, coadyuvando al cientifismo y a la objetivación de nuestra disciplina.

Por otro lado, un lenguaje documental se implica doblemente en el plano discursivo, lo que refuerza su posición ontológica: él mismo es un constructo artificial con una naturaleza y características propias a partir del lenguaje especializado, o sea se comporta como discurso, y manipula conceptos y ordenaciones de conceptos, enunciados, que deben proporcionar sentidos y coherencia en la representación de un discurso dado (comportamiento metadiscursivo).

A pesar de la importancia de lo expuesto, estamos en los albores de la aplicación del Análisis del Discurso en la Documentación: las disciplinas lingüísticas continúan siendo las principales abastecedoras de metalenguaje y anclaje epistémico. Puede vislumbrarse, desde esta posición concreta en la que nos situamos, un futuro inmediato de creciente y fructífero intercambio de parámetros discursivos y lingüísticos en la construcción de los lenguajes asociativos y de la Lingüística documental.

1.4.1 Discurso y lenguaje asociativo: disociaciones

Podríamos reducir la crítica a la historia de las clasificaciones y de los tesauros documentales en la disociación que mantienen estos constructos artificiales y teóricos respecto a los textos, a las demandas o, en general, a los discursos reales. Este insalvable abismo para la metodología convencional nos ha llevado a la pregunta siguiente: ¿Para qué/quién se construyen los lenguajes documentales? La irrupción de la lengua natural en las bases léxicas sirvió para ablandar los rigores de la codificación alfanumérica.

rica mostrando una incipiente y falaz naturalización de las estructuras. La sensación de familiaridad que transmiten los descriptores “naturalizados” termina en el extrañamiento o rechazo de las estructuras artificiales por parte de los usuarios. Es obvio que, para conseguir un entorno “amigable” en los lenguajes de representación, la transformación no sólo debe darse en términos de vocabulario sino también de estructura y discurso.

Los lenguajes asociativos, en general, presentan innovaciones tanto en el nivel léxico como en el organizativo con la pretensión de que el usuario del lenguaje se sienta cómodo con el vocabulario, con las relaciones teóricas y las combinaciones sintagmáticas y sintácticas previstas. La complejidad del instrumento hace necesario un software que facilite la manipulación desde un entorno gráfico en el que la intuición y un sencillo conjunto de reglas sean la principal brújula para la navegación.

Esta investigación tiene, por objetivo, la detección de las estructuras discursivas del Patrimonio histórico de Andalucía para su esquematización en una macroestructura categorial desde la que son regidos los escenarios en que se realizan las actividades, utilizando determinadas herramientas, para obtener un cierto producto contextualizado en un espacio y una época concreta merced a la acción antrópica. Del mismo modo, se asocian los agentes que ejecutan las técnicas a la materias primas que transforman y a los estilos en los que se insertan tales procedimientos. En estos itinerarios de organización discursiva introducidos en el **Thesaurus de Patrimonio Histórico Andaluz** de forma incipiente pero muy sugerente por un grupo de expertos en las diferentes disciplinas patrimoniales, radica el germen, la motivación y la fundamentación, del presente trabajo. Por lo tanto, sin la intervención del experto, del discurso, cualquier propuesta sobre lenguaje documental es inviable.

El debate especializado intragrupal debe establecerlo el metodólogo y extraer conclusiones del mismo que redunden en la consolidación de la metodología. A partir del análisis pormenorizado de las categorías, se construyen todas las combinaciones binarias posibles, todavía en el nivel macroestructural pero dentro de los pilares dis-

cursivos, para descender a la pragmática de la enunciación mediante la noción de escenario.

Sin embargo, la absorción de las estructuras discursivas en un lenguaje epistemográfico no es la única finalidad de estos: se abren nuevas fronteras para crear interfaces entre las metodologías de Análisis documental, campo de mayores retos y retraso investigador, a partir de las macroestructuras y escenarios de los lenguajes de modo que aquéllas no dependan de procedimientos ajenos o contradictorios respecto a los del sistema de representación. Así, el lenguaje asociativo, en su nivel macroestructural, debería estar presente en los campos de análisis de contenido como reflejo de los principales elementos descriptivos y analíticos supuestamente existentes en el discurso y, por tanto, en los textos y demandas que se generan en su interior.

1.5 Funciones: de la instrumentalidad al producto final

Como se observa empíricamente, la mayoría de los tesauros son considerados lenguajes documentales de estructura combinatoria o asociativa⁸ por el mero hecho de contar con un vocabulario en lengua natural se cumpla o no el principio esencial de esta tipología: estructuración horizontal. No podemos referirnos metonímicamente a los tesauros como lenguajes asociativos ni representantes, por excelencia, de sus características y naturaleza sin caer, involuntariamente, en algunos errores: identificar tesoro con lenguaje asociativo supone una reducción de la especie al tipo, afirmación avalada por las características de la mayoría de los tesauros existentes.

Pero, además, por las funciones de uno y otro, pueden establecerse nuevas diferencias. Genéricamente, podemos sostener que un lenguaje epistemográfico asume todas las funciones de un tesoro, es decir las teóricamente vinculadas a los lenguajes asociativos, pero no viceversa. Tale funciones son, a saber:

- normalizar el vocabulario en lengua natural mediante controles de orden morfosemántico y
- sugerir alternativas conceptuales al usuario mediante una red semántica preconstruida.

Esta doble función de todo tesoro y, por ende, de los lenguajes epistemográficos persigue el objetivo de marcar estrategias de formalización y representación en el resultado del análisis y en la demanda y, en último extremo, facilitar la comunicación o diálogo entre usuarios y sistemas.

Los tesauros tienen un evidente carácter mediador e instrumental. Pues bien, también los lenguajes que proponemos deben ser considerados como normalizadores de vocabulario, se les debe atribuir capacidad evocativa o sugestiva y sirven de puente entre el usuario final y la memoria, entre el documento y el analista, regulando las representaciones desde una u otra óptica. Por lo tanto, todo lenguaje epistemográfico se contruye para servir de herramienta de análisis y acceso a la información.

Los tesauros son mayoritariamente consultados como instrumentos para realizar una operación posterior, es decir, no se conciben como productos autónomos. El lenguaje epistemográfico, a diferencia de los anteriores, funciona también como base de conocimiento para la recepción y satisfacción finales de determinadas demandas de tipo factual. Así, un usuario puede buscar, en estos lenguajes, descriptores para acceder a la base de datos referencial (función instrumental) u obtener información acerca del tema que desea sin necesidad de efectuar nuevas consultas o utilizarlo como puente (función de producto final). Esta capacidad de los lenguajes epistemográficos se debe a la peculiaridad de sus estructuras lógico-semánticas, montadas desde la perspectiva discursiva como escenarios sintagmáticos, en las que se simulan situaciones textuales y comportamientos de búsqueda reales gracias a la intervención del experto en la confección de las mismas.

1.6 La IS 2788 o el anquilosamiento normativo

Cuando la norma IS2788 fue promulgada, por primera vez, en 1974⁹ establecía el reconocimiento y la importancia, para la comunidad documentológica internacional, de los tesauros en la gestión de la información. Si bien podemos considerar positiva esta normalización en relación al

estado de las bases de datos de la época, su reedición de 1986 suponía la consolidación de una reglamentación que condenaba, tanto a los tesauros regidos por ella como a la investigación sobre la materia, a la diatriba de construir lenguajes tradicionales o bien establecerse en la marginalidad de la utopía, en plena era del I3R y de la aplicación de los sistemas expertos en el control de información.

La industria de los SGD y de los propios gestores de tesauros, consolidaban y aun lo hacen, ese arbitrio internacional obsoleto convirtiéndolo, a los ojos de los documentalistas, en un principio y una referencia incontestables. Todo ello tiene una importantísima incidencia en la vertiente formativa, en la que los documentalistas adoptan la regla a sabiendas de sus lagunas, cuando no de su inutilidad.

Dos críticas generales podemos hacer a la IS2788, y a sus versiones nacionales, que afectan la base de los procedimientos que propone:

- las limitaciones e imposiciones de orden léxico que merman la adecuada representación de múltiples conceptos y realidades textuales y audio/visuales, por ejemplo, impidiendo expresar las acciones o movimientos, soslayando las lagunas léxicas en los idiomas, etc. desconsiderando la capacidad expresiva de adjetivos, gerundios o participios¹⁰.
- la insuficiencia del desarrollo conceptual y metodológico sobre la estructuración del vocabulario, basándose en criterios arcaicos pero modernizados con la apariencia de operadores jerárquicos que siguen significando lo mismo que un árbol o un tabulado a la vez que se relega, la inicial vocación asociativa de la IS2788, a un Término relacionado -TR- responsable de las asociaciones y reducido a cajón de sastre de poca utilidad y mal definido en la teoría y en la práctica¹¹.

La superación de estas dos reducciones por las que aboga la norma internacional debe proceder, obviamente, de posiciones contrarias y desreguladoras:

- a. admitiendo todas las categorías léxicas

que comporten significado autónomo, a juicio de los constructores, y favoreciendo la presencia de recursos morfológicos *a posteriori* además de mantener las gramáticas lógicas en la recuperación de datos junto a una mayor presencia de gramáticas naturales controladas.

b. adoptando nuevas estructuras basadas en la horizontalidad de las construcciones cognitivas en un discurso dado, definidas por la pericia de los expertos involucrados en el equipo que elabora el lenguaje.

Para la primera liberalización deben crearse reglas y marcos que canalicen el vocabulario (género, número, composiciones, prefijos, desinencias, identificadores). Para la segunda, debe formularse un inventario de relaciones teóricas y empíricas mediante categorías y vectores que permitan una concepción del mapa epistemográfico construido desde ópticas pertenecientes a la realidad de los discursos, del lenguaje y de los usuarios.

1.6.1 Evaluación de la norma desde la evolución de los sistemas de información

Procedamos, en este epígrafe, a evaluar los elementos y fases que propone la normativa vigente a fin de deducir los fundamentos que imponen una urgente modificación o, en su caso, superación de la misma. Nuestra exigente referencia parte de la situación actual de la tecnología y de los sistemas de información que convierten en obsoleta una norma que se viene defendiendo y aplicando. Sin embargo, nuestra posición no es utópica: nos referimos a tesauros técnicamente viables lo que se ha demostrado concretamente en la construcción del Tesauro de Patrimonio histórico andaluz, lenguaje convencional que contiene recursos y reglas que nos sirven de marco empírico para realizar, con precisión y objetividad, la presente crítica.

El análisis lo efectuamos sobre la publicación del Proyecto de Norma Española PNE 50-106, aprobado por el comité técnico científico de normalización de AENOR, en la Revista Española de Documentación Científica de los años 89/90, tiempo antes de su edición como norma española¹².

Ya en sus primeros párrafos (introducción, objeto y campo de aplicación), encontramos restricciones notables: "la norma se ocupa especialmente de relaciones *a priori*"¹³ que dan la impresión de escorar la lógica constructiva hacia las relaciones hiponímicas y partitivas. Además, "se limita a los centros que emplean indizadores humanos"¹⁴ expresión poco afortunada que establece una brecha respecto a las técnicas de análisis automático; de ahí, la reducida apuesta y la escasa confianza en la poscoordinación (en realidad, la norma debería establecerse en términos de pre o poscoordinación). Más adelante, continua: "trata principalmente de procedimientos para indizar colecciones de documentos incluidos en catálogos o bibliografías"¹⁵, es decir, (dejando de lado la terminología bibliotecológica empleada de poco anclaje documental) focaliza sus propuestas sobre documentos de cualquier naturaleza que divide, en un acto de anacronismo patente, en "impresos o no" (la negación cubre a todos los demás) omitiendo cualquier comentario sobre las características e incidencias de imágenes (fijas o en movimiento) y objetos en el análisis y su representación.

Define, por un lado, el "término de indización" como "la representación de un concepto, preferentemente en forma de sustantivo o frase nominal" (teóricamente la definición de descriptor) y, por otro, el "término preferente" (que en pura lógica es lo mismo que el anterior) como "aquél que se utiliza sistemáticamente en la indización para representar un concepto determinado, también llamado descriptor"¹⁶. Las confusiones desprendidas son palmarias: no sabemos si término de indización es un genérico de descriptor o su sinónimo y también ignoramos si descriptor es equivalente al "concepto determinado" o es mero significante (la representación del concepto). Esta última situación se clarifica, más adelante, al definir el "término no preferente" como "no descriptor".

Al hacer alusión a los "términos compuestos"¹⁷ se refiere a sintagmas nominales encabezados por un núcleo sustantivo y un modificador adjetivo, en todo caso contruidos *a priori* como parte del vocabulario y no de la potencia enunciativa del lenguaje.

Sobre los “términos de indización” la norma hace una categorización poco clarificadora (citamos los ejemplos que juzgamos más confusos o contradictorios):

a. Entes concretos que divide, inadmisiblemente, en:

- Objetos y sus partes físicas (pájaros, miembros (polisemia), regiones montañosas);
- Materiales (caucho, titanio).

b. entes abstractos, en los que incluye:

- Actividades y acontecimientos (golf),
- Entidades abstractas (repite la especie en el género) y propiedades (noticias, velocidad)
- Disciplinas
- Unidades de medida

c. Entes individuales expresados por nombres propios (Argentina). A mi juicio, la categorización utilizada por los normalizadores es ambigua, incoherente y de poca ayuda para organizar el vocabulario a pesar de la nota de 6.1.2¹⁸ en la que se advierte la necesidad de reconocer la existencia de estas clases. Se observa, también, un compromiso con el etiquetado categorial que no tiene contrapartida con el dedicado a la estructuración asociativa.

Respecto a los términos, abunda sobre su presentación en 6.2, insistiendo en el uso de sustantivos y frases nominales poniendo, como ejemplo, una precoordinación innecesaria como /hospitales para niños/ en la que mezcla conceptos de distinta naturaleza mediante un nexo final (para) en lugar de proponer, en todo caso, /hospitales infantiles/. Al hablar del papel de los adjetivos, sentencia que “debe evitarse su uso excepto en circunstancias especiales” (refrendando las bondades del sustantivo). En adverbios y verbos la prohibición de uso es tajante y sin referencia alguna a las lagunas léxicas que hacen que algunos conceptos (acciones o técnicas, por ejemplo) no puedan ser transportados en sustantivos sin sacrificar parcial o totalmente su significado: pasear o caminar(acción) en sustantivo generan polisemia o falso sentido: paseo, camino; hervir (procedimiento: hervido, puesto que /hervor/ contiene semas de efecto); cocinar (procedimiento: cocinado, ya que /cocina/ es espacio y objeto). Tampoco existe margen en la norma para expresar estados o situaciones estáticas:

destruido, sentado, herido o muerto (estado).

Al referirse a los “términos compuestos”¹⁹ la norma introduce una nueva confusión: “como regla general, puede establecerse que los términos deben representar *en lo posible*²⁰ conceptos simples o unitarios”. ¿Quiere esto decir que se admitiría una polisemia o varios conceptos incluidos bajo el mismo descriptor? En efecto, la norma debiera ser más contundente sobre la uniconceptualidad del descriptor (indicando la excepción de la polijerarquías), sobre todo por tener como marco un sistema tradicional: con la nueva tecnología, la polisemia se deshace en los contextos que ya reconocen los sistemas expertos dejando, en breve, de ser un problema para convertirse en una importante reductora de vocabulario²¹.

En el referente tecnológico de la muestra debe hablarse de un solo concepto por descriptor sea éste simple o compuesto lo que no es incompatible con el uso de adjetivos que generalizan pero son tan unisémicos como pudieran serlo los sustantivos. Más adelante, ahondando en la contradicción²², la norma se refiere a los “conceptos compuestos” (?) y da como ejemplo /contaminación atmosférica/ caso de concepto unitario representado por un descriptor sintagmático. Seguramente estas confusiones se producen porque la normativa no se compromete con una definición pragmática de lo que entiende por concepto o significado del mismo modo que lo hace con los significantes. En cualquier caso estas nociones, junto a la de tema o asunto, no han sido nunca suficiente ni satisfactoriamente esclarecidas por la Documentología a partir de las perspectivas establecidas por la Filosofía, la Semántica, la Lexicografía o la Terminología.

En lo que concierne a los criterios de poscoordinación de términos, la norma permite que el indizador, en un sistema poscoordinado, decida:

1. utilizar más términos compuestos para reducir el número de los asignados a un documento, produciendo un incremento notable en el tesoro y
2. reducir el vocabulario, fraccionando los términos, asignando mayor número de ellos por documento.

A pesar de esta libertad de acción se proponen algunas recomendaciones que justifican el

fraccionamiento, con las que concordamos, y otras, que no lo justifican y, con cuyos ejemplos discrepamos ya que no se darían en situaciones de mayor desregulación.

Desde nuestro punto de vista, sólo se justifica la precoordinación en el caso de expresiones peculiares cuyo fraccionamiento provocase incoherencia en alguna de las partes o en su combinación *a posteriori*: /escalera de caracol/ ejemplo de descriptor sintagmático de origen metafórico, o en los casos de consolidación del sintagma: /juez de paz/ en Agentes, o en Objetos, elementos estructurales concretos /cerchas españolas/, que deben permanecer con esas formas tras decisión de los documentalistas.

Más adelante²³, la norma coincide plenamente con nuestra visión del vocabulario poscoordinado: el fraccionamiento de términos compuestos en otros simples (unitérminos) asignados a diferentes categorías del tesoro excepto cuando se sacrifica en exceso la significación (procedimiento aplicado en el TPHA).

En una 2ª parte se publicaron, en la misma revista²⁴, las reglas relativas a la construcción de la estructura de los tesauros. La norma española en proyecto, PNE, consideraba que “la relación teórica que marca la diferencia entre un tesoro sistemático y una lista no estructurada de términos” es la relación jerárquica²⁵ cuando debiera ser la relación asociativa puesto que esa distinción ya se apreciaba en las clasificaciones. En cuanto a las asociaciones, precisamente, la norma sentenció: “se define con más facilidad en términos de negación que en términos positivos. Incumbe a las relaciones que establecen términos que no son equivalentes y que no pueden ser relacionados tampoco jerárquicamente”²⁶. Esta definición *ab negatio* de la estructura asociativa revela un claro tratamiento secundario de la misma en la norma. Precisamente, una descripción positiva, mediante el establecimiento de las reglas y parámetros que construyen las relaciones horizontales de un tesoro, zanjaría la indefinición e imprecisión en que se encuentran los TR, operadores connaturales de los lenguajes combinatorios.

Además, la norma condena los TR a un nuevo rol paradigmático (de sustitución), desnaturalizando su función inicial, con la siguiente afir-

mación: “la conexión entre ellos (términos) debe hacerse explícita en un tesoro sobre la base de que tal conexión podría revelar términos *alternativos* que podrían ser útiles en la indización o la recuperación”²⁷. El interés de la norma en las asociaciones es, como se ve, puramente formalista y reductor. De esa manera, el engranaje más rico, teóricamente, en un tesoro, se resume en un gran cajón de sastre donde todo cabe. Esto tiene varias consecuencias:

1. la mayoría de los tesauros que se aplican a la normativa son sistemas clasificatorios jerárquicos con algunas asociaciones protocolarias;
2. los TR de rango paradigmático son escasos y poco útiles;
3. no se cumplen los principios teóricos que fundamentan la aparición de los lenguajes combinatorios como ruptura y reacción contra los férreos esquemas clasificatorios que predominaban en los depósitos documentales desde la Edad Media.

En definitiva, la norma resuelve con varias reglas formales la asociación lógico-semántica estableciendo algunas reciprocidades tan mecánicas como inútiles. En el ejemplo de clases de /equinos/ mantiene el TR repetitivo e ineficaz entre los cuatro tipos formantes del mismo campo semántico (asnos, caballos, budeganos y mulos)²⁸. De más interés, pero también de gran imprecisión ya que son propuestas indicativas y algunas poco fundamentadas, son “las situaciones relacionales típicas halladas en la práctica”²⁹: disciplinas/objetos que estudian, operación/agente, etc. Este cuadro, que ha sido ya objeto de reproducciones y múltiples comentarios³⁰ aparece como la única y tímida incursión de la Norma en la estructura asociativa, relegándola al carácter de ejemplo y dejando por elaborar la reglamentación de mayor interés.

NOTAS

1. Autor francés que se autodefine como "epistemólogo práctico" cuando materializa las construcciones del discurso arqueológico en una base de conocimiento regida por el sistema experto Snark. Vid algunas experiencias en Gardin, J.C. et al.: *Systèmes experts et Sciences humaines*. Paris: Eyrolles, 1987. 269 p.

2. Modelo defendido, como marco de nuestras descripciones y postulados, en García Gutiérrez, A.: *Análisis documental del Discurso periodístico*. Madrid: CTD, 1992. 160 p.

3. En la concepción de este autor holandés. Vid su obra: Van Dijk, Teun: *La noticia como discurso: comprensión, estructura y producción de la información*. Barcelona: Paidós, 1990. 293 p.

4. El Análisis del discurso se abre paso en la, hasta ahora, hegemónica presencia de la disciplinas lingüísticas en la formación de la Lingüística documental, concretamente, en las aplicaciones de los procedimientos de análisis y representación.

5. Así lo han estudiado Schank y Abelson, entre otros. Ibid. p.42

6. Ibid. p.44

7. La lingüista brasileña Ana Mª Cintra describe la presencia de la lectura cognitiva o inconsciente y metacognitiva o estratégica en relación a los procedimientos de Análisis documental en Cintra, A.M.: *Estratégias de leitura em Documentação*. In: J. Smit (org): *Análise documental: a análise da síntese*. Brasília: Ibict, 1987. p.27-35

8. Así lo hacen la normativa ISO, los autores clásicos en la materia (Aitchison, Van Slype, Viet, Chaumier, etc.) y las introducciones de los grandes tesauros nacionales e internacionales.

9. IS2788: Principes directeurs pour l'établissement et le développement de thesaurus monolingues. Genève, International Standard Organization, 1974. III + 14 p. (2ª ed. 1986).

10. Constatación efectuada en las prácticas de análisis documental de fotografías y audiovisuales llevadas a cabo en la materia "Documentación periodística" de la Universidad de Sevilla. En los ejercicios de representación se han introducido, experimentalmente, nuevas categorías léxicas que respetan, en mayor medida, la integridad y expresividad de la imagen.

11. Asimismo se han emprendido experiencias de organización horizontal del vocabulario obteniendo representaciones aceptables ordenadas por nuevos operadores basados en vectores. La metodología para estos ensayos se tomó de García Gutiérrez, A.: *Estructura lingüística de la Documentación: teoría y método*. Murcia: Servicio de Publicaciones de la Universidad de Murcia, 1990. 166 p.

12. Documentación. Directrices para el establecimiento y desarrollo de tesauros monolingües (PNE- 50106, ISO2788-1986 E). Revista Española de Documentación Científica, parte 1: 12, 4, 1989 , p.463-483; parte 2: 13, 1, 1990, p.601-629.

13. Ibid., parte 1, p.464

14. Ibid., parte 1, p.465

15. Ibid., parte 1, p.465

16. Ibid., parte 1, p.466

17. Ibid., parte 1, epíg. 3.7

18. Ibid., parte 1, p.469

19. Ibid., parte 1, p.477

20. Subrayado nuestro.

21. Véase, en este sentido, como un sencillo software doméstico de reconocimiento de la voz, Viavoice de IBM, supera notablemente en 1997, y a nivel fónico, esas limitaciones gráficas de la polisemia.

22. Ibid., parte 1, p.478

23. Ibid., parte 1, epígrafe 7.3.3

24. Op. Cit., parte 2, 13, 1, 1990

25. Ibid., parte 2, epígrafe 8.3.1

26. Ibid., parte 2, epígrafe 8.4.1

27. Subrayado nuestro. Ibid., parte 2, epígrafe 8.4.1

28. Ejemplo recogido en Ibid., parte 2, epígrafe 8.4.2.3

29. Ibid., parte 2, epígrafe 8.4.3

30. Véanse, por ejemplo los trabajos de Aitchison, J. and Gilchrist, A.: *Thesaurus Construction: a Practical Manual*. 2nd ed.. London: Aslib, 1987. 173 p.; Currás, Emilia: *Tesauros: lenguajes terminológicos*. Madrid: Paraninfo, 1991. 284 p. Chan, L.M. et al: *Theory of subject analysis: a sourcebook*. Littleton (Co.): Libraries Unlimited, 1985. 431 p. o mi trabajo, ya citado, *Estructura lingüística de la Documentación op.cit.*