

Edita: Laboratorio de Tecnologías de la Información y Nuevos Análisis de Comunicación Social

Depósito Legal: TF-135-98 / ISSN: 1138-5820

Año 1º – Director: **Dr. José Manuel de Pablos Coello**, catedrático de Periodismo

Facultad de Ciencias de la Información: Pirámide del Campus de Guajara - [Universidad de La Laguna](http://www.unilaguna.es) 38200 La Laguna (Tenerife, Canarias; España)

Teléfonos: (34) 922 31 72 31 / 41 - Fax: (34) 922 31 72 54

[Junio de 1998]

## Elementos de lingüística en sistemas de información y documentación

(7.287 palabras - 15 páginas)

**Dr. Antonio García Gutiérrez** ©

Catedrático de la Universidad de Sevilla. España

[agarcia@cica.es](mailto:agarcia@cica.es)

### 1. Introducción

Las bibliotecas, los archivos, los servicios de documentación y las redes de información son instituciones que tienen ya una larga trayectoria práctica, de varios siglos a varios decenios. La tecnología ha impuesto cambios y hábitos que, potenciados por la globalización, han permitido la transformación de pequeños ficheros manuales en potentes memorias de datos consultables remotamente. Sin embargo, la verdadera revolución de los sistemas de información no proviene esencialmente del factor tecnológico, a pesar de la importancia del mismo. En ellos, la materia prima, o el producto transportado y almacenado, es la información misma y su manipulación no ha acusado una transformación semejante a la operada en los soportes. A tenor de las observaciones en algunas "redes de información", a mi entender apenas redes telemáticas, podemos concluir que el cambio debe provenir de la aproximación documentológica, más concretamente, de la lingüística documental (LD). (1)

A cualquier usuario de Internet, antes beneficiario privilegiado de bases de datos "on line" suministradas por multinacionales situadas en los países occidentales más industrializados, no se le escapa que, si bien se abre ante sus ojos un prometedor escaparate de posibilidades de obtención de información, ora la promesa es efectivamente sólo un escaparate sin trastienda, ora surgen conflictos lógico-semánticos y sintácticos en la localización de datos pertinentes, ora el nivel de ruido se dispara en relación a la demanda planteada. Muchos (los creyentes de la panacea tecnológica) atribuyen los problemas a la corta edad de la tele-informática (para otros "cortedad") aunque los errores en la era del ordenador multimedia en lo que se refiere a recuperación de información y satisfacción del usuario son idénticos a los conocidos hace treinta años. Nuevamente estamos ante una ausencia de aprovechamiento de los recursos de la lingüística aplicados a la documentación.

Nos hallamos, en consecuencia, ante el reto histórico de acompañar el necesario desarrollo tecnológico arbitrado por ingenieros y tecnólogos con modelos, aparato conceptual y metodologías aportadas por humanistas y científicos sociales. En este contexto, debemos resaltar la importancia de la consolidación y expansión de la lingüística documental como disciplina que enraíza sus fundamentos en los postulados de las ciencias del lenguaje, semánticos y gramaticales, esencialmente, en los cruces habidos con campos afines como análisis del discurso, análisis de contenido y, en general, en las denominadas, y en construcción, ciencias cognitivas.

En efecto, el objetivo de la ciencia de la información / documentación es la sistematización de principios de operación sobre el conocimiento con la finalidad pragmática de organizarlo, representarlo y ponerlo al alcance de la mayor cantidad posible de usuarios (entroncando ahí con la tecnología, si bien me refiero al interfaz amigable del lenguaje de representación y búsqueda y no al entorno informático). Pues bien, parece que los vientos apuntan en otra dirección: la aparente familiaridad de los software, el trabajoso desmenuzamiento del sistema intuitivo llevado ya a la drástica reducción icónica (todo ello muy beneficioso en el estricto campo de la relación con la computadora) crean la falacia de un, igualmente, acceso familiar e intuitivo a la información contenida en los ordenadores. Se confunde, por tanto, una vez más, el soporte con el contenido y se suscita, entre los usuarios menos avezados, la falsa realidad de que recuperar información es trivial porque, en todo caso, se recupera mucha información (incluso más de la solicitada).

El problema de la recuperación de información en Internet ha sido parcialmente evidenciado por las mismas máquinas que han pretendido su solución. Los llamados "robots" o motores de búsqueda como Yahoo, Altavista, Olé, Webcrawler, etc. que han pasado del rastreo sistemático a convertirse en una puerta a la que las ofertas web deben llamar, dirigen la más simple pregunta hacia masas incoherentes de información incluso con hiperbotones que cacarean una "advanced research" y que no es más que el viejo operador booleano usado desde los albores de la computación. Sobre este tema, particularmente, es urgente realizar

una investigación que formule con precisión el problema y apunte soluciones. En ese caso, también, el marco teórico sería fundamentalmente lingüístico.

## 2. Modelos y principios

Los investigadores que, a principios de los ochenta, decidimos romper con el paradigma tecnocrático imperante en la bibliografía norteamericana y europea, vinculado a la estadística, la empresa y la evaluación de sistemas, forma de entender la información derivada del modelo conservacionista originario y evolucionando hacia lo que algunos denominan hoy el paradigma digital, en definitiva el triunfo del soporte sobre el contenido, nos agrupamos (sin conciencia grupal desde luego) en torno a los análisis semánticos y sus aplicaciones a la documentación. Esta perspectiva fue aprovechada tanto por quienes veníamos de los problemas de la información como por lógicos y lingüistas interesados en los mismos, buscando una salida aplicada a sus conocimientos. Inicialmente, como es obvio, se produjeron desencuentros: excesivo celo en los postulados, falta de visión global de los procesos, extrañamiento respecto a los fines pragmáticos de la documentología... Como puede deducirse, posiciones propias de la desconexión de investigadores entre ellos y respecto al objeto, de un lado, actitud rupturista con el statu quo oficial de las investigaciones y crisis normal en el nacimiento de un nuevo enfoque.

El arqueólogo francés Jean Claude Gardin llevaba muchos años dando pistas sobre el camino a seguir en obras publicadas en los sesenta, por lo que se le puede considerar verdadero precursor de la nueva aproximación, especialmente, en su trabajo *Les analyses de discours* (2) de 1974 parcialmente publicado como artículo en inglés: *Document analysis and Linguistic Theory* (3). Esta línea le lleva a materializar sus postulados en aplicaciones en el área de conocimiento en la se halla especializado su grupo de investigación y, como consecuencia, publica conjuntamente con sus colaboradores varios libros en los ochenta que consolidan pragmática y magistralmente la aproximación que vengo exponiendo: *Systemes experts et Sciences humaines* (4), *La Logique du Plausible* (5) y, ya en los noventa, *Le calcul et la raison* (6), entre otros. Esta obra viene a confirmar la concepción de una epistemología práctica como sinónimo de documentología y su imbricación, desde la teoría lingüística, en el nuevo paradigma cultural / cognitivo que impregna a los que nos reconocemos en esta corriente de pensamiento científico.

De la apropiación del modelo estructuralista del signo lingüístico por la documentación, podemos decir que surgen las nuevas tendencias que observan los procesos documentales, formulan los problemas y proponen procedimientos de manera distinta a la oficialista si bien ensamblando, en esta nueva forma de pensar la información, los métodos, las normas y los autores clásicos de nuestra disciplina, bien con consideraciones muy críticas, bien reconduciendo y aprovechando ciertos bagajes. Como ilustración de lo expuesto baste mencionar las aportaciones esenciales de autores como Ranganathan y Vickery en cuanto a la clasificación del conocimiento o Salton y Ellis en recuperación, y en el lado opuesto el anquilosamiento de la norma Iso 2788 sobre elaboración de thesaurus o la ausencia alarmante de técnicas basadas en reglas para la objetivación del análisis del contenido de los documentos.

A pesar de la limitada bibliografía sobre la concepción que propugno, escasos pero fundamentales libros, tesis doctorales o contribuciones en revistas han hecho consistente la idea de que los problemas derivados de la obtención de información en los nuevos sistemas de información son problemas de lenguaje y, por tanto, la solución a los mismos proviene de las disciplinas que se ocupan tradicionalmente de estos, por emplear una expresión, las semánticas y gramáticas aplicadas a la gestión de la información y, por poner una etiqueta, la lingüística documental.

De esta forma, la LD se ocupa de ordenar los procedimientos de captación de los mensajes (lectura), de las transformaciones resultantes de la actividad anterior y de la organización y estructuración de dispositivos de representación a fin de que la obtención de conocimiento se dé eficaz y satisfactoriamente. Para ello, la disciplina introduce elementos de actuación (reglas procedimentales) y mecanismos de explicitación de los raciocinios, condición indispensable para que los procedimientos adquieran fiabilidad, sean verificables y, en consecuencia, científicos. La LD se distancia de las normativas, que persiguen el mismo fin normalizador, al conferir credibilidad científica a sus propuestas enmarcándolas en la lógica del proyecto investigador.

Así, tanto la elaboración de una técnica de análisis documental deberá ser montada sobre corpus rigurosamente verificados y validados experimentalmente, como la construcción de un thesaurus debe realizarse lanzando hipótesis metodológicas y epistemológicas, describiendo las variables consideradas en el vocabulario, en la estructura o predeterminables en el uso. Con ello, las herramientas de organización y representación citadas constituyen artefactos científicamente consistentes y sus aplicaciones o actualizaciones no quedan a merced de la arbitrariedad ni dependen de la casualidad, la memoria, la intuición, la experiencia o la lógica personales.

Veamos los principios básicos de trabajo que debemos considerar al desarrollar un procedimiento o lenguaje en un sistema de información:

### Aplicación

La documentación implica especialización o aplicación temática, tanto a nivel de investigación como de práctica. El documentólogo debe manejar un referente temático como marco real en el que se confirman o rechazan sus conjeturas y, del mismo modo, el documentalista trabaja sobre contenidos acotables. Excepcionalmente, el documentalista de prensa es un generalista, en el sentido de abarcar un área enciclopédica pero el discurso periodístico, siendo peculiar, responde a estructuras de producción sistemáticas y cualquier producto periodístico es reconocible como tal por lo que, en su caso, la aplicación viene determinada por la compensación de la intensidad o profundidad en favor de la extensión temática.

En consecuencia, todo documentalista, incluido el de prensa, debe ser formado en el discurso sobre el que pretende trabajar así como instrumentado con los métodos y teorías que le ayudan a entender las claves y elementos propios del mismo. La documentación general debe formarse como constructo teórico a partir de las teorías parciales obtenidas y aplicadas sobre discursos especializados.

## **Experimentación**

El método prioritario de observación y descripción en investigación documentológica es el experimental.

En cuanto al método empírico, se justifica por la necesidad de encontrar soluciones a universos de datos que precisan una canalización bien sea a través de mecanismos de selección, bien de análisis o de representación. Puesto que el problema habitual del documentalista es metodológico en relación a datos manipulables, puede construir "observables" artificiales mediante el muestreo y la simulación. En todo caso, las extrapolaciones generales son inviables y el grado de aprovechamiento de un discurso a otro estará sometido a un escrupuloso aumento de las muestras, de tal forma que el nuevo universo asuma los procedimientos extrapolados con todas las garantías.

Puesto que la experimentación vincula excesivamente un método a un corpus, las modelizaciones de más alto nivel resultan impracticables o insuficientes. Así, es de poca utilidad modelizar los principios de selección o los usuarios del discurso químico y establecer extrapolaciones hacia el discurso sociológico y, de éste, al discurso periodístico. Dentro de este último, los elementos de un método empírico de lectura, por ejemplo, deben ser modificados según la variable de género o extensión. En ese sentido, el reconocimiento artificial de estructuras sintácticas del discurso doxológico de la prensa se asemeja más al aplicado sobre el discurso de la argumentación y del saber científico, a pesar de la brevedad de un editorial por ejemplo, que a otros géneros de su propio ámbito: noticias, entrevistas o reportajes.

No solamente la necesidad de trabajar en corpus reales o simulados nos impone el método empírico. El obligado marco tecnológico sin el cual los procesos documentales modernos no son posibles (transmisión de millones de datos desde / hacia millones de usuarios potenciales) marca, también, el método a seguir. De hecho, la mayoría de los procedimientos de índole metodológica o reglada como lectura, síntesis o representación convergen necesariamente en una tecnologías que los hace viables o inviables.

## **Pragmatismo**

El objetivo de la documentación es de orden pragmático, es decir, todos los esfuerzos se dirigen a la obtención de un producto. Este hecho, que afecta a las vías de construcción teórica y a la misma epistemología documentológica, obedece al carácter históricamente práctico y manual de las actividades documentales: organización, ordenación, dosificación, representación, difusión, recopilación, son palabras claves del universo del documentalista y, por tanto, también de la perspectiva de su investigación.

La documentación es disciplina instrumental o auxiliar de otras ciencias o discursos, lo mismo que la terminología o la normalización. Así, la documentación hace suyo el objetivo de organizar y divulgar los conocimientos en otros campos y, en consecuencia, la instrumentalidad determina nuevamente su carácter pragmático.

En nombre del pragmatismo, pues, se investiga la documentación, se buscan y recortan contribuciones de otras disciplinas, se edifica una superestructura epistemológica a prueba del sismo constante que provoca la praxis, hasta tal punto, que la misma naturaleza de esa estructura central se basa en una renovada transformación.

## **Validación**

El método experimental sobre muestras hace necesaria una metodología de validación de resultados. La validación debe producirse sobre corpus en los que cualquier elemento del universo discursivo en cuestión haya tenido la misma oportunidad de participar. La evaluación de métodos documentales ha conocido un gran desarrollo en su vertiente tecnológica, merced al interés de las multinacionales por el rendimiento de los módulos de consultas de las bases de datos. Sin embargo, estos procedimientos verifican el sistema y sus prestaciones y no las relaciones que mantiene el discurso matriz con su representación documental y el rol que desempeñan productores, mediadores y usuarios en el proceso.

En consecuencia, es necesaria la investigación de métodos de validación documentológica y la elaboración de los mismos para los discursos específicos en el trabajo científico a la vez que, en el docente, el estudiante de documentación aplicada debe conocer metodologías evaluadoras para rectificar procedimientos profesionales y estar en condiciones de modificar las conductas en la adquisición de materiales y en la actualización de fondos.

## **3. Documentación y comunicación social**

Además de sus vitales relaciones con la lingüística, también la documentación mantiene conexiones con otras disciplinas del mismo ámbito, que enumeraré más adelante, además de su entronque con las ciencias de la comunicación de las cuales se declara partícipe. Nuestra disciplina es, para la mayoría de tratadistas y en la mayoría de las lenguas científicamente relevantes, la ciencia de la información (Information Science). En consecuencia, las referencias que sugiere el epígrafe imponen una matización al pluralizar: las denominadas en España ciencias de la información (más conocidas como ciencias de la comunicación, al menos, en Europa y América), o conjunto de disciplinas que tienen por objeto la descripción y la extracción de los postulados y leyes que rigen los procesos comunicativos promovidos por los mass media, su evolución, causas y efectos. La ciencia de la información, o documentación, tiene por objeto el establecimiento de metodologías y la explicación de los procesos de comunicación en los que interviene la información documental (información registrada reutilizable) y es un área de conocimiento de las ciencias de la comunicación. Hoy, las redes telemáticas dan a la documentación el carácter de mass media que poco tiempo atrás se reservaba a la prensa o la TV.

Un sistema documental comprende unos modos y unos medios de tratamiento y circulación de la información contenida en documentos. El objetivo esencial del sistema es informar sobre contenidos localizables en documentos de cualquier tipología. Tal vez, las diferencias más notables entre los especialistas de cualquier materia y sus documentalistas, serían resumibles en

dos puntos:

1) la condición de permanencia del soporte como elemento indispensable para la selección de información (en consecuencia, no son documentación: hechos, observaciones, reflexiones, deducciones, discursos, gritar, interpelar, dialogar, etc., fuente y discurso, salvando las distancias, propios de periodistas, historiadores, científicos, juristas, etc.).

2) el aprovechamiento derivado del proceso que realiza el documentalista. A diferencia de los especialistas que leen y observan para su propio conocimiento y producción, el documentalista es un delegado informativo que lee para otros, en sentido análogo al periodista: la captación de la información cobra sentido si hay inmediata difusión y recepción.

La documentación se ocupa del proceso de un discurso fragmentado en unidades físicas (soportes) y no tiene, por tanto, constancia directa de los hechos ni de la realidad. El soporte permite la manipulación de datos para su proceso y es, en consecuencia, un anclaje del conocimiento pero el documentalista abandona su suerte a la credibilidad y fiabilidad de la fuente.

Puesto que el proceso documental no tiene sentido sino es para culminar un ciclo comunicativo (al igual que el periodístico) dotando al usuario de información sobre fuentes que han sido intervenidas en diversos momentos por distintos agentes (políticos, agencias, redactores, analistas, clasificadores) se genera una mediación y el canal transmisor impone una codificación y decodificación de señal, tanto en el sentido semántico como técnico o telemático. Véanse las analogías del proceso documental y el periodístico en una sociedad moderna.

Es más, el proceso documental es un tipo de proceso comunicativo en el que los documentalistas son los emisores (persuasores en palabras de Lozano) (7), el mensaje es el producto que genera (resúmenes, datos factuales, índices...) o discurso documental, el código es el lenguaje de representación (semántico) y la señal del módem (telecomunicación), el medio es la infraestructura telemática y los receptores son los usuarios (interpretadores), habitualmente especializados.

Así, la documentación se inscribe en los modelos generales de la comunicación a la vez que produce sus propias teorías parciales. Documentación es un modo informativo que materializa sus productos a través de medios convencionales: libros, revistas, ordenadores, discos ópticos, páginas web, auxiliado por un necesario marketing que dé a conocer la oferta de información que proporciona un centro documental.

Shannon y Weaver publican en 1949 su modelo matemático (8) fijando el concepto de entropía sobre la suma de información requerida en una situación dada para eliminar la incertidumbre. Se aplica inicialmente sobre los procesos de transmisión electrónica, aspecto que incumbe esencialmente a ingenieros y tecnólogos más interesados en la capacidad de transmisión del canal que en la información transmitida.

Para De Bonville, "a partir de las teorías de Shannon y Weaver cristalizan los modelos comunicativos aportando un cuadro conceptual en el que sería reducido el conjunto de fenómenos de la comunicación humana" (9). El investigador canadiense examina el paradigma haciendo extrapolaciones hacia la documentación y describe sus cinco componentes: fuente que produce el mensaje, transmisor que adapta el mensaje de la fuente haciéndolo compatible con el canal, canal que transporta la señal, receptor que interpreta el mensaje mediante la captación de formas transformadas en datos para ofrecerlo al usuario, y usuario a quien se destina el mensaje. Concordamos, con De Bonville, en que este modelo fundamentado en los procesos de telecomunicación es perfectamente ajustable a las necesidades documentológicas por lo que nuestra disciplina se halla inserta en la epistemología comunicativa.

A pesar de la extrapolación elemental, el proceso documental genera sus propios instrumentos y métodos creando una idiosincrasia que, para los detractores de este modelo, proporciona en la documentación otras vinculaciones (con la ciencia de la ciencia y la epistemología). Zunde, por ejemplo, no critica la extrapolación, pero señala que la documentación debe tener un mejor conocimiento de sus propias leyes y de los fenómenos medibles antes de aceptar el modelo: "el objeto de estudio de la ciencia de la información son fenómenos empíricos asociados con procesos de información tales como la generación, transmisión, transformación, condensación, almacenamiento y recuperación. El objetivo último consiste en alcanzar mejor comprensión sobre la naturaleza de la información. Comenzando como hacen todas las disciplinas empíricas -con una descripción de los fenómenos en el dominio de su interés- "la ciencia de la información pretende establecer principios generales a través de los cuales puedan explicarse fenómenos observados" (10).

Con esta afirmación, Zunde reduce en gran medida el carácter auxiliar y aplicado de la documentación poniendo como objeto de estudio los procesos generados por la misma. Esto es razonable, como ocurre con el caso de las ciencias publicitarias, ya que el producto que genera el proceso documental, el discurso documental, ha sido poco observado y analizado por los teóricos de la documentación. Este discurso se compone de elementos o constructos elaborados para la comunicación y de él, nos interesan las condiciones de producción (actitudes, posición y limitaciones de los agentes productores), las estrategias y estructuras transmitidas, las interacciones y los efectos del propio discurso documental y, relacionado con el discurso "natural" que representa, las distorsiones, reducciones y simplificaciones, que lleva a cabo y los mecanismos (modos y medios) que emplea para su realización. Esto es un campo de investigación específicamente documentológico.

Ahora bien, no podemos olvidar el fin prioritario que tiene asignada la documentación en el conjunto de las ciencias instrumentales: generar procesos de organización y circulación de todo tipo de conocimientos. El mensaje debe ser codificado y decodificado en su recorrido, de tal forma, que puede producirse ruido en la recepción. El concepto de ruido es uno de los esenciales incorporados por la documentación en el control de su proceso (11).

Junto a su carácter de disciplina instrumental para el desarrollo científico pensamos, con De Bonville, que la función social del documentalista no se limita a crear y organizar memorias sino que, fundamentalmente, tiende a dar a conocer, a poner en circulación esos fondos sobre los que se pueden establecer consideraciones de índole cognitiva (en cuanto que son motores de nuevo conocimiento, no sólo soportes de conocimiento) y social (en cuanto el proceso alcanza cotas de difusión pública,

La dependencia documentológica respecto a los medios de comunicación ya fue enunciada por Otlet en 1934: documento es, para quien consolidó la documentación como disciplina académica, un sistema de signos sobre un soporte que se elabora con vistas a su transmisión (12). El documento nace en sociedad y a ella va destinado lo que indica el carácter social de la documentación, y en una dimensión inferior, Otlet reflexiona sobre la dependencia funcional de tres factores: lectores, libros y autores, lo que introduce un aspecto psicológico y psicosociológico (psicología bibliológica) que debe estudiar esta dependencia entre "perceptores, agentes y medios (tiempo y espacio)" (13).

La relatividad de los procesos documentales la establece Otlet sobre la pragmática receptiva, del mismo modo que Wittgenstein desde una concepción funcionalista sobre el uso del significado (14), aproximación en la que coinciden Eco, Foucault, Sartre y muchos otros pensadores y lingüistas (15). Para Otlet, "el libro no existe más que en función del lector, es decir, lo que no percibe el lector no existe para él, por tanto, su contenido desde la perspectiva de la recepción no es más que la expresión de las facultades del lector" (16). Algo que ya dijera Platón en *El banquete* muchos siglos antes.

Este carácter individualista del documento, como entidad social, es uno de los elementos esenciales del proceso documental, puesto que su objetivación, en teoría, desvincula el contenido de los usuarios. Sin embargo, documentación implica "socialización" y, en consecuencia, máxima objetivación de los procedimientos a fin de atender a mayores audiencias. En este sentido, las metodologías documentales se rigen por el principio del pragmatismo, se vinculan con los fines sociales del proyecto y se inscriben en los modelos generales que se ocupan del proceso de la comunicación en sociedad.

#### 4. Documentación y mass media

La presencia de medios o instrumentos usados para la difusión masiva en el ámbito documental es lo que dota a esta disciplina de su dimensión massmediática. Es la transmisión y uso social (Mijailov) lo que convierte definitivamente a la documentación en disciplina social y la ajusta a los postulados de las teorías comunicativas. El individuo ya no necesita buscar información documental porque ésta le sale al paso en carteles, periódicos, teletexto, vídeo o microordenadores ligados a Internet. Esta mutación de la posición del usuario de la documentación, de activo a pasivo, le convierte en foco de consumo de datos y, en consecuencia, en el objeto de gran parte de los estudios de aproximación psicosociológica y comunicológica imperantes hace veinte años en otras disciplinas. En este sentido, y salvo algunas peculiaridades propias, la documentología no tiene que inventar nuevos métodos de observación sino, por el momento, extrapolar y adaptar los experimentados por los científicos sociales, en general, y los de la comunicación, en particular, sobre sus distintos intereses. No cabe duda, que el hecho que culmina la transformación de la documentación tradicional (práctica antiquísima) en disciplina moderna del ámbito de la comunicación de masas, es la nueva tecnología de conservación y transmisión de conocimiento y el nuevo campo de posibilidades, y también de nuevos problemas, que comporta.

La documentación adopta dos modalidades de expresión social:

1) a través de los medios considerados masivos como prensa, radio, TV en los que se halla mezclada con datos informativos. Incluyamos también en esta categoría el libro y las revistas especializadas y científicas de cierta circulación. La información documental obtenida por un usuario, básicamente pasivo, presenta altas cotas de elaboración (y en consecuencia de mediación).

2) a través de medios de difusión individualizada, ante los que el usuario adopta un rol aparentemente activo, como videotex, redes telemáticas, CD-ROM, es decir, medios que permiten la interacción y que presentan los datos en un estado falaz de materia prima pero, no por ello, menos mediados. Los perfiles, la difusión selectiva de información -DSI- y la revista electrónica son ejemplos de servicios documentales a la carta, si bien Internet esta revolucionando en los últimos años las distintas concepciones y actuaciones documentológicas.

Ambos sistema de difusión documental tienen espacios públicos reservados y, lejos de hacerse competencia, se complementan y refuerzan incluso entre los más homogéneos: revistas y prensa, televisión y videotex, creando distintos espacios de consumo de información para los nuevos media. En la heterogeneidad de medios documentales, observamos como en el caso publicitario, el refuerzo que generan mutuamente: las redes referenciales apoyan al sector librero o de revistas, la documentación de congresos y reuniones realiza un marketing de determinadas publicaciones...

Así, la información documental es utilizada por muchos medios como un sistema publicitario más, además de servir como producto con valor comercial propio. No es de extrañar la existencia de pleitos millonarios sobre la propiedad y los derechos de autor en documentación (ej. Le Monde contra una sociedad canadiense que resumió y vendió los resúmenes de textos del diario en los ochenta) o el proteccionismo y advertencias legales de algunos medios sobre las transformaciones y venta de sus contenidos transformados, algo que no se observaba veinte años atrás.

#### 5. Documentación y análisis del discurso

El documentalista es un lector de textos, realiza una lectura dirigida (o que debe estar dirigida) por unas reglas específicas para la obtención de un resultado: la esencia del discurso o macroproposición global del productor. Por lo tanto, el objeto de su lectura no es el nivel de palabra o frase, sino el de texto o discurso. Este cambio de orientación disciplinar para el análisis documental ha supuesto una revolución en las investigaciones y en las prácticas y ha estado motivado por la incapacidad de la máquina de entender sentidos a partir de palabras fuera de contexto.

A pesar de la dificultad de obtener significados globales de forma mecánica es posible la convivencia de mecanismos reductores aplicados por el ser humano y el reconocimiento automático de los productos obtenidos hasta conseguir una liberalización de la lectura simulada, toda vez que exista una extraordinaria memoria empírica que dicte al motor de inferencia los comportamientos

a seguir en función de elementos y construcciones memorizadas miles de veces.

Este problema a resolver en los próximos años no es, sin embargo, prioritario puesto que existe una necesidad previa que paso a describir: el documentalista no es capaz de extraer las mismas conclusiones de un texto que otro colega que se aplica a la lectura del mismo texto. Incluso advertimos disparidad en los resultados obtenidos por el mismo agente lector a partir de un mismo texto en épocas distintas. Esta afirmación es constatable en cualquier centro de documentación.

Si bien el sentido común y la experiencia contribuyen a la construcción de reglas virtuales e intuitivas que el documentalista y sus colegas aplican mecánicamente, los resultados siguen presentando importantes divergencias además de otros problemas: las reglas empíricas se adaptan a textos específicos con mecanismos difíciles de explicitar lo que imposibilita la adaptación de un nuevo lector o equipo al esquema de trabajo y dificulta el acercamiento del usuario al sistema, desconocedor de los modos de segmentación textual.

El análisis del discurso proporciona un importante instrumental a la documentación para la resolución de algunos de estos problemas. Puesto que la misma disciplina está impregnada de un rico cruce interdisciplinar en el que intervienen teorías que explican los procesos mentales de la interpretación de la realidad (desde el cognitvismo), las estrategias de producción de textos y los contextos comunicativos y socioculturales en los que se desarrolla el discurso (teoría de la comunicación y pragmática discursiva), conecta con el mismo corpus epistemológico que constituye la documentación.

Al explicarnos, en consecuencia, cómo se produce y usa el texto junto a las condiciones y contextos involucrados a la vez que nos facilita herramientas para la detección de las estrategias discursivas que ocultan o refuerzan determinados elementos a la vez que afloran las proposiciones del autor de entre cientos de lexias y estructuras gramaticales de superficie, el análisis del discurso presta un auxilio de máxima importancia a la documentación, lo que ya ha sido demostrado en varias investigaciones teórico-prácticas (17).

## 6. Documentación y análisis de contenido

El análisis de contenido (AC) aporta a la documentación una larga experiencia en descripciones pragmáticas (18), muchas de ellas de vinculación social, de universos cargados de significados, los cuales, debidamente depurados y sometidos a referentes contruidos (tablas de indicadores) permiten hacer inferencias y extrapolaciones sustentadas en sólidos métodos de validación procedentes de la socioestadística.

El análisis documental es una metodología de lectura o captación de elementos a partir de textos (descripción y universo pragmáticos) que pretende la representación de los mismos en lenguajes controlados (tablas de descriptores) para posibilitar la recuperación ulterior. Vemos que la mayor divergencia acontece en los objetivos: inferir (AC) y recuperar (AD).

En cualquier caso, ambas disciplinas coinciden durante un largo trayecto común y poco explotado, de ricos y posibles intercambios: si el AC nos enseña cómo fabricar muestras, unidades operativas y métodos de observación y verificación, el AD le ofrece métodos de construcción, organización y ordenación de bases de datos, registros y campos, normalización semántica del vocabulario y amplias conexiones con otras disciplinas recortadas por el análisis documental y aprovechables para el AC.

Pero, a pesar de las aportaciones referidas, tal vez la más importante para la documentación, por la ausencia en sus investigaciones, sea la práctica constante de la validación en los trabajos sometidos al análisis de contenido (en sentido documental). En efecto, tanto en la construcción de métodos de lectura como de representación documentales, se trabaja sobre muestras (texto, vocabulario) compuestas por unidades menores. Según observamos en la bibliografía documentológica, las metodologías de creación de muestras y de distribución aleatoria que garanticen los resultados, así como los procedimientos de validación final que hagan fiables las conclusiones y pronostiquen posibles extrapolaciones, brillan por su ausencia. Por lo tanto, y dada la cercanía de ambas áreas, es necesario recortar la experiencia validadora del AC en aras de la consolidación que comportaría para la investigación documental.

## 7. Lógica y documentación

La documentación tiene una necesidad imperiosa, en su engarce con la tecnología, de formalización de elementos y enunciados tanto en el nivel de entrada como en el de salida y proceso de datos en un sistema. En determinadas operaciones de laboratorio, el investigador no está interesado en el significado real de los términos sino en su verosimilitud dentro del corpus que utiliza para la simulación. En este caso, la lógica proposicional o enunciativa es útil en cuanto que ha alcanzado grandes cotas de formalización en la representación de sentencias declarativas.

La declaración supone una reducción de la estructura sintáctica natural pero, como dice Allwood, "hay categorías morfosintácticas que no tienen contrapartida lógica" (19). La lógica predicativa, por ejemplo, no tiene en cuenta los enunciados imperativos o las interrogaciones a pesar de que teorías como la "hipótesis performativa" defiende que bajo estas formas subyace una afirmación en sus estructuras profundas y, por tanto, son objeto de análisis lógico. Por el momento, en las aplicaciones lógicas adoptadas por tecnólogos y, probablemente, a la espera de la confirmación de nuevos logros (especialmente de la Fuzzy Logic o Lógica difusa), la enunciación tópica de los sistemas expertos es declarativa (al menos en aquellos sistemas que ofrecen garantías).

Los investigadores de la llamada "semántica lógica" (Lewis, por ejemplo) trabajan para aplicar el análisis lógico a la lengua natural. Ese es el mayor punto de confluencia de la terna documentación - lógica - tecnología. La formalización pasa por la reducción, a inventarios controlados, de todas las equivalencias de cualquier categoría léxica posible. Claro está que la lengua natural, en un campo especializado, ofrece una morfosintaxis, distinta a la de la lengua coloquial, facilitando su simbolización.

La lógica construye lenguajes formales para evitar la vaguedad, la ambigüedad y la dependencia del contexto haciéndolos

exactos y unívocos. Cualquier constructor de lenguaje documental sabe que esos mismos son los objetivos que deben cumplir los vocabularios, en consecuencia, la lógica formal contribuye específicamente a la elaboración de lenguajes desambiguados.

En cuanto a la creación de prototipos inteligentes para la gestión documental hemos de recordar que la lógica estudia las reglas de inducción y deducción de elementos no necesariamente reales pero, a pesar del desinterés del lógico por la realidad psicosemántica, extraemos un importante aparato de inferencias posibles y extrapolables a enunciados reales, a fin de constituir en la máquina una base de reglas, es decir, un conjunto de procedimientos inferenciales humanos simulados.

El análisis que efectúa el documentalista sobre los textos se rige por dos lógicas: la lógica general, en cuanto organiza los procesos de adquisición del conocimiento científico, la construcción de hipótesis, de las leyes y teorías y la lógica formal, en cuanto nos informa de cómo están montados los razonamientos desde el punto de vista formal. Nótese que me refiero a discursos científicos cuya estructura responde, desde la primera concepción y por sus objetivos, a un alto grado de formalización. En ese sentido, la epistemología científica debe ser parte de la formación de los documentalistas puesto que les ayuda a comprender el discurso "logicista" (en palabras de Gardin) de la ciencia y a "mapear" las construcciones específicas del conocimiento.

Finalmente, la lógica matemática y, concretamente, la teoría de conjuntos y las aplicaciones del álgebra de De Boole ha sido de gran utilidad en los sistemas de recuperación de las bases de datos convencionales sobre conocimiento científico básico y experimental aunque de poca eficacia sobre los discursos humanos y sociales expresados en lengua natural y con sintaxis de cierta complejidad. La reducción de los operadores lógico-matemáticos en la combinatoria de búsqueda de datos es uno de los mayores problemas que debe resolver la documentación a partir de la superación de los mismos por símbolos formalizados que expresen todos los sentidos de los enunciados naturales. Este problema se ha potenciado al masificarse los datos y las demandas, los sistemas y los analistas en redes telemáticas.

## **8. El caso del thesaurus de Patrimonio Histórico andaluz: una aplicación de la teoría lingüística a los sistemas de información**

Las ventajas de la concepción científica de los procedimientos documentales defendida aquí no radican exclusivamente en el más que beneficioso fin de la objetivación que nos lleva a la posibilidad de programación y a la indispensable fiabilidad. Además, pueden darse infinidad de circunstancias positivas derivadas de la conversión de la raigambre científica de unas técnicas consideradas, en su aplicación además de en sus fundamentos, meramente profesionales.

Por ejemplo, lo expuesto puede ser ilustrado por el proyecto de construcción del thesaurus andaluz de Patrimonio Histórico, encargado a quien suscribe por el Instituto Andaluz de Patrimonio Histórico de la Consejería de Cultura, organismo de la comunidad autónoma de Andalucía (20).

En ese proyecto, no solamente han debido superarse las reducciones recogidas por la normativa internacional sobre construcción de thesaurus dado el calado y la extensión del objeto, la complejidad y multiplicidad de actores involucrados, sino que, además, la consistencia de las metodologías ya arbitradas por la LD ha hecho posible salir de un atasco corporativista a historiadores del arte, arqueólogos, arquitectos, antropólogos, conservadores y restauradores merced a la redistribución del mapa conceptual que nos propone la gramática de casos (o teoría de los casos universales) aplicada a la elaboración de lenguajes documentales. En efecto, el mismo thesaurus se construye a partir de unas conjeturas previas que afectan a su contenido y a su estructuración y que devienen hipótesis en un momento dado, junto a un inventario de variables que se erige como sistema de anclaje del constructo en elaboración a una realidad determinada, concebida como observable (aún en el nivel de confección teórica de la herramienta) para más adelante ser desarrollado sujeto a un método, la gramática de casos conceptuales enunciada por Fillmore y Pottier y recortada por Cunha, a su vez reconducido y modificado en constantes contrastaciones con los corpus iniciales.

Tras varios meses de cotejo de los casos de la gramática (nivel macroestructural) con la base léxica (nivel microestructural), se procede a realizar una división del trabajo por encima de las especialidades de los diez componentes del grupo de trabajo (especialistas en las distintas disciplinas que confluyen en el patrimonio histórico, tanto en su vertiente de investigación - universidad- como de conserva y explotación -museos-). La estrategia metodológica urdida por la gramática de casos nos ha permitido no solamente proceder dentro de un marco científico y por lo tanto fiable, si bien sujeto al condicionamiento de la variables, y en ese sentido hemos obtenido resultados convencionales pero no arbitrarios, sino además dejar de lado los desencuentros de las disciplinas mencionadas incapaces de articularse por sí mismas en un todo global pragmático (agruparse en un mismo foro) superado por la realidad: ya existían bases de datos interdisciplinares que urgían la sistematización de un vocabulario común y la normalización de las formas de análisis y acceso a la información.

En la concepción auténticamente patrimonialista de los bienes históricos y de interés cultural, la LD a través de uno de sus dispositivos, la metodología de casos conceptuales ha sido de crucial importancia para dar al traste con disputas sectaristas que afectan la recuperación global de información y el interés de los usuarios de las bases de conocimiento sobre patrimonio histórico y hacer posible un lenguaje integrador de todas las disciplinas concernidas.

### **Reflexión final**

Hemos visto, a lo largo de esta exposición, algunas frases y términos, que fuera de contexto nos harían parecer que no hablamos de documentación: macroestructura, microestructura, semántica, gramática, anclaje, base léxica, representación y todos los que de ellos dependen no citados en una comunicación con las pretensiones de ésta: sema, enunciado, archisemema, lexia, infraconceptos, estructura lógico- semántica, eje paradigmático y sintomático, acepción, lexicografía, terminología, proposiciones lógicas, etc. Todos ellos, combinados con el vocabulario más tradicionalmente documental: descripción bibliográfica, análisis y lenguajes documentales, búsqueda y recuperación, usuario, demanda, etc. constituyen el mapa conceptual de la lingüística documental, o genéricamente análisis documental para mis colegas del Departamento de

Biblioteconomía de la Universidad de Sao Paulo. Ahora bien, muchas teorías lingüísticas abandonadas o en desuso o plenamente vigentes pueden ser recortadas y aplicadas a nuestros fines siempre que haya indicios de utilidad: desde la documentalmente sobre- explotada semántica hasta las inexploradas sintaxis, lexicología y, de interés más reciente para los documentalistas por las máquinas captadoras y emisoras de fonemas que se nos avecinan, fonología.

Desde los postulados clásicos hasta los generativistas, y las derivaciones como la lógica semántica, el análisis del discurso, la semántica estructural o la semiótica textual por citar algunos campos en los que se han realizado incursiones o los intuimos prometedores, se constituyen los límites de esta vasta disciplina, en simbiosis con las que se ocupan de cómo se construye el raciocinio, su representación (ciencias cognitivas) y sus procesos de transferencia masiva mediante artefactos mecánicos (informática y telecomunicaciones). Sin la presencia sintética y simultánea de todo ese marco multidisciplinar en la mente del investigador de la documentación, generador de procedimientos e instrumentos útiles y pragmáticos, la organización y el acceso ordenado en los depósitos de conocimiento actuales nunca alcanzará mayores niveles de fiabilidad y satisfacción que en épocas pasadas.

## Notas

\* texto de la comunicación presentada en el V Simposio internacional sobre Comunicación social celebrado en Santiago de Cuba, del 21 a 25 de enero de 1997. Actualizado en junio de 1998.

1. Argumentos sostenidos en mis trabajos: Lingüística documental. Aplicación a la comunicación social. – Barcelona: Mitre, 1984. – 279 p.; Estructura lingüística de la documentación: teoría y método. – Murcia: Universidad de Murcia, 1990. – 166 p.; Análisis documental del discurso periodístico. – Madrid: CTD, 1992. – 160 p. y Procedimientos de análisis documental automático: estudio de caso. – Sevilla: Instituto Andaluz de Patrimonio Histórico, 1996. – 88 p.

2. Obra que no ha perdido vigencia, vid. Gardin, J.C.: Les analyses de discours. – Neuchâtel: Delachaux et Niestlé, 1974. – 178 p.

3. Véase ídem: Document Analysis and Linguistic Theory. – In: Journal of Documentation, v.29, 2, 1973. – p. 137-168 y Document Analysis and Information Retrieval. –In: Bol. Unesco bibliotecas, v.16, 1, 1960. – p.2-5

4. Idem: Systèmes experts et Sciences humaines. – Paris: Eyrolles, 1987. – 269 p.

5. Idem: La Logique du plausible. Essais d'Epistémologie pratique. – 2ème ed. –Paris: Maison des Sciences de l'Homme, 1987. – 330 p.

6. Le calcul et la raison:essais sur la formalisation du discours savant. – París: Ecole des Hautes Etudes en Sciences Sociales, 1991. – 293 p.

7. Lozano; Jorge: El discurso histórico. – Madrid: Alianza Editorial, 1987. – 223 p.

8. Vid. la clásica obra de Shannon, C y Weaver, W.: Teoría matemática de la comunicación. – Madrid: Forja, 1981. – 159 p.

9. El investigador quebequés de la Universidad Laval Jean de Bonville sienta, en su artículo, las bases de la adopción del modelo en documentación: Application du Paradigme du Shannon à la Bibliothéconomie et à la Documentation. – In: Revue canadienne des Sciences de l'Information". – v.3, mai 1978. – p.13-27

10. Zunde, P.: Information Theory and Information Science. – In: Information Processing and Management, 17, 6, 1981. – p.341

11. García Gutiérrez, A.: Lingüística documental... op. cit.

12. Consúltese la obra imprescindible de Otlet, Paul: Traité de Documentation. – Bruxelles: Mundaneum, 1934. -- p. 426a

13. Ibid., p. 34b

14. Wittgenstein apud Geckeler, Horst: Semántica estructural y teoría del campo léxico. – Madrid: Gredos, 1984. – 389 p.

15. Todos ellos se refieren, en sus respectivos ámbitos, a la participación del receptor en la construcción del significado, principio elemental de la documentación.

16. También Otlet destaca la figura del usuario como pieza clave en op. cit. p.33b

17. Por ejemplo, en la investigación realizada por mi colega de la Universidad de Sao Paulo Regina Obata: Contribução da Análise do Discurso para à análise documentária: o caso da documentação jornalística. – Sao Paulo: Escola de Comunicações e Artes de la USP, 1991. – 87 p. y anexos.

18. Véanse, al respecto, los trabajos de Bardin, L.: Análisis de contenido. – Madrid: Akal, 1986. –183 p. o Krippendorf, Klaus: Metodología de análisis de contenido: teoría y práctica. – Barcelona: Paidós, 1990. – 279 p.

19. Véase el excelente recorte teórico conceptual que realizan en su obra: Allwood, Jens; Lars, Gunnar y Dahl, Osten: Lógica para lingüistas. – Madrid: Paraninfo, 1981. – p. 183

20. Lenguaje construido desde distintas disciplinas para servir como herramienta de análisis de objetos (ánforas, indumentaria, armas, medallas, monedas o cualquier artefacto móvil), inmuebles (plantas de edificios, fachadas, cubiertas, cerchas, motivos ornamentales, túmulos, etc.) imágenes (audiovisuales y fotografías de objetos o realidades de interés patrimonial), textos (bibliografía sobre patrimonio) con el fin de ofrecer un instrumento central de referencia para los bienes históricos de Andalucía. Vid. Thesaurus de Patrimonio histórico andaluz. –Sevilla: IAPH, Consejería de Cultura, 1998 y su evaluación, junto a nuevas propuestas metodológicas en García Gutiérrez, A.: Principios de lenguaje epistemográfico: la representación del conocimiento sobre patrimonio histórico andaluz. – Sevilla: IAPH, 1998.

**FORMA DE CITAR ESTE TRABAJO EN BIBLIOGRAFÍAS:**

García Gutiérrez, Antonio (1998): Elementos de lingüística en sistemas de información y documentación. Revista Latina de Comunicación Social, 7. Recuperado el x de xxxx de 200x de:  
<http://www.ull.es/publicaciones/latina/a/66ant.htm>