

# LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

Josep AGUILÓ REGLA\*

## ABSTRACT

The aim of this paper is twofold. On the one hand, it tries to analyze and sistematize some of the concepts used when approaching the linguistic problems concerning the legal documentation automatic systems: legal language, documentary language, linguistic coordination (*a priori* and *a posteriori*) and thesaurus. On the other hand, it intends to carry out a detailed study of the legal thesauri, focusing basically on its two structural elements: the vocabulary and the paradigmatic relationships.

## 1. INTRODUCCIÓN

En este trabajo me propongo realizar una aproximación a la problemática lingüística que gira en torno a los sistemas automáticos de documentación jurídica. Para ello, en primer lugar, trataré de analizar con cierto detalle qué cabe entender por lenguaje jurídico en contextos documentales y de mostrar los distintos subconjuntos de términos que pueden distinguirse en función del papel y de la relevancia de los mismos a la hora de construir un lenguaje documental en el ámbito del Derecho. A continuación, procuraré precisar la noción de lenguaje documental y mostrar las distintas posibilidades de coordinación lingüísticas que ofrecen estos lenguajes. Todo ello, finalmente, desembocará en un estudio pormenorizado de los tesauros en el que, aparte de intentar aclarar su concepto y de mostrar las clasificaciones más importantes, me detendré en sus dos elementos estructurales: el vocabulario y las relaciones paradigmáticas.

Ahora bien, antes de abordar estos temas parece conveniente detenerse brevemente en la noción de sistema automático de documentación jurídica con el fin de delimitar de manera más precisa el ámbito de este trabajo. Por tales sistemas entiendo aquellos sistemas de información en los que los **datos** son el conjunto de los documentos jurídicos (legislación, jurisprudencia o doctrina) almacenados en sus ficheros, el **proceso** consiste en acciones de búsqueda (destinadas a localizar unidades documento-registros) y operaciones lógicas (destinadas a seleccionar exclusivamente los documentos que satisfagan las condiciones establecidas por el usuario en la ecuación de consulta), y la **información** es el conjunto de los documentos que el proceso consigue recuperar en cada búsqueda. Además, en esta caracterización conviene señalar que el usuario arquetípico de tales sistemas es el jurista que desea documentarse sobre un determinado caso con el fin de poder afrontar a continuación las tareas típicas de interpretación y aplicación; quiero decir con ello que este usuario es un demandante de "documentación automática" y no un demandante de "interpretación automática", lo que en mi opinión, sitúa a estos sistemas automáticos fuera de la órbita de los llamados "sistemas expertos".

## 2. EL LENGUAJE JURIDICO

Es sobradamente conocido - y creo que no controvertido - que el Derecho o, si se prefiere, el lenguaje jurídico forma parte de los lenguajes naturales, entendidos éstos - en oposición a los lenguajes artificiales - como sistemas sociales de comunicación históricamente dados y transmitidos de generación en generación como vehículo principal de interacción entre los individuos de una comunidad <sup>1</sup>. Ahora bien, no está de más preguntarse qué cabe entender por lenguaje jurídico. Desde un punto de vista lexicográfico - que es el que aquí nos va a interesar - creo que pueden darse dos respuestas. La primera supone considerar que el léxico del lenguaje jurídico viene constituido por el conjunto de términos técnicos, más o menos monosemánticos, privativos del Derecho (entendido éste como normas, como ciencia, como técnica, etc.). Esta primera respuesta esta en la línea de lo que ha sido la principal tarea intelectual en torno al Derecho; me refiero a la dogmática jurídica tradicional, preocupada especialmente por la defini-

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

ción, conceptualización, sistematización, etc. de las categorías e instituciones jurídicas. En el trasfondo de todo ello está la consideración de que dentro del léxico del lenguaje natural cabe distinguir, por un lado, un subconjunto de términos, que denominaremos **grupo fundamental**, compuesto por las palabras que designan conceptos para cuya comprensión y manejo no se requiere ninguna formación específica o profesional (en definitiva, que no se refieren a un campo específico de la actividad humana); y, por otro, **grupos específicos** que, en oposición al anterior, están compuestos por palabras que sí se refieren a campos específicos de la actividad humana<sup>2</sup>, como, por ejemplo, la medicina, el derecho, etc.

La segunda respuesta supone considerar que el léxico del lenguaje jurídico está constituido por todas las palabras contenidas en los documentos jurídicos. A los efectos de la documentación automática de legislación, jurisprudencia y doctrina esta respuesta resulta, sin duda, más pertinente, pues la problemática que dicha documentación plantea no puede abordarse simplemente a partir del **grupo específico**. A pesar de ello, aquí se tomarán en cuenta ambas respuestas, dado que resultará provechoso para nuestro análisis, si bien se utilizarán expresiones distintas para referirse a cada una de ellas. Podrán, pues, desde un punto de vista lexicográfico, distinguirse:

- el lenguaje natural (A): que abarca todo el patrimonio léxico de una determinada lengua;
- el lenguaje de los documentos jurídicos (B): que abarca todo el léxico incluido en los mismos;
- el lenguaje jurídico (C): que abarca el grupo de términos técnicos específicos del Derecho.

Esta claro que entre estos tres léxicos se da una relación de inclusión<sup>3</sup>

C ⊂ B ⊂ A

que, por otro lado, se impone por la propia definición de los conjuntos de términos implicados.

Por otro lado, el léxico del lenguaje de los documentos jurídicos contiene:

- términos que designan conceptos jurídicos, esto es, términos

que pertenecen al lenguaje jurídico entendido como léxico (grupo) específico o técnico;

- términos que designan un concepto jurídico, pero que, al mismo tiempo, tienen un significado común, general, exotérico y que, bajo este aspecto, entran dentro del grupo fundamental;

- términos que constituyen un tipo neutro, esto es, se utilizan en los documentos jurídicos, pero no designan ningún concepto jurídico. Evidentemente, estos términos pertenecen al grupo fundamental<sup>4</sup>.

A partir de todo esto, puede concluirse, en primer lugar, que el segundo tipo de términos contenidos en el léxico de los documentos jurídicos constituye la zona de penumbra entre los confines del grupo fundamental y del grupo técnico propio del Derecho. Y, en segundo lugar, que tomando como referencia todo el léxico del lenguaje natural, pueden formarse los siguientes subconjuntos:

I. El subconjunto compuesto por las palabras del grupo fundamental que no forman parte del léxico de los documentos jurídicos. Hay que advertir, como sabe cualquier persona que haya tenido alguna experiencia en el manejo de textos jurídicos que hay que ser muy cauteloso a la hora de pretender determinar **a priori** estos términos, dado que el Derecho tiende a regular todas las esferas de la vida. Sin embargo, por ejemplo, es poco probable que nos encontremos con palabras tales como "reír", "llorar", "cielo", "estrella", etc.

II. En subconjunto formado por las palabras incluidas en los documentos jurídicos, pero que pertenecen al grupo fundamental. Dentro de este subconjunto, a su vez, pueden distinguirse otros dos:

IIA. El que comprende las palabras que antes se han denominado neutras, como por ejemplo "circunstancia", "todo", "resultar", "análisis", "cuerpo", etc.

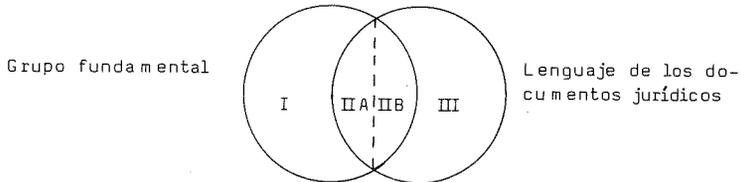
IIB. El compuesto por las palabras que tienen un significado jurídico, pero que pertenecen al grupo fundamental, como por ejemplo "divorcio", "matrimonio", "parlamento", "contrato", etc.

III. El subconjunto formado por las palabras que designan exclusivamente conceptos jurídicos y que, por tanto, no forman parte

## LENGUAJE JURIDICO, LENGUAJE DOCUMENTAL Y TESAURO

del grupo fundamental, como "antfresis", "enfiteusis", "litispendencia", "guarentigio", "falcidia", etc.

Aún a sabiendas de que las fronteras no son siempre claras y que tampoco vienen dadas **a priori**, la división anterior puede representarse del siguiente modo<sup>5</sup>:



A partir de esta división<sup>6</sup>, conviene decir algo a propósito de los conocidos problemas de los lenguajes naturales (ambigüedad y vaguedad) en relación con el lenguaje de los documentos jurídicos. Respecto del sector III, compuesto por términos que designan conceptos exclusivamente jurídicos, hay que decir que, por lo general, el fenómeno de la ambigüedad tendrá una presencia mucho más reducida que en los términos que pertenecen al grupo fundamental. Ello se debe a la pretensión de cualquier grupo específico de alcanzar el ideal de la univocidad. Por tanto, puede afirmarse que la ambigüedad no va a suponer un problema importante para este sector del lenguaje de los documentos jurídicos. No puede decirse lo mismo, sin embargo, respecto de la vaguedad de los conceptos designados por los términos contenidos en este sector. En efecto, no hay que olvidar que, por ejemplo, las eternas discusiones de los juristas a propósito de las naturalezas jurídicas pueden verse como discusiones en torno a la intensión (con sus correspondientes consecuencias extensionales) de los conceptos jurídicos. Y ello conlleva ciertos problemas a la hora de establecer, atendiendo a criterios semánticos, las relaciones paradigmáticas entre palabras de este género.

En relación con el sector IIA, podría pensarse - erróneamente - que no es relevante a la hora de crear un lenguaje documental en el ámbito del Derecho, dado que recoge los términos que se han definido como neutros. Sin embargo, razonar así supone, en cierto sentido, incurrir en una petición de principio - puesto que el problema que surge en torno a estos términos es básicamente el de su determinación - y sortear lo que, en mi opinión, es la cuestión central: la consideración

o no de tipo neutro está en relación directa con la finalidad específica que se asigne a cada lenguaje documental. Piénsese, por ejemplo, en palabras tales como "antibiótico", "tabaco", "automóvil", "hortaliza", "ordenador", etc.; no cabe duda de que ninguna de ellas designa un concepto jurídico, pero todas ellas designan entidades que pueden ser (y de hecho son) objeto de regulación jurídica. ¿Esta circunstancia hace que no pertenezcan al grupo de palabras consideradas neutras? Y, sean o no neutras, ¿son relevantes a la hora de determinar el vocabulario de un lenguaje documental en el ámbito del Derecho? En mi opinión, la respuesta a estas cuestiones sólo puede alcanzarse atendiendo a la finalidad particular que se pretende conseguir con cada lenguaje documental y al tipo de documentación implicada en el mismo. Al margen de lo anterior, no debe olvidarse de que este sector se extraerá el núcleo principal de las llamadas palabras vacías.

Pero el sector más problemático, y con mucho, es el IIB, que está compuesto por los términos que expresan un concepto jurídico y que además tienen un significado común. Ello se debe a que, desde la perspectiva documental, es aquí donde se concentran las principales dificultades en términos de ruido y de silencio.

Atendiendo a la doble naturaleza y a los significados de estos términos, puede decirse que caben dos posibilidades: A) Que exista concordancia entre el significado jurídico y el significado común, de forma que éstos coincidan o sean muy similares, como por ejemplo "matrimonio", "contrato", "arrendamiento", "padre", etc. Esta situación no resultará documentalmente problemática. B) Que exista divergencia, esto es, que los significados ni coincidan ni sean similares, como por ejemplo "auto", "causa", "contencioso", "capacidad", "competencia", etc.<sup>7</sup>. Estas ambigüedades, como es sabido, obedecen a los fenómenos de la homonimia y de la polisemia. Ahora bien, sería un error pensar que estos términos dejan de ser ambiguos en la medida en que se atienda exclusivamente a sus significados jurídicos (es decir, a sus significados posibles en contextos jurídicos); por el contrario, es muy común que incorporen varios significados jurídicos.

La cuestión se complica algo más si tenemos en cuenta que las dos alternativas de relación entre los significados de estos términos no son excluyentes sino que pueden concurrir conjuntamente. Veámoslo

## LENGUAJE JURIDICO, LENGUAJE DOCUMENTAL Y TESAURO

recurriendo al ejemplo del "socorrido" término "acción". El término "acción" en su sentido común significa algo así como el ejercicio de una potencia o el efecto de hacer algo. En el sentido jurídico puede designar al menos las siguientes nociones: 1. derecho que se tiene de una cosa; 2. facultad de acudir a los tribunales y recabar una decisión; 3. acto humano voluntario; 4. porción en que se divide el capital de una sociedad. A la vista de ello, no es difícil comprobar que cabe hablar de concordancia del sentido común del término con el significado jurídico 3.; y de divergencia en los otros casos.

### 3. EL LENGUAJE DOCUMENTAL

Responder a la cuestión de qué cabe entender por lenguaje documental no resulta del todo sencillo, dado que, por un lado, muchos autores prescinden totalmente de esta expresión; y, por otro, los que la utilizan tienden a identificarla con la de "tesauro" operando - en mi opinión - una cierta reducción, pues entre ambas puede definirse una relación de género a especie. La única forma útil - creo - de usar esta expresión es la de considerar que lenguaje documental es todo vocabulario susceptible de coordinación lingüística y cuyas unidades son palabras clave (palabras que sirven de clave de acceso a registros en lenguaje natural contenidos en un fichero directo). Así pues, para caracterizar a estos lenguajes resulta irrelevante tanto la cuestión de la técnica utilizada para seleccionar las palabras clave (indexación manual o indexación automática) como la de en qué momento se realice la coordinación lingüística (coordinación **a priori** - antes de la ecuación de consulta - y coordinación **a posteriori** - en la ecuación de consulta - ). Los autores que participan de la noción restrictiva de lenguaje documental (en ocasiones llamado también lenguaje descriptor) consideran que éste es un lenguaje artificial compuesto por palabras cuyo significado ha sido normalizado **a priori**<sup>a</sup>, lo que lleva necesariamente a identificarlo con tesauro. En mi opinión, el origen de esta identificación radica en la consideración de que por "palabras clave" sólo cabe entender aquellas que se introducen en el fichero inverso como resultado de una indexación manual (pues en otro caso no se entiende cómo podría normalizarse **a priori** su significado). En el fondo, lo que subyace a este planteamiento es la presuposición de que los sistemas de indexación

automática difícilmente pueden dar respuesta a la problemática lingüística que presenta la documentación de textos en lenguaje natural.

Para mostrar la inconveniencia de la reducción de lenguaje documental a tesoro, nada mejor que ver las extraordinarias posibilidades de coordinación lingüística que ofrece un sistema de documentación que ha sido indexado automáticamente y que dispone de los elementos estándar para formar la ecuación de búsqueda. El usuario que conozca bien las posibilidades que le ofrezca el sistema podrá realizar, entre otras, las siguientes coordinaciones lingüísticas:

- a) establecer las relaciones de sinonimia (verdadera o falsa) que estime oportuno utilizando el operador lógico "Y/O";
- b) romper la polisemia y la homonimia de determinadas palabras asociándoles, mediante el operador lógico "Y", otras que especifiquen el sentido en que se desean utilizar;
- c) reproducir los sintagmas contenidos en los textos recurriendo a los operadores sintácticos (con todas sus variantes);
- d) evitar sintagmas indeseados contenidos en los textos mediante el operador "SALVO";
- e) ignorar, mediante las funciones de enmascaramiento que permiten recuperar a partir de un lexema o morfema léxico, prefijos y/o sufijos;
- f) y, por último, si el sistema dispone de un generador gramatical o lematizador, recuperar a partir de un determinado lema todas las formas flexivas del mismo contenidas en los documentos<sup>9</sup>.

Desde la perspectiva que aquí nos interesa, la única diferencia radica en que los sistemas de indexación automática remiten al momento de la interrogación toda la problemática de establecer las relaciones de coordinación lingüística, mientras que los de indexación manual resuelven **en parte** dicha coordinación **a priori**, bien antes de la indexación, bien después de ella.

#### 4. EL TESAURO

##### 4.1. CONCEPTO

La palabra "tesoro"<sup>10</sup> se ha utilizado en ocasiones como título para léxicos, diccionarios exhaustivos, catálogos, etc.<sup>11</sup>. Sin embargo,

## LINGÜAJE JURÍDICO, LINGÜAJE DOCUMENTAL Y TESAURO

en los contextos documentales en que nos movemos suele utilizarse en un sentido más específico que el de simple elenco de palabras que puede abarcar desde un diccionario a un índice de materias. En efecto, en la literatura relativa a estas cuestiones pueden encontrarse infinidad de definiciones de "tesauro", todas ellas más específicas que la anterior noción. Atendiendo que unas tienden a resaltar el aspecto estructural, otras el funcional y otras a combinar ambos aspectos, pueden mostrarse, por ejemplo, las siguientes definiciones:

### a) Definiciones estructurales:

- *"Es un vocabulario controlado y dinámico de términos, relacionados semántica o genéricamente, que abarca totalmente un campo específico de conocimiento"*<sup>12</sup>.
- *"Conjunto estructurado de descriptores entre los que están reflejadas sus relaciones semánticas"*<sup>13</sup>.
- *"Es un lenguaje documental, esto es, un lenguaje artificial elaborado sobre la base del lenguaje natural, compuesto por palabras clave con un significado normalizado"*<sup>14</sup>.
- *Es "un léxico jerarquizado que comprende una red de interconexiones, de exclusiones, de discriminaciones y de proximidades semánticas bajo la forma de listas de sustitutos, de contrarios, de términos cercanos o genéricos, etc."*<sup>15</sup>.

### b) Definiciones funcionales:

- *"Es un instrumento de control de la terminología utilizado para trasladar en un lenguaje más estricto (lenguaje documental o lenguaje de indización) el lenguaje natural utilizado en los documentos"*<sup>16</sup>.

### c) Definiciones estructural-funcionales:

- *Es "una organización estructurada de palabras clave que en un sistema automatizado de documentación tiene la doble misión de identificar los documentos en el momento de su registro y de recuperación en la fase de búsqueda"*<sup>17</sup>.
- *"Es un conjunto ordenado de descriptores que constituye un sistema abierto destinado a la clasificación y organización de términos por materias o problemas o por cualquier otro criterio"*<sup>18</sup>.
- *"Cuando se usan las palabras clave, se genera un thesaurus, reuniendo las palabras clave de cada documento memorizado en una única lista*

## Josep AGUILÓ REGLA

*alfabética. El documentalista puede intervenir sobre esta lista de palabras clave, creando ligámenes entre los sinónimos o distinguiendo los homónimos y los homógrafos, para evitar cruces entre conjuntos de documentos heterogéneos*<sup>19</sup>.

Pues bien, a partir de este catálogo de definiciones, todas ellas aceptables en términos generales, intentaré aproximarme a la intensión del concepto "tesauro" a partir de lo que podría llamarse el denominador común a todas ellas o, si se prefiere, indicando las condiciones mínimas necesarias recogidas en las mismas.

Desde una perspectiva estructural, puede decirse que todas las definiciones concuerdan en las dos siguientes notas: a) un tesauro está compuesto por un vocabulario (un léxico, elenco de palabras clave, descriptores, ...) y b) entre las unidades de ese vocabulario se establecen ciertas relaciones. En relación con esta segunda nota, hay que decir que la definición de Losano parece dar a entender que el establecimiento de las relaciones entre las unidades del vocabulario, lejos de ser un elemento estructural del tesauro, es más bien opcional, contingente, etc. Ello se debe, en mi opinión, a necesidades impuestas por la terminología utilizada por este autor, pues al definir al tesauro o tesauro positivo en oposición al tesauro negativo, se ve forzado a incluir en el ámbito de aquella expresión a cualquier fichero inverso. Pero si se atiende al uso común que los documentalistas y los juristas informáticos hacen del término "tesauro", entonces hay que reconocer que la nota de integrar una estructura, un sistema, etc. es esencial para su definición.

Desde una perspectiva funcional debe resaltarse, en primer lugar, que las que he llamado "definiciones estructurales" incluyen naturalmente, aunque sea implícitamente, algunas notas de tipo funcional (las referencias a palabras clave, descriptores, lenguaje documental, etc. expresaban ya la función que el tesauro podía desempeñar). En cualquier caso, me parece claro que el tesauro es un instrumento lingüístico para ser utilizado en el ámbito de los sistemas automáticos de documentación. Algunas de las definiciones antes transcritas, sin embargo, operan una reducción del concepto de "tesauro" al establecer que una de sus funciones necesarias es la de ser un instrumento útil para la indexación

## LENGUAJE JURIDICO, LENGUAJE DOCUMENTAL Y TESAURO

de documentos; esta reducción, en mi opinión, resulta injustificada, pues supone considerar que sólo los tesauros **a priori** (ya se verá más adelante en qué consisten éstos) son "verdaderos" tesauros.

De lo anterior puede concluirse que el tesoro es un instrumento lingüístico, utilizado en el ámbito de la recuperación automática de documentación, compuesto por un vocabulario entre cuyas unidades se han definido ciertas relaciones de coordinación lingüística. Esta definición todavía puede especificarse más de la manera siguiente: A) La unidad léxica de ese vocabulario es la palabra clave (entendida en sentido amplio y que, por tanto, incluye palabras y expresiones compuestas - sintagmas -). B) El significado de las palabras clave no tiene por qué coincidir con el significado natural, pues aquel vendrá determinado no sólo por éste, sino también por las relaciones que se definan entre esa palabra clave y las restantes. C) La coordinación lingüística que supone el tesoro requiere siempre la intervención de los analistas o documentalistas responsables del sistema de documentación, dado que el ordenador no puede realizar operaciones a nivel semántico. Y D) las relaciones de coordinación lingüística entre palabras clave expresadas por el usuario bajo la forma de condiciones establecidas en la ecuación de consulta, no forman parte del tesoro; de ahí que anteriormente, cuando me refería a los lenguajes documentales, estableciera la distinción entre coordinación **a priori** y **a posteriori** (lo cual no debe llevar a pensar que la existencia de un tesoro elimina la coordinación **a posteriori**).

### 4.2. CLASES DE TESAUROS

Existen, evidentemente, infinidad de criterios para clasificar los tesauros; pero, evidentemente también, no todos ellos son relevantes en la misma medida. Aquí voy a centrarme sólo en el criterio de la forma de utilización y en el del momento de su construcción.

Por la forma de utilización puede distinguirse entre tesauros manuales y tesauros automáticos. Los primeros son los que están impresos y no forman parte de la base de datos (ni como instrumento de búsqueda ni como elemento auxiliar o informativo). En consecuencia, el usuario deberá consultarlos manualmente antes de la interrogación para conseguir diseñar una adecuada estrategia de consulta. En oposición a éstos, los tesauros automáticos están incorporados al sistema documental, bien

para ser utilizados como instrumento informativo que puede ser invocado por el usuario para guiarse en la interrogación, bien como instrumento de búsqueda que le permite extender el campo cubierto por la ecuación de consulta a todas las palabras relacionadas con las incluidas en dicha ecuación.

Por otro lado, atendiendo al criterio del momento de su construcción cabe distinguir entre tesauros **a priori** y tesauros **a posteriori**. Las expresiones **a priori** y **a posteriori** van referidas aquí a un momento distinto de aquél que marcaba el punto de inflexión entre las dos posibilidades de coordinación lingüística. En efecto, si allí el punto de inflexión se situaba en el momento de la interrogación, aquí se fija en el momento de la indexación de los documentos. Así, un tesoro **a priori** es el que se ha construído para representar todo el campo semántico de una determinada rama del saber y que se utiliza no sólo para recuperar información, sino también para indexar los documentos; lo que implica que su construcción es anterior a la elaboración de la base de datos y que, en la fase de indexación, los documentalistas deben elegir las palabras clave que identifiquen al documento de entre las contenidas en el tesoro. Este tipo de tesauros suelen caracterizarse como sistemas cerrados de palabras clave, aunque, por su propia naturaleza instrumental, evidentemente deben estar abiertos a posibles modificaciones o actualizaciones. Por el contrario, suele decirse que son **a posteriori** los tesauros que se construyen sobre ficheros inversos ya existentes con independencia de que éstos sean el resultado de la indexación manual o automática de los documentos.

Si atendemos a las diferentes clasificaciones que se han introducido hasta ahora, podemos formar los tres siguientes grupos de sistemas de documentación en función del tipo de lenguaje documental:

- a) Sistemas en los que la indexación de los documentos se realiza automáticamente y en los que toda coordinación lingüística es **a posteriori**; por tanto se trata de sistemas que no disponen de ningún tipo de tesoro.
- b) Sistemas que disponen de un tesoro **a priori** (lo que implica una coordinación lingüística **a priori**) que determina la subsiguiente actividad manual de indexación y que, además, en el momento de la interrogación permiten al usuario realizar una coordinación



una vez creado, de una u otra forma, éste interacciona sobre todas las bases del proceso documental. Así, por ejemplo, en las actualizaciones de un determinado sistema que dispone de un tesoro **a posteriori**, éste determinará la actividad de indexación de las nuevas unidades documento, llegando a comportarse para estas unidades como un tesoro **a priori**; del mismo modo, un tesoro **a priori** aunque sea anterior a la indexación se verá siempre en algún grado afectado por ésta dado el carácter dinámico del proceso documental.

Pero si lo anterior reduce la aparente distancia entre estos dos tipos de tesoros, hay que decir que si se consideran algunas consecuencias que afectan al vocabulario de los mismos, las distancias vuelven a agrandarse. No es difícil observar que el vocabulario de un tesoro **a priori** - al ser el resultado de una reconstrucción conceptual de una determinada rama del saber al margen del vocabulario concreto de los documentos - en el campo del Derecho tenderá a identificarse con el grupo específico de términos jurídicos; aunque, naturalmente, puede incluir también a los términos que designan un concepto jurídico y que simultáneamente tienen un sentido común. Pero lo que resulta casi imposible es que pueda abarcar a los términos que se han denominado neutros, porque para que el tesoro pueda aprehenderlos es necesario que tome contacto con el vocabulario de los documentos jurídicos. Este hecho es el que permite trazar una semejanza entre estos tesoros y lo que tradicionalmente se han llamado "esquemas de clasificación". Por el contrario, los tesoros **a posteriori** se caracterizan por mantener una relación fuerte entre su vocabulario y el de los documentos que tratan de representar. Estas diferencias son las que llevan generalmente a afirmar que los tesoros **a priori** son especialmente aptos para usarse en la documentación de la doctrina jurídica, dado que la distancia entre el vocabulario del tesoro y el de los documentos no resulta empobrecedora sino que se presenta, por decirlo de algún modo, como necesaria<sup>20</sup>. Ello no ocurre ni con la legislación ni con la jurisprudencia, por lo que resulta inadecuado desaprovechar las excelentes posibilidades documentales que ofrece el lenguaje de los propios documentos (o de los sustitutivos de los mismos).

## 5. EL VOCABULARIO DEL TESAURO

### 5.1. LA SELECCION DE LAS PALABRAS CLAVE

Como ya se ha dicho, una de las características de los tesauros **a posteriori** consiste en que toman como base las palabras contenidas en un fichero inverso que, evidentemente, ha sido el resultado de la indexación de los documentos. En este sentido, se puede decir - siguiendo a V. Knapp - que en estos casos *"la indexación de los documentos jurídicos se realiza dos veces: la primera vez con la intención de coleccionar las expresiones necesarias para el establecimiento del tesoro, la segunda con la finalidad de traducir los documentos jurídicos a un lenguaje documental"*<sup>21</sup>. En efecto, la primera indexación no se limita a ser una simple colección de todas las palabras contenidas en los documentos, puesto que de ella se excluyen las palabras vacías en sentido estricto, esto es, las que carecen de significado autónomo y que son determinables **a priori** (como, por ejemplo, los artículos, los verbos auxiliares, las preposiciones, etc.). La consideración de vacías de ciertas palabras les cierra la puerta para llegar a formar parte el tesoro.

Una vez que se dispone de esta colección de palabras significativas se realiza la segunda indexación que consiste en seleccionar las palabras que pasarán a formar parte el tesoro. Ahora bien, en mi opinión, resulta casi imposible establecer en abstracto las reglas generales que deben regir esta segunda indexación. Al igual que ocurría cuando nos preguntábamos si las palabras neutras del lenguaje de los documentos jurídicos debían formar parte del vocabulario documental o no, aquí la cuestión de las reglas debe resolverse también atendiendo a la finalidad particular que se asigne al sistema y al material documental de que se trate. No obstante ello, puede decirse que es común considerar que las unidades del vocabulario del tesoro deben ser selectivas<sup>22</sup>, **mono-semánticas y no sinónimas**.

### 5.2. LAS EXPRESIONES COMPUESTAS (SINTAGMAS)

Se ha dicho ya que en el ámbito del tesoro las palabras clave pueden ser palabras individuales o expresiones compuestas. La función de las expresiones compuestas en el tesoro no es otra que la de representar las relaciones sintagmáticas<sup>23</sup> que se dan en los documentos jurídicos. Ahora bien, no hay que olvidar que, dado el carácter lineal

del discurso, todas las palabras contenidas en un documento entablan relaciones sintagmáticas tanto con las que las preceden como con las que las suceden. Y, ciertamente, resulta obvio que no todos estos sintagmas pueden incorporarse al tesoro, pues éste adquiriría unas dimensiones intolerables. Por esta razón, la exposición sobre las expresiones compuestas debe centrarse en torno a los criterios que permiten resolver cuándo un determinado sintagma debe ser incorporado como palabra clave al tesoro<sup>24</sup>. Así, y teniendo en cuenta la dificultad que implica establecer en abstracto y a *priori* las reglas que deben regir estas cuestiones, pueden citarse los siguientes casos en que deberían incluirse:

- a) Cuando se trata de locuciones, sintagmas o frases hechas que resulta imposible reconstruir en una coordinación a *posteriori*. Piénsese, por ejemplo, en "*non bis in idem*", "*prior in tempore, potior in iure*", etc.
- b) Cuando la separación de dos palabras haga perder el verdadero sentido del sintagma y su mantenimiento permita eliminar polise-mias. Por ejemplo, "prisión menor" (es distinto de "prisión" y de "menor"), "local de negocio", etc.
- c) Cuando de no construirse el sintagma, alguna de las palabras que lo componen difícilmente podría utilizarse en algún contexto distinto al propio sintagma. Por ejemplo, "lucro cesante", "letra de cambio", "vicios ocultos", etc.

### 5.3. REGLAS GRAMATICALES

El vocabulario del tesoro debe seguir las dos siguientes reglas gramaticales:

- a) Las palabras individuales deben ser lemas, por tanto las únicas combinaciones admitidas entre morfemas léxicos (o lexemas) y morfemas gramaticales serán las determinadas por el singular para los sustantivos, el singular para los adjetivos que no tienen marcas formales para el género, el masculino singular para los adjetivos que sí tienen dichas marcas formales y el infinitivo para los verbos.

Esta regla general presenta, sin embargo, dos excepciones relacionadas con el singular. En efecto, el uso del plural puede admitirse en los dos siguientes casos: 1. Cuando en el uso común del lenguaje jurídico una determinada palabra se utilice siempre bajo la forma del plural, como por ejemplo, "aguas", "alimentos", "arras", "expensas",

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

"frutos", etc. 2. Cuando la introducción del plural, o si se prefiere, la distinción entre el plural y el singular permite sortear el peligro del ruido, dado que el plural adquiere un significado diferente (no sólo por el número) del singular como, por ejemplo: "padre" (ascendiente masculino de primer grado) y "padres" (ascendientes masculino y femenino de primer grado); "auto" (resolución judicial motivada) y "autos" (conjunto de las diferentes piezas que componen una causa civil o criminal); "capitulación" (convenio militar o político en el cual se estipula la entrega o rendición de una plaza, ejército o lugar fortificado) y "capitulaciones" (contrato matrimonial por el cual se establecen las futuras condiciones del régimen patrimonial de la sociedad conyugal); "diligencia" (cuidado o celo en el desempeño de una función o en el cumplimiento de una obligación) y "diligencias" (ciertas actuaciones judiciales con el concurso o no de las partes).

b) La segunda regla consiste en que las expresiones compuestas, en la medida en que pretenden reproducir algunos sintagmas contenidos en los textos, deben asumir la forma habitual en que éstos aparecen en los mismos y en que son utilizados en el uso común del lenguaje jurídico. Así, por ejemplo, "vicios ocultos" y "abusos deshonestos" se incorporarán al tesauro en la forma del plural; sin embargo, otras expresiones como "local de negocio" o "locales de negocio", que pueden aparecer en los textos tanto en singular como en plural, adoptarán en el tesauro la forma del singular. Asimismo, las expresiones compuestas que se incorporen al tesauro deberán respetar el orden de las palabras asignado en el uso común del lenguaje jurídico.

### 6. LAS RELACIONES PARADIGMATICAS

Al estudiar el concepto de tesauro se ha visto que existía un amplio consenso en considerar que éste consistía en un vocabulario entre cuyas unidades (palabras clave) se definían ciertas relaciones. Asimismo, es sabido que en lingüística general suelen considerarse básicamente dos tipos de relaciones: las sintagmáticas y las paradigmáticas. Pues bien, la relevancia de las relaciones sintagmáticas para los sistemas automáticos de documentación ha sido ya apuntada, así como las dos formas de afrontarlas desde el sistema: bien incorporando el resultado de las mismas (sintagmas) al vocabulario documental (expresiones

compuestas) a través de una coordinación lingüística **a priori**; bien representando dichas relaciones en la ecuación de consulta a través de una coordinación lingüística **a posteriori**. Ahora conviene detenerse, pues, en las relaciones paradigmáticas<sup>25</sup>, en aquellas que para Saussure "*forman parte del tesoro interior que constituye la lengua de cada individuo*"<sup>26</sup> y que para nosotros - **mutatis mutandi** - integran el tesoro (tesauro) interior de algunos sistemas de documentación.

Ahora bien, no es difícil observar que bajo el rótulo de relación paradigmática se incluyen asociaciones de muy diversa naturaleza entre palabras. Conviene, pues, aclarar que sólo nos vamos a detener en aquéllas que son relevantes en el ámbito del tesoro y que, por tanto, la exposición se centrará en primer lugar en las relaciones que constituyen el denominador común a un conjunto de tesauros concretos, para a continuación proceder a la clasificación y definición estricta de cada una de las referidas relaciones.

Si se comparan entre sí un número suficiente de tesauros, no es difícil comprobar que, dada la recurrencia de ciertas relaciones, existe un amplio consenso en considerar a algunas de ellas como fundamentales (el consenso se extiende también a la notación de las mismas). Así, en esta primera aproximación, hay que considerar las relaciones de sinonimia, jerarquía y asociación, aunque, como se verá, con ellas no se agota el catálogo de las mismas.

a) Sinonimia. Dejando para más adelante la definición precisa de sinonimia, hay que decir que para la implementación de esta relación es muy común recurrir a la noción de equivalente, esto es, una palabra que se define como sinónima de una palabra clave, pero cuyo uso no está autorizado en el ámbito del tesoro. Se generan así los dos siguientes tipos de relaciones:

- 1) **US = Use**; **A US B**: el término A no es palabra clave sino equivalente y en su lugar debe utilizarse la palabra clave B. Esta relación suele ser considerada como asimétrica, irreflexiva e intransitiva.
- 2) **UF = Used for**; **B UF A**: B es palabra clave y se utiliza en lugar de su equivalente A.

No es difícil observar que:  $A \text{ US } B \leftrightarrow B \text{ UF } A$

No siempre la relación de sinonimia se implementa de esta manera, pues puede definirse también entre dos palabras clave subsistiendo ambas

como tales<sup>27</sup>.

b) Jerarquía. Dentro de lo que aquí se llaman relaciones de jerarquía se incluyen especies bastante diferentes que sólo se verán con precisión más adelante. Baste aquí con indicar que las jerarquías suelen implementarse mediante las dos siguientes relaciones:

1) **BT = Broader Term**; A BT B: la palabra clave A tiene como superior jerárquico a la palabra clave B. Las propiedades de esta relación están en función del uso que de la misma se haga en cada tesauro, pues, por ejemplo, no es lo mismo utilizarla para representar una relación de género a especie que una relación de todo-parte.

2) **NT = Narrower Term**; B NT A: la palabra clave B tiene como inferior jerárquico a la palabra clave A. Así, es claro que

$$A \text{ BT } B \leftrightarrow B \text{ NT } A$$

c) Asociación. A propósito de este género de relaciones hay que decir que a su amparo se representan especies muy diversas. De momento bastará con fijar su notación y la nota que permite diferenciarlo de las anteriores:

**RT = Related Term**; A RT B: la palabra clave A está en una relación de asociación no jerárquica con la palabra clave B. Evidentemente, se trata de una relación simétrica.

Hasta aquí se ha mostrado lo que es común a la práctica totalidad de los tesauros. Ahora conviene entrar a analizar con más detalle las relaciones que pueden representarse en un tesauro y para ello, en primer lugar, voy a proponer una clasificación de las mismas para, a continuación, procurar definir de manera estricta cada una de ellas.

### 6.1. CLASIFICACION DE LAS RELACIONES PARADIGMATICAS

Como se sabe, las palabras clave son signos lingüísticos y, como tales, están compuestos por un significante (forma gráfica) y un significado (concepto designado). Por ello, podemos clasificar las relaciones paradigmáticas entre palabras clave según que la relación sea el resultado de asociar dichas palabras bien por sus significantes, bien por sus significados. A las primeras las denominaremos relaciones morfológicas. A las segundas, relaciones semánticas. Asimismo, dentro de las relaciones

semánticas podemos distinguir entre relaciones basadas en estructuras lógicas y relaciones basadas en estructuras de lo real. Soy consciente de que la introducción de esta distinción puede resultar problemática y de que en contextos exclusivamente lingüísticos tal vez fuera más apropiado hablar sólomente de relaciones de sentido. Pero en el ámbito del tesoro me parece que su presencia se justifica en términos de utilidad. Con ella se trata de mostrar la diferencia que existe entre relacionar palabras por considerar exclusivamente relaciones intensionales y extensionales entre los conceptos implicados, y relacionarlas por considerar relevantes ciertas relaciones que se dan en el mundo externo entre las entidades referenciadas por tales palabras. Así pues, las primeras responden básicamente a cómo está estructurado el lenguaje sin necesidad de tomar en consideración al mundo externo; y las segundas responden básicamente a cómo es el mundo y no a cómo está estructurado el lenguaje.

A partir de todo ello, las relaciones paradigmáticas pueden clasificarse bajo la forma de un índice de jerarquías:

Relaciones paradigmáticas.

. Relaciones semánticas.

.. Relaciones basadas en estructuras lógicas.

... Sinonimia o equivalencia semántica.

... Jerarquía.

.... Hiponimia.

.... Cohiponimia.

... Contraste.

.... Incompatibilidad.

.... Oposición.

..... Complementariedad.

..... Antonimia.

..... Inversión.

.. Relaciones basadas en estructura de lo real.

... Relaciones partitivas.

... Relaciones (verticales) todo-parte o parte-todo.

... Relaciones (horizontales) parte-parte.

... Relaciones proceso-agente.

. Relaciones morfológicas.

- .. Homografía.
- .. Equivalencia sintáctica.
- .. Comparación.

## 6.2. RELACIONES BASADAS EN ESTRUCTURAS LÓGICAS

### A) SINONIMIA

La relación de sinonimia se utiliza para designar la identidad de significados entre las unidades del vocabulario documental. Ahora bien, ¿cuándo son sinónimas dos palabras?. La respuesta en términos abstractos no es difícil, aunque sí lo es resolver los problemas concretos que se suscitan cuando las definiciones de sinonimia se proyectan sobre casos particulares. En primer lugar, se dirá que dos términos son sinónimos cuando los conceptos por ellos designados tengan la misma extensión<sup>28</sup> o también la misma intensión<sup>29</sup>. A partir de la intensión puede definirse la sinonimia considerando la conjunción de dos implicaciones<sup>30</sup> recíprocas. Así, si **a** y **b** son dos términos, se dirá que son sinónimos si y sólo si **a** implica **b** y **b** implica a **a**:

(Sin. a, b)  $\leftrightarrow$  (a  $\leftrightarrow$  b)  $\leftrightarrow$  (a  $\rightarrow$  b & b  $\rightarrow$  a)

Adoptando una perspectiva extensional, podrá definirse por relaciones entre clases. Consideremos que **x** e **y** son dos términos y que **A** y **B** son las clases o extensiones de **x** e **y**, se dirá que **x** e **y** son sinónimos si y sólo si **A** está contenida en **B** y **B** está contenida en **A**, esto es, si y sólo si **A** = **B**<sup>31</sup>:

$$(A = B) \leftrightarrow [(A \subset B) \& (B \subset A)]$$

Por otra parte, en lingüística es muy común, para definir la sinonimia, recurrir a la noción de coimplicación o equivalencia entre oraciones. Así, si tenemos dos oraciones **O**<sub>1</sub> y **O**<sub>2</sub> que sólo se diferencian por el hecho de que donde **O**<sub>1</sub> tiene una unidad **x** **O**<sub>2</sub> tiene una unidad **y** (esto es, tienen la misma estructura) y si **O**<sub>1</sub> implica a **O**<sub>2</sub> y **O**<sub>2</sub> implica a **O**<sub>1</sub> (esto es, **O**<sub>1</sub>  $\leftrightarrow$  **O**<sub>2</sub>), podemos decir que **x** e **y** son sinónimos<sup>32</sup>.

En cualquier caso, en la construcción de un tesauro es muy común distinguir entre dos tipos de sinonimia: la sinonimia absoluta, perfecta, verdadera o total, y la sinonimia relativa, imperfecta, falsa, parcial o cuasisinonimia<sup>33</sup>. Las dos definiciones anteriores, en términos inten-

sionales y extensionales, lo son de sinonimia total. En cuanto a la definición lingüística, se dirá que los términos  $x$  e  $y$  son sinónimos totales si y sólo si en todas las oraciones en las que aparezca  $x$ ,  $x$  puede ser sustituido por  $y$ , y viceversa. Por el contrario, se dirá que son sinónimos parciales cuando sólo en algunas de las oraciones en las que aparezca  $x$ ,  $x$  puede ser sustituido por  $y$ <sup>34</sup>.

La sinonimia parcial puede definirse también en términos extensionales, a partir de la intersección entre clases. Así, si  $x$  e  $y$  son términos y  $A$  y  $B$  sus clases o extensiones, se dirá que  $x$  e  $y$  son sinónimos parciales si se verifican conjuntamente las siguientes condiciones: a)  $A \neq B$ ; b)  $A \not\subset B$ ; c)  $B \not\subset A$ ; d)  $A \cap B \neq \emptyset$ <sup>35</sup>.

A la vista de lo anterior, puede decirse que conceptualmente la sinonimia total es una relación en la que se verifican las propiedades de reflexividad, simetría y transitividad; mientras que la sinonimia parcial es una relación reflexiva, simétrica y no transitiva<sup>36</sup>.

Por último, hay que resaltar que en la implementación documental de estas relaciones es muy frecuente no respetar - por razones documentales y de oportunidad - las propiedades que conceptualmente se han atribuido a cada una de ellas. Respecto de la sinonimia total basta comparar lo dicho aquí con lo dicho más arriba a propósito de las relaciones **US (Use)** y **UF (Used for)** o **EQU (Equivalent)**. Y respecto de la sinonimia parcial hay que decir que suele implementarse bien a partir de la relación **ET (Environmental Term)**<sup>37</sup>, bien a partir de la relación **RT (Related Term)**<sup>38</sup>.

## B) JERARQUIA

Aquí vamos a referirnos a las relaciones de género a especie y viceversa y las relaciones de especie a especie por referencia a un género común, y para ello conviene introducir las nociones lingüísticas de hiponimia y cohiponimia.

- Hiponimia. La relación paradigmática de hiponimia se da entre dos términos cuando uno designa un concepto más general y el otro uno más específico. En ocasiones, para designar a esta relación se ha utilizado la expresión "inclusión"<sup>39</sup>; sin embargo, es preferible la de hiponimia<sup>40</sup>, pues aquélla resulta ambigua<sup>41</sup>. Para comprobarlo, recurramos a título de ejemplo a las expresiones "derecho subjetivo"

y "derecho de crédito". La ambigüedad de "inclusión" viene dada por el hecho de que si dicha relación se determina a partir de un análisis intensional, entonces se dirá que "derecho subjetivo" está incluido en "derecho de crédito". Por el contrario, si se determina a partir de un análisis extensional, lo correcto será decir que "derecho de crédito" está incluido en "derecho subjetivo". La explicación de lo anterior radica sencillamente en que la intensión y la extensión guardan entre sí una relación inversa. Por eso, es preferible hablar de "hiponimia" y definirla tanto intensional como extensionalmente.

Desde la perspectiva intensional, la hiponimia se define a partir de la implicación semántica unilateral. Así, se dirá que entre dos términos **a** y **b** se da una relación de hiponimia si y sólo si el término **a** (hipónimo o subordinado) implica al término **b** (hiperónimo o superior-ordenado) y **b** no implica a **a**. En otros términos, la relación de hiponimia exige que si la implicación directa  $a \rightarrow b$  es verdadera, la conversa  $b \rightarrow a$  sea falsa, y viceversa<sup>42</sup>. Lo anterior puede representarse del siguiente modo:

$$(\text{Hip. } a, b) \leftrightarrow [a \rightarrow b \ \& \ -(b \rightarrow a)]$$

Así, por ejemplo, podría decirse que puesto que "homicidio" implica "delito" y "delito" no implica "homicidio", entre estos términos se da una relación de hiponimia donde "homicidio" es el hipónimo y "delito" el hiperónimo.

La hiponimia puede definirse también extensionalmente. Sean **x** e **y** dos términos y **A** y **B** sus clases o extensiones. Se dirá que **x** es hipónimo de **y** si y sólo si concurren conjuntamente las siguientes condiciones: a)  $A \subset B$ ; b)  $B \not\subset A$ ; c)  $A \neq B$ <sup>43</sup>.

La relación de hiponimia es, pues, una relación no reflexiva, no simétrica y transitiva.

En la implementación documental de la relación de hiponimia suele recurrirse a dos relaciones inversas: **BT (Broader Term)** o hiponimia ascendente (especie a género) y **NT (Narrower Term)** o hiponimia descendente (género a especie). Una vez más hay que decir que en la implementación de las relaciones es común no respetar las propiedades que conceptualmente se definen de cada una de ellas. Así, en este caso, por regla general suele considerarse como una relación no transitiva. La explicación de ello radica en que las relaciones **BT** y **NT** se

utilizan, aparte de para representar las dos formas de hiponimia vistas, para representar relaciones verticales del tipo parte-todo o todo-parte, que no son relaciones ni transitivas ni intransitivas, sino más bien no transitivas<sup>44</sup>.

- Cohiponimia. En la estructuración de un tesoro la cohiponimia juega un papel fundamental. Con ella se expresa la relación que existe entre dos términos que tienen un hiperónimo común.

Intensionalmente la cohiponimia se definirá a partir de la conjunción de dos hiponimias, donde el término hiperónimo sea común a ambas. Puede, por tanto, representarse del siguiente modo:

$$(\text{Cohip. } a, b) \leftrightarrow [\bar{a} \rightarrow c \ \& \ -(c \rightarrow a)] \ \& \ [b \rightarrow c \ \& \ -(c \rightarrow b)]$$

Así, se dirá que dos términos son cohipónimos si y sólo si existe un tercer término tal que sea hiperónimo de los dos anteriores. Por ejemplo, "robo" y "hurto" son cohipónimos en la medida que tienen a "delito contra la propiedad" como hiperónimo común<sup>45</sup>.

También puede definirse en términos extensionales. Sean  $x$ ,  $y$ ,  $z$  tres términos y  $A$ ,  $B$ ,  $C$ , sus clases o extensiones; se dirá que  $x$  e  $y$  son cohipónimos si y sólo si se verifican las siguientes condiciones: a)  $C/(A \ C \ \& \ B \ A)$ ; y b)  $A \ B = \emptyset$ .

La relación de cohiponimia es no reflexiva, simétrica y no transitiva. En la implementación documental de la misma se recurre a la relación ya conocida **RT (Related Term)**, que se utilizaba también para representar la cuasisinonimia. Sin embargo, la diferencia entre la cohiponimia y la cuasisinonimia radica en que mientras que en la primera la intersección entre las clases de los conceptos afectados es igual al conjunto vacío, en la segunda, por definición, es distinta del conjunto vacío.

### C) CONTRASTES

Siguiendo a Lyons, dentro de los contrastes - "término más general y libre de implicaciones en cuanto al número de elementos contrastantes en un conjunto dado"<sup>46</sup> - distinguiremos la incompatibilidad - contraste no binario - de la oposición - contraste binario - y dentro de ésta última, a su vez, distinguiremos la complementariedad, la antonimia y la inversión.

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

- Incompatibilidad. La incompatibilidad suele predicarse de los contrastes no binarios (no dicotómicos) entre conjuntos de términos de muchos miembros<sup>47</sup>. Para definir esta relación los lingüistas suelen recurrir a las nociones de aserción y de implicación semántica. Así, se dirá que dos términos son incompatibles cuando la aserción de uno implica la negación del otro, pero la negación de uno no implica la aserción del otro<sup>48</sup>, esto es, **a** implica **-b**, pero **-a** no implica **b**. Por ejemplo, "robo" implica "no-hurto", pero "no-robo" no implica "hurto". Ello puede representarse del siguiente modo:

$$(\text{Incom. } a, b) \leftrightarrow (a \rightarrow -b) \ \& \ -( -a \rightarrow b).$$

La explicación de ello radica en que es un contraste no binario.

Extensionalmente puede definirse del siguiente modo. Sean **x** e **y** dos términos, **A** y **B** sus clases o extensiones, y **z** un elemento. Se dirá que **x** e **y** son incompatibles si la pertenencia de **z** a **A** implica que **z** no pertenece a **B**, y la no pertenencia de **z** a **A** no implica que **z** pertenezca a **B**. En otras palabras, **x** e **y** son incompatibles cuando se verifican conjuntamente estas dos condiciones: a) **A**  $\rightarrow$  **-B**; y b) **-B**  $\rightarrow$  **A**.

En ocasiones se ha pretendido explicar (incluso equiparar) la incompatibilidad a través de la contrariedad lógica. Sin embargo, aquí es fundamental resaltar el papel que las nociones de campo conceptual y de campo semántico juegan en la estructuración del tesauro. Y, en este sentido, - como escribe Lyons - "*lo importante es que la incompatibilidad como relación léxica se basa en un contraste dentro de la similitud. Así, 'rosa' y 'cerdo' son contrarios, pero no merece la pena discutir su condición de incompatibles desde el momento en que uno denota una flor y otro un animal y, por tanto, sus sentidos respectivos no se delimitan recíprocamente*"<sup>50</sup>.

Esta relación de incompatibilidad, ciertamente, no juega un papel autónomo excesivamente relevante dentro de la estructuración de un tesauro y, además, su implementación documental suele realizarse a través de la relación de cohiponimia.

- Complementariedad. La complementariedad es un contraste binario - y, por tanto, una oposición - no graduable entre términos. En lingüística suele definirse a partir de la implicación semántica bilateral. Se dirá,

pues, que dos términos son complementarios cuando concurren estas dos condiciones: a) la aserción de uno implica la negación del otro; y b) la negación de uno implica la aserción del otro. En definitiva, la complementariedad puede definirse como la conjunción de dos implicaciones (directa y conversas) o equivalencia<sup>51</sup>; es decir:  
(comp. a, b)  $\leftrightarrow (a \rightarrow -b \ \& \ -b \rightarrow a) \leftrightarrow (a \leftrightarrow -b)$ .

Por ejemplo, "culpable" e "inocente" son complementarios, pues ser culpable implica no ser inocente y no ser inocente implica ser culpable; dicho de otro modo: ser culpable equivale a no ser inocente.

Lo anterior puede expresarse también en términos extensionales: si  $x$  e  $y$  son términos,  $A$  y  $B$  sus clases, y  $z$  un elemento, se dirá que  $x$  e  $y$  son complementarios cuando si  $z$  pertenece a  $A$ , entonces  $z$  no pertenece a  $B$  y si  $z$  no pertenece a  $A$ , entonces  $z$  pertenece a  $B$ . Como se ve, para que dos términos sean complementarios se requiere que la clase de uno sea igual a la negación de la clase del otro<sup>52</sup>.

En ocasiones se ha pretendido explicar la complementariedad en términos de contradicción lógica. Sin embargo, aquí tampoco conviene olvidar que la complementariedad como relación léxica se ve también afectada por la noción de campo conceptual; y, así, mientras que en ningún caso podríamos decir que entre "hombre" y "mujer" se da una relación de contradicción, sí podríamos tratarlos en el tesoro como complementarios, dado su carácter dicotómico dentro de un determinado campo conceptual.

En el tesoro experimentado por el **I.D.G.** de Florencia la implementación documental de esta relación se realiza a través de **CO** (**Complementary Term**) que se define como no reflexiva, simétrica y no transitiva<sup>53</sup>.

- Antonimia. La antonimia se predica de los pares e opuestos (contraste binario o dicotómico) que son graduables. La graduación supone siempre comparación entre términos a partir de una determinada propiedad<sup>54</sup>. En otras palabras, puede decirse que son antónimos los términos que designan conceptos que constituyen los extremos (lo que les confiere el carácter de contraste dicotómico) de una escala ininterrumpida, continua, o no discreta (lo que les confiere el carácter de graduables) de valores. Como ejemplos jurídicos de antónimos podemos

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

mencionar los siguientes pares de opuestos: "grave" y "leve", "general" y "especial", "superior" e "inferior", "anterior" y "posterior", etc.

El hecho de que en muchas ocasiones se utilicen antónimos sin una regla explícita de graduación puede llevar a confundirlos con los complementarios. Sin embargo, la diferencia radica en que las cualidades representadas por los antónimos no son independientes y, además, son relacionales. Y, así, mientras que entre los complementarios pueden hacerse deducciones del siguiente tenor "si X es un ciudadano inocente, entonces X es una persona inocente"; sin embargo, éstas no funcionan con los antónimos, pues pueden suponer una violación de la regla implícita de comparación como, por ejemplo, pasar de "Y es una falta grave" a "Y es un acto ilícito grave". Otra forma de percatarse de la diferencia entre unos y otros es la siguiente: consideremos estas dos oraciones "X es una norma más general que Y" e "Y es una norma más general que Z"; si tratáramos la oposición entre "general" y "especial" como si fueran complementarios, entonces llegaríamos a la paradójica conclusión de que Y es "general" y "especial", o lo que es lo mismo, que es "general" y "no general".

- Inversión. La inversión es la relación de oposición que se da, por ejemplo, entre "marido" y "esposa", "acreedor" y "deudor", "comprar" y "vender", etc. donde cada término de cada par constituye el inverso del otro. Esta relación se diferencia de la complementariedad en que aquí no puede decirse que la negación de uno implique la afirmación del otro; y se diferencia también de la antonimia en que la inversión expresa una oposición no graduable. Ahora bien, lo anterior no es suficiente para caracterizar a esta relación. Lo que en última instancia caracteriza a la inversión - como lo pone de manifiesto Lyons - es que los predicados implicados en ella sólo pueden interpretarse como predicados de dos lugares (diádicos) y, así, a partir de la noción lógica de relación inversa, puede mostrarse que si R es una relación de dos lugares y R' su inversa, entonces puede sustituirse R por R' y transponer los términos de la relación para obtener una equivalencia<sup>55</sup>:

$$(\text{Inv. } x, y) \leftrightarrow [\overline{[R(x, y) \leftrightarrow R'(y, x)]}]$$

Por ejemplo, si "x es acreedor de y" equivale a "y es deudor de x", entonces "acreedor" y "deudor" son términos inversos.

En el tesoro experimentado por el I.D.G de Florencia dicha relación se implementa a través de **INV (Inverse Term)**, que se define como no reflexiva, simétrica y no transitiva.

### 6.3 LAS RELACIONES BASADAS EN LAS ESTRUCTURAS DE LO REAL

Dentro de este género de relaciones, que en ocasiones se han denominado sistemáticas<sup>56</sup> o, también, asociativas<sup>57</sup>, voy a referirme a las dos más frecuentes: la relación partitiva y la relación proceso/agente.

Por relación partitiva suele entenderse tanto la relación vertical todo/parte o parte/todo como la relación horizontal parte/parte por referencia a un todo común. Estas relaciones presentan grandes analogías con las de jerarquía (hiponimia y cohiponimia) antes vistas. De hecho, ambos tipos de relaciones suelen implementarse documentalmente de la misma forma, a partir de las relaciones ya conocidas **BT (parte/todo)**, **NT (todo/parte)** y **RT (parte/parte)**. Un ejemplo de este tipo de relaciones puede ser el siguiente: "edificio escolar" puede considerarse un todo respecto de "aulas", "biblioteca", "laboratorio", etc., que serían partes relacionadas con ese todo y entre sí.

Por otro lado, dentro de las relaciones basadas en estructuras de lo real también juega un papel relevante la oposición proceso/agente. Es común implementarla a partir de la relación horizontal **RT (Related Term)**. Ejemplos de este tipo de relación podrían ser "contaminación atmosférica" (proceso) y "humos" o "emanaciones" (agentes).

### 6.4. LAS RELACIONES MORFOLOGICAS

Dentro de las relaciones morfológicas, que son el resultado de asociar palabras por sus significantes, vamos a detenernos en la homografía y en la equivalencia sintáctica.

La relación de homografía permite superar los problemas documentales que plantean los fenómenos de la polisemia y de la homografía. Es común implementarla de las dos siguientes formas: bien a partir de la relación **SC (Scope Note)**, bien a partir de la relación **HOM (Homograph)**. Por ejemplo:

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

- Causa **SC n**. Donde **n** puede representar una nota aclaratoria del sentido en que se utiliza dicha palabra si el tesoro es manual o un número que sirva para recuperar esa nota si el tesoro es automático.
- Causa **HOM Causal** (antecedente necesario que origina un efecto), Causa2 (fundamento por el cual se adquiere un derecho).

Finalmente, la equivalencia sintáctica es la relación que se establece entre palabras que, teniendo el mismo morfema léxico (la misma raíz), pero morfemas gramaticales diferentes (son, por tanto, morfológicamente diferentes) cumplen, sin embargo, la misma función documental. Esta relación es especialmente útil para los sistemas de indexación automática, dado que en ellos pueden subsistir, por ejemplo, un sustantivo, un verbo y un adjetivo que en realidad cumplan la misma función documental. Por ejemplo: "juez" y "judicial"; "adopción", "adoptar", "adoptivo", etc. Esta relación se implementa a partir de **SE (Syntactical Equivalent)**<sup>59</sup>. En ocasiones, en oposición a esta relación se establece la relación **CFR (Compare)**, para relacionar términos morfológicamente diversos que tienen el mismo morfema léxico y que por razones documentales no se definen como equivalentes sintácticos<sup>59</sup>.

### NOTAS

- <sup>1</sup> Cfr. entrada "Lengua natural" de la voz "Lengua" en CERDA MASSO, RAMON (Coordinador): **Diccionario de lingüística**, Ed. Anaya, Madrid, 1986, pág. 175.
- <sup>2</sup> Tomo las expresiones "grupo fundamental" y "grupos específicos" de KNAPP, VIKTOR: "Alcuni problemi relativi alla costruzione di un thesaurus giuridico, en **Informatica e Diritto**, enero-marzo 1979, pág. 178. En adelante voy a tener muy en cuenta este artículo de Knapp.
- <sup>3</sup> Cfr. **Ibidem**, pág. 178.
- <sup>4</sup> Cfr. **Ibidem**, págs. 178 y 179.
- <sup>5</sup> Cfr. **Ibidem**, págs. 179 y 180.

6 En este punto, tal vez resulte clarificador introducir la distinción que dentro de los lenguajes normativos hace Scarpelli entre conceptos fácticos (*fattuali*) y conceptos calificadorios (*qualificatori*). Según este autor, los primeros designan simples cosas, propiedades o relaciones, mientras que los segundos designan cosas, propiedades o relaciones calificadas por normas. Pues bien, si proyectamos esta distinción sobre la división hecha en el texto, no es difícil observar que, en términos generales, el conjunto IIA contiene términos que designan conceptos fácticos, mientras que los conjuntos IIB y III contienen términos que designan conceptos calificadorios. Sobre la distinción entre conceptos fácticos y calificadorios véase SCARPELLI, U. : **Contributo alla semantica dil linguaggio normativo**, Giuffrè, Milán, 1985 (segunda edición), págs. 123 y ss.

7 Cfr. *Ibidem*, págs. 182 y 183.

8 En este sentido pueden verse *Ibidem*, pág. 177.

9 Las ventajas de los lematizadores o generadores gramaticales frente a las funciones de enmascaramiento a la derecha, se ven claras si se piensa, por ejemplo, en verbos, no todos ellos ajenos al Derecho, como los siguientes: "errar", "dormir", "morir", "asir", "caer", "dar", "ir", "oir", "poder", "pudrir" o "podrir", "tener", "venir", "ver", "yacer", etc.

10 Originaria del latín "thesaurus". Cfr. la voz 'Tesoro' del **Diccionario de la Real Academia de la Lengua Española**.

11 Cfr. en la voz 'Thesaurus' de la **Gran Enciclopedia Larousse**, Ed. Planeta, Barcelona, 1984.

12 Esta definición ha sido establecida por la I.S.O. (Organización Internacional para la Estandarización) en **Documentation-principes directeurs pour l'etablissement et le développement de thesaurus monolingues destinés à la recherche documentaire**, 1972; como la cita de GIANNANTONIO, ETTORE: **Introduzione all'informatica giuridica**, Ed. Giuffrè, Milán, 1984, pág. 184; y de AMAT NOGUEIRA, N. : **Documentación Científica y Nuevas Tecnologías de la Información**, Ed. Pirámide, Madrid, 1987, pág. 212.

## LENGUAJE JURIDICO, LENGUAJE DOCUMENTAL Y TESAURO

- <sup>13</sup> En BAUER-BERNET, HELENE: "Legal Thesauri and Data Processing", en MARTINO, A. A. (Compilador): **Deontic Logic, Computational Linguistics and Legal Information Systems**, Ed. North-Holland, Amsterdam-New York-Oxford, 1982, pág. 237.
- <sup>14</sup> En KNAPP, V. : ob. cit., pág. 177.
- <sup>15</sup> Definición atribuida sin referencia bibliográfica a Linant, Xavier en Elhazaz Molina, A. : "La legislación educativa y la informática jurídica. La base de datos LEDA.", en Rivero, A. M. y Santodomingo, A. (Compiladores): **Introducción a la informática jurídica**, Ed. Fundesco, Madrid, 1986, pág. 67.
- <sup>16</sup> Definición propuesta por la I.S.O., véase la nota 12.
- <sup>17</sup> Elhazaz Molina, A. : ob. cit., pág. 67.
- <sup>18</sup> Reisinger, L. : "Rechtsinformatik" (W. de Gruyter, Berlin, 1977); tomo la cita de Taddei Elmi, G. : "Logic, Linguistic and legal thesauri", en Martino, A.A. (Compilador): ob. cit., pág. 292.
- <sup>19</sup> Losano, M. G. : **Curso de informática jurídica** (trad. M. Atienza, J. Ruiz Manero y J. Aguiló Regla), Ed. Tecnos, Madrid, 1987, pág. 228.
- <sup>20</sup> No hay que olvidar que en el ámbito de la doctrina es imposible establecer una relación fuerte entre el lenguaje documental y el de los documentos, puesto que la unidad documento (el registro) no contiene generalmente zona de texto, y si la contiene (resumen) nunca puede cumplir una función sustitutoria respecto del documento fuente.
- <sup>21</sup> Knapp, V. : "Legal Thesauri", en Martino, A. A. : ob. cit., pág. 284.
- <sup>22</sup> Esta característica implica que de la colección de palabras deben eliminarse aquéllas que a pesar de poseer un significado autónomo no resultan útiles por su altísima frecuencia. No eliminarlas supondría que cada vez que el usuario las utilizara en el sistema tendría que realizar casi un vaciado completo del fichero directo.
- <sup>23</sup> Saussure explica del siguiente modo en qué consisten las relaciones sintagmáticas: *"(...) en el discurso, las palabras contraen entre sí, en virtud de su encadenamiento, relaciones fundadas sobre*

*el carácter lineal de la lengua, que excluye la posibilidad de pronunciar dos elementos a la vez. Estos se alinean unos detrás de otros en la cadena del habla. Estas combinaciones que tienen por soporte la extensión pueden ser llamados sintagmas. El sintagma se compone siempre, por tanto, de dos o más unidades consecutivas (...) Situado en un sintagma, un término adquiere su valor sólo porque se opone al que le precede o al que le sigue, o a los dos".* En Saussure, F. : **Curso de lingüística general**, Ed. Akal, Madrid, 1980, pág. 172.

- 24 En cualquier caso, debe recordarse que el usuario podrá representar en la relación de consulta, mediante una coordinación **a posteriori**, sintagmas no contemplados en el tesoro.
- 25 Saussure explica del siguiente modo en qué consisten las relaciones paradigmáticas, si bien él las denomina relaciones asociativas: "*(...) al margen del discurso, las palabras que ofrecen algo en común se asocian en la memoria, y se ofrecen así grupos en cuyo seno reinan relaciones muy diversas (...) Como vemos estas coordinaciones son de una especie completamente distinta a las primeras (las sintagmáticas). No tienen por apoyo la extensión; su sede está en el cerebro.*" En Saussure, F. : ob. cit., págs. 172-173.
- 26 **Ibidem**, pág. 173.
- 27 En el TLS/Thesaurus experimentado por el **Istituto per la Documentazione Giuridica** (I.D.G.) de Florencia, la relación de sinonimia se implementa a partir de la relación **EQU (Equivalent)** y se define entre dos palabras clave y no entre una palabra clave y su equivalente. En este caso, naturalmente, la relación es simétrica. Cfr. Archi Stanta, A. y otros: "A linguistic tool for legal data retrieval", en Martino, A. A. (Compilador): ob. cit., pág. 300 y ss.
- 28 Cfr. Taddei Elmi, G. : ob. cit., pág. 290. Para M. Bunge no es suficiente definir la sinonimia a partir sólo de la identidad entre la extensión de dos conceptos, porque - dice - puede darse conceptos que tengan la misma extensión y que, sin embargo, tengan

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

intensiones diferentes (ej. "triángulo equilátero" y "triángulo equiángulo"). Por tanto, este autor exige la identidad entre las intensiones y las extensiones de dos conceptos. Cfr. Bunge, M. : **La investigación científica**, Ed. Ariel, Barcelona, 1981. Sin embargo, en el ámbito del tesauro no es común introducir restricciones de este género en la noción de sinonimia.

29 Cfr. Archi Stanta, A. y otros : ob. cit., pág. 329.

30 En este punto tal vez convenga advertir que el término implicación es ambiguo. En lógica suele distinguirse entre "implicación material" e "implicación formal". La primera expresión se reserva para referirse a una relación tal que sólo resulta falsa cuando el antecedente es verdadero y el consecuente es falso; es decir, se utiliza para aludir a los condicionales contingentes. En oposición a ella, la expresión "implicación formal" se utiliza para referirse a aquellas fórmulas en las que la verdad del antecedente asegura formalmente la verdad del consecuente; es decir, se utiliza para aludir a los condicionales tautológicos. Sin embargo, los lingüistas suelen hablar de "implicación material" para referirse a una relación en la que no es necesario que exista entre el antecedente y el consecuente una conexión de significado; y a esta noción suelen oponer la de "implicación semántica" (también llamada "implicación estricta", entrafiamiento o vinculación) para referirse a una relación en la que por razones semánticas no es posible que sea verdadero el antecedente y que el consecuente sea falso. En este trabajo, el término implicación se utilizará en este último sentido y para su representación se recurrirá al signo " $\rightarrow$ ". Sobre las distinciones aquí apuntadas véanse, por ejemplo, Guibourg, Ricardo A. y otros: **Lógica, proposición y norma**, Ed. Astrea, Buenos Aires, 1983, págs. 56-57, 73-78; y Lyons, John: **Semántica**, Ed. Teide, Barcelona, 1980, págs. 140 y 159.

31 Cfr. Taddei Elmi, G. : ob. cit., pág. 294.

32 Cfr. la voz "Sinonimia" en Dubois, Jean y otros: **Diccionario de lingüística**, Ed. Alianza, Madrid, 1983, pág. 330.

33 Cfr. Knapp, V. : "Alcuni problemi...", ob. cit., pág. 188; y "Legal THesauri", ob. cit., pág. 279.

Josep AGUILÓ REGLA

- <sup>34</sup> Cfr. Taddei Elmi, G. : ob. cit., pág. 294.
- <sup>35</sup> Cfr. **Ibidem**, págs. 294 y 295.
- <sup>36</sup> Cfr. **Ibidem**, pág. 302.
- <sup>37</sup> Cfr. Archi Stanta, A. y otros: ob. cit., págs. 330 y 336.
- <sup>38</sup> Cfr. Ciampi, C. y otros: "THES/BID: A Computer-based Thesaurus of Terminology in Computers and the Law", en **Informatica e Diritto**, enero-abril, 1982, pág. X.
- <sup>39</sup> Cfr. Lyons, John : **Semántica**, Ed. Teide, Barcelona, 1980, pág. 273.
- <sup>40</sup> Esta palabra fue acuñada por analogía con sinonimia, antinomia, etc. Cfr. **Ibidem**, pág. 273.
- <sup>41</sup> Sobre la ambigüedad de "inclusión", véase Taddei Elmi, G. : ob. cit., pág. 296; y Archi Stanta, A. y otros : ob. cit., págs. 330-331.
- <sup>42</sup> Cfr. Lyons, John : ob. cit., pág. 274; Taddei Elmi, G. : ob. cit., pág. 296; y Archi Stanta, A. y otros : ob. cit., pág. 331.
- <sup>43</sup> La tercera condición es redundante, dado que si verifican las dos primeras resulta imposible que **A** pueda ser igual a **B**. Sin embargo, esta redundancia no resulta del todo inútil. Se ha introducido para evitar confundir esta noción de hiponimia con la de hiponimia impropia, bilateral o simétrica a la que algunos autores recurren para definir a la sinonimia total, perfecta o absoluta. En este sentido, pueden verse Lyons, John : ob. cit., págs. 274-275; y Taddei Elmi, G : ob. cit., pág. 297.
- <sup>44</sup> Cfr. Lyons. John : ob. cit., págs. 295 y 296.
- <sup>45</sup> No se entiende qué razones pueden haber llevado a Taddei Elmi y a Archi Stanta a sostener que la cohiponimia puede definirse en términos de una implicación donde el primer término es la conjunción de dos implicaciones unilaterales y el segundo una implicación bilateral o coimplicación. Esto es:
- $$[(a \rightarrow c) \& (b \rightarrow c)] \rightarrow (a \leftrightarrow b)$$
- Cfr. Taddei Elmi, G. : ob. cit., pág. 298; y Archi Stanta, A. y otros : ob. cit., pág.332.

## LENGUAJE JURÍDICO, LENGUAJE DOCUMENTAL Y TESAURO

- 46 Lyons, John : ob. cit., pág. 261.
- 47 Cfr. **Ibidem**, pág. 270.
- 48 Cfr. Archi Stanta, A. y otros : ob. cit., pág. 332.
- 49 Cfr. Taddei Elmi, G. : ob. cit., págs. 299-300.
- 50 Lyons, J. : ob. cit. pág. 271.
- 51 Cfr. Archi Stanta, A. y otros : ob. cit., pág. 333.
- 52 Cfr. Taddei Elmi, G. : ob. cit., págs. 300-301.
- 53 Cfr. Archi Stanta, A. y otros : ob. cit., pág. 336.
- 54 Cfr. Lyons, J. : ob. cit., pág. 254.
- 55 Cfr. Lyons, J. : ob. cit., págs. 262-263.
- 56 Cfr. Archi Stanta, A. y otros : ob. cit., pág. 334.
- 57 Cfr. Aitchinson, Jean (Preparado por) : **Thesauro de la Unesco**, Ed. Unesco, 1984, pág. XIV.
- 58 Cfr. Archi Stanta, A. y otros : ob. cit., pág. 335.
- 59 Cfr. **Ibidem**, pág. 329.

\*Seminario de Filosofía del Derecho. Universidad de Alicante