

# 3

## CAPITULO

**METODO E  
INSTRUMENTOS  
CONCEPTUALES**

## Capítulo 3º

### METODO E INSTRUMENTOS CONCEPTUALES

#### 3. METODO E INSTRUMENTOS CONCEPTUALES

El método general utilizado, para canalizar las hipótesis de esta investigación, es de base experimental. A partir de la observación de fragmentos periodísticos, vertebraremos inferencias que, una vez depuradas, constituirán un modelo explicativo y una metodología de análisis y representación.

El método se compone de elementos procedimentales, tomados prestados de otras disciplinas, que son aplicados en momentos adecuados de la investigación de tal manera que, mediante un correcto engarce, nos lleva a la comprobación de nuestras conjeturas. Pero, en primer lugar, es necesario exponer la justificación que nos ha permitido fabricar las primeras hipótesis de rango metodológico.

El objeto, sobre el que actuamos para proponer una solución de procedimiento, es el proceso documental del discurso periodístico, en definitiva, el tratamiento que el documentalista da a una noticia para que ésta sea localizada entre otras similares, y en relación a esas otras, dentro de un complejo sistema en el que fluctúan miles de referencias y millones de datos. El objetivo es instrumentar al documentalista de prensa para facilitar y optimizar su tarea, a la vez que, colateralmente, redefinir su rol a medio plazo.

Para proceder a observar el corpus y demostrar que el discurso periodístico tomado sectorialmente presenta un crecimiento sostenido de bases léxica y estructural, establecemos la metodología, primero sobre el eje del paradigma y, a continuación, sobre el eje del sintagma para, finalmente, analizar los resultados conjuntamente.

#### 3.1 EJE PARADIGMATICO

Dentro del Paradigma (eje de las selecciones posibles) seguimos un camino jerárquico, de lo específico a lo genérico, mediante el procedimiento analítico propio de la construcción de esquemas globales. Los términos del enunciado son sustituidos por sinónimos y/o clasemas ascendentes o descendentes, según el caso, con el fin de uniformar la muestra conceptualmente y poder ordenarla en función de una estructura global. El objetivo a cumplir, por este camino, es demostrar que en la información periodística existe un componente lógico-estructural invariable (en su percepción diacrónica a corto plazo) sustentado en estratos o capas que presentan más variación cuanto más específica es la terminología pero que, aún así, la variación de la base semántica es asumible por el modelo general sin ser modificado.

En el eje paradigmático se procede, por la misma vía, a la esquematización de las acciones, con el fin de conseguir un inventario estable. El método de las modalidades ha sido aplicado, sin grandes resultados, para obtener una estructura jerárquica de los datos y, sin mayor fortuna, excepto por el auxilio en la organización, en la jerarquización verbal. Como resultados prácticos, hemos obtenido una aproximación empírica a la ordenación, necesaria y posible, de conceptos y acciones en primitivas.

Hemos elegido la metodología del Análisis del discurso, propuesta por Van Dijk para el caso de la información de actualidad (1) y recordada por Amaro (2) en su perspectiva documental. Consiste en la sustitución de un texto por una macroproposición global en el proceso analítico-sintético. El procedimiento de Van Dijk se ha practicado tanto en la obtención de la muestra (véase el cap. 4) como para proponer al documentalista de prensa una parte del aparato instrumental que precisa para reconocer, desmontar y representar el texto periodístico.

#### 3.2 EJE SINTAGMATICO

En el Sintagma (eje de las elecciones realizadas) se observa el comportamiento funcional de

las categorías léxicas, esto es, las estructuras que forman entre sí y la relación que mantienen respecto a la totalidad del enunciado. Para la explicitación de la estructura y la resolución posible de algunos problemas, he tomado la Gramática de Casos, enunciada por Fillmore y Pottier (3), adaptada y aplicada por Cunha a la construcción de lenguajes documentales (4) y analizada por mí mismo (5) como procedimiento de interrogación textual.

La teoría de casos supera el análisis sintáctico convencional (analiza síntesis virtuales de macroproposiciones globales) y es útil para identificar los componentes de la enunciación en un texto toda vez que el texto adquiera proporciones reducidas. Con el enunciado-texto, obtenido mediante el procedimiento de Van Dijk, merced a las contribuciones de Amaro, podemos operar en el nivel práctico (de documentalista de prensa) y abrimos una vía para la simulación automática. En consecuencia, superamos el bloqueo que imponen el texto y los enunciados compuestos al operar con estos como sumatoria de enunciados simples. El enunciado simple consta de una acción introducida por un verbo y de elementos periféricos.

El camino seguido comienza con el estudio empírico de enunciados simples y reales de la muestra construida. Supuestamente, todos los elementos significativos de los enunciados responden a patrones de acción y conceptuales. Por lo tanto, en torno a la acción se interroga el enunciado con los casos conceptuales para detectar cuáles de ellos son habituales en los enunciados, qué posiciones ocupan y qué nexos los introducen (conectores). Una vez detectadas todas las estructuras sintagmáticas, se procede a la reducción. Con ello se obtiene un patrón de estructura funcional que puede constar, en el nivel práctico, de varios formatos posibles y altamente formalizados, que utilizaría el documentalista para alimentar la memoria automática. En la andadura, obtenemos un esquema de las primitivas que subsume el inventario de nexos, nuevamente en el eje paradigmático.

Sin perjuicio de exponer, más adelante, las modificaciones de denominación y extensión de algunos casos en virtud de las necesidades

peculiares de la construcción periodística, describimos, a continuación, el modelo original de Fillmore /Pottier junto a la reducción y adaptación de Cunha, como método de construcción de vocabulario científico (6):

#### **Casos de Fillmore/ Pottier      Casos de Cunha**

|               |             |
|---------------|-------------|
| Causativo     |             |
| Instrumental  | Instrumento |
| Agentivo      | Agente      |
| Dativo        |             |
| Nominativo    |             |
| Ergativo      |             |
| Acusativo     | Objeto      |
| Sociativo     |             |
| Locativo      | Lugar       |
| Beneficiativo | Producto    |
| Final         | Finalidad   |

La aplicación de los casos (adaptados por Cunha) en la estructura enunciativa obtenida por el procedimiento de Van Dijk (adaptado por Amaro) a fin de obtener ángulos de búsqueda (con interrogadores naturales) no convencionales que respondan a las necesidades de los usuarios de la Documentación periodística es nuestra contribución al estado de cuestión. Esto supone partir de unas condiciones y llegar a unas modificaciones metodológicas.

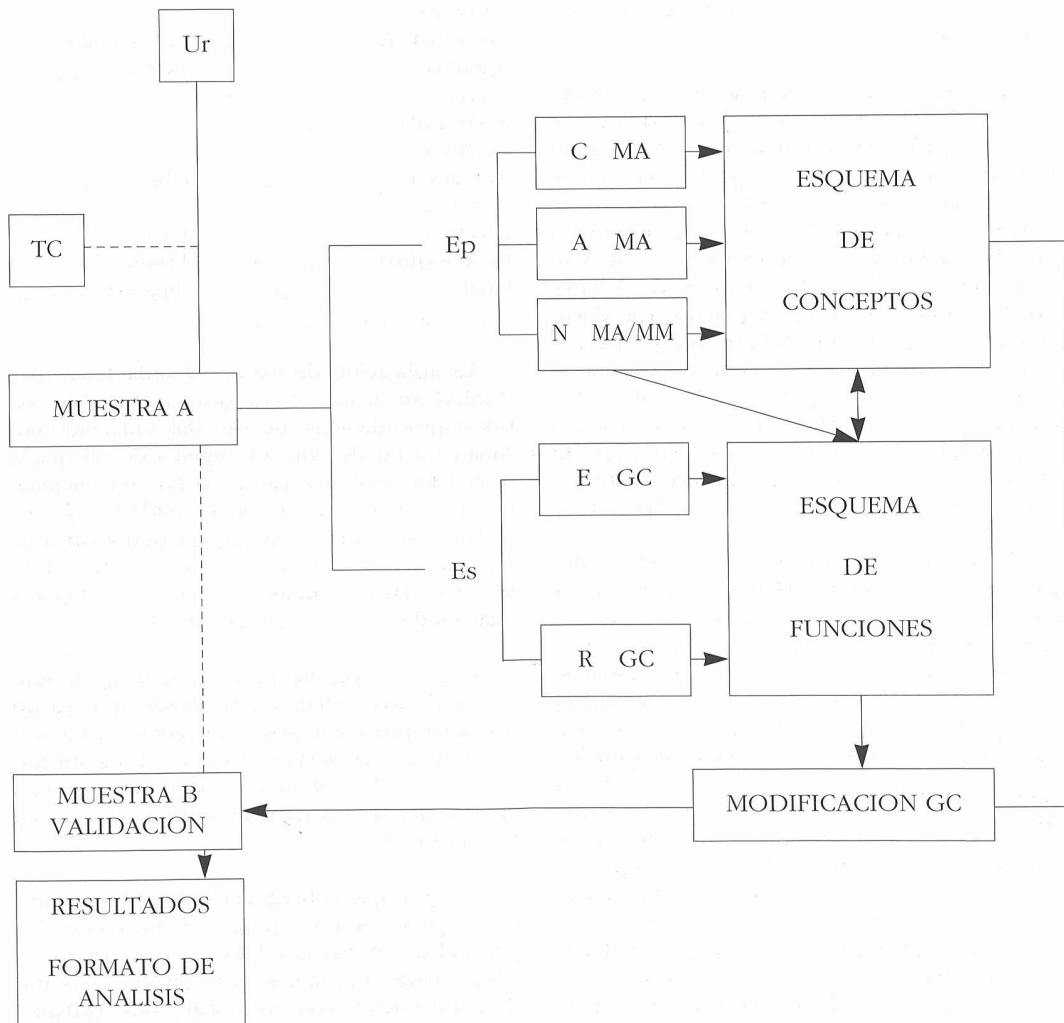
El caso conceptual produce un desglose funcional de los conceptos presentes, que ya no nos interesan como dato sino como elementos funcionales. La desvinculación del significado hace posible la simbolización del enunciado con las iniciales de los casos y su formalización en lenguaje lógico.

Una vez que conseguimos transportar el discurso periodístico a lenguaje simbólico, los resultados son mecanizables y programables, es decir, puede conseguirse el reconocimiento artificial de roles y conceptos naturales mediante una tecnología avanzada. El formalismo podría ser el propuesto por el lenguaje Prolog (concretamente por el intérprete Arity) ya ajeno al marco de investigación documentológica (7).

### 3.3 ESQUEMA DEL ARBOL METODOLOGICO

A continuación, esquematizamos el momento de la actuación de los diferentes recortes

metodológicos empleados sobre el corpus y la relación y ordenación que mantienen:



Ur: Universo de referencia  
 C: repertorio de conceptos  
 A: repertorio de acciones  
 N: repertorio de conectores  
 MA: método analítico

MM: método de modalidades  
 GC: método de casos  
 Ep: eje paradigmático  
 Es: eje sintagmático  
 E: estructuras sintácticas  
 R: roles conceptuales

### 3.4 INSTRUMENTOS CONCEPTUALES

A tenor de los instrumentos teóricos elegidos, y de la utilización de métodos ajenos, surgen conceptos esenciales que nos sirven para explicar determinados fenómenos o problemas de nuestro objeto, corpus y objetivos.

El aparato conceptual teórico tiene dos precedencias. Del propio universo práctico de la Documentación periodística, nos llegan términos que han adquirido cierta consolidación entre los profesionales e investigadores; digamos que son términos propios. De los análisis del significado, tomados globalmente, y de las Gramáticas, nos servimos de algunos conceptos-clave para operar en nuestra materia. Es necesario definir y justificar su presencia en esta investigación para entender el desarrollo específico del trabajo y contribuir a la consolidación del lenguaje epistemológico de la Documentación periodística. Veamos los conceptos más genéricos:

#### 3.4.1 Descriptores y temas

La metodología, aplicada al corpus, persigue dos objetivos: la captación de la representación más genérica posible de los temas aludidos en las unidades operativas y la detección de cómo se organizan funcionalmente esos temas a partir de las gramáticas específicas que utilizan los periodistas en la producción de sus textos.

Para ello, contamos con conceptos propiamente documentológicos y préstamos de otras disciplinas que nos permiten la reflexión en niveles virtuales. Procedentes del análisis efectuado por los documentalistas sobre los textos, es decir, pertenecientes al discurso documental, surgen los conceptos “descriptor” o “palabra-clave” y “tema” o “asunto”. Estos términos tienen un sentido estrictamente documentológico, si bien pueden ser adscritos a la Lexicología o a la Filosofía originariamente, por lo que son empleados en acepción especializada y pragmática.

El descriptor es una palabra clave controlada por un lenguaje documental (que implica reglas formales y semánticas). La palabra clave es, por tanto, un segmento o ángulo de búsqueda rele-

vante del texto en el estado previo a su normalización. A efectos prácticos, y mientras no se señale la distinción, ambas denominaciones serán utilizadas como sinónimas en este trabajo.

Los descriptores se obtienen por la vía del análisis textual, es decir, se infieren de un texto los conceptos que lo constituyen de acuerdo a un criterio pragmático: los conceptos deben tener suficiencia de tratamiento por parte del autor del texto a fin de que ofrezcan pertinencia en la recuperación. A esta operación se le denomina “indización” o descripción característica (8). Tanto el documentalista sobre su oferta (texto segmentado) como el usuario sobre su demanda (búsqueda segmentada) realizan un proceso indizador: aislar y combinar conceptos para representar conocimiento en la fase de entrada de datos del sistema, el primero, y en la fase de salida, el segundo.

Los temas o asuntos (asumidos por los documentalistas como evidencia empírica), por el contrario, se obtienen por la vía de síntesis textual. A esta operación se le llama “clasificación”. El texto, usado ahora como referencia, puede ser la producción del documentalista: tradicionalmente, el resumen o la ecuación de descriptores.

La operación analítico-sintética caracteriza el texto en sus dos dimensiones: global y específica y nos introduce en dos conceptos que van a priorizar la experimentación llevada a cabo: macroestructuras y superestructuras.

#### 3.4.2 Macroestructuras

Para la utilización de ambos conceptos, hemos acudido a Van Dijk quien los había definido de manera genérica en su obra “Texto y contexto” (9) y vuelve a hacerlo, de modo aplicado al discurso periodístico, en un libro de gran trascendencia documental: “La noticia como discurso...” (10). Para Van Dijk, los temas subyacen en la estructura del significado y, por lo tanto, pertenecen al macronivel global de la descripción del discurso. No se definen los temas en función de palabras u oraciones individuales sino a nivel textual, por ello, el concepto

teórico que propone para describir los asuntos o temas es el de macroestructura semántica (11).

Según el autor, las macroestructuras son un conjunto organizado de macroproposiciones. A cada texto le corresponde una macroproposición. Estas pueden subsumirse formando la macroestructura y, a la vez, subsumen otras proposiciones menores. Esta jerarquía, para Van Dijk, puede definirse mediante macrorreglas: reglas de transformación semántica, en sentido ascendente, que relacionan proposiciones de nivel más bajo con macroproposiciones de nivel más alto. Los temas derivan de los significados de un texto a través del resumen de macrorreglas (12).

La indización descompone el texto en proposiciones de autor bajo forma sustantiva. Es fácilmente ejecutable la ligazón gramatical de los descriptores en función del sentido general del texto. Si, para Van Dijk, en los textos periodísticos, a veces, la macroproposición viene recogida en titulares y sumario (13), incluso el titular puede expresar la macroproposición, se confirma nuestro procedimiento de contruir un corpus artificial a partir de titulares y descriptores, por lo tanto, nuestra **muestra de trabajo representa el universo global de macroproposiciones**. Siendo esto así, **es posible localizar y esquematizar la macroestructura** a partir de la construcción de unidades discretas que se componen de elementos del título, de los descriptores y del propio texto o resumen. Esta confirmación de procedimiento inicial para nuestra investigación se encuentra fundamentada en un consistente trabajo de Amaro (14).

### 3.4.3 Superestructuras

En la búsqueda de un ordenamiento de las categorías de los textos periodísticos, Van Dijk expone que "los esquemas periodísticos realmente existen y tanto los periodistas como los lectores los utilizan, al menos implícitamente, en la producción y comprensión de la noticia" (15). A pesar de que el investigador holandés lleva a cabo una comprobación empírica, proponiendo las categorías básicas de la superestructura, lo que realmente nos aporta a la Documentación

periodística es el concepto teórico de superestructura y la certeza de la localización de un punto de encuentro entre emisores y receptores de la base de datos (precisamente el esquema).

Para el autor, "la macroestructura (significado global) del discurso posee algo más que sus principios organizativos propios. Es necesario también algún tipo de sintaxis total, la que define las formas posibles en que los temas pueden insertarse y ordenarse en el texto real" (16). Esta superorganización se obtiene mediante la macrosintaxis que pretendemos obtener, en esta investigación, a partir de cientos de registros reales transportados a unidades de trabajo. Van Dijk reconoce, sin embargo, que "la teoría de los esquemas del discurso se halla aún en su infancia. La mayor parte del trabajo trata acerca de la organización de tipos de discursos específicos, como relatos y debates" (17).

El análisis documental se nutre de los avances de la teoría del discurso, eso es evidente y, a la vez, puede contribuir enormemente a su construcción, como un tipo de discurso más (discurso esquematizado). En su perspectiva más poderosa, la pragmática, esta investigación documentológica aplicada al texto periodístico utiliza el concepto de superestructura para describir, de manera global, los comportamientos de los enunciados producidos por el documentalista de prensa y formalizar estructuras parciales y totales para el reconocimiento mediante ordenador.

### 3.3.4 Primitivas

La primitiva es un concepto que tomamos de la bibliografía sobre Inteligencia artificial en su entronque con el lenguaje natural. Se trata de elementos que subsumen campos conceptuales y, en ese sentido, podemos establecer analogías con los macrodescriptores documentales y con las macroproposiciones discursivas.

Necesitamos primitivas para desencadenar procesos de raciocinio simulado que pendulan a través de los estratos conceptuales supraordenados, verificando y reconociendo adscripciones y dependencias. Schank (18) alude a los concep-

tos de guión y escenario (guiones complejos), es decir, situaciones concretas que suceden en el nivel actancial y que tienen que ser aprendidas por la máquina para poder proceder a la simulación del comportamiento de esos actores y de sus acciones.

En nuestro trabajo, la esquematización corresponde a una modelización empírica de primitivas en varias dimensiones enunciativas (conceptos, acciones, conectores y estructuras) que deben poder subsumir los textos de la muestra (guiones) para reproducir la lectura ponderada del documentalista de prensa.

### 3.3.5 Otros conceptos

Veamos, a título indicativo, términos que, aún siendo ajenos, comienzan a ser habituales en el propio lenguaje de la Documentología periodística. Son términos procedentes de la Morfosintaxis: morfema, grafema, predicado, enunciado, núcleo, elementos marginales, etc.; de la Semántica: sema, lexema, semema, archisemema; del Análisis de Contenido: indicadores, muestreo, unidades de trabajo; de la Tecnología: formato, campo, registro, fichero invertido, entre otros, todos ellos adscritos, esencialmente pero no en exclusiva, a esas disciplinas.

En los oportunos epígrafes, y por orden de aparición, serán recortados los sentidos generales de los mismos y, en su caso, las adaptaciones a que son sometidos en función de la aplicación.

## 3.5 LA BASE DE DATOS COMO DISCURSO

Una base de datos, o más específicamente, la información que contiene, presenta un comportamiento discursivo. Podemos, por tanto, considerarla como objeto de aplicación de conceptos y metodologías provenientes de los llamados análisis semánticos, especialmente, del análisis del discurso.

Una base de datos es un depósito de miles de textos que comporta tres dimensiones: semántica, sintáctica y de articulación intertex-

tual. En su dimensión semántica, la base de datos ofrece conceptos y direcciones sémicas que constituyen un entramado significativo y una coherencia que pueden ser observados desde perspectivas singulares o globales. El entramado semántico de la base de datos puede esquematizarse, lo que supone una normalización y una ordenación, dando lugar a la construcción de la estructura básica de un lenguaje documental.

La dimensión sintáctica de la base de datos aparece en el micronivel enunciativo, los registros, mostrando las relaciones que mantienen las categorías léxicas entre sí. Puesto que los registros son unidades altamente formalizadas como conjunto expresivo, por imposiciones tecnológicas y de la recuperación, es posible localizar sus sintaxis repetitivas y elevar y aislar, inductivamente, la gramática de gramáticas, es decir las relaciones funcionales que mantienen las macrocategorías léxicas entre sí. Esta inducción nos permite proponer un esquema general de roles válido, inicialmente, para el discurso establecido como campo de trabajo.

La articulación intertextual es la última dimensión de la estructura de contenido de una base de datos. La coherencia global se establece mediante el conjunto de articulaciones de registros, en nuestro caso textos periodísticos, formando relatos y unidades discursivas superiores. Como reflejo del discurso periodístico, disperso y asistemático en apariencia, la base de datos debe proponer la aglutinación de estructuras funcionales y significados como modelo de organización interna. La reflexión sobre las dimensiones macrosemántica y macrosintáctica de la base de datos nos avala, finalmente, para obtener un formato de trabajo integrado que debe ofrecer, en su aspecto práctico y en términos informáticos, la mayor familiaridad del interfaz al documentalista de prensa.

La noticia es un texto que debe ser reducido para su proceso documental, por lo tanto, es necesaria una metodología dirigida al análisis del discurso de actualidad con la finalidad de establecer las proposiciones del autor por encima de la superficie textual. Aunque la macroproposición coincide con el lead o sumario y los



titulares, habitualmente, esto no es una regla general, por lo que es preciso un método riguroso que nos permita explicitar, en todo momento, como se ha procedido en la transformación de los textos. La investigadora Regina Amaro ya había utilizado la metodología de Van Dijk para el análisis del discurso periodístico sobre un corpus de cinco textos con resultados y conclusiones que he debatido en otro lugar (19). La aportación de Amaro es esencial por la simplicidad del procedimiento y por el antecedente que supone para la Documentación periodística. En este trabajo, la metodología de Van Dijk se usa a nivel masivo sobre los registros del

universo de referencia con objeto de obtener unidades menores y más operativas. Es decir, la aplicación de este método es inicial, afectando a la construcción del corpus que va a ser analizado y pensamos que los productos que ofrecen van a tardar en poder ser generados automáticamente. Puesto que la dependencia de la tecnología es total y encontramos limitaciones excesivas para la reducción de textos, optamos por partir de los textos ya reducidos mediante un método que explicita la reducción, pero efectuada por el ser humano, a fin de avanzar en la propuesta de una rutina automática que agilice los archivos de recortes.

### **Notas bibliográficas**

- (1) Van Dijk, Teun: op. cit. p. 68-78.
- (2) Amaro, R.: **Contribuição da análise do discurso para à análise documentária: o caso da documentação jornalística**. São Paulo: ECA/USP, 1991. 87 p., anexos.
- (3) Pottier, B.: op. cit., p.142-147.
- (4) Cunha, I.: Estruturação de vocabulário. In: **Análise documentária: a análise da síntese**. /J. Smit (org.). Brasília: Ibict, 1987. p.61-85.
- (5) García Gutiérrez, A.: **Análisis...** op. cit. p. 140-143.
- (6) Idem: **Estructura lingüística de la Documentación: teoría y método**. Murcia: Servicio de Publicaciones de la Universidad de Murcia, 1990. p. 115.
- (7) Prolog ya ha mostrado, suficientemente, posibilidades de asimilar nuestro método, aún imponiendo nuevos parámetros y limitaciones: véanse las recientes tesis doctorales de Simone Bastos Vieira y de Carmen López Rincón (cit. en 1.4), de distinta naturaleza, que vienen a demostrar ese extremo.
- (8) García Gutiérrez, A.: **Lingüística documental...** op. cit. p.101 ss.
- (9) Van Dijk, Teun: **Texto y contexto: Semántica y Pragmática del discurso**. Madrid: Cátedra, 1980. 357 p.
- (10) Idem: La noticia ...op. cit.
- (11) Ibid., p.54.
- (12) Ibid., p.56.
- (13) Ibid., p.61.
- (14) Amaro demuestra la coincidencia entre el resumen documental y la macroproposición global en op. cit. (conclusiones).
- (15) Van Dijk, Teun: **La noticia...** op. cit. p. 89.
- (16) Ibid., p.77.
- (17) Ibid., p.79.
- (18) Schank, R.: **El ordenador inteligente**. Barcelona: A. Bosch, 1986. 281 p.
- (19) Una valoración de las tesis de Amaro en García Gutiérrez, A.: **Análisis documental...** op. cit. p.145-147.