

# *Santo Tomás de Aquino y la Informática*

## *El “Index Thomisticus”*

El pasado año se celebró el centenario de la encíclica *Aeterni Patris*, de León XIII, sobre la doctrina de Sto. Tomás de Aquino, punto de partida de una fecunda revitalización de su pensamiento, y este año se conmemora el centenario de la muerte de S. Alberto Magno, su predecesor y maestro en tantos caminos de vida religiosa e investigación doctrinal. Son, pues, fechas oportunas para dar noticia de la que quizás sea la más grande obra —en algún aspecto al menos— dedicada a Sto. Tomás en estos últimos decenios. Nos referimos al *Index Thomisticus* realizado mediante computadora IBM<sup>1</sup>.

La idea de este *Index* nació en la mente del jesuita P. Roberto Busa, con motivo de su tesis doctoral en la Universidad Gregoriana de Roma<sup>2</sup>. Fueron dos —nos dice él mismo— los motivos determinantes: el primero, la necesidad de un elaborado estudio filosófico-gramatical sobre la terminología y el estilo de Sto. Tomás, antes de adentrarse en la interpretación de su pensamiento; el segundo, advertir cómo el estudio del lenguaje común nos sitúa junto al honta-

1. SANCTI THOMAE AQUINATIS OPERUM OMNIUM INDICES ET CONCORDANTIAE — in quibus verborum omnium et singulorum formae et lemmata cum suis frequentis et contextibus variis modis referuntur quaque — auspice PAULO VI Summo Pontifice — consociata plurimum opera atque electronico IBM automato usui digessit ROBERTUS BUSA S. I. in Gallaratensi Facultate Philosophica Aloisiani Collegii Professor. Frommann-Holzboog, Stuttgart (s. d.).

2. Publicada después con el título *La terminologia tomistica dell'interiorità. Saggi di metodo per un'interpretazione della metafisica della presenza* (Milano, Rocca, 1949).

nar del que surgen los llamados problemas filosóficos, y nos hace percibir una coincidencia, desde la verdad y hacia la verdad, entre los más diversos y contrarios pensadores a pesar de sus notables y notorias discrepancias. Por eso, una profundización en los basamentos filológico-gramaticales del lenguaje de Sto. Tomás serviría para actualizar la presencia de la sustancia de su doctrina en el pensamiento de hoy, al tiempo que ayudaría para armonizar la filosofía moderna con la trascendencia cristiana<sup>3</sup>.

Dada la extensión de la obra del Doctor Angélico y la ambición de los propósitos, sólo un método mecánico haría humanamente posible la empresa, reduciendo al mínimo alcanzable el tiempo de preparación de índices y concordancias, y garantizando el mayor grado de precisión y totalidad.

La idea fue madurando en los tres años siguientes, y con las adhesiones, ánimos, ilustraciones, consejos y colaboraciones recibidas de diverso tipo, cristalizó en proyecto definitivo. En 1949 se iniciaron en Gallarate (donde el P. Roberto Busa era profesor en la Facultad de Filosofía) los análisis de los textos tomistas con máquinas IBM; se continuaron durante dos años, a partir de 1967, en el Centro Nazionale Universitario di Calcolo Elettronico de Pisa; luego, durante quince meses, desde 1971, en el Centro di Ricerca della IBM Italia en Venecia. Con los técnicos y compiladores colaboraron miembros de la Comisión Leonina (autores de la edición crítica de las obras de Sto. Tomás), teólogos, filósofos y lingüistas.

En el aspecto más puramente cuantitativo diremos que el *Index Thomisticus* contiene 1.700.000 líneas de texto, para un total de 10.600.000 palabras, que recogen el vocabulario íntegro de lo que podemos llamar el *Corpus thomisticum*. Es decir, de todas las obras de Sto. Tomás, y de aquellas otras en que algún momento se le atribuyeron. Como es sabido, casi a raíz de la misma muerte del Sto. Doctor, y durante los siglos posteriores hasta el nuestro inclusive, se incluyeron, al ritmo de aparición de nuevos manuscritos o por con-

3. R. BUSA: *L' "Index Thomisticus"*, "Sapientia Aquinatis" (Communicationes IV Congressus Thomistici Internationalis, Romae 1955), pp. 601-602; *Id.*: *L' Index Thomisticus e l' informatica filosofica*, "Revue Internationale de Philosophie" 27 (1973) pp. 31-32; *Id.*: *L' Index Thomisticus. Contenuto, finalità, prospettive*, "La Civiltà Cattolica" (1974), I pp. 253-254.

veniencias didácticas, no sólo obras de autenticidad dudosa, sino otras de autores distintos, unos anónimos y otros cuya paternidad hoy está ya comprobada; o se completaba un libro o una exposición que Sto. Tomás dejó inacabados o que faltaban en un manuscrito, con un texto que un contemporáneo o discípulo escribía para este fin, o que se tomaba de otro autor. Todas ellas —auténticas, dudosas y ciertas de otros autores— suman un total de 179 obras.

De ellas, 100 son indudablemente auténticas de Sto. Tomás. (Cada uno de los cuatro libros de los *Comentarios a las Sentencias* de Pedro Lombardo, y cada una de las tres partes de la *Suma Teológica* más el *Suplemento*, se numeran independientemente, como obras distintas). Comprende este grupo obras propiamente dichas y opúsculos, comentarios y reportaciones (lecciones registradas por estenógrafos). Se recogen también las frases y palabras que el propio Sto. Tomás escribió y tachó en sus escritos y que se encuentran en los manuscritos originales que llegaron hasta nosotros; se agrupan, por comodidad organizativa, como si constituyesen una sola obra.

A todas ellas se añaden otras 18 obras cuya autenticidad tomista aún no está claramente determinada. Son muy breves: 80.000 palabras suman en su conjunto. Por eso no se analizaron separadamente, sino como un grupo último de las obras tomistas, cuyo total alcanza 8.736.000 palabras. Se ha seguido la edición crítica de la Comisión Leonina en cuanto a las obras ya publicadas por ésta o a punto de publicarse. En las que tal revisión crítica todavía no se ha llevado a cabo, se recurre a la edición más reciente. Recordándonos los autores del *Index* que las ediciones críticas no modificaron, en la casi totalidad de los casos, la sustancia de los textos.

De las 61 obras restantes, no auténticas, 31 son de autor conocido y 27 de autor desconocido. Suman en total 1.800.000 palabras, y su composición va desde el siglo IX al XVI: de Aimón de Auxerre (+850) a Tomás de Vio *Cayetano* (+1538)<sup>4</sup>.

Las técnicas de elaboración electrónica empleadas, sumariamente expuestas, son:

— codificación de cada palabra según el tipo de la frase.

4. R. BUSA: Presentación del *Index Thomisticus*, 2. Operis argumentum; Id.: *L' Index Thomisticus. Contenuto, finalità, prospettive*, pp. 251-252.

- su lematización.
- selección de los homógrafos.
- codificación de las palabras, según su “peso” para los varios tipos de concordancia.
- léxico electrónico.
- definición del sistema de homografía y de polimorfía.
- censo de todo el vocabulario, primero a nivel de formas, después a nivel de lemas.
- adaptación de los índices estadísticos a las concordancias.
- articulación de los índices.
- estructuración de los diferentes modos de “concordar” las palabras de un texto.
- contextualización “razonada” automáticamente.
- eliminación de las redundancias, elaborando *ad modum unius* las locuciones de varias palabras (*liberum arbitrium, Beatus Apostolus Paulus*, etc.)<sup>5</sup>.

Cada palabra está acompañada, además de la interpretación y la referencia, por indicaciones de códigos “especificativos”, que indican si se encuentra en citas literales o de sólo el sentido de otros autores; si Tto. Tomás se expresa allí directamente o hace referencia a otros lugares de sus escritos; o si la palabra se encuentra en niveles secundarios de la exposición (como títulos, conclusiones, expresiones de conexión o divisiones de la materia) o en una frase que representa el núcleo principal de lo que quiere comunicar. Pues para determinar cualitativamente el lenguaje de un autor se debe atender a estas condiciones, ya que el valor de las palabras en su vocabulario y en relación con su pensamiento, difieren según tales situaciones; juntamente con su frecuencia de uso, nos indica la importancia que alcanza en su sistema expresivo.

Cada palabra queda también “lematizada”, es decir, codificada según su morfología y designada por su lema: p. e., *est, sunt, erat, fuit* están acompañadas por códigos que indican la respectiva perso-

5. Presentación del *Index Thomisticus*, 4. *Quid contulerit automatatum usus.*

na y tiempo gramatical, y seguidas del código de indicación del verbo *sum*. Así, las tablas estadísticas presentan dos sumarios: uno a nivel de forma, y otro a nivel de lema. La homografía de cada palabra está igualmente definida y codificada (en sus diferentes direcciones: entre los lemas, con la onomástica y con las enclíticas): p. e., cuando *caro* viene de *caro-carnis* o si proviene de *carus-a-um*. También la polifonía está analizada y codificada.

Finalmente, todas las palabras han recibido un “código de concordancia” que las distingue según la relación entre su cualidad o tipo semántico y su frecuencia.

El *Index Thomisticus* consta de dos secciones principales: una de *índices* y otra de *concordancias*, respondiendo a los dos componentes —estadístico y analítico— de un completo inventario lingüístico. Los índices recogen todos los medios expresivos que articulan el vocabulario tomista en sus diversas dimensiones de léxico, morfología y sintaxis. Las concordancias son el elenco de todas las frases que contienen una misma palabra, condicionada ésta por la estructura y el sentido de la frase.

Los *índices*, a su vez, se dividen en dos grupos: *índices recapitulativos* e *índices particulares* de cada una de las obras. Los índices recapitulativos son: a) de distribución, b) de vocabulario y c) de frecuencia. El de vocabulario comprende cinco sumarios, y el de frecuencia, nueve agrupaciones.

La *sección de concordancias* se organiza en dos ciclos, que se refieren, respectivamente, a las 118 obras de Sto. Tomás, y a las 61 obras restantes de otros autores. Cada uno de estos ciclos ordena las palabras de sus textos en cinco tipos de concordancia, de tal modo que sólo esté presente en uno de ellos.

El *Index Thomisticus* ofrece así al estudioso una cantidad ingente de documentación en un amplio abanico de direcciones. Sirve, en primer lugar y sobre todo, para un mejor conocimiento de las obras y el pensamiento de Sto. Tomás de Aquino, fundamentado en el análisis filológico y léxico de sus textos. “Un medio —se nos dice— que permite traducir el vocabulario antes de traducir las obras”<sup>6</sup>. O di-

6. *Ib.*, 3. Operis propositum et utilitas, 3.4.1.

cho de otra manera: "El análisis léxico debe preceder a la interpretación filosófica"<sup>7</sup>. No sólo permite establecer el vocabulario tomista esencial de su filosofía y teología, sino también léxicos sectoriales de otras disciplinas, como las ciencias naturales y la pedagogía, por ejemplo. Pero no sólo de las obras de Sto. Tomás. Tratándose de materias idénticas puede resultar una ayuda importante para resolver cuestiones de crítica textual, autenticidad, cronología y estilística de otros autores. También permitirá profundizar, completar y sistematizar el estudio del latín culto medieval, y aún el del latín clásico. Una parte importante de este *Index* lo constituye el *Lexicum electronicum latinum*, cuya composición supuso registrar en banda magnética 90.000 lemas del *Lexicum Totius Latinitatis*, de Forcellini.

Todavía más. Puede servir —por razón de sus estructuras y de su método— en estudios tanto de lingüística pura como de lingüística aplicada, aun tratándose de categorías lingüísticas tan diferentes a las tomistas, como puedan ser los textos científicos contemporáneos. En unos y otros hay una parte no pequeña de coincidencias, que confirman las formulaciones comunes a todo lenguaje como tal. De hecho, al tiempo que se preparaba este *Index*, como labor corroboradora de sus métodos y estructuras, se realizaban trabajos colaterales en otras lenguas (hebreo, arameo, nabateo, griego, inglés, italiano, ruso, alemán), sobre textos que sobrepasaban los cuatro millones de palabras y que abarcaban desde los manuscritos del Qumran a los *Prolegomena* de Kant, textos italianos de diversos siglos, sumarios de física nuclear y artículos rusos de física y química.

A pesar de la abundancia y variedad de su documentación, el *Index* resulta fácil de consultar por su especial estructura: la diversificación de códigos, la lematización de las palabras, la pluralidad articulada de índices y concordancias se pensaron precisamente para hacer más accesible su uso en orden a una información hasta el detalle.

Este *Index Thomisticus* puede considerarse como el primer ejemplo de elaboración electrónica en el campo humanístico, el ingreso de la automación en el tratamiento de una información no numérica:

7. *L'Index Thomisticus. Contenuto, finalità ,prospettive*, p. 250.

---

concretamente en los análisis lingüísticos. Por primera vez se afrontaba y realizaba felizmente un trabajo de esta envergadura sobre un material de tan amplias dimensiones, como son las obras de Sto. Tomás de Aquino. Tal tipo de trabajos se multiplicarían después, ampliándolos a nuevas disciplinas. Aprovechando en muchas ocasiones las experiencias acumuladas en aquella empresa y aplicando los métodos allí empleados. Hoy la informática instrumenta prácticamente todos los ramos del saber.

El doctor Angélico se hizo presente así, como adelantado, en una de las técnicas más depuradas de nuestros días.

FERNANDO SORIA