

MANUAL DE ANÁLISIS INFOASISTIDO DE TEXTOS APLICADO AL TEATRO DE LOS SIGLOS DE ORO

Ricardo Serrano Deza

(Madrid: Universidad Nacional de Educación
a Distancia, 2001, 178 págs.)

Muchos y muy variados son los manuales dedicados al análisis de textos literarios desde diversas perspectivas metodológicas, pero son muy pocos los especializados en el análisis informatizado de textos en lengua castellana. Esta nueva contribución a un campo tan prometedor como aún inexplorado nos llega, una vez más, de la mano de la UNED. Una vez más, decimos, porque no es la primera ocasión en que esta Universidad se vuelca en el estudio de la relación entre literatura y nuevas tecnologías. Ya en el año 1996, el Instituto de Semiótica Literaria, Teatral y Nuevas Tecnologías, dirigido por el doctor Romera Castillo, dedicó su VI Seminario Internacional al tema *Literatura y multimedia*. El resultado de aquel encuentro fueron unas actas pioneras en este campo dentro del panorama de la literatura hispánica, recogidas en José Romera Castillo *et alii* (eds.), *Literatura y multimedia* (Madrid: Visor Libros, 1997).

Un segundo impulso prestado por la UNED a la investigación de las relaciones entre literatura y nuevas tecnologías, gracias a la iniciativa del profe-

sor Romera, lo constituyó la publicación del libro de Carlos Moreno Hernández, *Literatura e hipertexto. De la cultura manuscrita a la cultura electrónica* (Madrid: UNED, 1998). Y por último, este *Manual de análisis infoasistido de textos*, que esperamos venga a fomentar posteriores investigaciones en un campo en el que queda mucho por hacer.

Ricardo Serrano Deza es licenciado en filosofía del lenguaje por la Universidad de Salamanca, *maestro* en etnolingüística por la Universidad de Montréal y doctor en educación comparada por esta misma universidad. En la actualidad dirige el Departamento de Lenguas Modernas de la Universidad de Québec à Trois-Rivières. Su investigación se mueve fundamentalmente en dos direcciones: el teatro del Siglo de Oro y la aplicación de las nuevas tecnologías al estudio de la literatura. La confluencia de estas dos direcciones ha dado lugar al presente manual, cuyo origen se encuentra en unas conferencias impartidas en la Universidad de Salamanca (*Informática y lenguas modernas*, 1994, y *Análisis infoasistido de textos*, 1995) y en el Instituto de Estudios Superiores de Monterrey (1997). Este nacimiento entre seminarios se refleja en la organización del libro, repartido en cinco jornadas que aluden tanto a las jornadas de trabajo como a las de la división clásica del teatro áureo.

¿Cuáles son los beneficios de aplicar la informática al análisis del teatro del Siglo de Oro? En palabras del propio autor del libro, son fundamentalmente dos: *materialidad* y *cantidad*. Es decir, la posibilidad de tratar de entrada el texto material de manera global y la posibilidad de trabajar en bloque sobre un corpus amplio, como pueden ser un autor completo o un subgénero. Ventajas nada desdeñables si tenemos en cuenta la enorme extensión del corpus áureo en general y de la obra de algunos comediógrafos en particular, un corpus difícilmente abarcable en su totalidad si no se aborda con ayuda informática.

En cuanto a lo que respecta al manejo de este manual, se trata de una herramienta de trabajo compuesta por un disquete y un libro. El disquete contiene el programa *Brocense* (en dos versiones distintas, para Windows y para MacOS); el libro contiene las instrucciones necesarias para moverse por el programa y propone diversos ejercicios que se pueden realizar con los textos. No se asuste el lector ante la aparente complejidad que una primera mirada al texto puede hacer suponer: no se le pedirán conocimientos específicos más allá de los presentes en cualquier usuario habitual de un ordenador, ni tampoco un *software* especializado. Las instrucciones del autor y un mínimo de disposición llevarán al lector de la mano hasta donde quiera llegar.

Otro elemento de utilidad para quien no se haya adentrado anteriormente en este campo y desee profundizar en él son las referencias bibliográficas

básicas que Serrano Deza ofrece al final del libro. En ellas recoge textos clásicos del análisis estadístico de textos, como los de Muller o Brunet, complementados con posteriores aportaciones de diversos desarrollos metodológicos del análisis infoasistido de textos. También hace referencia a los principales programas de análisis de texto existentes: *Brocense*, herramienta base de este libro, *Hyperbase* y *FreeText*, remitiendo al lector a las direcciones donde dichos programas pueden descargarse. Toda una invitación a continuar nuestras investigaciones por este camino.

El programa *Brocense* viene equipado con un corpus de cinco textos de teatro áureo sobre los que comenzar a trabajar: piezas de Calderón, Cervantes, Tirso y Ruiz de Alarcón. Para ampliar este repertorio ofrece además una lista de direcciones en Internet que van desde la Real Academia Española hasta la Universidad de Québec o el Centro de Estudios Cervantinos de la Universidad de Alcalá, donde podemos proveernos de nuevos textos digitalizados.

Una vez instalado, el programa *Brocense* nos ofrece cuatro módulos en los que trabajar, bajo los nombres de *formatos*, *grupos*, *presencias* y *conversaciones*. En cada uno de estos módulos podemos insertar el texto que vamos a convertir en materia de análisis y someterlo a diversas operaciones. En todos ellos el texto aparece dispuesto en fichas numeradas por las que podemos movernos hacia adelante o hacia atrás.

El módulo *formatos*, que es el más sencillo, nos permite por ejemplo visualizar el texto a renglón corrido o separarlo por versos, numerar los versos o segmentar el texto según diversos criterios. Dos de los criterios de segmentación propuestos por Serrano Deza como alternativa a la segmentación tradicional por escenas, y que parecen particularmente sugestivos, son la segmentación por grupos de presencia y la segmentación por inicio de forma métrica. Es decir, una división basada en la reconfiguración del grupo de personajes en el juego de la escena, o bien en cada cambio significativo de forma métrica que presente cierta constancia.

Otra de las aplicaciones del módulo *formatos* nos da la opción de contabilizar el número de veces que una determinada forma léxica aparece a lo largo del texto, remitiéndonos de forma automática a los versos que la contienen. Partiendo de esta función sencilla podemos internarnos en análisis más complejos sobre la distribución de campos léxicos y semánticos en las obras. Es fácilmente comprobable, por poner un ejemplo, cómo en *La verdad sospechosa* el uso de la raíz «verdad» tiende a aumentar con el avance del texto, mientras que va disminuyendo el uso de la raíz «mentira».

Por otra parte, los módulos *presencias* y *conversaciones* nos permiten trabajar con series de personajes y con las relaciones que se dan entre ellos:

¿cuál es la frecuencia de réplicas de determinado personaje? ¿y su frecuencia de aparición? Podría darse el caso (y de hecho, se da) de que personajes aparentemente poco significativos respecto al número de palabras pronunciadas nos sorprendan por su gran importancia al pasar a la consideración de su presencia. Pero sigamos preguntándonos: ¿existe alguna correlación entre la presencia de unos personajes y la ausencia de otros? ¿Podemos establecer parejas conversacionales de hablante-oyente con una frecuencia de aparición significativa? En la reducción a datos estadísticos fiables de éstas y otras preguntas se encuentra la base de toda investigación rigurosa.

A partir de aquí, todo depende de la curiosidad del lector, o si se quiere, de su capacidad de hacerse preguntas: la herramienta para responderlas está en sus manos. Así pues, conecte el ordenador e investigue. Los resultados merecen la pena.

Irene Aragón González