

JOSSE DE KOCK

(en colaboración con Dirk Geens, Nicole Delbecque, Carmen Gómez
Molina y Robert de Keyser)

Elementos para una estilística computacional

Madrid, Edt. Coloquio, 1984. Col. "Temas y Monografías"

2 vols. (t. I 176 págs.; t. II 192 págs.)

Felipe Gómez Solís

Los estudios sobre la estilística cuantitativa están adquiriendo un desarrollo cada vez mayor, especialmente en la investigación anglosajona, que trabaja con métodos lingüísticos, cuantitativos y automatizados desde los años cincuenta.

Dentro del dominio románico, España —detrás de Francia e Italia— va iniciándose en la aplicación de las computadoras a los trabajos lingüísticos y literarios, sobre todo por la labor que ha realizado el Centro de Cálculo de la Univ. Complutense (CCUC) a través de Manuel Alvar Ezquerro, quien ha obtenido un organigrama de índices de rimas de manera automática y programado de una forma coherente la realización de concordancias de textos literarios medievales, como la *Vida de Santa María Egipcíaca*, *Vida de San Ildelfonso e Infancia y muerte de Jesús* (Vid., además, su *Proyecto de lexicografía española*, Barcelona, Planeta, 1976, pp. 203-230; "Obtención automática de índices de rimas y de sufijos", en *Homenaje a Vicente García de Diego*, *Revista de Dialectología y Tradiciones Populares*, XXXII, 1976, pp. 35-41, y "Algunos rasgos léxicos de Berceo y su cotejo con otros poemas hagiográficos", en *Anuario de Letras*, XVI, 1978, pp. 251-260). Trabajos, como el de Manuel Alvar (*Informática y lingüística*, Málaga, Librería Agora, 1984, *Cuadernos de Lingüística*, 1.), muestran las ventajas de la automatización (eficacia, rapidez y economía) en la obtención de un diccionario automatizado del español actual, en la geografía lingüística, como el proyecto del Atlas plurilingüe de Europa (ALE), o el *Atlas Lingüístico y Etnográfico para la provincia de Santander* (ALES), en colaboración con María Pilar Nuño. No hay que olvidar, por otra parte, los estudios de Julio Fernández-Sevilla ("Ordenadores electrónicos y atlas lingüísticos", en la *Utilización de ordenadores en problemas de lingüística*, *Revista de la Universidad Complutense de Madrid*, XXV, 1976, pp. 88-100, y *Formas y estructuras del léxico agrícola andaluz*, Madrid, 1975), E. García Camarero ("Lingüística matemática en el Centro de Cálculo de la Universidad de Madrid", en *RSEL*, I, 1, 1971, pp. 169-172, y "Algunos textos lingüísticos del CCUM", en *RSEL*, III, 2, 1973, pp. 462-463) y Víctor Sánchez de Zavala ("El se-

minario de "lingüística matemática" del Centro de Cálculo de la Universidad de Madrid", en *RSEL*, I, 1, 1971, pp. 173-176). Además, la revista *Lingüística Española Actual* (LEA) ha dedicado un número monográfico (III, 2, 1981) a la aplicación de las máquinas a la lingüística, como en la fonética y fonología (Antonio Quilis, "El empleo de los ordenadores en la investigación fonética", pp. 197-220; Josse De Kock en colaboración con A. R. Lecours, W. Brems y J. Klüssendorf, "La glosolalia. Análisis automatizado de los fonemas de una gran jerga", pp. 221-274), la lexicología, lexicografía y semántica (Amparo Morales de Walters, "Los diccionarios de frecuencia y las calculadoras electrónicas", pp. 275-284; E. García Camarero, "Algunos datos cuantitativos del español", pp. 337-358), y Jean Roudil, "Del tratamiento automático de los textos medievales españoles al análisis semántico y a la estructura del léxico: Pasos principales", pp. 285-304), la geografía lingüística (el citado trabajo de Manuel Alvar y María Pilar Nuño, "Un ejemplo de atlas lingüístico automatizado: El ALES", pp. 359-374) y la estilística (Josse De Kock, "Método para una posible automatización del análisis estilístico", pp. 305-336).

El libro de Josse De Kock, que ahora reseñamos, *Elementos para una estilística computacional* es una obra indispensable para conocer y, a la vez, profundizar en la metodología de la estilística cuantitativa y automatizada. Tal disciplina —que trabaja de forma totalmente diferente a la estilística clásica de Ch. Bally— ha dado lugar a fuertes reacciones dentro de la crítica literaria, ya que no puede llegar a fijar la definición de lo literario, y por tanto, no se ocupa del aspecto connotativo. J. de K., consciente de las limitaciones de los procedimientos de la estilística cuantitativa, reconoce tales límites y observa cómo la metodología, además, está lejos de su punto de desarrollo máximo, y con un buen criterio afirma: "Tampoco se puede pedir a la estilística cuantitativa lo que no puede dar" (t. II, p. 64) o "El instrumento estadístico es un medio de investigación primordial, a condición de no atribuirle más poder del que puede tener" (t. II, pp. 65-66). Frente a este humilde reconocimiento de las limitaciones, J. de K. piensa que el mayor riesgo del empleo de un método cuantitativo en el análisis estilístico sería que la estilística se dejara absorber por la lingüística, aunque debe haber colaboración entre disciplinas diferentes, como la lingüística, la literatura, la estadística y la informática.

Elementos para una estilística computacional constituye la culminación de una serie de estudios de Josse De Kock dedicados a los trabajos estadísticos dentro de la lingüística y la estilística (véanse su obra, ya clásica, *Introducción a la lingüística automática en las lenguas románicas*, Madrid, Gredos, 1974, los artículos "Por una estilística lingüística y cuantitativa con la ayuda del ordenador", en *Revista de la Universidad Complutense*, XXV, 102, 1976, pp. 113-129, en colaboración con D. Geens, y el citado de *Lingüística Española Actual*, o el reciente trabajo "La noción de auxiliaridad. Las construcciones intransitivas con adjetivo verbal", en *Philologica Hispaniensia (in honorem Alvar)*, II, Madrid, Gredos, 1985, pp. 283-304).

El libro consta de dos volúmenes totalmente diferenciados. En el primero, J. de K. ilustra un ejemplo práctico de la estilística basada en métodos cuantitativos y automatizados, tomado de Miguel de Unamuno. En el segundo, se contiene una enciclopedia, que es una clara metodología de tal estilística (pp. 9-66), con varios apéndices (se incluyen concordancias de poemas y artículos periodísticos de Unamuno del ejemplo práctico, t. I, pp. 67-134), una amplia y útil bibliografía (pp.

135-182) y una relación de textos españoles medievales, modernos y contemporáneos automatizados (pp. 183-187).

Empecemos, pues, con la enciclopedia que, a nuestra manera de entender, debería ir en el tomo I. Dentro de la metodología de la estilística cuantitativa, se explican, en primer lugar, algunos conceptos básicos, como la `variación`, la `selección` y la `desviación` en relación con la comparación de la dimensión extratextual (se hace una crítica del término `norma` y se prefiere hablar de `punto de comparación`); así, para J. de K. "la estilística, cuando se funda en comparaciones múltiples, no debe limitarse al lenguaje `literario`, puesto que difícilmente puede concebirse que éste último sea un lenguaje descriptible en sí; hay que estudiarlo refiriéndose a otras modalidades del lenguaje" (t. II, p. 13). En segundo lugar, J. de K. analiza los instrumentos básicos de la estilística cuantitativa, como son el *índice* (lista exhaustiva de formas con referencia al lugar de aparición en el texto), la *concordancia* (incluye una sección de línea, línea entera o incluso varias líneas, donde la palabra figura), la *lematización* (consiste en reducir las formas declinadas y conjugadas a formas básicas o radicales, artículos de diccionarios) y los *programas*. En tercer término, se exponen las aplicaciones de los índices y concordancias en el campo de la lexicología y semántica, morfosintaxis y métrica.

El mayor número de trabajos estilísticos automatizados se halla en el área del léxico y el estudio cuantitativo del vocabulario se llama *lexicometría*, que relaciona el vocabulario y la extensión de textos, para poder obtener la riqueza lexical de una obra. Por otra parte, se pueden analizar palabras claves de un tipo de poesía o prosa, tanto en textos modernos como antiguos (F. M. Waltman ha hecho un análisis lexicométrico del *Poema del Mío Cid* y afirma que del empleo de las fórmulas no se puede derivar que haya dos autores). Además, con los métodos de la estilística cuantitativa se puede profundizar en la teoría de los campos semánticos, en la atribución de obras (se han hecho grandes progresos y la mayoría de los trabajos de atribución conciernen a obras clásicas. La estilística cuantitativa sería muy útil para atribuir obras de nuestros siglos XVI y XVII) y en el influjo de un autor sobre otro.

Los estudios estilísticos lingüísticos cuantitativos, que se sitúan en el nivel morfosintáctico, son pocos frecuentes. Tan sólo se limitan a la cuantificación de las clases de palabras, sus funciones gramaticales, extensión y complejidad de la oración y la frase, y son realizados desde un punto de vista distribucionalista o generativista, especialmente en el dominio anglosajón. También la estilística cuantitativa se adentra en la lingüística del texto de Van Dijk (por ejemplo, R. Pellen aplica la lingüística del texto en el *Poema del Mío Cid*).

La métrica se somete al ordenador con facilidad. Así, se ha podido automatizar la métrica de las lenguas clásicas, que estriba en la posición de las vocales.

El ejemplo práctico consiste en comparar 1762 poemas de desigual extensión del *Cancionero* de Unamuno, compuestos durante los siete últimos años de su vida (1928-1936) y 212 artículos de periódicos, publicados entre junio de 1931 y julio de 1936 en *El Sol* (Madrid), *Ahora* (Madrid), *El Adelanto* (Salamanca), *El Norte de Castilla* (Valladolid), *Heraldo de Aragón* (Zaragoza), *La Voz de Guipúzcoa* (San Sebastián) y *Caras y Caretas* (Buenos Aires). Además, J. de K. compara sus resultados con los obtenidos por A. Juilland y E. Chang-Rodríguez en *Frequency Dictionary of Spanish Words* (La Haya, 1965). Aparecen listas alfabéticas de las 500 formas más frecuentes en *Cancionero*, artículos de periódicos y el diccionario de frecuencia de A. Juilland y E. Chang-Rodríguez con los valores de la diferencia de

la frecuencia relativa, se estudian verbos concretos, como *haber* y *ser* y se llegan a unas conclusiones, como la extremada recurrencia de la clase de los sustantivos, del pronombre personal de segunda persona (con la forma átona *te*) y la conjunción temporal *mientras (que)* en el *Cancionero*. Se hace, además, un análisis de los formantes constitutivos del verbo, como el tiempo, modo y el número: existe escasez del imperfecto, falta de formas del modo subjuntivo, ausencia de oraciones condicionales, frecuencia del indicativo, abundancia de la segunda persona singular del llamado modo imperativo (formas de cortesía) y frecuencia más alta de la segunda persona singular (la primera se reemplaza por la tercera y ésta por la segunda).

La estilística con métodos cuantitativos y automatizados no se opone a la estilística clásica, y en todo caso una y otra sirven para describir el hecho literario. No se excluyen, sino que se complementan. *Elementos para una estilística computacional* de Josse De Kock —en colaboración con Dirk Geens, Nicole Delbecque, Carmen Gómez Molina y Robert de Keyser— abre nuevos caminos de investigación en el campo de la estilística cuantitativa y automatizada, que en nuestro país está empezando a desarrollarse, como lo demuestran los análisis lingüístico-cuantitativos aplicados a escritores contemporáneos, como A. Machado, Pedro Salinas, L. Cernuda, César Vallejo, Juan Rulfo, Neruda, entre otros.