

El formalismo cuantitativo contemporáneo: entre la tradición y la innovación

Este año se cumplen diez años de la publicación de dos libros de referencia de los estudios literarios modernos: *Distant Reading* de Franco Moretti y *Macroanalysis: Digital Methods and Literary History* de Matthew Jockers. En estos textos, los dos miembros del emblemático Laboratorio Literario de Stanford hicieron una síntesis de su trabajo de años en el campo de los métodos digitales y cuantitativos aplicados al estudio de la literatura, reflexionaron sobre las perspectivas futuras de estos métodos y consolidaron las bases del nuevo “formalismo”, que aspiraba entonces a volverse una rama autónoma y poderosa de las humanidades digitales.

La propuesta metodológica de la lectura distante (*distant reading*, en oposición a la lectura cercana, *close reading*), formulada y promovida por Moretti, planteaba pasar del estudio meticuloso de unos pocos textos canónicos al análisis, a partir de grandes conjuntos de datos, de las unidades menores o mayores que un texto individual (ya sean los procedimientos, temas, tropos, géneros o sistemas literarios). El acceso a extensas colecciones de documentos digitalizados, acompañados por metadatos, debía permitir el descubrimiento y el análisis de patrones y estructuras y la identificación de proporciones y cambios, transformando así nuestra comprensión de la evolución literaria.

La propuesta de Matthew Jockers estaba cerca de la de Moretti, aunque Jockers prefería hablar de “macroanálisis” en lugar de la provocadora “lectura sin leer”. Según Jockers, los estudios literarios no podían ignorar la transformación digital. La capacidad de extraer información de miles de textos disponibles en bibliotecas digitales debía transformar nuestra forma de investigar y argumentar, pues las pruebas anecdóticas, las intuiciones y las generalizaciones basadas en pequeñas muestras ya no podían tenerse por suficientes.

En 2011, tanto Moretti como Jockers firmaron un manifiesto del Laboratorio Literario de Stanford (del cual eran cofundadores) llamado “Quantitative



Formalism: An Experiment”. La expresión “formalismo cuantitativo” señalaba la presencia de una influencia teórica del formalismo, pero, al mismo tiempo, recalca una innovación: el formalismo del Laboratorio Literario de Stanford era cuantitativo, mientras que se creía que el formalismo del siglo xx era cualitativo. La idea de que el formalismo cuantitativo era algo nuevo, producido por la revolución tecnológica, ha sido puesta en tela de juicio desde entonces varias veces: en los estudios literarios, la revolución digital hizo que los investigadores empezaran a redescubrir, a la luz de las nuevas tendencias, propuestas metodológicas de hace décadas que iban en la misma dirección, pero que no habían tenido mucha popularidad y difusión.

Este tema fue discutido por primera vez durante el congreso “Russian Formalism and the Digital Humanities”, que se celebró en la misma universidad de Stanford en abril del 2015. En 2019, el *Journal of Literary Theory* publicó un número monográfico titulado *Moscow Formalism and Literary History*, en el cual aparecieron traducidos al inglés importantes trabajos precursores del formalismo cuantitativo contemporáneo. En la introducción, los editores Frank Fischer, Marina Akimova y Boris Orekhov cuentan que la idea de publicar un volumen especial de textos inéditos en inglés de los formalistas rusos surge a raíz del creciente interés por la cuantificación en los estudios literarios digitales. En el verano del 2021, apareció el número monográfico *Digital Humanities and Russian and East European Studies* de la revista *Russian Literature* (Ámsterdam), en el cual los artículos de Ilya Kliger, Basil Lvoff y Pavel Arseniev trataban sobre la lectura distante y el legado del formalismo.

Este tema se discutió también en las páginas de esta revista. En el volumen 24 (núm. 1 del 2022), en su artículo “Formalismo cuantitativo ‘viejo’ y ‘nuevo’ (El Círculo Lingüístico de Moscú y el Laboratorio Literario de Stanford)”, uno de los editores del presente volumen exploró la relación entre la tradición del formalismo ruso y las humanidades digitales modernas. Se presentó por primera vez en lengua hispana el legado “cuantitativo” de Borís Yarzó (1889-1942) y Borís Tomashevski (1890-1957), miembros del Círculo Lingüístico de Moscú (1915-1924), en cuyas investigaciones se anticipan temas y metodologías del siglo XXI.

Consideramos, pues, que el formalismo cuantitativo es un fenómeno que puede considerarse “nuevo” y “viejo” a la vez. Por un lado, su surgimiento se debe, en gran medida, al giro digital y al desarrollo tecnológico de las últimas décadas, lo que ha permitido ampliar las posibilidades de análisis en este campo. Por otro lado, sus prácticas críticas tienen una larga historia

y muchos precursores, tanto dentro del formalismo y sus herederos teóricos (estructuralismo, semiótica cultural) como en disciplinas como la filología clásica, la lingüística computacional, la traductología, la versología y la estilometría.

El presente número monográfico no tiene como propósito hacer una arqueología de las ideas, abogar por los estudios literarios computacionales o defenderlos contra las críticas, que no son pocas (ver, por ejemplo, “The Computational Case against Computational Literary Studies” de Nan Z. Da, publicado en *Critical Inquiry* en 2019, y las múltiples respuestas a este texto). Su objetivo es, más bien, mostrar, a partir de una selección de artículos, lo que pueden hacer —y lo que hacen efectivamente— hoy los estudiosos de la literatura con el apoyo de los métodos cuantitativos y digitales. Aunque nuestro propósito no es polémico, creemos que los mismos trabajos podrán rebatir la idea según la cual el análisis cuantitativo solo es capaz de demostrar lo obvio. Después de diez años de la publicación de los libros de Moretti y Jockers, es hora de establecer los métodos de investigación digitales como parte integral de la caja de herramientas de los estudiosos de la literatura. Ya no se trata de una novedad, de una moda pasajera, sino de una práctica consolidada.

Al mismo tiempo, hay que reconocer que el formalismo cuantitativo literario aún constituye una novedad en Colombia y, en general, en América Latina. Los avances importantes en esta área, como, por ejemplo, el volumen *Humanidades Digitales 4: Corpus y Literatura en México*, editado por Ester Bautista e Ignacio Rodríguez y publicado en 2021, todavía no logran una gran difusión y un posicionamiento dentro del campo dinámico y diverso de las humanidades digitales. En este sentido, el objetivo de este número monográfico no solo es presentar un balance, sino también dinamizar el campo de los estudios literarios cuantitativos en América Latina.

El número, y los temas que en él se abordan, invita a reflexionar sobre algunos aspectos propios del nuevo formalismo cuantitativo. Uno de los rasgos más notables de este movimiento es su carácter internacional y transcultural. A diferencia de lo que sucede en muchos campos de la literatura y la filología, en los que las fronteras lingüísticas son poco permeables, los humanistas digitales buscan desarrollar metodologías aplicables a cualquier material literario, sin importar su origen geográfico o lingüístico. Este número es un claro ejemplo de ello, con la participación de investigadores que trabajan en Alemania, Argentina, Colombia, Estados Unidos, Estonia, Italia, República Checa, Rusia y Suiza.

El material literario analizado es también muy variado. Las contribuciones abarcan desde la literatura clásica, medieval y renacentista hasta el siglo xx, y cubren múltiples ámbitos lingüísticos, incluyendo el hispano, el portugués, el italiano, el ruso y el checo. Esta variedad cultural y lingüística demuestra que las humanidades digitales no necesariamente son una imposición del mundo angloamericano orientada a la producción masiva de *papers*, como se dice a menudo: de hecho, sus orígenes intelectuales están bien lejos de la academia anglófona y las investigaciones que se hacen en este campo pueden vincularse íntimamente con las tradiciones académicas nacionales, no obstante su carácter globalizado.

Aunque los textos que componen este número son diversos en sus temas y enfoques, existe una profunda conexión entre ellos. Esta conexión va más allá del uso de métodos digitales y cuantitativos, que si bien pueden tener un mayor o menor peso en cada contribución, no son la única ligazón que los une. La unión más profunda se encuentra en el compromiso de los investigadores con un programa teórico que se puede resumir a través de la siguiente serie de dicotomías:

texto	vs.	contexto / intertexto
textos individuales	vs.	conjuntos de textos / literatura en su conjunto
lectura cercana	vs.	lectura distante
enfoque cualitativo	vs.	enfoque cuantitativo

En este número monográfico, los elementos de la segunda columna son los que cobran protagonismo. Como veremos más adelante, en lugar de centrarse en obras individuales, los autores revisan conjuntos de textos o piensan la literatura en su conjunto, practican la lectura distante y buscan datos cuantitativos para respaldar sus análisis. Por supuesto, se trata de la orientación principal y no de una exclusión total de los enfoques cualitativos. A menudo los dos programas críticos se complementan, pues la lectura cercana y lo que la acompaña es fundamental para la lectura distante y viceversa.

Al momento de abrir la convocatoria, esperábamos recibir textos dedicados a cualquier género o forma discursiva. Sin embargo, es claro que prevalecieron contribuciones sobre textos en verso: es probable que los textos poéticos, orientados a la explotación y ordenación de diversos niveles lingüísticos y, por

tanto, más ordenados en su estructura, requieran, casi que de forma orgánica, de los métodos cuantitativos. Las investigaciones de este tipo se benefician, además, del conocimiento de las teorías formalistas y estructuralistas. El análisis cuantitativo de la prosa aún está descubriendo sus métodos, como el macroanálisis de los indicadores de etnia y género, el modelado temático, el “modelado predictivo” de la correlación entre el cambio del prestigio literario y el cambio de las formas literarias o la construcción de modelos de grafo de los géneros literarios.

Los tres artículos que abren el número tratan problemas de poética comparada y de traductología. En “¿Cómo evaluar la fidelidad de una traducción? Métodos cuantitativos en el estudio de las traducciones poéticas”, Vera Polilova se propone revisar y presentar los métodos de evaluación cuantitativa de la exactitud léxica y de la exactitud rítmica de las traducciones en verso, y describe varios enfoques prácticos desarrollados para afrontar este problema en la academia de habla rusa de los siglos xx y xxi. La investigadora también hace un análisis cuantitativo y comparativo de tres traducciones métricas al ruso del poema “Reyerta” de Federico García Lorca (de Valentín Parnaj, Anatoli Gueleskul y Pável Grushkó). El artículo incluye anexos que muestran los pormenores de la aplicación de los métodos empleados.

En su trabajo “La repetición en las traducciones homéricas: consideraciones teóricas y análisis de casos”, Alejandro Abritta escoge un procedimiento clave y revisa su funcionamiento en el texto original y en sus traducciones. Parte de la importancia fundamental de las repeticiones en la poética homérica para revisar, después, cómo se transmite este rasgo en distintas traducciones al español (de Luis Segalá y Estalella, Emilio Crespo Güemes y Javier Pérez). Luego de examinar tres modalidades distintas de repetición (pasajes completos de varios versos, fórmulas y una expresión recurrente) en tres traducciones, el autor identifica estrategias de cada traductor y caracteriza las traducciones.

Un procedimiento particular es también el protagonista del artículo “La disimilitud de lo similar: la rima de la octava real desde una perspectiva comparada, cuantitativa y funcional”. En él, Anastasia Belousova, Juan Sebastián Páramo Rueda y Paula Ruiz Charris estudian las características de la rima en las obras en octava real de Ludovico Ariosto y Torquato Tasso, y en sus traducciones al español (Juan de la Pezuela y Juan Sedeño) y al ruso (Aleksandr Triandafilidi y Román Dubrovkin). Este artículo es, además, un homenaje a la versología comparada moderna, que cumple un siglo en 2023 (en 1923, se publicaron el

libro de Antoine Meillet sobre el origen indoeuropeo de los metros griegos y *Sobre el verso checo, con particular referencia al verso ruso* de Roman Jakobson).

El siguiente texto, “Polimetría en el teatro de Calderón de la Barca: aproximación a un análisis métrico del conflicto dramático y los personajes en sus obras” de Carlos E. Acuña Feijoo, es un trabajo que se inscribe dentro la poética cuantitativa y se centra en uno de los principales rasgos característicos del teatro del Siglo de Oro: la polimetría. El autor se ocupa de la popularidad de las distintas formas métricas en cada drama, de su presencia en el discurso de los personajes y de la relación dinámica entre el cambio de formas métricas y el desarrollo del conflicto dramático.

El siguiente bloque de artículos pertenece al campo de la estilometría y la atribución de autoría. En “Reis melhor do que eu’: los heterónimos de Pessoa desde una perspectiva estilométrica”, Daniil Skorinkin y Boris Orekhov aplican el método Delta de John F. Burrows, que últimamente se ha impuesto como el principal método de atribución de autoría, a los textos escritos por Fernando Pessoa “en nombre propio” y a los que pertenecen a una serie de identidades ficticias que comúnmente se denominan “heterónimos”. Tras una serie de experimentos, se muestra que Delta identifica textos de distintos heterónimos de Pessoa como textos pertenecientes a diferentes autores, es decir, Pessoa logra “engañar” el algoritmo. Este resultado pone de manifiesto la necesidad de mayor prudencia al momento de establecer la autoría a partir del método Delta y nos invita a prestar mayor atención a sus limitaciones. El siguiente texto discute la posibilidad de utilizar datos sobre el ritmo poético para la atribución de autoría. Petr Plecháč y David J. Birnbaum llevan a cabo una serie de experimentos y llegan a la conclusión de que la organización rítmica (y, en concreto, la regularidad acentual) puede funcionar como un rasgo significativo para el reconocimiento de la autoría.

El bloque de artículos se cierra con “Unread, yet preserved: A case study on survival of the 19th century printed poetry”, en el cual Antonina Martynenko indaga en problemas relacionados con la construcción de los *corpus* para la investigación literaria. Con la lectura distante, Moretti pretende leer “el gran sin leer” (*the great unread*, expresión de Margaret Cohen). Sin embargo, el tamaño mismo del *corpus* no garantiza su representatividad. Martynenko medita en cómo podría recopilarse un *corpus* “completo” de los libros de poesía del siglo XIX, teniendo en cuenta los datos históricos y los posibles sesgos de supervivencia. El análisis permite establecer diferentes “tasas de pérdida”, según las distintas

tipologías de fuentes, y también nos hace conscientes de los problemas de representatividad que existen incluso en los *corpus* aparentemente exhaustivos.

En la sección de notas publicamos un sucinto manifiesto, en el cual Igor Pilshchikov presenta un plan de acción para el desarrollo de la poética digital en el campo de la eslavística. Pilshchikov sostiene que es necesario escribir manifiestos de este tipo no solo para coordinar esfuerzos y reflexionar sobre el estado de la disciplina, sino también para superar una situación particular: las humanidades digitales han trabajado principalmente con materiales en inglés, y esto ha llevado a la presencia de sesgos en muchos de sus métodos. Para evitar los errores del “angloglobalismo” es necesario desarrollar programas de humanidades digitales que tengan en cuenta las especificidades de las lenguas nacionales.

En la sección de entrevistas se publican conversaciones con Franco Moretti y con Emil Volek que invitan a una reflexión teórica y, al mismo tiempo, que revelan la dimensión humana, biográfica e histórica de las teorías.

El número se cierra con una traducción y dos reseñas que funcionan, en conjunto, como una especie de estado del arte de los estudios cuantitativos actuales. El artículo del humanista digital italiano Fabio Ciotti “Modelos y métodos computacionales para la crítica literaria: el estado del arte”, traducido por Elena Pérez Palacio, ofrece un panorama nítido y eficaz de la crítica literaria computacional. Presenta, además, respuestas a las críticas más frecuentes en contra de este enfoque. John Alexander Contreras Caicedo reseña el libro de Andrew Piper *Enumerations: Data and Literary Study* (University of Chicago Press, 2018), una de las contribuciones más importantes de los últimos años al campo del formalismo cuantitativo. Laura Camila Gómez Camelo y Valeria Muñoz Landínez, a su vez, reseñan el libro de Petr Plecháč *Versification and Authorship Attribution* (Karolinum/ICL, 2021; con participación de Artjoms Šeļa), una investigación pionera en el campo de la estilometría computacional por su intento de enriquecer con métodos versológicos los estudios de atribución de autoría.

Después de repasar las diversas investigaciones presentadas en este número, quizás a la lectora se le ocurra una pregunta legítima: ¿las humanidades digitales responden a preguntas y problemas ya planteados, o ellas mismas formulan preguntas imprevistas? Para estos editores, la mayoría de los textos que aquí presentamos son del primer tipo: solucionan (o tratan de solucionar) problemas que desde antaño han sido evidentes para los estudiosos de la literatura. Sin embargo, nuestra esperanza no puede ser otra que, a partir de la continua

discusión y elaboración alrededor de estos asuntos (y el presente número participa de ello), esas preguntas insospechadas terminen por aparecer.¹

Anastasia Belousova

Universidad Nacional de Colombia, Bogotá, Colombia

Igor Pilshchikov

Universidad de California, Los Ángeles, Estados Unidos

Universidad de Tallin, Tallin, Estonia

Editores invitados

¹ Una parte del trabajo de preparación de este número monográfico se ha realizado en el marco del proyecto de investigación PRG319, el cual ha recibido el apoyo del Consejo de Investigación de Estonia (ETAG).