## BOLETÍN

DE LA

### REAL ACADEMIA ESPAÑOLA

TOMO XCIX· CUADERNO CCCXIX· ENERO-JUNIO DE 2019

# LA GRACIOSA Y GRATUITA DISPUTA SOBRE LA AUTORÍA DE LA HISTORIA VERDADERA DEL INCONFUNDIBLE BERNAL DÍAZ DEL CASTILLO

RESUMEN: Dedico este trabajo al análisis de la lengua de *Historia verdadera*, de Bernal Díaz del Castillo, y de las *Cartas y Relaciones*, de Hernán Cortés, a fin de determinar si existe un parentesco entre ambos textos –como el antropólogo Christian Duverger conjetura– que nos permita establecer con un grado de probabilidad aceptable que ambos textos salieron de la misma mano. Para ello me serviré de distintos análisis cuantitativos (estilometría) y cualitativos de los dos textos en debate.

*Palabras clave:* Estilometría; Atribución; Determinación de autoría; Cronistas de Indias; Bernal Díaz del Castillo; Hernán Cortés.

# THE FUNNY AND GRATUITOUS DISPUTE OVER THE AUTHORSHIP OF *HISTORIA VERDADERA* OF THE UNMISTAKABLE BERNAL DÍAZ DEL CASTILLO

ABSTRACT: I devote this article to the analysis of the language used in the *Historia* verdadera by Bernal Díaz del Castillo and in the Cartas y Relaciones by Hernán Cortés in order to determine if there exists a relationship between both texts –as the anthropologist Duverger conjectures – which would allow us to establish with an acceptable degree of probability that both texts came from the same hand. For that, I will use a

quantitative analysis (stylometry) and a qualitative analysis (corpus linguistics) from both texts.

*Keywords*: Stylometry; Attribution; Authorship determination; Chroniclers of the Indies; Bernal Díaz del Castillo; Hernán Cortés.

#### i. Planteamiento filológico

Cada autor emplea la lengua común de una manera original, propia e inconfundible; cada hablante actualiza las posibilidades que un idioma ofrece de manera absolutamente individual y propia y, por lo tanto, reconocible; cada hablante tiene su propio idiolecto y éste se manifestará a través de elecciones idiosincrásicas en los textos que produce: un vocabulario preferente; unas concretas inclinaciones para seleccionar unos elementos en vez de otros o para elegir entre dos construcciones alternativas que la lengua ofrece para expresar lo mismo. Los hablantes tendemos a hacer, entre todas las posibilidades del vocabulario y gramática de una lengua, unas elecciones individualizadoras¹ (Hoey 2005). Y hoy (con la ayuda de herramientas informáticas aplicadas a la estilometría) es posible cuantificar los rasgos caracterizadores del idiolecto de un autor dado con índices de probabilidad muy altos en la determinación o atribución de autoría. Estas herramientas, si el procesamiento de los textos es el correcto, garantizan un alto grado de objetividad y constituyen un excelente punto de apoyo para análisis estilísticos más convencionales.

#### 2. Antecedentes

El objetivo de mi trabajo es analizar por procedimientos informáticos la autoría de la *Historia verdadera de la conquista de Nueva España*, obra tradicionalmente editada bajo la firma de Bernal Díaz del Castillo, pero puesta en duda la autoría real de la misma por los estudios de Duverger<sup>2</sup>, a partir de una serie de argumentos que voy a resumir siguiendo las reseñas que de la tesis

<sup>&</sup>lt;sup>1</sup> Michael Hoey, *Lexical Priming. A new theory of words and language*, London: Routledge, 2005.

<sup>&</sup>lt;sup>2</sup> Christian Duverger, *Crónica de la eternidad*, México, Taurus, 2012.

del antropólogo han realizado especialistas en historia y en literatura del siglo xvI tales como Guillermo Serés (impecable editor de Bernal)<sup>3</sup> o Esteban Mira Caballos<sup>4</sup>:

- a) Bernal Díaz era una persona iletrada con un conocimiento limitado para la escritura, e incluso para la firma, por lo que es complicado aceptar que escribiera una obra como la que nos ocupa en la que abundan las referencias letradas.
- b) La *Historia verdadera de la conquista de Nueva España* es obra de Hernán Cortés. El metellinense la redactó durante su estancia en Valladolid, entre 1542 y 1545 (por la mañana atendía sus asuntos en la chancillería y dictaba a Francisco López de Gómara su propia historia, aprovechando las noches, en secreto, para redactar su crónica que dejaría anónima).
- c) Supuestamente esconde su autoría para evitar que, como iba a ocurrir con sus *Cartas de Relación* o con la historia de López de Gómara, los celos de Carlos V impidiesen que el relato de sus hazañas viera la luz o fuera, como sus *Cartas*, retirado de la circulación.
- d) A su muerte, en 1547, su hijo Martín Cortés llevó consigo la *Historia* a México. Pero nunca se publicó.
- e) El manuscrito acabó en Guatemala en manos de Bernal Díaz y éste, con la ayuda de su hijo Francisco, manipuló el original y se atribuyó su autoría.
- <sup>3</sup> Bernal Díaz del Castillo, *Historia verdadera de la conquista de la Nueva España*, edición, estudio y notas de Guillermo Serés. Madrid: Real Academia Española Barcelona: Galaxia Gutenberg-Círculo de Lectores, 2011.
- <sup>4</sup> Esteban Mira Caballos, doctor en Historia de América y autor de una de las últimas biografías documentadas sobre Hernán Cortés (*Hernán Cortés. Mitos y leyendas del conquistador de Nueva España*, Badajoz, Fundación de los Pizarro, 2017), ha puesto en evidencia las contradicciones, falsedades y manipulaciones documentales de la obra de Duverger que atribuía la *Historia verdadera de la Conquista de Nueva España* a Hernán Cortés.

Estas conjeturas<sup>5</sup>, esgrimidas sin apoyo documental alguno, merecen una cuádruple consideración:

- a) por un lado, todas ellas se fundamentan en hipótesis difícilmente documentables;
- b) juntas, dan lugar a una ficción en relación con la cual, en principio, sólo puede afirmarse que ciertamente no deja mucho campo para un debate serio;
- c) tal y como se halla planteada la atribución de Duverger, ésta se convierte en una cuestión de fe, respecto a la cual el público lector necesariamente habrá de dividirse en agnósticos y creyentes. Y sobre las creencias de cada uno es imposible discutir en términos cartesianos, pues, ante la ausencia de documentos sobre los que sustentar argumentos, cualquier cosa que se predique se predica en el vacío. Bien es verdad que también las creencias pueden imponerse por una fuerza diferente a la de los argumentos, como Voltaire ciertamente sabía y como parece que Duverger –a juzgar por el prólogo que Guy Rozart y José Pantoja escriben para el volumen colectivo *El historiador de lo inverosímil* practica<sup>6</sup>;
- d) Para que una cuestión de fe, como la que Duverger demanda del autor, se traduzca en una cuestión de razón –y podamos discutir sobre ella– es preciso que las conjeturas presentadas por el antropólogo reciban o el respaldo objetivo del documento con valor probatorio, o la confirmación del único hecho objetivo incontestable con el que contamos: el texto.
- <sup>5</sup> Para una revisión diferente de la fabulada argumentación de Duverger a favor de la autoría de Cortés, remito al artículo de Hugh Thomas, "Una nueva historia de una conquista vieja", Letras Libres, 10 de junio 2013. Online: http://www.letraslibres.com/mexico-espana/bernal-o-cortes (visitado el 6 de junio 2018).
- <sup>6</sup> El historiador de lo inverosímil lleva un elocuente subtítulo: Para acabar con la impunidad de Duverger. La referencia completa es Guy Rozart y José Pantoja (editores), El historiador de lo inverosímil. Para acabar con la impunidad de Duverger, Xalapa, Veracruz, Libros de Graphen, no 2, 2015. Los editores en su prólogo explican con lujo de detalles tanto la manera de trabajar de Duverger como los métodos (más emparentados con las "intrigas" políticas que con el "debate" científico-académico) de que se sirve para la "trasferencia de conocimiento".

#### 3. Objetivo

Frente a las conjeturas, contamos con unos textos que muestran unos rasgos de estilo (y el estilo es el hombre) a partir de los cuales podemos establecer objetivamente, con la simple confrontación de la *Historia verdadera* y de las *Cartas y relaciones*, si las páginas que han pasado a la historia con la firma de Bernal y las que han llegado a nosotros con la firma de Cortés han salido efectivamente de la misma pluma.

El objeto de mi análisis pretende comprobar el grado de similitud textual existente entre el documento **dubitado** de la *Historia verdadera de la conquista de Nueva España* –cuya autoría se cuestiona– y las muestras **indubitadas** que ofrecen las *Cartas y relaciones* de Cortés, autor con el que las conjeturas de Duverger emparentan la *Historia verdadera*. Para ello, me voy a servir de herramientas informáticas que hoy han alcanzado notable desarrollo y excelentes resultados, caracterizadas por su capacidad para tratar grandísimas cantidades de información y presentar los resultados en forma de datos perfectamente cuantificables, lo que limita radicalmente la interferencia de la subjetividad en el análisis.

#### 4. Corpus de análisis

#### 4.1 Descripción

Para poder dar respuesta al objeto de este informe, trabajaré con los siguientes materiales:

A) Un documento **dubitado**, en tanto en cuanto se han expresado dudas sobre su autoría, la *Historia verdadera de la conquista de la Nueva España*, una crónica de Indias tradicionalmente atribuida a Bernal Díaz del Castillo. Los especialistas sostienen que la crónica en cuestión fue concluida en torno a principios

<sup>7</sup> El corpus que a continuación se describe, a fin de que el lector pueda comprobar que no ha sido manipulado por el analista, está depositado en https://github.com/jblasco/cronistas\_de\_indias, al alcance de quien quiera disponer de él.

de los 60 del siglo XVI<sup>8</sup>. Se divide en un preámbulo y doscientos catorce capítulos, de los cuales, la autoría de un número reducido de ellos, especialmente los situados en la parte final del libro, se sospecha que pertenece al hijo de Bernal Díaz del Castillo. El texto nos ha llegado a través de varios testimonios:

- a) El conocido como manuscrito Guatemala, (G);
- b) el impreso publicado por fray Alonso Remón para la edición princeps de Madrid, 1632 (M);
- c) la copia, póstuma, de G que hizo Francisco Díaz del Castillo, hijo de nuestro cronista, habitualmente designado manuscrito Alegría (A)9.

Hoy existe el consenso de que el testimonio que debe servir de base para cualquier edición rigurosa es el G, que constaba en un principio de 299 folios (29'50 x 43 cm). Fue restaurado en 1951 en la Biblioteca del Congreso de Washington y en la actualidad consta de 287 folios, al final de los cuales (antiguo 299) aparece la firma de Bernal Díaz. Dos tipos de tinta se distinguieron cuando

- <sup>8</sup> Fue citada en varias obras de la época, lo que quiere decir que tuvo cierto renombre y que entre 1560 y 1630 circularon varios manuscritos (con toda seguridad serían más de los dos que han llegado a nuestros días). Varios autores lo citan, y siempre con referencia a Bernal; así en el "Catálogo de autores que han escrito historias de Indias", de Alonso de Zorita (1585); en las *Décadas*, de Antonio de Herrera (1601-1615); en la *Monarquía Indiana*, de Juan de Torquemada (1592-1613); en la *Historia de Tlaxcala* de Diego Muñoz Camargo (1591); en los *Anales de Aragón*, de Bartolomé de Argensola, con cuatro referencias a Bernal Díaz del Castillo. Finalmente, la *Historia verdadera de la conquista de la Nueva España* fue impresa por Alonso Remón en 1632.
- <sup>9</sup> Este manuscrito de la *Historia verdadera de la conquista de la Nueva España* se halló en Murcia en 1930. En el margen inferior del fol. I lleva la siguiente anotación "De Ambrosio del Castillo, herencia única que ubo de su padre". Para más información véase, M. L Pardo Morote,. "El Códice Alegría en la Biblioteca Nacional", en *Boletín de la Dirección General de Archivos y Bibliotecas*, n. 51 (1959), p. 7 y ss. Y C. Sáenz de Santa María, "Importancia y sentido del manuscrito Alegría de la *Historia verdadera* de Bernal Díaz del Castillo", en *Revista de Indias*, n. XI (1951), pp. 123 -143. Este manuscrito es conocido también como "Códice Alegría" por ser José María Alegría su último propietario, hasta que sus descendientes lo vendieron en 1950 a la Biblioteca Nacional de España.

lo restauraron: negra y rojo pardo. Carmelo Sáenz de Santamaría cree identificar tres tipos de letra, que convencionalmente llama A, B y C y que dividen el texto en tres secciones, lo que le permite afirmar que G fue escrito por tres manos diferentes: asigna indubitablemente la sección C a Bernal y, con reparos, también las secciones A2 y B; el resto, A1 especialmente, los cree de mano ajena<sup>10</sup>.

- **B)** Varios documentos **indubitados**, en tanto en cuanto, en principio, no existen dudas sobre su autoría:
  - 1) *Cartas de relación*, Hernán Cortés, fechadas entre 10 de julio de 1519 y el 3 de septiembre 1526.
  - 2) Historia General de las Indias, Francisco López de Gómara (1552).
  - 3) Crónica de la conquista de Nueva España, Francisco López de Gómara (1554).
  - 4) Brevísima relación de la destrucción de las Indias, Fray Bartolomé de las Casas (1552)<sup>11</sup>.
- Acepto lo establecido por Sáenz de Santamaría [1951, 1959, 1966/1982 y 1984] y confirmado por Guillermo Serés [*op. cit.*]: el primer tipo lo divide en AI: folios 1-17, 98-101, 115-117, 266-267, 288, 298-299; y A2: 279, 284, 2851 y 293-295. La sección B1 abarcaría los folios 18-64, 91-971, 102-114, 118-265 y 268; la B2 la cree más antigua que la anterior y abarca los folios 65-90, excepto el 72 bis; y, en fin, la B3, los folios 272-273, 280-283 y 285v-286. La tercera sección, C, que comprende los folios 269-271, 274-278, 2851, 289-292 y 2951 es la que identifica como autógrafa, porque coincide con cartas conservadas de Bernal de los años 1552-1558. Muy interesante resulta la confrontación de estas observaciones con el excelente trabajo de Guillermo Turner R., "La memoria de Bernal Díaz del Castillo en el proceso de escritura de la *Historia verdadera de la conquista de la Nueva España*. La evidencia de las formas y contenidos del texto", en *Historia*, 82 (2012), pp. 25-43, en el que se rastrean en la escritura de Bernal las huellas de la cultura escrita y las de la cultura oral.
- Existen otros documentos relacionados con la vida de Bernal Díaz del Castillo, que no contemplo en mi estudio por tratarse, en un alto porcentaje, de textos ajenos o de documentos formales y, en consecuencia, muy lastrados por el abundante uso de fórmulas administrativas que contaminan los *usus scribendi* del protagonista: así el acta del matrimonio de Bernal Díez del Castillo con Teresa Becerra en 1544; el texto de dos encomiendas de Bernal de 1549; varios escritos de solicitud de favores dirigidos al propio Rey; 5 cartas de problemática autoría: una

- 5) La llamada *Carta inédita de Cortés* (1526)<sup>12</sup> y las *Cartas* (1526, 1530, 1532, 1533 y 1534) ajenas a las *Relaciones*, que recogen Chantal López y Omar Cortés<sup>13</sup>.
- 6) Documentos de Bernal Díaz del Castillo<sup>14</sup> como el testimonio de Probanza de Leonor de Alvarado, *Méritos y servicios, Pedro de Alvarado: Nueva España*, Archivo General de Indias, Patronato, 69, R1, fol. 50r-55r<sup>15</sup>.

Los documentos **indubitados** nos servirán, confrontados con el documento **dubitado** (esto es la *Historia verdadera*), para establecer la aproximación o

dirigida a Carlos V para reclamar tierras fechada en 1552; dos, fechadas en 1558, a Felipe II y a Bartolomé de las Casas; una, de 1567, solicitando un cargo del virreinato de la Nueva España; una última, de 1573, en la que reclama tierras, indios y una dote para sus hijas. Un documento de legitimación de Diego Díaz del Castillo, hijo bastardo de Bernal; un Aviso de expedición a España del manuscrito de *Historia verdadera de la conquista de la Nueva España* firmado por el presidente de Guatemala en 1575. Estos documentos son de escaso valor para confrontar con la *Historia*, pues se trata de documentos legales y por lo tanto formularios; además ofrecen justificadas dudas de autoría, lo que los inhabilita como testimonios indubitados.

- <sup>12</sup> Sigo la edición de Joaquín García Icazbalceta, en *Colección de documentos para la historia de México*, México, 1858, págs. 473 y ss.
- <sup>13</sup> Véase http://www.antorcha.net/biblioteca\_virtual/historia/relacion/indice.html (como el resto de referencias de internet que se citan en este trabajo, la totalidad de ellas han sido consultadas en 7 junio 2017).
- <sup>14</sup> Me refiero, por ejemplo, a la Cédula firmada por Hernán Cortés; Cédula de Encomienda del Licenciado Marcos de Aguilar, 7 de febrero 1527; Cédula Firmada por Alonso de Estrada, 3 de abril de 1528; Carta de Bernal Díaz al Virrey Mendoza, de 7 de febrero de 1539; Carta al Virrey Mendoza, de 22 de febrero de 1539; Carta de Hernán Cortés, de febrero de 1539; Cédula Real dirigida a Pedro de Alvarado para que le provea de encomiendas vacantes, de 19 de junio de 1540; e Información Ad perpetuam rei memoriam de los méritos y servicios de Bernal Díaz, de Febrero, 1539. La mayor parte de estos documentos fue transcrita por Barbón, en Bernal Díaz del Castillo, Historia Verdadera de la Conquista de la Nueva España, edición crítica de José Antonio Barbón Rodríguez, Universidad Nacional Autónoma de México, Servicio Alemán de Intercambio Académico, Agencia Española de Cooperación Internacional, 2005, especialmente páginas 815-835.
- <sup>15</sup> Agradezco el conocimiento de este documento a Nora Jiménez (Centro de Estudio de las Tradiciones. El Colegio de Michoacán), que lo comenta en su trabajo "Bernal Díaz. Regreso a la silla del Autor" (art. En prensa).

alejamiento del idiolecto de Cortés respecto al de Bernal. Pero para poder llevar a cabo tal confrontación es preciso proceder a la "canonicalización" de los textos, lo que en ningún modo habrá de suponer alteración por pequeña que sea de los mismos.

#### 4.2 Canonicalización

Antes de proceder al análisis de este corpus con las herramientas informáticas especialmente concebidas para la estilometría, es preciso prepararlo a partir de ediciones impresas y digitales que están a nuestro alcance y preparar estas para que puedan ser analizadas.

En primer lugar, se trata de disponer ediciones digitales fiables. Concretamente, para la *Historia verdadera de la conquista de la nueva España*, de Bernal, seguiré la edición de Guillermo Serés (que se basa, como sabemos en el manuscrito Guatemala); para las *Cartas y Relaciones* de Hernán Cortés tomo el texto digitalizado de la edición de Chantal López y Omar Cortés<sup>16</sup>; para la *Brevísima relación de la destruición de las Indias* sigo la edición digital de José Miguel Martínez Torrejón para la Biblioteca Virtual Miguel de Cervantes<sup>17</sup>; para la *Crónica de la conquista de Nueva España* sigo la edición de Jorge Gurría Lacroix para la Biblioteca Ayacucho<sup>18</sup>; para la *Historia general de las Indias* de López de Gómara me sirvo de la edición de Jorge Gurría Lacroix en la Biblioteca Virtual Miguel de Cervantes<sup>19</sup>.

Pero no es suficiente contar con unos textos digitales fiables. Posteriormente es necesario regularizar su ortografía y eliminar todos los metadatos que puedan contaminar el análisis. Procedemos para ello a preparar los textos en TXT

<sup>&</sup>lt;sup>16</sup> http://www.antorcha.net/biblioteca\_virtual/historia/relacion/indice.html (junio de 2017).

<sup>&</sup>lt;sup>17</sup> http://www.cervantesvirtual.com/obra-visor/brevsima-relacin-de-la-destruccin-de-las-indias-o/html/ (junio de 2017).

<sup>&</sup>lt;sup>18</sup> Caracas, Fundación Biblioteca Ayacucho, 2007.

<sup>&</sup>lt;sup>19</sup> http://www.cervantesvirtual.com/portales/juan\_cervera/obra-visor/historia-general-de-las-indias--o/html/ (junio de 2017).

(UTF-8)<sup>20</sup>, y eliminamos todas aquellas marcas que pudieran haberlos contaminado en la edición digital.

Trabajaremos con los textos completos, pero para obtener un resultado más perfilado procederemos también a dividir los textos en varios fragmentos a fin de trabajar siempre con porciones de textos que (dubitados o indubitados) cuenten al menos con otro texto salido de la misma mano, y así mismo cuidaremos trabajar siempre con un corpus más equilibrado en extensión (trabajamos con textos de unas 20 000 palabras), libres de etiquetas y de cualquier metadato que pudiere falsear los resultados de nuestro trabajo de cuantificación.

Como resultado de la totalidad de operaciones realizadas sobre los documentos objeto de nuestro trabajo, el corpus de análisis queda constituido como sigue:

DUBITADOS (DUB) (según el tipo de análisis, en unos casos trabajaremos exclusivamente con el texto completo de la Historia verdadera; en otros, con la totalidad de los fragmentos (20 000 palabras) de la misma<sup>21</sup>; en otros, con parte de ellos; en otros, finalmente, con la totalidad de dubitados). El corpus dubitado quedaría constituido así: Bernal\_Historia\_verdadera (el texto completo: 387 790 tokens); BER\_1 (Historia verdadera: 20 000 tokens); BER\_2 (Historia verdadera: 20 000 tokens); BER 3 (Historia verdadera: 20 000 tokens); BER\_4 (Historia verdadera: 20 000 tokens); BER\_5 (Historia verdadera: 20 000 tokens); BER\_6 (Historia verdadera: 20 000 tokens); BER\_7 (Historia verdadera: 20 000 tokens); BER\_8 (Historia verdadera: 20 000 tokens); BER 9 (Historia verdadera: 20 000 tokens); BER 10 (Historia verdadera: 20 000 tokens); BER\_11 (Historia verdadera: 20 000 tokens); BER\_12 (Historia verdadera: 20 000 tokens); BER\_13 (Historia verdadera: 20 000 tokens); BER\_14 (Historia verdadera: 20 000 tokens); BER\_15 (Historia verdadera: 20 000 tokens); BER\_16 (Historia verdadera: 20 000 tokens); BER\_17 (Historia verdadera: 20 000 tokens); BER\_18 (Historia verdadera: 20 000 tokens);

<sup>&</sup>lt;sup>20</sup> También es posible trabajar con XLM-TEI, para lo que se recomienda la herramienta OxGarage accesible desde la web del Consorcio TEI <a href="http://www.tei-c.org/oxgarage/">http://www.tei-c.org/oxgarage/</a> (junio de 2017).

<sup>&</sup>lt;sup>21</sup> La división del texto en fragmentos equivalentes de unas 20 000 palabras la puede realizar automáticamente el paquete 'stylo' de R.

BER\_19 (*Historia\_verdadera*: 20 000 tokens); BER\_20 (*Historia\_verdadera*: 20 000 tokens); BER\_21 (*Historia\_verdadera*: 20 000 tokens).

INDUBITADOS (según el tipo de análisis, trabajaremos exclusivamente con el texto completo de las *Relaciones*; en otros casos, con los fragmentos –20 000 tokens – de la obra de Cortés; en otros, finalmente con la obra completa y los fragmentos<sup>22</sup>, con parte de ellos, o con el conjunto de la totalidad de dubitados). Así quedaría inicialmente el corpus indubitado: Cortes\_Relaciones (texto completo de *Cartas y Relaciones*: 172 948 tokens); COR\_1 (*Cartas y relaciones*: 20 000 tokens); COR\_2 (*Cartas y relaciones*: 20 000 tokens); COR\_3 (*Cartas y relaciones*: 20 000 tokens); COR\_4 (*Cartas y relaciones*: 20 000 tokens); COR\_5 (*Cartas y relaciones*: 20 000 tokens); COR\_6 (*Cartas y relaciones*: 20 000 tokens); COR\_7 (*Cartas y relaciones*: 20 000 tokens); COR\_8 (*Cartas y relaciones*: 20 000 tokens); COR\_9 (*Cartas y relaciones*: 20 000 tokens); CAR\_1526 (4067 tokens); CARTA\_1530 (3145 tokens); CARTA\_1532 (1427 tokens); CARTA\_1533 (1080 tokens); CARTA\_1534 (463 tokens); CARTA\_1NÉ-DITA (6003 tokens)<sup>23</sup>; Probanza\_Leonor\_Alvarado (*Méritos y servicios, Pedro de Alvarado: Nueva España*, Archivo General de Indias, Patronato, 69, R1).

corpus de Referencia: A fin de comprobar si los textos que nos ocupan se apartan y en qué grado de la lengua estándar, es preciso contar con otros textos ajenos al caso objeto del dictamen, cuya confrontación con los textos de análisis nos pueden aportan datos de gran relevancia para el mismo. Para confrontar los resultados obtenidos de nuestro análisis con lo que sucede en un corpus más amplio que el de los dos textos en cuestión, nos hemos servido

<sup>&</sup>lt;sup>22</sup> Como en el caso anterior, la división del texto en fragmentos equivalentes de unas 20 000 palabras la realiza automáticamente el paquete 'stylo' de R, sin intervención del analista en la elección del lugar de fragmentación.

<sup>&</sup>lt;sup>23</sup> Recuerdo que la conocida como "Carta inédita" remite al texto recogido con ese título por Joaquín García Icazbalceta en su *Colección de documentos para la historia de México*, México, Porrúa, 1980. Para las cartas de 1526, 1530, 1532, 1533 y 1534 sigo la edición digital de Chantal López y Omar Cortés (ed. cit.). (http://www.antorcha.net/biblioteca\_virtual/historia/relacion/indice.html). (junio de 2017).

también de los siguientes textos<sup>24</sup>: Herrera\_Historia\_General (Historia\_General: 122 313 tokens)<sup>25</sup>; Antonio\_Solis\_Historia\_Conquista (Historia de la Conquista de México: 197 802 tokens); Casas\_Brevisima\_Relación\_1 (Brevísima\_relación: 15 000 tokens); Casas\_Brevisima\_Relación\_2 (Brevísima\_relación: 15 000 tokens); Cervantes\_Salazar\_Crónica (Crónica de la Nueva España: 309 080 tokens); Gómara\_Historia\_Conquista\_Mexico (Historia\_Conquista\_Mexico: 180 205 tokens); Gómara\_Historia\_General: (Historia\_General\_Indias: 158 392 tokens).

Todos los textos que conforman nuestro corpus de análisis pertenecen a la esfera de "historia de Indias" (cronistas e historiadores) y tienen un rango temporal bastante cercano (entre 1520 y 1601). El más alejado de la fecha de redacción de la obra que se nos ha transmitido con el nombre de Bernal, es Herrera y Tordesillas (1601). No obstante, he querido incluirlo porque en vida se le criticó al cronista Mayor de Castilla y de las Indias que se aprovechaba de materiales previos de otros cronistas que habían sido testigos de la conquista. De la misma manera nos interesaba contar con Gómara, porque según se ha afirmado escribió su historia muy cerca de Cortés, aunque no me consta que fuera confesor de Cortés como Duverger afirma sin aportar documentación al respecto.

Finalmente, para nuestro análisis contaremos además con bases de datos de textos de referencia del español coetáneo (*Corpus diacrónico del Español*, CORDE, de la Real Academia Española o el *Corpus del español* de Mark Davies de BYU), y otros textos accesibles por buscadores online (Google, Yahoo, etc.)

<sup>&</sup>lt;sup>24</sup> Para los textos del corpus de referencia no siempre hemos manejado textos completos, ya que ellos no son objeto de nuestro análisis, sino de confrontación respecto a lo hallado en los textos que sí se han sometido a análisis.

<sup>&</sup>lt;sup>25</sup> Selecciono aquellos capítulos de la *Historia General de los hechos de los castellanos en las islas y tierra firme del mar océano* que se refieren a México.

#### 5. La estilometría

La estilometría<sup>26</sup> estudia el estilo lingüístico de los documentos (por lo general, de los documentos escritos). Los análisis estilométricos se utilizan a menudo para determinar la autoría de documentos que han llegado a nosotros como anónimos. El trabajo se realiza a partir de la comparación estadística de ciertos valores, como la longitud de la oración, la riqueza léxica, las *colocaciones* de palabras o el uso de términos inusuales, entre otros muchos más parámetros que pueden cuantificarse (frecuencia relativa de vocales, de sustantivos o de verbos; n-grams de caracteres, n-grams de palabras, etc.).

En tanto que en el pasado se hizo hincapié en los elementos más raros o más llamativos de un texto, las técnicas contemporáneas de determinación de autoría pueden trabajar incluso en la identificación de patrones de escritura inconscientes o de dificilísimo control por parte de quien produce, imita o plagia un documento. Es más, la estilometría se ocupa casi exclusivamente en el análisis de fenómenos de producción textual que el autor, por mucho que tenga voluntad de estilo o por mucho que sea consciente de la lengua que usa, no puede controlar.

El método estilométrico se basa en la invariante *idiolecto*: todos los documentos producidos por un mismo autor ofrecerán rasgos mensurables –por ejemplo, la frecuencia léxica – con resultados estadísticamente significativos. Y, en este sentido, dos textos con un comportamiento similar en varios rasgos mensurables evidencian que, o han sido producidos por un mismo autor (autoría común) o uno de los autores ha copiado al otro (plagio lingüístico).

Las valoraciones estadísticas ofrecidas por el análisis estilométrico permiten medir la distancia lingüística que existe entre diferentes muestras textuales, que se realizan atendiendo a cuatro grupos de eventos [Véase Stamatatos: 2000, 2014 y 2015]:

<sup>26</sup> En el contexto español es de obligada referencia José Manuel Fradejas Rueda, "El análisis estilométrico aplicado a la literatura española: las novelas policíacas e históricas", *Caracteres*, 5, 2 (2016), 196-264. http://revistacaracteres.net/revista/vol5n2noviembre2016/analisis-estilometrico/. Véase también Olivia Fialho y Sonia Zyngier, "Quantitative Methodological Approaches to Stylistics", *The Routledge Handbook of Stylistics*, ed. Michael Burke, Londres, Routledge, 2014.

- a. Léxico: frecuencia de palabras, riqueza y densidad de léxico, *n-grams* de palabra para distribución de palabras, errores, palabras de función.
- b. Caracteres: *n-grams* de caracteres, longitud de palabra, puntuación.
- c. Sintaxis: estructura de las frases, tipos de subordinadas, predominio de coordinación o subordinación, distribución de las palabras más frecuentes.
- d. Semántica: sinonimia, polisemia (implica un análisis aún más sofisticado, ya que utiliza capas de conocimiento que no se revelan en el propio texto).

Para procesar cuantitativamente las muestras textuales recurro a múltiples programas informáticos y herramientas de análisis textual (como CopyCath, Voyant, Textalyser, el paquete o librería 'stylo' de R o JGAAP). Pero aquí me voy a limitar a dar cuenta de los resultados que ofrece 'stylo' que trabaja sobre todo a partir de la frecuencia del léxico.

#### 5.1 'stylo'

El paquete 'stylo' de R<sup>27</sup> da acceso a un método de análisis de léxico absolutamente objetivo, ya que trabaja comparando entre varios textos muy altas

<sup>27</sup> Brevemente explicaré cómo funciona y qué hace 'stylo': el analista establece el corpus objeto de análisis e indica el rango de palabras más frecuentes (MFW) con las que va a trabajar (por ejemplo, desde 100 hasta 1000 con un incremento de 100 palabras hasta completar el texto de manera íntegra). Posteriormente el analista elige entre el cálculo de la distancia euclidiana <a href="https://sites.google.com/site/computationalstylistics/stylo">https://sites.google.com/site/computationalstylistics/stylo</a>, que refleja la mayor semejanza o desemejanza entre textos del corpus atendiendo a la frecuencia léxica o el árbol de consenso con Classic Delta y Eder Simple (José Calvo Tello: "Entendiendo Delta desde las Humanidades", *Caracteres*, 5, 1 (2016), 140-176. http://revistacaracteres.net/revista/vol5n1mayo2016/entendiendo-delta/). El dendrograma de distancia euclidiana refleja las conexiones directas de cada texto con aquel otro que le es más próximo: el dendrograma de consenso combina los resultados de las MFW en rango 1-100, 101-200, y así sucesivamente hasta llegar a 1000 (2000 o 5000, según elección del analista). Para más información puede verse detenidamente la metodología expuesta por Matthew L Jockers, *Analysis with R for Students of Literature*, New York, Springer, 2014, p. 109), y testada por múltiples analistas con textos de diferentes lenguas y distintas épocas y los artículos seminales de Maciej Eder, "Visualization in Stylometry: Cluster

cifras de palabras frecuentes (en adelante MFW, 'Most Frequent Words'). Opera del siguiente modo: analiza las palabras más frecuentes en el rango que el analista determine (500, 1000, 2000, 3000 o 5000 palabras). En función de los resultados de tales análisis, establece agrupaciones entre los textos y asocia cada uno de los textos con aquel(los) otro(s) que le resulte(n) lingüísticamente más próximo(s).

'stylo' calculará la distancia cuantitativa existente entre los distintos textos del corpus (en donde cero es la identidad absoluta entre dos textos y cualquier otra medida superior a cero indica la distancia entre esos dos textos).

En una primera aproximación, sometemos a análisis la totalidad de los textos que conforman nuestro corpus y que 'stylo' cuantifica numéricamente marcando la distancia entre cada par de textos. Para visualizar mejor los resultados muméricos obtenidos de medir la distancia, la herramienta ofrece a los profanos en estadística un dendrograma, en el que se visualiza gráficamente en forma de árbol lo que las cifras numéricas que operan en el algoritmo Delta Eder<sup>28</sup> (que es el que hemos empleado) significan: los textos objeto de nuestro análisis se organizan en subcategorías, que se van ramificando hasta llegar al nivel de detalle deseado. Estos dendrogramas ayudan a apreciar con mayor facilidad las relaciones de agrupación entre los textos analizados.

Así, el dendrograma (o gráfico) resultante de confrontar con 'stylo' la totalidad de los textos del corpus, sobre la base de las 1000 MFW, se dibuja de la siguiente manera:

(Véase Figura 1: Dendrograma 1)

Creo que la imagen es suficientemente elocuente, destacaré no obstante varias observaciones que me parecen muy relevantes:

I. La herramienta de 'stylo' ha distinguido dos grandes ramas: en una de ellas se sitúa el texto completo de la *Historia verdadera* y los distintos fragmentos de 20 000 palabras en que lo hemos descompuesto para el

Analyses Using Networks", *Digital Scholarship in the Humanities*, 3 (2015), 1-15, y Jan Rybicki & Mike Kestemont, "'Stylo': a package for stylometric analyses". Computational Stylistic Group, 20-02-2017. https://sites.google.com/site/computationalstylistics/stylo (junio 2017)

<sup>28</sup> John Burrows, "Delta: A Measure of Stylistic Difference and a Guide to likely Authorship", *Literary and Linguistic Computing*, 17, 3 (2002), 267-287.

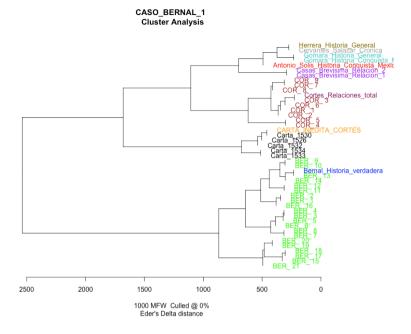


Figura 1: Dendrograma 1.

análisis; en la otra rama se agrupan la totalidad de crónicas objeto de nuestro análisis: las obras de Solís, de Gómara, de Bartolomé de las Casas, de Cortés, sin confundir nunca unas con otras y sin mezclar ninguna de ellas con los fragmentos de Bernal.

2. En la rama de Bernal, el fragmento que mejor recoge el "estilo" de la *Historia verdadera* parece ser el fragmento 13 (BER\_13), que contiene texto de los capítulos CLVIII y CLIX, pero la coincidencia de todos los fragmentos en una distancia próxima del cero (desde luego por debajo de 0,5) es muestra inequívoca de que la totalidad del texto de la *Historia verdadera* pertenece a una sola mano. Si otras manos ajenas a la mano principal han intervenido en su redacción, la participación de las mismas ha de considerarse marginal, limitada a leves correcciones o añadidos que el análisis apenas detecta: sólo el fragmento BER\_21 (parte del

CAP. CCXIV) parece alejarse, bien que levemente, del eje vertical en que se agrupa el resto de fragmentos.

- 3. Tanto el texto completo de las *Cartas y Relaciones*, de Cortés, como cada uno de los nueve fragmentos (COR\_1-9), en que lo hemos dividido para el análisis, se asocian en una misma rama secundaria dentro de la rama mayor que agrupa a la totalidad de los cronistas analizados. De idéntica manera que en el caso anterior, la proximidad que todos los fragmentos mantienen respecto al cero, y entre sí, parece apuntar a una sola mano. El fragmento más alejado del denominador común de todos ellos parece ser el fragmento signado como COR\_8, que pertenece al cuerpo central de la *Quinta relación*. Sin embargo, tanto el texto de la llamada *Carta inédita*, como el de las otras cinco cartas (1526, 1530, 1532, 1533 y 1534) ajenas a las *Relaciones* se agrupan entre sí en una rama secundaria diferente a la de las *Cartas y relaciones*, lo que (con el apoyo de alguna otra medición que viniese a confirmar el distanciamiento) podría ser indicio de haber sido redactadas por mano ajena a la de Cortés.
- 4. Sin embargo, este análisis no refleja la independencia de la *Primera Carta y relación*, que según algunos críticos se perdió y fue reelaborada por la Junta de regimiento de la ciudad de Veracruz. Lo que refleja 'stylo' es que, si es cierto que esta carta se reconstruyó por manos ajenas a Cortés, éste la revisó y uniformó su estilo verbal con el de las otras cuatro cartas.
- 5. De todos los cronistas analizados, el estilo de Bartolomé de las Casas es el que más se aproxima al estilo de Cortés.

Como corolario de los puntos anteriores, se puede afirmar que ni en su totalidad ni en parte alguna de la *Historia verdadera* se perciben huellas de ninguno de los cronistas que hemos analizado. Ni Cortés, ni Bartolomé de las Casas, ni López de Gómara, ni Solís, ni Cervantes de Salazar se aproximan al estilo de Bernal. Si Bernal no fuese el autor de la *Historia verdadera*, desde luego ninguno de los autores considerados en mi estudio podría ser candidato serio a su autoría.

Para corroborar o poner en debate los datos anteriores procedemos a repetir el análisis reduciendo ahora el corpus de análisis a los 21 fragmentos de la *Historia verdadera*. Mantenemos la búsqueda sobre las 1000 MFW. El resultado no varía sustancialmente, pero ahora se hacen más evidentes algunas de las observaciones arriba apuntadas:

(Véase Figura 2: Dendrograma 2)

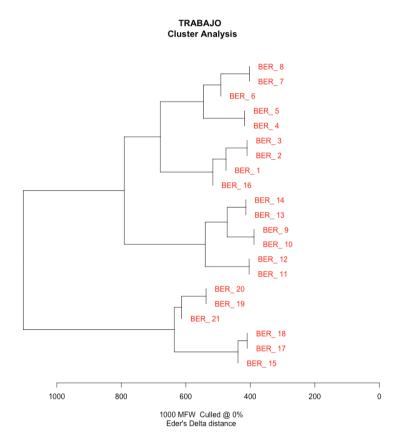


Figura 2: Dendrograma 2.

#### Interpretación:

- I. Las dos grandes ramas en que se agrupan ahora todos los fragmentos dejan en la rama inferior los fragmentos últimos de la *Historia verdadera* (desde BER\_15, que recoge parte del capítulo CLXVI y se extiende hasta el cap. CLXXIII), con la salvedad del fragmento BER\_16 (desde cap. CLXXIII al CLXXXI), lo que ahora sí nos permitiría considerar la posibilidad de la existencia de una segunda mano al menos para BER\_19, 20 y 21, que no sólo se agrupan en una tercera rama, sino que además se apartan más que el resto del cero.
- 2. Con los mismos argumentos, también podríamos considerar la presencia de otras manos (¿la misma segunda mano anterior? ¿Una tercera mano?) para BER\_1, 6 Y 16.

En cualquier caso y, como corolario de la interpretación de este nuevo gráfico, ninguno de los fragmentos de la *Historia verdadera* se acerca al uno (en este caso al 1000), lo que hace muy remota la posibilidad de que esas segundas o terceras manos hayan tenido, en ninguno de los fragmentos, un papel relevante, más allá de posibles notas interpoladas que en ningún caso superarían las 1000 palabras. Desde luego, sería muy interesante comprobar si aquellos fragmentos que más se alejan del eje vertical del cero guardan alguna correspondencia con las tres manos que a partir de un análisis del ms. G había establecido Sáenz de Santamaría. Pero no me detendré ahora en esta cuestión.

Demos un nuevo paso. Ampliamos ahora nuestro corpus de análisis, reuniendo los fragmentos cortesianos junto a los de Bernal. Si Hernán Cortés hubiera tenido algo que ver con la autoría del texto de Bernal –en todo o en parte– los fragmentos COR\_1-9 se mezclarían con los fragmentos BER\_1-21 mostrando la contaminación estilística naturalmente derivada de la común autoría de los mismos. Pero, antes de ofrecer los resultados, quiero dejar constancia de que 'stylo' trabaja en la cuantificación de elementos (1000 o 2000 palabras más frecuentes, por ejemplo) que no pueden impostarse ni controlarse voluntariamente por parte de un autor. Quiero decir que no existe la posibilidad de que Cortés, autor de las *Cartas y relaciones*, impostase un estilo

totalmente diferente al suyo para dar forma a la *Historia verdadera*. Se pueden imitar expresiones características de un autor, preferencias léxicas, etc., pero resulta imposible suponer que un autor como Cortés escribiera un texto de la amplitud de la *Historia*, manteniendo idénticos porcentajes de frecuencia sustantivos/adjetivos, o asociaciones de palabras, o caracteres, a lo largo de un texto tan voluminoso como el que nos ocupa.

Este es el resultado del nuevo análisis:

(Véase Figura 3: Dendrograma 3)

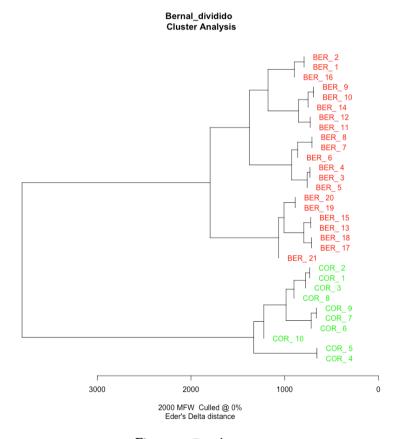


Figura 3: Dendrograma 3.

BRAE · t. XCIX· c. CCCXIX· 2019· págs. 5-44· ISSN 210-4822

Todos los textos de Bernal Díaz del Castillo se asocian en una rama independiente y los de Cortés, en otra, lo que deja en evidencia lo siguiente: la separación absolutamente inequívoca, en dos ramas claramente diferenciadas, de los fragmentos de Bernal y los de Cortés indica –sin ningún tipo de ambigüedades– que los fragmentos de la *Historia verdadera* (BER\_1-21) en que hemos dividido el texto de Bernal proceden de un idiolecto radicalmente diferenciado del que caracteriza la totalidad de los fragmentos en que se dividió el texto de Cortés. No hay contaminación cortesiana, ni en todo ni en parte, en ninguno de los segmentos de la *Historia verdadera*<sup>29</sup>.

En otras palabra, y concluyendo, a la vista de estos datos, resulta altamente improbable que ni en todo ni en parte Cortés haya tenido la más mínima intervención en la redacción de la *Historia* que la tradición, desde el mismísimo siglo xvI, nos ha transmitido siempre unida al nombre de Bernal. No obstante, vamos a realizar nuevas pruebas, incluyendo ahora en el corpus el texto procedente de la información evacuada por Bernal en la Probanza de Leonor de Alvarado, sobre "Méritos y servicios de Pedro de Alvarado: Nueva España". A pesar de la naturaleza híbrida (digámoslo así) de los textos que contienen una información como esta³o, el resultado es de nuevo elocuente:

(Véase Figura 4: Dendrograma 4)

<sup>&</sup>lt;sup>29</sup> Aunque no apunta al objetivo que me he planteado en este estudio, no puedo dejar de notar, a la vista de lo que refleja el *dendrograma* 3, que el fragmento 10 de Cortés (Cor\_10), parece apartarse de la homogeneidad que refleja el resto de los textos.

<sup>&</sup>lt;sup>30</sup> El texto de las respuestas puede estar condicionado o inducido, en parte al menos, por la formulación de la pregunta, tanto como por la mediación del escribano.

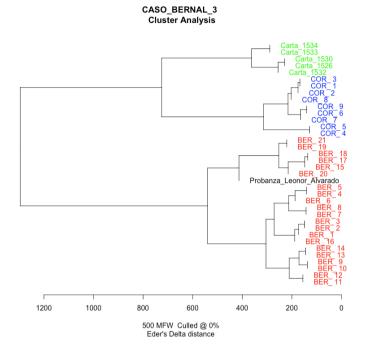


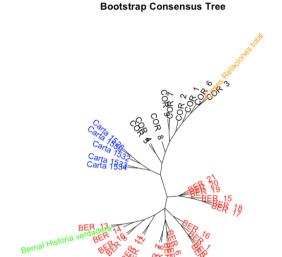
Figura 4: Dendrograma 4.

Si consideramos que el texto de la *Provanza* (con independencia de su naturaleza) es obra del soldado Bernal (casi analfabeto según Duverger), con este texto vienen a asociarse la totalidad de las porciones de la *Historia natural*, mostrando que todos ellos tienen un mismo idiolecto que ese documento del que se hace responsable el susodicho soldado Bernal Díaz del Castillo, en tanto que esta misma Probanza se aleja indubitadamente del idiolecto cortesiano.

Todavía procederemos a un nuevo análisis recurriendo a la aplicación ahora de la fórmula Eder Delta con el *Cosensus Tree* (aplicando de acuerdo con las recomendaciones aceptadas por la comunidad académica un parámetro del 0.5) que refleja la suma de los resultados en análisis, con un incremento que va de 100 a 1000 MFW):

(Véase Figura 5: Dendrograma 5)

CASO\_BERNAL\_3



100-1000 MFW Culled @ 0% Eder's Delta distance Consensus 0.5

Figura 5: Dendrograma 5.

De nuevo, el resultado es inequívoco y el protocolo con el que trabajamos nos dice que la lengua de los textos (el léxico, la frecuencia de types y tokens) está muy lejos de confirmar la supuesta autoría de Cortés para la *Historia verdadera*.

#### 5.2 Otras formas de análisis

Los dendrogramas y los árboles de Consensus 0.5 (que discriminan inequívocamente las posibles asociaciones entre los textos de nuestro corpus y que alejan totalmente el texto dubitado de la *Historia verdadera de la conquista de* 

Nueva España del idiolecto de Cortés) no nos ofrecen otra información que los porcentajes de distancia estadística; información que no nos deja estudiar los usos de selección léxica (morfológica o sintáctica) en el corpus: cuestiones como la densidad léxica, los campos léxicos o semánticos, los conectores, la gramática textual; en definitiva, todo aquello que puede servirnos para trazar el perfil lingüístico y marcas que nos permitan hacer una determinación de autoría.

Si unas pocas líneas de texto de un autor indubitado son suficientes al menos para establecer muy serias dudas sobre la paternidad del mismo en relación con otro texto dubitado, observemos la siguiente muestra que he tomado de cada uno de los textos en debate, acudiendo a las primeras palabras de los mismos:

	Cortés	Bernal
Muestra	Bien creemos que vuestras ma-	Notando estado cómo los
primera	jestades, por letras de Diego Veláz-	muy afamados coronistas, an-
página	quez, teniente de almirante en la	tes que comiencen a escrebir
	isla Fernandina, habrán sido infor-	sus historias, hacen prime-
	mados de una tierra nueva que	ro su <mark>prólogo y preámbulo</mark>
	puede haber dos años poco más o	con razones y retórica muy
	menos que en estas partes fue des-	
	cubierta, que al principio fue inti-	subida, para dar <mark>luz y crédito</mark>
	tulada por nombre Cozumel y des-	a sus razones, porque los cu-
	pués la nombraron Yucatán, sin ser	riosos letores que las leyeren
	lo uno ni lo otro, como por esta	tomen melodía y sabor de-
	nuestra relación vuestras reales al-	llas . Y yo, como no soy latino,
	tezas mandarán ver; y porque las	no me atrevo a hacer preám-
	relaciones que hasta ahora a vues-	bulo ni prólogo dello, porque
	tras majestades de esta tierra se han	ha menester para sublimar los
	hecho, así de la manera y riquezas	heroicos hechos y hazañas que
	de ella como de la forma en que	hecimos cuando ganamos la
	fue descubierta y otras cosas que	Nueva España y sus provincias
	de ella se han dicho, no son ni han	en compañía del
	podido ser ciertas porque nadie	valeroso y esforzado capitán
	(6: 1 / : :	,

(Sigue en la página siguiente)

Cortés Bernal don Hernando Cortés. hasta ahora las ha sabido como será que ésta que nosotros a vuestras reales después, el tiempo andando, por sus heroicos hechos fue Maraltezas escribimos y contaremos aquí desde el principio que fue qués del Valle. Y para podello escrebir tan sublimadamente descubierta de esta tierra hasta el como es dino, fuera menester estado en que al presente está, porque vuestras majestades seotra elocuencia y retórica mejor que no la mía; mas lo que pan la tierra que es, la gente que la posee y la manera de su yo vi y me hallé en ello peleando, vivir y el rito y ceremonias, seccomo buen testigo de vista, vo lo escrebiré, con el ayuda de Dios, ta o ley que tienen, y el feumuy llanamente, sin torcer a una do que en ella vuestras reales alparte ni a otra. Y porque soy tezas podrán hacer y de ella poviejo de más de ochenta y cuatro drán recibir y de quién en ella parte ni a otra. Y porque viejo vuestras majestades adelante han de más de ochenta y cuatro años sido servidos, porque en todo y he perdido la vista v el oír, vuestras reales altezas puedan hay por mi ventura no tengo cer lo que más servidos serán; otra riqueza que dejar a mis y la cierta y muy verdadera relahijos y decendientes salvo esta ción es en esta manera. (233 palabras) mi verdadera y notable relación, como en ella verán, no tocaré por agora en más de decir y dar razón de mi patria y dónde soy natural y qué año salí de Castilla y en capitanes compañía de qué anduve militando y dónde agora tengo mi asiento y vivienda .

Tabla 1

(244 palabras)

Apenas 250 palabras (las 250 primeras palabras del texto de cada uno de los documentos confrontados) son suficientes para descubrir, con una lectura superficial, unos usos verbales que en principio hacen muy difícil sostener ni siquiera la duda de una común autoría. En efecto, señalaré solamente unos pocos fenómenos (que marco en rojo, verde y amarillo para facilitar la observación):

- 1. Las 233 palabras del texto de Cortés revelan una sintaxis bastante compleja, en la que se puede observar cómo se suceden las oraciones de relativo dependientes de otra de relativo; o las oraciones subordinadas causales o finales dependientes de otra causal:
  - a) habrán sido informados de una tierra nueva que puede haber dos años poco más o menos que en estas partes fue descubierta, que al principio fue intitulada por nombre Cozume
  - b) porque las relaciones que hasta ahora a vuestras majestades de esta tierra se han hecho, así de la manera y riquezas de ella como de la forma en que fue descubierta y otras cosas que de ella se han dicho, no son ni han podido ser ciertas porque nadie hasta ahora las ha sabido como será ésta que nosotros a vuestras reales altezas escribimos y contaremos aquí desde el principio que fue descubierta de esta tierra hasta el estado en que al presente está, porque vuestras majestades sepan la tierra que es...
- 2. De la misma manera, en las 244 palabras del texto que siempre se atribuyó a Bernal, además de una sintaxis notablemente más clara y ordenada, se observa un gusto por lo que los gramáticos de la época llamaban "sinónimos voluntarios", esto es, coordinación de términos o sintagmas que tienen un valor semántico muy próximo:

prólogo y preámbulo, razones y retórica, luz y crédito, melodía y sabor, hechos y hazañas, Nueva España y sus provincias, valeroso y esforzado, elocuencia y retórica, yo vi y me hallé, la vista y el oír,

hijos y decendientes, verdadera y notable, decir y dar razón, mi patria y dónde soy natural, asiento y vivienda.

Nada más y nada menos que una quinta parte del texto de la *Historia verdadera de la conquista de la Nueva España* que hemos tomado como banco de pruebas, lo conforman estos sinónimos.

Siguiendo también una metodología tradicional, basada en la observación de fenómenos verbales en el texto, he analizado las seis primeras páginas de la obra de Cortés, con la pretensión de descubrir algunos usos no muy comunes (usos con poca fortuna en la lengua estándar de la época). Ofrezco en la siguiente tabla los resultados de la consulta en CORDE<sup>31</sup> (1.º de junio de 2017) acerca de fenómenos observados en esas seis páginas:

	Cortés	Bernal
bien + adjetivo	Usado con valor de [correcta-	Usado con valor de
	mente]	[muy]
puede haber + numeral	puede haber dos años poco más	
	o menos (2); puede haber tres le-	
	guas; puede haber veinte leguas	
intitular (por 'nombrar'	fue intitulada por nombre Co-	
o 'llamar') (9 casos)	zumel; a la dicha tierra intitu-	
	lada de Yucatán; se puede in-	
	titular de nuevo emperador de	
	ella; la ciudad de Cempoal, que	
	yo intitulé Sevilla; se intitulaba	
	capitán general; que se intitula-	
	ba de gobernador de aquella tie-	
	rra,	

(Sigue en la página siguiente)

<sup>&</sup>lt;sup>31</sup> Todas las búsquedas se realizan con la acotación de fechas 1520 y 1568, que son las fechas con las que trabaja su hipótesis Monsieur Duverger.

	Cortés	Bernal
en tal manera que (15 ca-	pelearon otro día en la mañana	[un solo caso]
sos)	con ellos, en tal manera que mu-	
	rieron veinte y seis españoles; co-	
	menzaron luego a flecharlos en	
	tal manera que hirieron a veinte	
	españoles; cercados con su tierra,	
	en tal manera que con ningunas	
	provincias tiene contratación; $Y$	
	aquel día acabamos de ganar to-	
	da la calle de Tacuba y de ado-	
	bar los malos pasos de ella, en tal	
	manera que; los cueros de ellas,	
	curados en tal manera que mu-	
	chos de ellos se conocieron	
adellos	tierra muy rica de oro, porque a	
	todos los naturales de ella los ha-	
	bían visto traer puesto adellos en	
	las narices, adellos en las orejas	
	y en otras partes	
hasta tanto que (11 ca-	hemos servido a vuestras reales	
sos)	altezas y serviremos hasta tanto	
	que la vida nos dure.	
de más y allende	acordamos de escribir a vuestras	
	majestades y les enviar todo el	
	oro y plata y joyas que en esta	
	tierra habemos habido, de más	
	y allende de la quinta parte que	
	de sus rentas y derechos reales	

(Sigue en la página siguiente)

	Cortés	Bernal
cosa alguna ('nada')	[24 casos]	[2 casos]
mejor que no	I caso	(6 casos): otra elocuen-
		cia y retórica mejor que
		no la mía; Y desque su-
		po nuestro intento, dijo
		que era mejor que no el
		suyo, etc.
con el ayuda de Dios		(7 casos)

Tabla 2

La conclusión no ofrece dudas: los usos de escritura de Cortés no se ven reflejados en la textualidad de la *Historia verdadera*, de Bernal, de una manera significativa.

#### 5.3 Voyant

Pero no soy partidario de fiarme de intuiciones. Por ello acudiré a una nueva herramienta como Voyant<sup>32</sup> para obtener ciertos datos que puedan objetivar cuantitativamente la confrontación que nos ocupa entre la las *Cartas y relaciones* y la *Historia verdadera*. No tendré en cuenta las observaciones derivadas del análisis en Voyant que se basan en la extensión de la frase en uno y otro texto ya que se podrá argüir que ello depende de la puntuación (lo cual no es cierto del todo) y la puntuación en un texto de los siglos xVI y xVII no es responsabilidad del autor, sino del editor moderno. Otra información que Voyant nos proporciona a partir del léxico sí puede resultar relevante. En efecto, sometidos al análisis de Voyant los textos de nuestro corpus, la herramienta nos devuelve resultados como los que muestra la figura (Ver *Figura 6*).

Muchas conclusiones podríamos obtener (por ejemplo, respecto al diferente uso de indigenismos en uno y otro texto; a la persona gramatical de los verbos empleados y al protagonismo de la primera persona; al uso de nombres propios de españoles; a la visión del mundo para uno y otro: en el primero

<sup>32</sup> https://voyant-tools.org/



Figura 6: Voyant.

aparecen como voces distintivas soldados, guerreros y batallas vs. las que resultan distintivas en Cortés: majestades, altezas, sacra, mezquitas). Curioso y muy significativo es que Bernal recurra a nombres de jerarquía eclesiástica católica para nombrar entidades indígenas en las que Cortés subrayará su carácter pagano recurriendo a nombres vinculados con el mundo musulmán. Así, Cortes hablará de mezquitas, en tanto que Bernal se referirá a los sacerdotes paganos como papas.

También nos permite **Voyant** hacer alguna incursión en la sintaxis de los textos o en otros usos de escritura. Veamos, por ejemplo, cuáles son los usos de uno y otro autor a la hora de apuntar al nombre propio de las personas sobre las que se habla. Cortés usa con cierta frecuencia el verbo *intitular* con el valor de "poner nombre", en tanto que Bernal Díaz del Castillo prefiere siempre el verbo *nombrar*. Así, en tanto que en Bernal se refiere a un "puerto que se dice e nombra en lengua de indios Axaruco", Cortés escribe: "una tierra nueva que pueda haber dos años poco más o menos que en estas partes fue descubierta, que al principio fue intitulada por nombre Cozumel. Es verdad que Cortés también emplea nombrar, pero en cambio Bernal nunca emplea *intitular*.

#### 5.4 AntConc

Me sirvo ahora de la herramienta informática AntConc<sup>33</sup> para obtener los casos de *verbatim* más significativos, entendiendo por *verbatim* la reproducción exacta de una secuencia de texto, palabra a palabra, de modo que todos los elementos de la secuencia aparezcan en el mismo orden, sin paráfrasis, declinación, conjugación, sustitución o abreviación de cualquier tipo, y sin realizar siquiera un cambio trivial que pueda alterar el significado.

Los *verbatim*, o calcos lingüísticos presentes en dos textos distintos, pueden significar: (a) que son expresiones comunes de esa lengua en ese momento determinado (frases hechas, expresiones habituales, etc.), y por tanto no idiosincrásicas; o (b) que son usos individuales y exclusivos de una persona, y por tanto idiosincrásicas, lo que implicaría que han sido producidas por el mismo autor o que uno de los dos textos es plagio del otro.

Recordemos que M. Coulthard<sup>34</sup> establecía que el número de ítems –número de palabras iguales, en el mismo orden– que pueden reproducir dos hablantes distintos se reduce drásticamente en función de la longitud de la frase. Por tanto, cuantos más elementos tenga un *verbatim*, cuanto más largo sea, mayor probabilidad existe de que sea una marca idiosincrásica. No obstante, debemos cotejar siempre esa marca en bases de datos de referencia para cerciorarnos que no sea una expresión común del lenguaje, compartida por múltiples hablantes.

Así procedemos en primer lugar a extraer, con 'AntConc', los casos de *verbatim* más llamativos en la obra de Bernal Díaz del Castillo, para ver si su recurrencia se produce también en los escritos de Cortés. Posteriormente, para poder determinar el grado de singularidad de esos *verbatim*, rastrearemos su búsqueda en las bases de datos de referencia del español de la época (exactamente en el CORDE). Si los *verbatim* tienen presencia en estas bases de datos de referencia, significa que serán marcas comunes que forman parte del léxi-

<sup>&</sup>lt;sup>33</sup> Laurence ANTHONY, "AntConc [Manual]", 2016, (http://www.laurenceanthony.net/software/antconc/releases/AntConc344/help.pdf) (consultado febrero 2017).

<sup>&</sup>lt;sup>34</sup> M. Coulthard, "Identificación de la autoría, idiolecto y singularidad lingüística", *Applied Linguistics* 25, 4 (2004), 431-477. (http://dx.doi.org/10.1093/applin/25.4.431).

co propio del lenguaje contemporáneo de estos autores; si por el contrario no tienen presencia en estas bases de datos, significa que son marcas altamente exclusivas (y por tanto idiosincrásicas), y que constituyen una especie de firma inconsciente del que escribió el texto que ha llegado a nosotros con el nombre de Bernal Díaz del Castillo.

Trabajamos inicialmente con algunos casos de *verbatim* limitándonos a comparar lo que ocurre en Bernal y en Cortés. En las *Tabla 3* hemos recogido los casos de *verbatim* presentes en el texto firmado por Bernal, con su correspondencia en Cortés:

	CASOS DE VERBATIM	BERNAL DÍAZ	CORTÉS
		DEL CASTILLO	
I	en aquella sazón	246 ocurrencias	
2	todos los más	127	7
3	pueblo que se dice	72	9
4	lo que sobre ello	60	I
5	fator y veedor	31	
6	se me acuerda	28	
7	aquella sazón era	2.1	
8	dejemos desto y digamos	2.1	
9	gran copia de	17	I
Ю	a dar mandado	18	I
II	velas y escuchas	14	4
12	que en ninguna manera		24
13	traer a la memoria	16	
14	dende a pocos días	15	I
15	con buen concierto	IO	
16	dejemos esto y volvamos	II	
17	porque en aquella sazón	27	

Tabla 3

Nadie podrá negar que resulta muy extraño que alguien que usa determinadas expresiones como una *muletilla*<sup>35</sup> hasta el punto de repetirla hasta 246 en una misma obra (es el caso de "en aquella sazón") escriba otra en la que dicha expresión no aparezca ni una sola vez. Y, al revés, "que en ninguna manera", tan del gusto de Cortés (24 ocurrencias en sus *Cartas*), no aparezca ni una sola vez en la *Historia Verdadera* no se explicaría si ambas obras hubieran salido de la misma pluma.

Aunque las cifras que refleja la *Tabla 3* son elocuentes, los casos de *verbatim* analizados en dicha tabla, con las discordancias que reflejan entre los autores objeto de estudio, no son muy significativos y, por lo tanto, no podemos otorgarles valor probatorio, pues se documentan también en bastantes otros autores del momento (tales como Pedro Mejía, Fernández de Oviedo, fray Luis de Granada, etc.), de modo que podrían llegar a considerarse expresiones de época. Pero, si ahora buscamos mediante el CORDE en cuántos autores se dan juntas todas estas expresiones, las cosas se concretan mucho más. Buscamos inicialmente (por razones de capacidad de la herramienta) en qué autores se producen a la vez, por ejemplo, *en aquella sazón, todos los más, pueblo que se dice, lo que sobre ello, gran copia de.* El resultado es curioso (Ver *Figura 7*).

No deja de ser interesante que el CORDE sólo documente la conjunción de estos 5 elementos en 548 casos, en sólo dos documentos: la *Historia verdadera de la conquista de la Nueva España*, Bernal Díaz del Castillo (498 ocurrencias), y la *Crónica de la Nueva España*, de Cervantes Salazar (70 ocurrencias).

Hagamos ahora la misma comprobación en el CORDE con traer a la memoria, dende a pocos días, con buen concierto, y dejemos esto y volvamos. El CORDE sólo registra un único documento (con 42 ocurrencias) en el que todos ellos se den conjuntamente: la Historia verdadera de la conquista de la Nueva España, lo que otorga a la suma de estas cuatro expresiones un alto valor idiosincrático y la convierte en un rasgo idiolectal notablemente elocuente (Ver Figura 8).

<sup>&</sup>lt;sup>35</sup> Según la segunda acepción de la palabra en el DRAE: f. *Voz o frase que se repite mucho por hábito*.

Resultado de la consulta al banco de datos Cómo citar el CORPUS				
Consulta:	en aquella sazón y todos los más y pueblo que se dice y lo que sobre ello			
Resultado:	y gran copia de, en 1520-1568, en todos los medios, en CORDE  548 casos en 2 documentos.			
Ver estadística	340 Casus en 2 documentos.			
voi obtadistioa				
Filtros:	Casos			
Ratio:	10			
Mantener document	os (Solo para filtro sobre casos).			
OBTENCIÓN DE EJ	EMPLOS			
Recuperar	Concordancias Normal.			
Clasificación:	<b>○</b>			
Agrupación:	Marcas:			
Figura 7: CORPUS.				
Cómo citar el CORPUS	risulta ai balico de datos			
Consulta:	traer a la memoria y dende a pocos días y con buen concierto y dejemos			
Resultado:	esto 'y' volvamos, en 1520-1568, en todos los medios, en CORDE			
Ver estadística				
Filtros:	Casos			
Ratio:	10			
☐ Mantener documentos (Solo para filtro sobre casos).  Filtrar				
OBTENCIÓN DE EJEMPLOS				
Recuperar	Concordancias Normal.			

Figura 8: CORPUS.

Hechas estas comprobaciones, parece evidente que las expresiones recogidas en la *Tabla 3* tienen un valor idiosincrático mucho más relevante del que inicialmente estábamos dispuestos a otorgarles. Pero los casos de *verbatim* comentados más arriba no son los únicos que 'AntConc' nos ofrece. Muy al

contrario, la herramienta nos permite documentar muchos otros casos que el CORDE nos muestra que son usos muy exclusivos, y que están lejos de poder ser considerados expresiones de un sociolecto de época. Lo refleja muy bien la *Tabla 4* (la primera columna recoge el caso de *verbatim;* la segunda el número de ocurrencias en Bernal Díaz del Castilo; la tercera el número de ocurrencias en Cortés; y la cuarta, el número de ocurrencias en otros autores que escriben entre 1520 y 1568):

CASOS DE VERBATIM	BERNAL	CORTÉS	CORDE
	DÍAZ DEL		
	CASTILLO		
			12: Zurita, Cervantes
soldado que se decía	128		Salazar, Cieza de León
			4: Anónimo, Zurita,
murió de su muerte	98		Gutiérrez de Santa Clara
			Fenández de Oviedo (1),
			López de Gómara (1),
			Juan de Timoneda (1),
ensillados y enfrenados	8		Romance anónimo (1)
no se me acuerda el nombre	8		Casas(1), Anónimo (1)
los nuevamente venidos de	8		López de Gómara (1)
cosas tocantes a nuestra	2.2		Casas (1)
a llamar de paz	15		Anónimo (1)
después de ganado México	22		Cervantes de Salazar (2)
			Pedro de Medina (1),
como dicho tengo en	25		Alonso de Santa Cruz (1)
dejemos de hablar en	24		Jorge de Montemayor (1)
			Cervantes Salazar (9),
			Cieza de León (25),
lo que más pasó	25		Gómez de Toledo (1)

(Sigue en la página siguiente)

CASOS DE VERBATIM	BERNAL	CORTÉS	CORDE
	DÍAZ DEL		
	CASTILLO		
			Cieza de León (1),
todos los más soldados	28		Fernández de Oviedo (1)
			Casas (9), Alonso de
			Santa Cruz (2),
			Fernández Oviedo (8),
			Betanzos (2),
desde allí adelante	36	2	Juan de Jarava (4)
y volvamos a decir	17	I	
			Casas (17), Cieza
			de León (2), Cervantes
			Salazar (1), Jerónimo
			de Urrea (1), Fernández
el tiempo andando	30	I	de Oviedo (1)
			Fernández de Oviedo (2),
			Zurita (3), Casas (1),
corredores del campo	46		López de Gómara (1)
-			Casas (1), Cieza
			de León (2),
vecino que fue	39		Borregán (1)
			1: Gutiérrez de
murió en poder de	68		Santa Clara
			Pedro Mejía (2),
			Fernández de Oviedo (12),
			Juan de Molina (6),
que en aquella sazón	67		Zurita (9)

Tabla 4

Para considerar el valor discriminatorio que ofrecen los casos de *verbatim* de la *Tabla 4*, merece la pena que prestemos atención a las elevadas cifras con las que los mismos se documentan en Bernal, a la extensión (cuatro o cinco pa-

labras) de los mismos, y a la exclusividad con la que se producen en el conjunto total de la literatura de los años que estudiamos<sup>36</sup>.

Y los casos de *verbatim* detectados en la *Historia verdadera* no se acaban aquí. La *Tabla 5* ofrece nuevos casos todavía más exclusivos, pues no sólo no se documentan en ningún otro autor del momento que estudiamos (1520-1568), sino que tampoco aparecen, en una búsqueda actual, entre al menos las 30 primeras entradas de Google<sup>37</sup>.

CASOS DE VERBATIM	BERNAL	CORTÉS	CORDE
	DÍAZ DEL		
	CASTILLO		
			NINGÚN OTRO CASO
			EN LAS 30 PRIMERAS
			ENTRADAS
los caciques y papas	13		DEL CORDE
obra de cuatro leguas	8		NINGÚN OTRO CASO
en todos tres reales	8		NINGÚN OTRO CASO
quedarse ha aquí	13		NINGÚN OTRO CASO
lo cual diré adelante	II		NINGÚN OTRO CASO
lo teníamos de costumbre	14		NINGÚN OTRO CASO
desque aquello vio	3		NINGÚN OTRO CASO
vara y piedra	25		NINGÚN OTRO CASO
que se decía fulano	26		NINGÚN OTRO CASO

(Sigue en la página siguiente)

<sup>&</sup>lt;sup>36</sup> El lector debe considerar la riqueza de la base de textual sobre la que trabaja la herramienta del CORDE. En cualquier caso este trabajo puede contrastarlo el lector con el excelente estudio de Guillermo Turner R. ("La memoria de Bernal Díaz del Castillo...", art. cit.) sobre las huellas de fuentes escritas y huellas de fuentes orales en la verbalización de la *Historia verdadera*.

<sup>&</sup>lt;sup>37</sup> Las excepciones a la afirmación a la que responde esta nota son muy pocas. En algunos casos (muy pocos) uno puede encontrar ejemplos en Google, distintos a Bernal, a partir del ítem 15. En cualquier caso, la secuencia verbal de todos los *verbatim* de la *Tabla 15* tiene un muy alto valor idiosincrático como cualquier lector entenderá.

CASOS DE VERBATIM	BERNAL	CORTÉS	CORDE
	DÍAZ DEL		
	CASTILLO		
pasó otro buen soldado	27		NINGÚN OTRO CASO
volvamos a nuestra relación	10		NINGÚN OTRO CASO
otro buen soldado que	32		NINGÚN OTRO CASO
por mí memorado	6		NINGÚN OTRO CASO
otras veces por mí nombrado	9		NINGÚN OTRO CASO
capítulo que dello habla	56		NINGÚN OTRO CASO

Tabla 5

Los *verbatim* detectados por AntConc, y que yo acabo de comentar, muestran que es tal el gusto de la *Historia verdadera* por ciertas expresiones, y que las repite su autor con índices de frecuencia tan altos, que resulta de todo punto imposible imaginar que al menos alguna de ellas no apareciese también en otros escritos del mismo autor. Si Cortés fuese, como se ha afirmado con impostada solemnidad, el autor de la *Historia verdadera*, alguno, al menos alguno, de los casos de *verbatim* documentados en esta última habrían dejado huella en las *Cartas y relaciones*. Y ello no es así.

#### 6. Conclusiones

Un análisis cuantitativo de los textos dice que es altamente improbable que las *Cartas y relaciones* y la *Historia verdadera* hayan salido de la misma mano. Es la estadística aplicada al texto la que así lo indica. Yo no he hecho otra cosa que poner ante los ojos del lector los resultados. En todo caso, el lector es el juez único de este proceso.

Ante la ausencia de documentos sobre los que rigurosamente afirmar, o negar, la atribución del texto firmado por Bernal, la pura textualidad se convierte en el único escenario en que el debate tiene sentido. A mi análisis del habla de Bernal quise llegar libre de prejuicios y me enfrenté con los textos de las *Cartas y relaciones* y con la *Historia verdadera* sin haber analizado el libro

de Monsieur Duverger y sin conocer (o conocer sólo en superficie) algunas de las críticas que el mismo había recibido. Es lo recomendable a fin de que ni el enfoque ni los procesos de análisis queden contaminados por la subjetividad del analista. Sólo una vez concluido mi trabajo de análisis leí detenidamente la propuesta de autoría (es verdad que no pude tragarme el epílogo, pero ya al llegar a tal punto me había divertido bastante) y las reseñas que había provocado el mismo. No se excede León Portilla cuando, sin más ayuda que el olfato de lector avezado, califica la tesis que aquí se discute de "fantasías de la temeridad", tras sentar un presupuesto con el que estoy totalmente de acuerdo y que merece un recuerdo: "Hay libros que por dar a conocer algo que mucho importaba o por sostener una tesis merecedora de atención, hacen ruido. Hay otros que también lo hacen, aun cuando sea por poco tiempo, por ofrecer exactamente lo contrario"<sup>38</sup>. Tampoco se excede Nora Jiménez en un documentado trabajo que oportunamente califica desde el título la propuesta de autoría que discutimos: "El 'historiador' de lo Inverosímil. Para acabar con la impunidad de Duverger"39; y algo parecido se deduce de las tesis de Guillermo Serés, el más riguroso editor moderno de Bernal. Yo sólo añadiré que la diferencia entre el rigor académico y la superchería reside en que la academia pone por delante la verdad al ego. La superchería funciona con otros parámetros.

Atribuir la *Historia verdadera* a Cortés me parece un error, pero puedo equivocarme. Y así lo reconoceré si se me demuestra con documentos o con argumentos (y no con conjeturas por muy pomposa y solemnemente que se prediquen en español o en francés). El rigor no garantiza el acierto. Pero la fantasía, en historia, sí garantiza el error. Pienso (y esto alienta en mis palabras) que una obligación que debería ir en el sueldo de cualquier académico es la de la denuncia de aquellos escritos que (por ignorancia, por pereza, por diletantismo o por hinchazón del ego) pretenden hacer pasar la ficción gratuita por historia documentada, lo que en castellano tiene un nombre, *superchería*, y en relación con esta Cervantes ya dijo todo cuanto hay que decir: "caballeros,

<sup>&</sup>lt;sup>38</sup> "Fantasías de la temeridad", *Nexos*, 1, abril, 2013.

<sup>&</sup>lt;sup>39</sup> Trabajo leído en el congreso "Miradas Historiográficas actuales sobre la Conquista Americana. El revisionismo en la obra de Christian Duverger", recogido en las *Actas* del mismo (Published on Feb 12, 2014) y que puede leerse en https://issuu.com/guyrozat/docs/actas\_duverger (junio 2017).

¡que se tengan, y miren que no es bien usar de superchería con nadie!". Pero a veces los académicos nos ponemos estupendos y, a fuer de neutrales y corteses, preferimos mirar a otro lado en vez de denunciar en puro castellano usos más propios de mentalidades autoritarias del feudal mundo del medievo que de la modernidad que inician (con el racionalismo, que ya no traga la verdad impuesta) Cervantes, en la ficción, y Descartes, en la filosofía.

FRANCISCO J. BLASCO Universidad de Valladolid