



PROGETTO
MAMBRINO

HISTORIAS FINGIDAS



Los modelos de HTR *Silves1549_BNE* y *Spanish Gothic* como herramientas de la labor ecdótica

Giada Blasut

(Università di Verona)

Abstract

Esta contribución informa de la aplicación a textos áureos impresos en letra gótica de dos modelos de transcripción automatizada realizados dentro de la plataforma Transkribus. Serán objeto de estudio el modelo individual *Silves1549_BNE* y el modelo extendido *Spanish Gothic* ideados por Stefano Bazzaco en colaboración con diferentes miembros de las universidades de Lyon, Oporto, Verona y Zaragoza¹. Palabras clave: libros de caballerías castellanos; *Silves de la Selva*; Transkribus; HTR (*Handwritten Text Recognition*); modelo HTR *Spanish Gothic*

This work reports the results of the application of two Transkribus print models based on Spanish books published in Gothic typeface between the 15th and the 16th centuries: the *Silves1549_BNE* and the extended *Spanish Gothic* model. The two models were created by a group of researchers from the Universities of Lyon, Oporto, Verona and Zaragoza.

Key words: Spanish Romances of Chivalry; *Silves de la Selva*; Transkribus; HTR (*Handwritten Text Recognition*); HTR model *Spanish Gothic*



¹ Han formado parte del grupo de investigación coordinado por Stefano Bazzaco los siguientes investigadores: Nuria Aranda García (École Normale Supérieure de Lyon), Ángela Torralba Ruberte (Universidad de Zaragoza), Ana-Milagros Jiménez (Universidad de Zaragoza), Pedro Monteiro (Universidade do Porto), Federica Zoppi (Università di Verona) y quien firma estas líneas.

Hacia la primera edición crítica del *Silves de la Selva*

El libro de caballerías castellano *Silves de la Selva* es el duodécimo y último de la saga narrativa *Amadís de Gaula*². Hoy en día, esta novela constituye la única publicación del ciclo amadisiano que carece todavía de una edición moderna del texto. Por esta y otras razones que he tenido ocasión de describir recientemente (Blasut, 2021), el *Silves de la Selva* es el objeto de estudio de mi tesis doctoral en la que aspiro a ofrecer lo siguiente: la biografía actualizada de su autor, Pedro de Luján; un estudio literario que informe de las estructuras narrativas de la novela; y la edición crítica del texto³. Como es de esperar, la realización del tercer objetivo, la edición de la novela, cuenta con las dos fases fundamentales de la labor filológica: la *recensio* y la *constitutio textus* (Blecua, 2001, 33).

La tradición textual del *Silves de la Selva* está constituida por dos ediciones publicadas en Sevilla en el taller de Dominico de Robertis en 1546 y 1549, respectivamente (Eisenberg y Marín Pina, 2000, 259-269). Aun así y como ha descubierto la comunidad científica, Dominico de Robertis ya había muerto cuando en 1549 salió la segunda publicación del *Silves de la Selva*. Las investigaciones revelan también que quien se estaba haciendo cargo de la gestión del taller era Pedro de Luján, sobrino político de Dominico de Robertis, al ser su madre Isabel Álvarez de Rivas, hermana de la mujer del impresor (Romero Tabares, 1998; Maillard Álvarez, 2007; Álvarez Márquez, 2009; Gestoso y Pérez, 1924, 118 y 118 n. 1; Hazañas y la Rúa, 1949, 90 y 102 n. 3)⁴. De modo que, con toda probabilidad, el autor del *Silves de la Selva*, Pedro de Luján, fue también el impresor de la segunda edición del texto; lo que lo convierte, junto con Andrea Pescioni, en uno de los dos tipógrafos sevillanos que «llevaron su

² En verdad, si se consideran como parte integrante del género de los libros de caballerías castellanos también las obras redactadas originariamente en otras lenguas europeas, habría que considerar el *Esferamundi* como última entrega del ciclo amadisiano. El libro fue escrito en italiano por Mambrino Roseo da Fabriano y publicado en seis partes en los talleres venecianos de Michele Tramezzino entre 1558 y 1565. Sobre la vida y la obra de Mambrino Roseo da Fabriano, consúltense Bognolo (2010) y Bognolo-Cara-Neri (2013).

³ Para la biografía de Pedro de Luján véase: Romero Tabares (1998), Maillard Álvarez (2007), Álvarez Márquez (2009), Castillejo Benavente (2019) y Bazzaco (2020).

⁴ Fue Francisco Escudero y Perroso quien, por primera vez, avanzó la hipótesis de que Pedro de Luján debía ser el sucesor de Dominico de Robertis (1894, 25 y 236).

destreza en el manejo de la pluma a la composición de obras de propia creación o a la traducción al castellano de otras ajenas» (Álvarez Márquez, 2007, 13-14)⁵.

En un estudio reciente he tenido ocasión de indicar que las ediciones del *Silves de la Selva* no son idénticas, a pesar de que se publicaron en la misma casa tipográfica. En realidad, el cotejo revela que la segunda derivó de la primera –al compartir errores comunes conjuntivos–, pero que se diferencia notablemente de la *princeps* por las abundantes variantes textuales que transmite (Blasut, 2021). En la base del cotejo efectuado y consciente de que en 1549 Pedro de Luján estaba gestionando la tipografía que había fundado Robertis, he elegido la segunda publicación de la novela (Sevilla, 1549) como texto base de la edición que propongo en mi tesis.

Por consiguiente, necesitaba fuera la transcripción del texto de 1549 para proseguir en la labor filológica. Sin embargo, el *Silves de la Selva* es, al igual que los demás libros de caballerías castellanos, un texto largo; su longitud ocupa 150 folios (recto y verso) escritos a doble columna, de unas 46 líneas cada una. Esto implica disponer de mucho tiempo para transcribirlo manualmente a través de un normal procesador de textos como, por ejemplo, Microsoft Word.

Por este motivo y a raíz de los avances que la aplicación de modelos de transcripción automatizada habían aportado en el campo ecdótico, decidí recurrir a una herramienta digital que permitiera reducir el tiempo de trabajo sin renunciar a la calidad del mismo. En virtud de ello, Stefano Bazzaco me aconsejó que utilizara la plataforma Trankribus; lo que me ofrecía también la ocasión de poder valorar la aportación de la plataforma a la tradicional labor filológica. ¿Puede Trankribus revelarse una herramienta no solo útil, sino también fundamental para realizar una edición crítica de un texto español del Siglo de Oro?; ¿qué ventajas ofrece respecto a una transcripción manual?; e incluso, ¿puede el filólogo utilizar Trankribus sin renunciar al contacto directo con el ejemplar que transmite el texto?

⁵ El impresor, editor y mercader de libros Andrea Pescioni tradujo del francés al castellano la obra de Pedro Bovistau, Claudio Tesserant y Francisco Beleforest *Historias prodigiosas y maravillosas de diversos sucesos acaecidos en el mundo* (Francisco del Canto, 1586, Medina del Campo) (Álvarez Márquez, 2007, 14).

La necesidad de un modelo de transcripción para el *Silves de la Selva*

Cuando, en 2020, me dispuse a utilizar Transkribus para la transcripción automatizada del *Silves de la Selva*, todavía no estaba disponible un modelo aplicable a textos áureos impresos en letra gótica. El Progetto Mambrino (grupo de investigación de la Universidad de Verona coordinado por Anna Bognolo y Stefano Neri), ya había desarrollado un modelo de transcripción automatizada apto para la aplicación a libros de caballerías escritos en lengua italiana a imitación de los castellanos⁶. Sin embargo, por tratarse de impresos en letra cursiva, y no gótica como el *Silves de la Selva*, su modelo no podía aplicarse al texto objeto de mi interés (Bazzaco, 2018). Con todo, los excelentes resultados de la transcripción automatizada de las novelas italianas demostraban, de manera eficiente, la fiabilidad de Transkribus y las ventajas de su aplicación: decrecimiento del tiempo útil a la transcripción, fiabilidad de la plataforma Transkribus, reducción de las tareas del editor del texto (Bazzaco, 2020). Por eso, y a falta de un modelo aplicable a los textos castellanos impresos en letra gótica, se hacía necesario crear un modelo *ex profeso* para transcribir el *Silves de la Selva*; lo que realicé bajo la coordinación de Stefano Bazzaco.

Como es sabido, en la plataforma se pueden crear dos tipos de modelos: los individuales, cuya aplicación queda reservada a un solo texto; y los modelos extendidos aptos para textos diferentes, pero parecidos en su conformación tipográfica. Ambos tipos precisan de una fase inicial denominada *Golden Transcription* (transcripción manual) mediante la cual Transkribus aprende, sirviéndose de técnicas de *machine learning*, la correlación entre la representación visual de los caracteres de la imagen y

⁶ En el siglo XVI, los libros de caballerías castellanos se tradujeron a las principales lenguas europeas (francés, italiano, alemán, inglés y holandés). Incluso, algunos libros del ciclo narrativo del *Amadís de Gaula* contaron también con traducción al hebreo. En Italia, Francia y Alemania se produjeron además continuaciones originales o *aggiunte* que fueron sucesivamente traducidas al castellano. La cuestión de la recepción de los libros de caballerías en Europa puede leerse en Bognolo (2020a; 2020b), Bognolo-Cara-Neri (2013), Neri (2008a; 2008b; 2013), Schlusemann y Wierzbicka-Trwoga (2020), Sánchez-Martí (2019) y Wilson (2014). Para la difusión en Italia consúltese Bognolo (1984; 2008; 2011) y Bognolo-Cara-Neri (2013).

su transcripción⁷. Aun así, la labor de transcripción manual difiere por la cantidad de textos transcritos y, por ende, por el número de páginas sobre las que se construyen. Si, por un lado, los modelos individuales se basan en un único texto, del que se transcribe tan solo una veintena de páginas; por el otro los modelos extendidos disponen de un corpus más vasto de textos y por consiguiente de un número más copioso de páginas transcritas manualmente. De tal manera que las principales diferencias que atañen la génesis de los modelos son dos. En primer lugar, el número de textos empleados: uno en el caso del modelo individual, mientras que varios en el caso del modelo extendido; y, en segundo, la cantidad de caracteres que el software tiene la oportunidad de aprender mediante la transcripción manual. En el caso del modelo extendido es mayor el número de páginas transcritas manualmente, por lo que aumenta la posibilidad de que la plataforma vea y aprenda nuevos caracteres (mayúsculas, minúsculas, signos de puntuación u otros especiales). Del mismo modo, la transcripción de un elevado número de páginas conlleva, evidentemente, la reiteración de muchos caracteres que la plataforma ve una y otra vez, aumentando así su capacidad mnemotécnica. Como tendré ocasión de profundizar a continuación, además de la diferenciación ya señalada, hay otro punto fundamental de divergencia entre los dos modelos y que se refleja incluso en su denominación: su finalidad. Frente al modelo individual que se puede aplicar al solo texto que ha sido empleado durante su creación, el modelo extendido aspira a transcribir automáticamente también textos externos a su corpus, es decir, a obras que no se han utilizado como *dataset* para su producción.

En lo que atañe a la creación del modelo individual del *Silves de la Selva*, he transcrito manualmente dentro de la plataforma las primeras veinte páginas de la novela siguiendo los criterios que se indican a continuación:

- u/v/i/j/y aparecen como en el texto;
- letras mayúsculas y minúsculas aparecen como en el texto;
- longitud de las líneas es fiel al texto. Un guion indica la división

⁷ Ofrezco aquí solamente una descripción somera de las principales fases de creación del modelo individual del *Silves de la Selva*. Para una explicación detallada del flujo de trabajo que supone Transkribus véase Bazzaco (2020).

- de las palabras cortadas al final de línea incluso cuando este no se halla en el ejemplar;
- la unión y la separación de las palabras refleja el uso actual, con la excepción de algunos vocablos como «toda via» y «tan/m bien» que se mantienen separados;
 - toda abreviación se ha desarrollado y en ningún caso se han acentuado las palabras, ni se han introducido apóstrofes al fin de subrayar la fusión de dos vocales («dellos» y no «d'ellos»);
 - los errores de imprenta no se han corregido;
 - no se han introducido signos de puntuación que no estén en el texto.

En definitiva, se trata de criterios destinados a obtener un modelo de transcripción diplomática del *Silves de la Selva*, al que denominaré *Silves1549_BNE*.

Valoración del modelo individual del *Silves de la Selva*

La creación y la aplicación del modelo individual de transcripción automatizada del *Silves de la Selva* se han revelado fructuosas y de gran utilidad. Entre las ventajas más destacadas resalta, sin duda, la rapidez de ejecución de la plataforma. Transkribus necesita solamente de algunos minutos para transcribir la novela entera. Pero es más, la aplicación de Transkribus no solo implica un ahorro en términos de tiempo, sino también de deberes; el editor encomienda la transcripción del texto que necesite a la plataforma y se reserva para sí mismo solamente la fase sucesiva de su labor, es decir, la corrección de la transcripción. Esta puede realizarse de distintas maneras, bien dentro de la plataforma, bien fuera de ella. Si bien la primera se revela de gran utilidad, ya que permite visualizar simultáneamente tanto la imagen como su transcripción automatizada, son muchas las ventajas que aporta la corrección realizada fuera de la plataforma. En primer lugar, es preciso señalar que es posible exportar la transcripción en múltiples formatos (Transkribus Document, PDF, TEI, DOCX, Simple TXT, Excel e IOB), modificables, por lo general, por parte del editor. Este aspecto es una de las características más destacadas y

apreciadas de Transkribus, puesto que el editor puede emendar la transcripción automatizada directamente dentro del documento y, posteriormente, utilizar este mismo como esqueleto de la edición que pretende realizar. Lamentablemente, una vez revisada la transcripción, el texto no puede volver a subirse a la plataforma; lo que permitiría, en mi opinión, no solamente disponer de una transcripción correcta incluso dentro de la plataforma, sino también contribuir al *entrenamiento* de Transkribus.

En lo que a la valoración de la aplicación del modelo individual del *Silves de la Selva* se refiere, el estudio revela lo siguiente: se han contabilizado catorce errores por página, espacios incluidos. Cabe señalar que la mayoría de los fallos cometidos por la sistema de HTR concierne a la errónea unión y separación de palabras, cuando no su acentuación. Por lo tanto, la naturaleza de tales errores determina una mínima intervención por parte del editor, quien se limita a colocar o eliminar algún espacio en blanco o bien a añadir los acentos necesarios. La adición de un espacio en blanco se puede ejemplificar con el caso en el que la conjunción copulativa «y» y el nombre del personaje que lo sucede, «Anaxartes», vienen concebidos por la plataforma como si de un solo término se tratara «yanaxartes». De tal manera que en la fase de revisión se hace necesario separar los dos términos para restablecer la exacta lectura del texto. Al contrario, un poco más adelante, la plataforma sobreinterpreta la información separando la palabra «tierra» en dos partes sin significado alguno, «tier» y «ra». También en este caso, el quehacer del editor es rápido y limitado, si bien fundamental; solamente hace falta eliminar el espacio en blanco que Transkribus ha colocado entre los dos fragmentos. Cabe destacar que no se trata de errores excesivamente graves, ya que no afectan seriamente al significado del texto. El editor puede, además, subsanarlos fácilmente incluso sin recurrir a la imagen del texto.

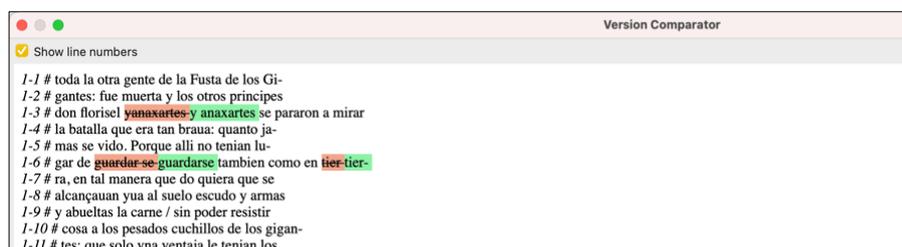


Fig. 1. Errores de transcripción cometidos por Transkribus

Dado que esta tipología de error representa la mayoría, el editor ya sabe dónde puede hallar tales incorrecciones. Como se puede apreciar (Fig. 1 y 2), dichas incongruencias suelen darse, por lo general, al finalizar una línea; lo que las convierte en errores muy previsibles y fácilmente subsanables.

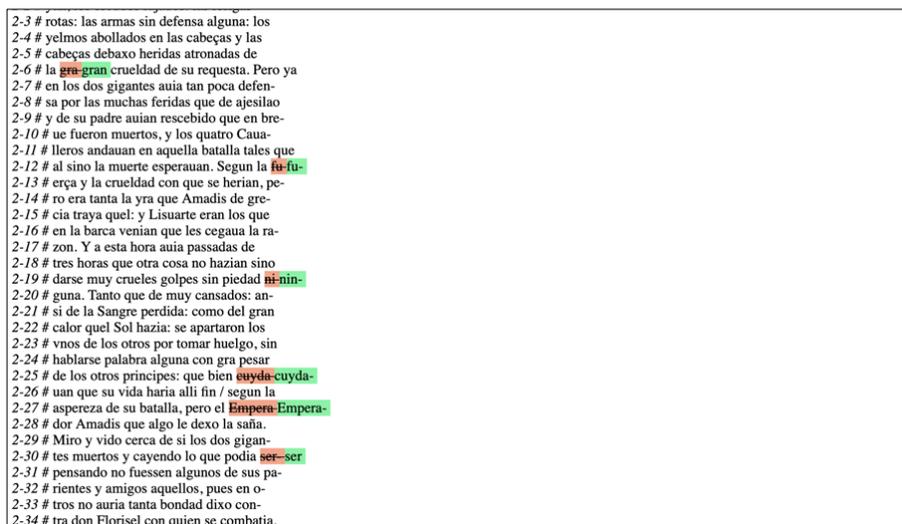


Fig. 2. Ejemplos de errores más frecuentes en una página transcrita

Se han localizado también situaciones en las que Transkribus se equivoca en el reconocimiento de un carácter que transcribe sustituyéndolo por otro, o más de uno, como ocurre, por ejemplo, con el término «cuñada» donde Transkribus convierte equivocadamente una «ñ» en una «u», o bien el caso del verbo «Hablo» cuya letra inicial viene por

error interpretada como una «N». Según se verá también analizando el modelo extendido, es bastante frecuente que la plataforma se equivoque con las letras mayúsculas. Esto se debe a que tales caracteres suelen escasear en los textos de la época, lo que implica que Transkribus tiene pocas ocasiones de verlos y aprenderlos durante la transcripción manual necesaria para su aprendizaje.

Valoración del modelo extendido *Spanish Gothic* (Siglos XV-XVI)

La cronología de publicación de los títulos elegidos para crear el modelo *Spanish Gothic* (nombre en Transkribus: *SpanishGothic_XV-XVI_extended*) abraza un marco temporal muy amplio que empieza en 1487 y finaliza en 1563, fechas en la que se publicaron las ediciones de *Doctrinal de caballeros* y *Libro del conde Partinuplés*, respectivamente, con las que se ha trabajado⁸. Mediante este corpus se aspira a realizar un modelo destinado a la transcripción de textos impresos en castellanos durante el Siglo de Oro español y sus albores. Análogamente, se ha querido cubrir también una vasta geografía que refleje la multitud de talleres activos en la publicación de libros en castellano cuya localización trasciende incluso, amén de los confines de los reinos españoles, las fronteras peninsulares como en el caso de la *Tragicomedia de Calisto y Melibea* impresa en 1515 en Roma, en el taller de Marcellus Silber. A pesar de haberse impreso en lugares y décadas diferentes, todas estas obras comparten unas características editoriales que las convierten en un corpus homogéneo de

⁸ Las obras y las ediciones utilizadas para la creación del modelo *Spanish Gothic 15th-16th Century* son las siguientes: *Doctrinal de los Caballeros* de Alonso de Cartagena (Burgos, Fadrique Biel de Basilea, 1487), *La Fiameta* de Juan Boccaccio (Salamanca, [Impresor de la Gramática de Nebrija], 1497), *Crónica del Rey Don Rodrigo* (Crónica Sarracina) de Pedro de Corral ([Sevilla], [Meinardo Ungut y Estanislao Polono], 1499); *Retablo de la Vida de Cristo* de Juan de Padilla (Sevilla, Juan Cromberger, 1510); *Tragicomedia de Calisto y Melibea* (Roma, Marcelo Silber, 1515); *Historia de la linda Magalona* (Sevilla, Jacobo Cromberger, 1519); *Libro del conde Partinuplés* (Sevilla, Jacobo Cromberger, 1519); *Lisuarte de Grecia* de Juan Díaz (Sevilla, Jacobo y Juan Cromberger, 1526); *Historia del rey Canamor* (Valencia, Jorge Costilla, 1527); *Florando de Inglaterra* (Lisboa, Germán Gallarde, 1545); *Silves de la Selva* de Pedro de Luján (Sevilla, Herederos de Dominico de Robertis), 1549; *Lisuarte de Grecia* de Feliciano de Silva (Sevilla, Jácome Cromberger 1550); *Historia de la reina Sebilla* (Burgos, Felipe de Junta, 1551); *Libro del conde Partinuplés* (Burgos, Herederos de Juan de Junta, 1558); *Leandro el Bel* de Pedro de Luján (Toledo, Miguel Ferrer, 1563); *Libro del conde Partinuplés* (Burgos, Felipe de Junta, 1563).

textos apto para la creación de un modelo de transcripción automatizada. La *conditio sine qua non* es la tipología de tipos móviles utilizados para su impresión: la letra gótica. Al mismo tiempo, es de suma importancia la calidad de la reproducción digital de los textos, sea a color o en blanco y negro. Gracias a las peculiaridades del corpus que se han descrito anteriormente es posible aplicar el modelo a textos publicados no solo en diferentes décadas del Siglo de Oro español, sino también en lugares y talleres distintos. Considérese además que la transcripción manual de las obras antes citadas se ha realizado a partir de reproducciones digitales extraídas de varios portales y que las disparidades en el formato y en la calidad de la imagen han garantizado, contrariamente a lo que se podría pensar, un entrenamiento más objetivo de Transkribus. El software puede no solamente trabajar con reproducciones a color y en blanco y negro, sino también, como tendré ocasión de demostrar, con imágenes más o menos borrosas.

Para valorar la eficacia del modelo *Spanish Gothic* se han concebido dos tipos de pruebas. Primero, y como en el caso del modelo individual antes descrito, se ha aplicado el modelo a un texto utilizado para su creación y, en segundo lugar, se ha experimentado en obras ajenas que no pertenecen al corpus de textos transcritos manualmente para crear dicho modelo. Dado que entre los textos utilizados para la creación del modelo extendido figura también el *Silves de la Selva*, o sea, la misma obra para la cual se había realizado previamente un modelo individual de transcripción automatizada, se ha decidido aplicar también el modelo extendido a este texto a fin de evaluar cuál de los dos modelos, el individual creado *ex profeso* para esta novela, o el extendido –de cuyo corpus forma parte también el *Silves*– se revela mejor para su transcripción. De tal manera que la aplicación del modelo extendido al *Silves de la Selva* permite comparar directamente los dos modelos y realizar un análisis de su eficacia.

La aplicación del modelo extendido *Spanish Gothic* para la transcripción del *Silves de la Selva* ha dado resultados sorprendentes. La plataforma ha cometido solamente siete errores por página, es decir, la mitad exacta de los que había cometido sirviéndose del modelo individual. En la imagen que se propone a continuación es posible comparar la cantidad y la calidad de los errores (Fig. 3). Según se aprecia, el modelo

extendido no deja de cometer, si bien en raras ocasiones, los mismos errores producidos por el modelo individual, como ocurre, por ejemplo, con el verbo «cuidauan» que en ambos casos no se reconoce como una única palabra sino como dos, «cuyda» y «uan». Adviértase que en esta, como en otras situaciones, el error se produce al final de una línea, punto crítico de la transcripción debido a que sus cortes inducen evidentemente a error a la plataforma. Al contrario, por lo que concierne a la unión y separación de palabras colocadas en el interior y no al final de una línea, el *Spanish Gothic* resulta más idóneo para su transcripción. Tan solo por poner un ejemplo, considérese la palabra «guardarse» que cuando se aplica el modelo extendido se escribe como corresponde, mientras que se separa indebidamente en el individual. Sin embargo, en la transcripción del modelo extendido pueden apreciarse errores que no se dan en el modelo individual, como ocurre con la abreviatura de la palabra «dom» que no se desarrolla, pero se trata de raros casos.

Como conclusión del cotejo de las dos transcripciones puede verse que el modelo extendido aplicado a un texto utilizado para su creación es mucho más fiable que un modelo individual creado para el mismo texto. Gracias a la utilización del modelo *Spanish Gothic* los errores cometidos por la plataforma han disminuido notablemente determinando una exactitud de la transcripción automatizada del 99,78%; la plataforma se equivoca solamente 0,22 veces cada cien caracteres, espacios incluidos.

Así pues, falta por analizar la eficacia del modelo extendido cuando se aplica a los textos para los que ha sido creado: las obras que no formaban parte del corpus de textos utilizados para su creación.

<p>1-1 # toda la otra gente de la Fusta de los Gi-</p> <p>1-2 # gantes: fue muerta y los otros principes</p> <p>1-3 # don florisel yanaxartes-y anaxartes se pararon a mirar</p> <p>1-4 # la batalla que era tan braua: quanto ja-</p> <p>1-5 # mas se vido. Porque alli no tenian lu-</p> <p>1-6 # gar de guardar-se guardarse tambien como en tier-tier.</p> <p>1-7 # ra, en tal manera que do quiera que se</p> <p>1-8 # alcancauan yua al suelo escudo y armas</p> <p>1-9 # y abueltas la carne / sin poder resistir</p> <p>1-10 # cosa a los pesados cuchillos de los gigan-</p> <p>1-11 # tes: que solo vna ventaja le tenian los</p> <p>1-12 # principes y era ser muy diestros en el es-</p> <p>1-13 # grima con que reparauan y herian a sus</p> <p>1-14 # contrarios lo mejor que podian tanto que</p> <p>1-15 # al cabo de dos horas que la batalla tura-</p> <p>1-16 # ua. Ya los gigantes dauan señal de su ven-</p> <p>1-17 # cimiento y de la mejoría que los princi-</p> <p>1-18 # pes sobrellos tenian: mas no sin falta de</p> <p>1-19 # sangre que todo estaua lleno della tanto</p> <p>1-20 # que por los embornales de la fusta pas-</p> <p>1-21 # saua, saliendo a dar diferente matiz a la-</p> <p>1-22 # cercana mar-mar. Estando pues la batalla</p> <p>1-23 # en el estado que aueys oydo, parescio por</p> <p>1-24 # la mar vna barca en que dos Caualleros</p> <p>1-25 # venian: los quales como la fusta vieron</p> <p>1-26 # el vno embraçando el escudo, dio vn salto</p> <p>1-27 # en la fusta: la espada en la mano y tras el</p> <p>1-28 # el otro cauallero, pero el principe Don</p> <p>1-29 # Elorisel-Elorisel de niquea y Anaxartes que a-que-que-</p> <p>1-30 # llo vieron, pensando que en fauor de los</p> <p>1-31 # gigantes venian: se le pusieron delante sus</p> <p>1-32 # espadas en las manos: y comiença vna</p> <p>1-33 # cruel y aspera batalla quanto jamas se-se</p> <p>1-34 # vido. Dandose-Dandose mortales golpes que a</p> <p>1-35 # vezes arrodillauan: a vezes cayan: que</p> <p>1-36 # os dire sino que la fusta toda cruxia con</p> <p>1-37 # los desmesurados golpes que se dauan</p> <p>1-38 # resonando el hecho delllos con la ma-</p> <p>1-39 # yor sonoridad del mundo por las con-</p> <p>1-40 # cauidades de la mar. Pero a esta sazón</p> <p>1-41 # ya el principe don rogel auia dado con</p> <p>1-42 # el gigante con quien peleaua a sus pies</p> <p>1-43 # cortandole las enlazaduras del yelmo</p>	<p>1-1 # toda la otra gente de la Fusta de los Gi-</p> <p>1-2 # gantes: fue muerta y los otros principes</p> <p>1-3 # do-don florisel yanaxartes-y anaxartes se pararon a mirar</p> <p>1-4 # la batalla que era tan braua: quanto ja-</p> <p>1-5 # mas se vido. Porque alli no tenian lu-</p> <p>1-6 # gar de guardarse tambien como en tier-</p> <p>1-7 # ra, en tal manera que do quiera que se</p> <p>1-8 # alcancauan yua al suelo escudo y armas</p> <p>1-9 # y abueltas la carne / sin poder resistir</p> <p>1-10 # cosa a los pesados cuchillos de los gigan-</p> <p>1-11 # tes: que solo vna ventaja le tenian los</p> <p>1-12 # principes y era ser muy diestros en el es-</p> <p>1-13 # grima con que reparauan y herian a sus</p> <p>1-14 # contrarios lo mejor que podian tanto que</p> <p>1-15 # al cabo de dos horas que la batalla tura-</p> <p>1-16 # ua. Y a los gigantes dauan señal de su ven-</p> <p>1-17 # cimiento y de la mejoría que los princi-</p> <p>1-18 # pes sobrellos tenian: mas no sin falta de</p> <p>1-19 # sangre que todo estaua lleno della tanto</p> <p>1-20 # que por los embornales de la fusta pas-</p> <p>1-21 # saua, saliendo a dar diferente matiz a la</p> <p>1-22 # cercana mar. Estando pues la batalla</p> <p>1-23 # en el estado que aueys oydo, parescio por</p> <p>1-24 # la mar vna barca en que dos Caualleros</p> <p>1-25 # venian: los quales como la fusta vieron</p> <p>1-26 # el vno embraçando el escudo, dio vn salto</p> <p>1-27 # en la fusta: la espada en la mano y tras el</p> <p>1-28 # el otro cauallero, pero el principe-principe Don</p> <p>1-29 # Elorisel de niquea y Anaxartes que aque-</p> <p>1-30 # llo vieron, pensando que en fauor de los</p> <p>1-31 # gigantes venian: se le pusieron delante sus</p> <p>1-32 # espadas en las manos: y comiença-comiençan vna</p> <p>1-33 # cruel y aspera batalla quanto jamas se</p> <p>1-34 # vido. Dandose mortales golpes que a</p> <p>1-35 # vezes arrodillauan: a vezes cayan: que</p> <p>1-36 # os dire sino que la fusta toda cruxia con</p> <p>1-37 # los desmesurados golpes que se dauan</p> <p>1-38 # resonando el hecho delllos con la ma-</p> <p>1-39 # yor sonoridad del mundo por las con-</p> <p>1-40 # cauidades de la mar. Pero a esta sazón</p> <p>1-41 # ya el principe don rogel auia dado con</p> <p>1-42 # el gigante con quien peleaua a sus pies</p> <p>1-43 # cortandole las enlazaduras del yelmo</p>
<p>1-44 # y la cabeça. Se paro a ver la batalla que</p> <p>1-45 # entre los otros passaua: espantado de</p> <p>1-46 # ver cosa tan fiera quanto jamas lo viera</p> <p>2-1 # porque ya los quatro Caualleros tra-</p> <p>2-2 # yan, los escudos rajados: las lorigas</p> <p>2-3 # rotas: las armas sin defensa alguna: los</p> <p>2-4 # yelmos abollados en las cabeçaç y las</p> <p>2-5 # cabeçaç debaxo heridas atronadas de</p> <p>2-6 # la gra-gran crueldad de su requesta. Pero ya</p> <p>2-7 # en los dos gigantes auia tan poca defen-</p> <p>2-8 # sa por las muchas feridas que de ajessiao</p> <p>2-9 # y de su padre auian rescebido que en bre-</p> <p>2-10 # ue fueron muertos, y los quatro Caua-</p> <p>2-11 # lleros andauan en aquella batalla tales que</p> <p>2-12 # al sino la muerte esperauan. Segun la fu-fu-</p> <p>2-13 # erça y la crueldad con que se herian, pe-</p> <p>2-14 # ro era tanta la yra que Amadis de gre-</p> <p>2-15 # cia traya quel: y Lisuarte eran los que</p> <p>2-16 # en la barca venian que les cegaua la ra-</p> <p>2-17 # zon. Y a esta hora auia passadas de</p> <p>2-18 # tres horas que otra cosa no hazian sino</p> <p>2-19 # darse muy crueldes golpes sin piedad ni-nin-</p> <p>2-20 # guna. Tanto que de muy cansados: an-</p> <p>2-21 # si de la Sangre perdida: como del gran</p> <p>2-22 # calor quel Sol hazia: se apartaron los</p> <p>2-23 # vnos de los otros por tomar huelgo, sin</p> <p>2-24 # hablarse palabra alguna con gra pesar</p> <p>2-25 # de los otros principes: que bien cuyda-cuyda-</p> <p>2-26 # uan que su vida haria alli fin / segun la</p> <p>2-27 # aspereza de su batalla, pero el Empera-Empera-</p> <p>2-28 # dor Amadis que algo le dexo la saña.</p> <p>2-29 # Miro y vido cerca de si los dos gigan-</p> <p>2-30 # tes muertos y cayendo lo que podia ser-ser</p> <p>2-31 # pensando no fuesen algunos de sus pa-</p> <p>2-32 # rientes y amigos aquellos, pues en o-</p> <p>2-33 # tros no auria tanta bondad dixo con-</p> <p>2-34 # tra don Florisel con quien se combatia.</p> <p>2-35 # Por dios cauallero que me digays co-</p> <p>2-36 # mo passa esta auentura: porque quiça no</p> <p>2-37 # hierre contra vosotros: que mucho me</p> <p>2-38 # da el coraçon que soys personas con-</p> <p>2-39 # tra quien no deua alçar el espada: don</p>	<p>1-44 # y la cabeça. Se paro a ver la batalla que</p> <p>1-45 # entre los otros passaua: espantado de</p> <p>1-46 # ver cosa tan fiera quanto jamas lo viera</p> <p>2-1 # porque ya los quatro Caualleros tra-</p> <p>2-2 # yan, los escudos rajados: las lorigas</p> <p>2-3 # rotas: las armas sin defensa alguna: los</p> <p>2-4 # yelmos abollados en las cabeçaç y las</p> <p>2-5 # cabeçaç debaxo heridas atronadas de</p> <p>2-6 # la gran crueldad de su requesta. Pero ya</p> <p>2-7 # en los dos gigantes auia tan poca defen-</p> <p>2-8 # sa por las muchas feridas que de ajessiao</p> <p>2-9 # y de su padre auian rescebido que en bre-</p> <p>2-10 # ue fueron muertos, y los quatro Caua-</p> <p>2-11 # lleros andauan en aquella batalla tales que</p> <p>2-12 # al sino la muerte esperauan. Segun la fu-</p> <p>2-13 # erça y la crueldad con que se herian, pe-</p> <p>2-14 # ro era tanta la yra que Amadis de gre-</p> <p>2-15 # cia traya quel: y Lisuarte eran los que</p> <p>2-16 # en la barca venian que les cegaua la ra-</p> <p>2-17 # zon. Y a esta hora auia passadas de</p> <p>2-18 # tres horas que otra cosa no hazian sino</p> <p>2-19 # darse muy crueldes golpes sin piedad ni-nin-</p> <p>2-20 # guna. Tanto que de muy cansados: an-</p> <p>2-21 # si de la Sangre perdida: como del gran</p> <p>2-22 # calor quel Sol hazia: se apartaron los</p> <p>2-23 # vnos de los otros por tomar huelgo, sin</p> <p>2-24 # hablarse palabra alguna con gran pesar</p> <p>2-25 # de los otros principes: que bien cuyda-cuyda-</p> <p>2-26 # uan que su vida haria alli fin / segun la</p> <p>2-27 # aspereza de su batalla, pero el Empera-</p> <p>2-28 # dor Amadis que algo le dexo la saña.</p> <p>2-29 # Miro y vido cerca de si los dos gigan-</p> <p>2-30 # tes muertos y cayendo lo que podia ser</p> <p>2-31 # pensando no fuesen algunos de sus pa-</p> <p>2-32 # rientes y amigos aquellos, pues en o-</p> <p>2-33 # tros no auria tanta bondad dixo con-</p> <p>2-34 # tra don Florisel con quien se combatia.</p> <p>2-35 # Por dios cauallero que me digays co-</p> <p>2-36 # mo passa esta auentura: porque quiça no</p> <p>2-37 # hierre contra vosotros: que mucho me</p> <p>2-38 # da el coraçon-coraçon que soys personas con-</p> <p>2-39 # tra quien no deua alçar el espada: don</p>

Fig. 3. Errores de transcripción en unas páginas del *Silves*

Con el cometido de investigar más en profundidad las ventajas que el modelo extendido *Spanish Gothic* puede ofrecer a la comunidad científica y, sobre todo, a los editores de textos, se ha efectuado un segundo tipo de prueba. Esta última consiste en la aplicación del modelo a un corpus externo de textos que no hayan sido previamente utilizados para su creación. Se pretende simular con este examen la labor de un editor interesado en transcripción automatizada, a fin de ofrecer un análisis objetivo de la aplicación del modelo extendido y de su valoración. Las obras utilizadas en esta prueba se han seleccionado casualmente dentro del portal Biblioteca Digital Hispánica, donde están disponibles numerosas reproducciones digitales de diferente formato y calidad, en blanco y negro o a color. Los títulos seleccionados y que, evidentemente, cumplen con las condiciones necesarias para la aplicación del modelo –la tipología de carácter móvil, la letra gótica; la lengua, el castellano; y el periodo de publicación, los siglos XV-XVI– son los que se enumeran a continuación:

- *Felixmarte* (Valladolid, Francisco Fernández de Córdoba, 1556);
- *Reprobación de las supersticiones y hechicerías* (Salamanca, Pedro Tovans, 1540).

La primera de las obras que se ha sometido a esta prueba ha sido el libro de caballerías castellano *Felixmarte* (Valladolid, Francisco Fernández de Córdoba, 1556), cuyo aspecto tipográfico guarda un estrecho parecido con el *Silves de la Selva* al pertenecer ambas al mismo género literario y editorial. De igual manera, también su reproducción digital es, al igual que la anterior, de buena calidad y a color.

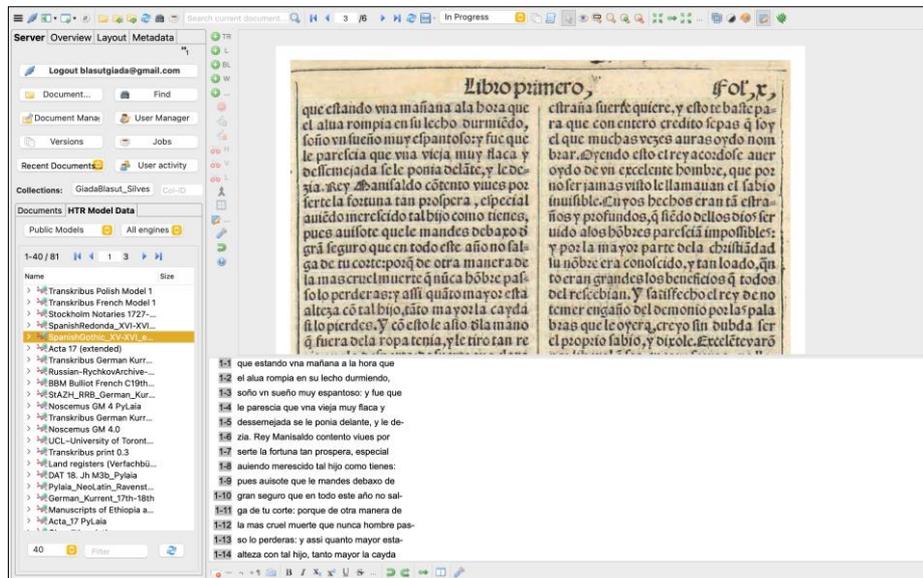


Fig. 5. Prueba 1: *Felixmarte de Hircania* (1556)

Quando se aplica el modelo extendido al *Felixmarte*, los errores cometidos por la transcripción automatizada aumentan levemente, pasando de los siete del *Silves de la Selva* a los diez de este libro. La clasificación de los errores coincide con las precedentes, al mismo tiempo, destacan las omisiones y adiciones impropias de guiones al finalizar una línea. Por poner unos ejemplos, considérense la palabra «christiano» que la plataforma no reconoce como una única palabra, o, al contrario, los términos «esta alteza» que se unen por error mediante la añadidura de un guion. En esta misma estela se insertan también aquellos puntos donde sobran o faltan espacios en blanco entre dos palabras que aparecen así erróneamente unidas o separadas, como ocurre, por ejemplo, con la expresión «Jesuchristo» donde el software une dos palabras. Además, en unas pocas situaciones, aparece un carácter en lugar de otro, como la «m» en la lugar de la «n» en la palabra «con», o un punto en lugar de la coma en la frase «gran sobresalto puso al rey. y a todos los que la vieron». Incluso, en algunas ocasiones, se aprecia la omisión de la letra «n» cuando se trata de una abreviación colocada en el interior o al final de una palabra, como se da en las palabras «con» y «dilacion». A pesar de estos pequeños

y escasos errores, hay que reconocer que también la aplicación del modelo a un texto externo ofrece resultados satisfactorios; tan solo unos diez errores por página, lo que significa que Transkribus se equivoca solamente 0,29 veces cada cien caracteres o, lo que es lo mismo, que la traducción automatizada es correcta en el 99,71% de los casos. Se trata, de toda manera, de un éxito mayor del que puede obtenerse aplicando un determinado modelo individual. Estos datos demuestran, por un lado, la validez del modelo extendido *Spanish Gothic* y, por otro, su enorme utilidad y practicidad; ya no es necesario que el usuario desarrolle un modelo individual para cada obra que le interesa, dado que su aplicación no solo requeriría mucho más tiempo, sino que sería también menos eficaz que el modelo extendido.

El estudio de la aplicación del modelo a obras externas al corpus prosigue con el análisis de un libro didáctico de Pedro Ciruelo, la *Reprobación de las supersticiones y hechicerías* (Salamanca, Pedro Tovans, 1540) –más conocido como *Supersticiones*–, que no está impreso a doble columna como los demás que se han considerado hasta ahora, sino a línea tirada, y cuya reproducción digital está disponible en blanco y negro. Téngase además en cuenta la siguiente variante: la calidad de estas imágenes no es tan nítida como las que se han utilizado anteriormente. Con todo, los resultados de la aplicación del modelo a las *Supersticiones* continúan dando prueba de la validez del modelo extendido. El número de errores sigue siendo escaso, once errores por página, lo que implica que la transcripción es exacta al 99,39%, dato que adquiere aún más relevancia si se considera que este texto tiene una decena de líneas menos que los demás textos analizados.

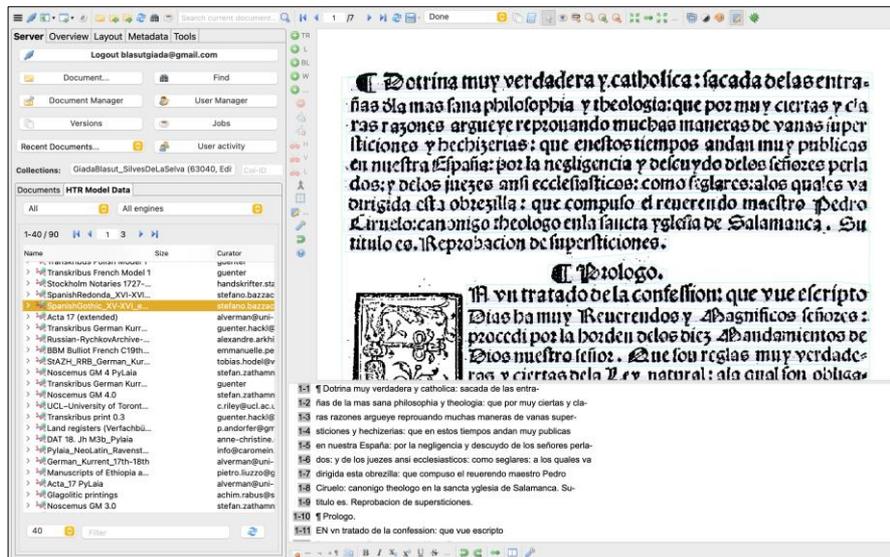


Fig. 5. Prueba 2: *Reprobación de las supersticiones y hechicerías* (1540)

Según el número de errores detectados, se infiere que el modelo *Spanish Gothic* trabaja mejor con las reproducciones a color y con imágenes de buena o calidad media. De este modo, el éxito del reconocimiento de los caracteres por parte del software está directamente relacionado con la calidad de la imagen. Asimismo, la transcripción de las *Supersticiones* realizada a partir de esta peculiar digitalización ofrece la posibilidad de destacar otra de las ventajas más exitosas de Transkribus. Realmente, incluso en los casos de difícil o laboriosa lectura, la plataforma proporciona siempre una interpretación del texto. Dicho de otro modo, aun cuando Transkribus no consigue leer debidamente algún carácter a causa, generalmente, de la mala calidad de la imagen, ofrece siempre una solución al editor, es decir, una posible lectura del texto. En tales casos, la plataforma interpreta los caracteres basándose en el parecido que tienen con otros que ya conoce o a los que está más acostumbrada. Esto hace de Transkribus y sobre todo del modelo *Spanish Gothic* una herramienta puntual y cuidada que ofrece incluso, como en el caso de las *Supersticiones*, una importante ayuda para el editor cuando la calidad de la reproducción digital no es la mejor y dificulta la exacta interpretación del texto..

Bibliografía citada

- Álvarez Márquez, María del Carmen, *Impresores, librerías y mercaderes de libros en la Sevilla del quinientos*, Zaragoza, Libros Pórtico, 2009, vol. I.
- , *La impresión y el comercio de libros en la Sevilla del Quinientos*, Sevilla, Secretariado de Publicaciones de la Universidad, 2007.
- Bazzaco, Stefano, «El Progetto Mambrino y las tecnologías OCR: estado de la cuestión», *Historias fingidas*, 6 (2018), pp. 257-272. DOI: <<https://doi.org/10.13136/2284-2667/89>> (cons. 15/10/2021).
- , «El reconocimiento automático de textos en letra gótica del Siglo de Oro: creación de un modelo HTR basado en libros de caballerías del siglo XVI en la plataforma Transkribus», *Janus. Estudios sobre el Siglo de Oro*, 9 (2020), pp. 534-561. URL: <<https://www.janusdigital.es/articulo.htm?id=160>> (cons. 15/10/2021).
- , «Introducción», en *Leandro el Bel*, ed. Stefano Bazzaco, Alcalá de Henares Editorial, Universidad de Alcalá, 2020.
- Blecu, Alberto, *Manual de crítica textual*, Editorial Castalia, Madrid, 2001.
- Bognolo, Anna, «La prima traduzione italiana dell'*Amadís de Gaula*: Venezia 1546», *Annali della Facoltà di Lingue e Letterature Straniere di Ca' Foscari*, Bulzoni, XXIII, 1, 1984, pp. 1-29.
- , «Libros de caballerías en Italia», en «*Amadís de Gaula*», 1508: quinientos años de libros de caballerías, ed. José Manuel Lucía Megías, Madrid, Biblioteca Nacional de España, Sociedad Española de Conmemoraciones Culturales, 2008, pp. 333-341.
- , «Vida y obra de Mambrino Roseo da Fabriano, autor de libros de caballerías», *eHumanista. Journal of Iberian Studies*, 16 (2010), pp. 77-98.
- , «Il romanzo cavalleresco spagnolo in Italia e la collezione di *Amadís* della Biblioteca Civica di Verona», en *L'età di Carlo V. La Spagna e l'Europa. Sesto quaderno del Dottorato in Letterature Straniere e Scienze della Letteratura Università di Verona*, ed. Silvia Monti, Verona, Edizioni Fiorini, 2011, pp. 125-145.
- , «Il romanzo spagnolo del Rinascimento», en *Forme di romanzi dall'Antico alle soglie del Moderno. Le forme e la storia*, ed. Antonio Pioletti, XIII/2, Soveria Mannelli, Rubbettino, 2020a, pp. 163-186.

- , «Libros de caballerías: ficciones españolas para el Renacimiento europeo», en *En la villa y corte. Trigesima Aurea. Actas del XI Congreso de la Asociación Internacional Siglo de Oro (Madrid, 10-14 de julio de 2017)*, eds. Ana Martínez Pereira, María Dolores Martos Pérez, Esther Borrego Gutiérrez y Inmaculada Osuna Rodríguez, Madrid, UNED-Universidad Complutense, 2020b, pp. 53-82.
- Bognolo, Anna y Bazzaco, Stefano, «Tra Spagna e Italia: per un'edizione digitale del Progetto Mambrino», *eHumanista/IVITRA*, 16 (2019), pp. 20-36.
- Bognolo, Anna, Giovanni Cara y Stefano Neri, *Repertorio delle continuazioni italiane ai romanzi cavallereschi spagnoli. Ciclo di Amadis di Gaula*, Roma, Bulzoni, 2013.
- Castillejo Benavente, Arcadio y Cipriano López Lorenzo, *La imprenta en Sevilla en el siglo XVI (1500-1600)*, Córdoba, Sevilla, Editorial Universidad de Córdoba, Editorial Universidad de Sevilla, 2019.
- Eisenberg, Daniel y M^a del Carmen Marín Pina, *Bibliografía de los libros de caballerías castellanos*, Zaragoza, Prensas Universitarias de Zaragoza, 2000.
- Escudero y Perosso, Francisco, *Tipografía hispalense: anales bibliográficos de la ciudad de Sevilla desde el establecimiento de la imprenta hasta fines del siglo XVIII*, Madrid, Establecimiento tipográfico «Sucesores de Rivadeneyra», 1894.
- Hazañas y la Rúa, Joaquín, *La imprenta en Sevilla: Noticias inéditas de sus impresores desde la introducción del arte tipográfico en esta ciudad hasta el siglo XIX*, Sevilla, Diputación Provincial, 1945-1949, 2 vols.
- Maillard Álvarez, Natalia, *Difusión y circulación de la cultura escrita en Sevilla. 1550-1600*, Tesis doctoral, dir. Carlos Alberto González Sánchez, Sevilla, Universidad de Sevilla, 2007.
- Neri, Stefano, «Cuadro europeo de la difusión del ciclo del Amadís de Gaula (siglos XVI-XVII)», en *Amadís de Gaula: quinientos años después. Estudios en homenaje a Juan Manuel Cacho Blecua*, eds. José Manuel Lucía Megías y María Carmen Marín Pina, Alcalá de Henares, Centro de Estudios Cervantinos, 2008a, pp. 565-591.

- , «El Progetto Mambrino. Estado de la cuestión», en *“Tus obras los rincones de la tierra descubren”*. *Actas del VI Congreso Internacional de la Asociación de Cervantistas*. Alcalá de Henares, 13 al 16 de diciembre de 2006, eds. Alexia Dotras Bravo, José Manuel Lucía Megías, Elisabet Magro García y José Montero Reguero, Madrid, Asociación de Cervantistas; Centro de Estudios Cervantinos, 2008b, pp. 577-589.
- , «Cuadro de la difusión europea del ciclo palmeriniano (siglos XVI-XVII)», en *Palmerín y sus libros: 500 años*, eds. Aurelio González, Axayácatl Campos García Rojas, Karla Xiomara Luna Mariscal, Carlos Rubio Pacho, México D.F., El Colegio de México - Centro de Estudios Lingüísticos y Literarios, 2013, pp. 285-314.
- Romero Tabares, Isabel, *La mujer casada y la amazona. Un modelo femenino renacentista en la obra de Pedro de Luján*, Sevilla, Universidad de Sevilla, 1998.
- Sánchez-Martí, Jordi, «The Printed Popularization of The Iberian Books of Chivalry Across Sixteenth-Century Europe», en *Crossing Borders, Crossing Cultures*, eds. Massimo Rospoche, Jeroen Salman and Hannu Salmi, Berlin, Boston, De Gruyter Oldenbourg, 2019, pp. 159-180.
- Schlusemann, Rita y Krystyna Wierzbicka-Trwoga, «Narrative Fiction in Early Modern Europe», *Quaerendo*, 51/1-2 (2021), pp. 160-188.
- Wilson, Louise, «The Publication of Iberian Romance in Early Modern Europe», en *Translation and the Book Trade in Early Modern Europe*, eds. J. Pérez Fernández y E. Wilson-Lee, Cambridge, Cambridge University Press, 2014, pp. 201-216.