

**Esteban Romero Frías y
María Sánchez González (editores)**
Prólogos de Paul Spence y de Nuria Rodríguez Ortega

Esteban Romero Frías; Octavio Rubiera; Belén Álvarez; Carlos Arcila; Mabel Calderín Cruz; Luis Núñez; Ysabel Briceño; Erika Ortega; Silvia Eunice Gutiérrez; María Sánchez González; Daniel Villar-Onrubia; Juan M. García Sierra; Manuel D. Pérez Heredia; Javier Jiménez Cuadros; Patricia Lázaro Pernias; Daniel Tena Parera; David Roca Correa; Josep María Blanco Pont; Alfonso González Quesada; Álvaro Baraibar; Victoria Tur-Viñes; María del Carmen Quiles-Soler; José Manuel de Pablos; Ana María Rodera Bermúdez; Ana María González Ramos; Ernesto Priani Saisó

Ciencias Sociales y Humanidades Digitales

**Técnicas, herramientas y experiencias de
e-Research e investigación en colaboración**

CAC, Cuadernos Artesanos de Comunicación / 61



Universidad
de La Laguna



Universitat d'Alacant
Universidad de Alicante



UNIVERSIDAD
DE MÁLAGA



Sociedad Latina de
Comunicación Social



Marcar y leer TEI, la experiencia de un taller

Ernesto Priani Saisó. Universidad Nacional Autónoma de México
epriani@filos.unam.mx

1. Introducción

LA BIBLIOTECA DIGITAL del Pensamiento Novohispano (BdPn) (www.bdpn.unam.mx) es resultado de un proyecto de investigación que en el año 2006 comenzó a experimentar con algunas herramientas de edición digital para producir una publicación académica de textos novohispanos accesible vía Internet. Se trataba, ante todo, de identificar sistemas y métodos que ayudaran a crear una edición que sirviera como instrumento para estudiar las obras, más allá de ponerlas a disposición en línea. En un inicio, se comenzó a trabajar con el *Discurso ethereológico del nuevo cometa* de fray Diego Rodríguez, un documento que ha sido señalado como fundamental para la formación del pensamiento científico en México (Trabulse, 2001: 7-8), pero que por diversas circunstancias no contaba con una edición contemporánea, lo que añadía valor a la empresa, pues no sólo se trabajaba en la creación de una herramienta para facilitar el estudio de la obra a través de Internet, sino que además se recuperaba un texto importante.

El punto de partida era hacer una transcripción digital del texto y, a partir de ésta, buscar la forma para que se pudieran generar automáticamente algunos índices, así como diseñar un mecanismo

para trabajar colaborativamente en la identificación de la fuentes clásicas y medievales del *Discurso*, una labor que no había sido aún emprendida por quienes han investigado la obra, así como un sistema para hacer comentarios al texto.

Como resultado de estas intenciones se formó un pequeño grupo de trabajo de tres personas, compuesto por el autor y dos alumnos de la licenciatura en filosofía. Se buscó financiamiento por parte de la Universidad Nacional Autónoma de México y se estableció un convenio de colaboración con la Dirección General de Servicios de Cómputo Académico (DGSCA) de la Universidad (hoy Dirección General de Tecnologías de la Información). De ahí, se procedió a trabajar en dos direcciones: en el estudio de la obra, sus fuentes y sus repercusiones, y en la parte técnica junto con personal de la DGSCA. Esta consistió en el diseño de la Definición de Tipo de Documento (DTD por sus siglas en inglés), que es una descripción de estructura y sintaxis de un documento XML, en la cual estableceríamos las categorías con las que queríamos formar los índices. Paralelo a esto comenzamos a analizar y ejecutar el proceso de marcado del texto con las categorías anteriores, así como el sistema de procesamiento del documento XML para la generación de los índices y las herramientas para hacer identificar fuentes y hacer comentarios.

Conforme avanzamos, tanto en el conocimiento del texto como en las posibilidades operativas del sistema que estábamos desarrollando, fuimos ampliando nuestra perspectiva y simultáneamente, reconociendo algunos problemas con las decisiones técnicas tomadas.

En la parte del conocimiento del texto, se observó que se trataba de una obra vinculada a un conjunto de documentos y, junto con ellos, constituía el testimonio de una controversia sobre la naturaleza de los cometas en la Nueva España, originada por el paso de dos de estos fenómenos celestes, uno en 1652 y otro en 1680, visibles desde América. Esas obras, al igual que el *Discurso ethereológico*, carecían de una edición moderna -con la única excepción de la *Libra filosófica y astrológica* de Carlos de Sigüenza y Góngora- y de ello surgió la idea de diseñar una biblioteca digital, como un laboratorio de experimentación.

He analizado ya en otra parte los problemas que detectamos en el desarrollo tecnológico del sistema (Priani, 2012: 101). Estos eran

básicamente dos. Por un lado se desarrolló la base de datos que generaba los índices automatizados, para ser alimentada de forma independiente de los archivos XML marcados con nuestras categorías, lo que significaba duplicar el trabajo, en vez de simplificarlo. Por otro lado, y este es el tema más relevante de este capítulo, las categorías XML con que habíamos creado nuestro DTD nos estaban causando muchas dificultades. Habíamos definido más de cien categorías, muchas de las cuales eran muy ambiguas, o bien se duplicaban sin quedar claro cuál era la diferencia entre algunas; en conjunto, representaban un instrumento sumamente difícil de manejar.

Tras una discusión del proyecto con el *Virtual Humanities Lab* de la Universidad de Brown en 2008, decidimos abandonar el intento de generar un DTD y unas categorías propias y comenzar a trabajar con las marcas generadas por la *Text Encoding Initiative* (TEI)¹⁸² tanto porque simplificaba el trabajo, como porque nos hacía trabajar con estándares internacionales.

Dejamos también el proceso de generar la base de datos por un lado y los XML por otro para trabajar en un sistema que extrajera directamente los datos de los XML, interrumpimos la colaboración con la DGSCA y comenzamos prácticamente un nuevo proyecto. Todo esto nos enfrentó a diversos retos, pero uno de ellos fue el de nuestra formación y el de la formación de los colaboradores de la Biblioteca Digital del Pensamiento Novohispano en el uso de TEI para el marcado de los textos.

2. Planteamiento

El objetivo de estas líneas es reflexionar sobre la experiencia de formación en el uso de las guías del *Text Encoding Initiative* (TEI) dentro de la Biblioteca Digital del Pensamiento Novohispano (BdPn), de estudiantes y profesores de las licenciaturas en Filosofía, Historia y Letras de la Facultad de Filosofía y Letras de la Universidad Nacional Autónoma de México, y de la Facultad de Letras de la Universidad Veracruzana.

¹⁸²<http://www.tei-c.org/Guidelines/>

Puesto que trataremos de dar cuenta de la experiencia pedagógica en el proyecto, el artículo será ante todo descriptivo; de los objetivos, alcances y procedimientos seguidos en la BdPn, como de la utilización de las guías TEI para el marcado de texto y de la aplicación específica de éstas dentro de la edición digital de obras novohispanas. Por otro lado, describirá la forma en que fue enseñado el uso de las guías TEI entre los estudiantes de dos instituciones, de los ejercicios de aplicación, del proceso de auto-aprendizaje implicado a lo largo del proyecto y de los resultados de esa experiencia pedagógica; los errores más frecuentes, las dificultades más grandes para aprender a marcar los textos, así como los resultados más notables del aprendizaje; el marcado de las obras pero también la discusión sobre el etiquetado como método de lectura y como sistema hermenéutico.

Las limitaciones de esta exposición están dadas por su propia naturaleza: se trata de una experiencia particular dentro de un proyecto documentando un caso específico que, aunque comparable con otras experiencias similares, guarda su singularidad tanto por el modo de adaptación y utilización de las guías TEI, como por las condiciones específicas en que fue impartido dentro de un marco institucional en el que la enseñanza del uso de herramientas digitales no es impulsada pero sí reconocida (Priani, 2012: 99).

En este sentido, el capítulo tendrá al final un carácter reflexivo sobre las implicaciones tanto metodológicas como pedagógicas de la capacitación en el uso de una herramienta técnica dentro de las humanidades. El interés es subrayar que las guías TEI se pueden comprender como un método de lectura y de análisis provechoso más allá del campo específico de su aplicación técnica, para la enseñanza de las humanidades en general. Esto con la finalidad de sostener que *e-Research* como la Biblioteca Digital del Pensamiento Novohispano hace aportaciones específicas al campo de la enseñanza de las humanidades con el uso de herramientas digitales dentro de los procesos pedagógicos.

Nuestro objetivo, una vez que redefinimos los alcances y la nueva forma de la biblioteca, fue crear una colección de ediciones académicas de textos novohispano. Al tratarse sobre todo de impresos y manuscritos de los que se conserva en muchos casos una sola copia, pensamos que una edición académica debía entenderse

sobre todo como la preservación del texto según lo propone Tanselle (1995: 10). Esta tiene como meta mantener y transmitir el texto en su forma original sobre la base de una práctica editorial plenamente sustentada. Pero, en la medida en que este texto se presenta dentro de una edición electrónica, se espera que su preservación se vea acompañada y enriquecida por las posibilidades del medio para explorar los textos e integrar su conexión con otros textos (Schillingsburg, 1996: 165):

“(An electronic scholarly edition) it is a tool for students of a work, returning to it with the intention of exploring its history, its connectives tissues, its roots and ramifications. Such students want to be able to search texts from stern to stern electronically, to ove directly from one passage in a work its corresponding passages in other versions or from to its sources or context in other works. They want annotations, textual variants, variorums of critical commentaries...”.

Asumimos, pues, que una edición académica contempla tanto el rigor con el que el texto es reproducido como la presentación de introducciones, notas y el aparato textual fruto de su estudio (Price, 2008). Así, en la BdPn las transcripciones diplomáticas y paleográficas que son la base de la digitalización de los textos, reúnen plenamente los requerimientos de una edición académica. Los criterios¹⁸³ seguidos están ajustados a la red del Corpus Hispánico y Americano en la Red: Texto Antiguos (CHARTA)¹⁸⁴.

El estudio, la interpretación de las obras así como las notas críticas se encuentran presentes en dos lugares. Uno, la herramienta para hacer anotaciones filológicas e interpretativas que, al momento de escribir estas líneas, septiembre de 2013, se encuentra deshabilitada para su rediseño. Dos y sobre todo a través de las etiquetas de marcado TEI, que se han convertido en la herramienta principal para la exploración y conocimiento de los textos de parte del grupo de trabajo.

¹⁸³ <http://www.bdpn.unam.mx/criteria>

¹⁸⁴ <http://www.charta.es/criterios-de-edicion-/>

3. Las marcas TEI

El valor de anotación interpretativa de las etiquetas es una conclusión a la que hemos llegado después de un proceso de práctica y enseñanza del marcado textual. Como ya se señaló, la decisión de utilizar las guías TEI tenía la intención de resolver un problema técnico, pero también la de ajustar el proceso de trabajo a estándares internacionales. Sin haberlo nunca hecho explícito, los miembros del grupo de la BdPn tomamos inicialmente las etiquetas sólo como instrucciones para ser interpretadas por la computadora y hasta que comenzamos a explicar a otros cómo llevar a cabo las anotaciones no nos dimos cuenta de que "... el marcado viene a hacer explícitas ciertas características de aquel texto original en papel, lo exhibe trayéndolo a la visibilidad dentro del texto. El marcado es, por tanto, esencialmente notacional. Afecta la expresión del texto, ambas, digital y bibliográfica, añadiendo cierto tipo de estructura a los dos" (Buzzetti y McGann, 2004). Y esas estructuras, como notaremos más adelante, son esencialmente hermenéuticas.

El equipo de trabajo de la BdPn aprendió a marcar con las guías TEI de manera autónoma, utilizando la documentación disponible en el sitio www.tei.org. Parte de la decisión de utilizar la versión TEI *Lite*¹⁸⁵ se debió a que está construida con un sentido más pedagógico y la explicación de cómo se genera un documento es, así mismo, un ejemplo de lo que se puede hacer con las marcas. Además, el esquema utilizado por TEI *Lite* abarcaba la totalidad de las marcas tanto descriptivas como analíticas en las que estábamos interesados, así que decidimos hacer una adaptación de ella para los fines específicos de la biblioteca.

Como ya ha sido reiterado, nuestro interés estaba centrado en la automatización de unos índices de nombres y de lugares que decidimos extender por las facilidades que nos ofrecían las marcas TEI a índices de referencias, fechas, términos y abreviaturas. Sólo con posterioridad nos dimos cuenta que buscábamos marcar la evidencia de la cultura compartida por todos los autores de los textos comprendidos en la biblioteca, así como las fuentes generales utilizadas por todos ellos, sin distinción entre modernas y antiguas. Cosa, por lo demás, que había sido el ánimo que orientó, desde su

¹⁸⁵http://www.tei-c.org/release/doc/tei-p5-exemplars/html/tei_lite.doc.html

inicio, nuestra decisión de hacer una herramienta de estudio de textos digitales. Por supuesto, incluimos también marcas que pudieran describir el documento como tal tanto en su parte bibliográfica como estructural, pues eran necesarias en su identificación como en su posterior representación en HTML. De esta forma la versión del TEI *Lite* adaptada por nosotros, es la siguiente:

a) Marcas para describir digitalmente las obras:

<text></text> Declara el inicio y el fin del texto marcado.

<front></front> Describe la portada o frente del documento impreso.

<titlePage></titlePage> Declara que se trata de la página de título.

<docTitle></docTitle> Marca el título del documento.

<docImprint></docImprint> Declara el lugar y el impresor del texto.

<docDate></docDate> Marca la fecha de impresión del texto.

<body></body> Corresponde al cuerpo del texto. Abre y cierra la totalidad de éste.

<p></p> Describe el inicio y el final de cada párrafo.

<div></div> Declara el inicio y cierre de una división. En la biblioteca se utiliza para establecer el inicio y el final de un folio. El folio es, junto con el párrafo, la unidad de referencia de los textos de la biblioteca.

<pb></pb> Marca el numeral del folio o el número de página, si existe.

b) Marcas de interpretación:

<term></term> Marca una palabra, varias palabras, una designación simbólica o el uso de un término técnico. Es una categoría de marcado central, porque permite identificar cierto tipo de palabras y expresiones en el texto cuyo estudio y recuperación es relevante para la investigación. Acepta 7 atributos (type): Médico (Med), Astrológico (Astro),

Astronómico (Astron), Filosófico (Philo), Matemático (Mathe), Teológico (Theolo) y Español antiguo (OldSp).

`<foreign></foreign>` Identifica términos o grupos de términos en un idioma diferente al de la redacción del texto. Acepta el atributo lang (Idioma). Para el idioma que más aparece, el latín, se utiliza el atributo "lat".

`<date></date>` Sirve para marcar una fecha.

Como la gran mayoría de los cometas son nombrados por la fecha en que pasan. Utilizamos "type" y "comet" para señalar que se trata de un cometa.

`<abbr></abbr>` Marca una abreviatura. Las abreviaturas no se regularizan ni se expanden, se etiqueta tal cual está en el texto.

`<name></name>` Marca los nombres propios de personas, dioses, lugares, etcétera. Acepta el uso del atributo "type" con los siguientes elementos: persona (person); planeta (planet); Dios (god); lugar (place); institución (inst); constelación (constelation); signo del zodiaco (zodiacal); divinidad católica (CathDiv). Utiliza "reg" para estandarizar el nombre a su uso contemporáneo.

`<q></q>` Identifica las citas textuales dentro de los textos, en las que no aparece la referencia bibliográfica. Deben colocarse al iniciar y al terminar la cita.

`<cit></cit>` Se utiliza para citas textuales claramente identificadas con su referencia bibliográfica.

`<refTitle></refTitle>` Se utiliza para marcar el título de una obra dentro de un texto. Acepta "reg" para estandarizar el nombre moderno de la obra.

Con la definición plena de este sistema de marcas y con algunos textos ya marcados e incorporados a la biblioteca, el grupo de trabajo comenzó a diseñar un taller de marcado TEI para formar nuevos miembros del grupo, sobre todo estudiantes de las carreras de Filosofía, Historia y Letras de la Facultad de Filosofía y Letras de la Universidad Nacional Autónoma de México, pudiendo incorporarse para realizar su servicio social trabajando en la biblioteca. El taller sirvió, además, para formar dos grupos más de trabajo:

- 1) El grupo de investigación de Escritos de Mujeres, coordinado por la Dra. Clara Inés Ramírez, investigadora del Instituto de Investigaciones sobre la Educación y la Universidad, conformado en su mayoría por estudiantes del posgrado de Historia de la UNAM quienes marcaron para la BdPn *La fundación del convento de Inés de la Cruz* y que actualmente trabajan en otros más para constituir la colección de Escritos de Mujeres en la biblioteca.
- 2) El grupo de marcado de textos coordinado por la Dra. Guadalupe Rodríguez de la Facultad de Letras de la Universidad Veracruzana, constituido por estudiantes de la carrera de esa Facultad que marcaron colaborativamente el *Reportorio de los tiempos, y historia natural desta Nueva España* de Enrico Martínez que actualmente forma parte de la biblioteca.

4. Experiencia pedagógica del marcado en TEI

Enseñar a marcar con TEI es relativamente sencillo. La comprensión de la mecánica del marcado: abrir y cerrar la marca, señalar la categoría, agregar el tipo, meter una regularización, etc., lleva realmente muy poco tiempo. Es mucho más importante la práctica, que es el espacio donde suelen aparecer las dificultades a la hora de reflexionar sobre el texto. En realidad, las dudas son pocas con respecto de la sintaxis del XML, pero muchas y muy variadas en cuanto a la interpretación del documento

El curso-taller que se diseñó tiene una duración de 8 horas, usualmente impartidas en dos jornadas de 4 horas cada una. El contenido del taller contempla dos horas iniciales en que se expone el proyecto de la Biblioteca, se hace una introducción general al XML y se explican las marcas utilizadas en la BdPn, así como el software para marcar XML; normalmente Notepad++¹⁸⁶ en caso de Windows y TextWrangler¹⁸⁷ en caso de Mac, ambas por ser gratuitas y útiles para el propósito.

Después se da paso a una sesión práctica en la que los asistentes que ya cuentan con un texto previamente transcrito, son orientados por el

¹⁸⁶ <http://notepad-plus-plus.org/>

¹⁸⁷ <http://www.barebones.com/products/textwrangler/>

ponente y algún asistente en la formación de su primer documento XML utilizando las marcas TEI que se aplican dentro de la BdPn.

La segunda sesión, dedicada completamente al ejercicio del marcado, comienza con un breve resumen de lo visto y un análisis de los problemas que se hayan presentado en la jornada anterior. El taller finaliza con la validación del documento que cada asistente preparó a través de una herramienta creada ex profeso para ello y que llamamos *TEI validator*¹⁸⁸ y con la corrección final de los errores identificados.

5. Las herramientas didácticas

Para el taller, la BdPn es un recurso didáctico. Se utiliza como referente general, pero también como herramienta de consulta. La documentación de la biblioteca¹⁸⁹, que fue desarrollada para llevar un registro de las decisiones tomadas dentro del proyecto, fue modificada para que cumpliera el rol de material para el taller. Una función semejante tienen los criterios editoriales¹⁹⁰ en cuanto a la transcripción.

Desde que comenzamos a marcar los textos de la BdPn, nos dimos cuenta de la necesidad de contar con una herramienta para poder validar las sintaxis del XML. Dado que el marcado no se hace manualmente, lo usual es que se cometan algunos errores. Por ejemplo, abrir con una categoría y cerrar con otra, `<term>Sangrías</name>`, o bien no escribir correctamente la categoría `<trm>`, omitir el cerrado de la categoría `<name>Pablo<name>`, por sólo citar las más frecuentes.

Fue en función de esa necesidad que se desarrolló un *parser* (analizador semántico) externo a la biblioteca que indicara los errores de la sintaxis del XML. Ese es nuestro *TEI validator* que con el tiempo se convirtió en una herramienta muy útil dentro de los cursos de TEI porque validaba el marcado, señalaba los errores cometidos y posteriormente ofrecía una vista del resultado.

¹⁸⁸<http://siedrix.com/work/tei/>

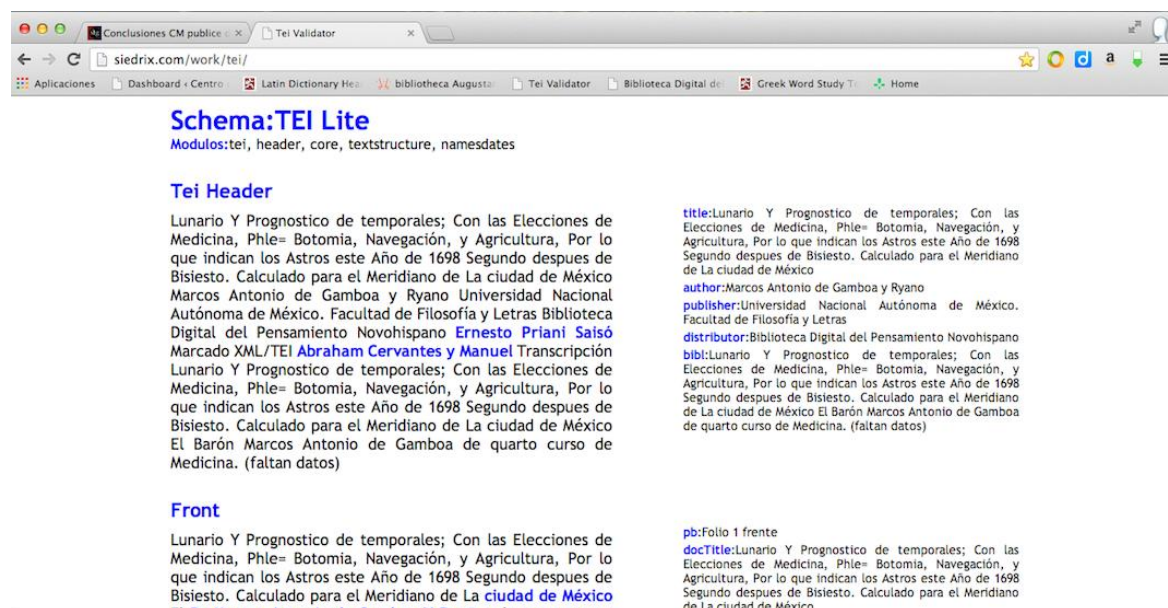
¹⁸⁹<http://www.bdpn.unam.mx/documentation>

¹⁹⁰<http://www.bdpn.unam.mx/criteria>

Figura 1. Error en el TEI Validator.



Figura 2. Texto validado.



6. La experiencia del etiquetado

La riqueza del taller se encuentra, sin embargo, en la experiencia pedagógica misma, visible a través de las dificultades y los problemas que se hacen manifiestos en las sesiones y que nos han permitido reflexionar, con mayor profundidad, en las implicaciones de enseñar a marcar y de las cuales destacaríamos dos: el aprendizaje de un modo de lectura a partir de categorías dadas y el conocimiento de la forma en como se procesa computacionalmente un texto.

En términos generales, lo que se aprende a marcar para la BdPn son dos cosas: primero, los elementos estructurales del texto; el cuerpo del texto, la portada, el título, los autores, la división por folios que utilizamos y la división por párrafos. Esta parte normalmente no ofrece mucha dificultad y se hace mecánica muy fácilmente una vez que se entendió cuáles son los elementos estructurales que interesa describir.

No ocurre lo mismo, sin embargo, con las marcas interpretativas y de contenido que enfrentan al estudiante no con la estructura del texto, que es algo relativamente conocido, sino con un texto que no conoce a profundidad y con la necesidad de reflexionar sobre él con una intención hermenéutica.

Cuando se proponen ciertas categorías para etiquetar palabras dentro de un texto, el proceso de lectura se altera porque hay que dirigirlo a identificar aquellos términos que coinciden con los criterios establecidos para describirlos (marcas). Así, cuando buscamos nombres y la categoría <nombre> tiene atributos que distinguen los nombres propios del de los lugares, el que marca tiene que detenerse a reflexionar si, por ejemplo, San Agustín es un nombre propio o el nombre de una calle o de un convento, para después tomar una decisión. Así, cuando leemos para marcar, vamos al texto guiados por una búsqueda de palabras/frases marcables que nos obligan a reflexionar sobre el texto.

En la experiencia de los talleres, la categorización de las palabras, es decir, la identificación de los términos y los nombres según sus atributos, -como términos médicos o astrológicos, o los nombres de lugares y propios- se dificulta tanto por problemas generales en la comprensión del proceso de marcado, como por las dudas particulares que surgen en la atribución de etiquetas.

Dos son las dificultades iniciales: el marcar en exceso y el no marcar todas las apariciones de las palabras. Normalmente, al enfrentarse al texto con una herramienta formada por etiquetas para clasificar elementos en él, se tiende a tratar de marcar el mayor número de palabras, sin que sea necesario. Como marcar implica hacer una lectura que discrimina cosas dentro del texto, y no una que intente comprenderlo, como usualmente se hace, este problema revela que aún no se ha entendido la naturaleza de la lectura que se hace. Es, pues, una manifestación del cambio en el orden de la lectura y una oportunidad para explicar con mayor precisión, en qué consiste leer para marcar.

Paradójicamente, ocurre también en la primera etapa que cuando se encuentra una palabra que se repite numerosas veces, como podría ser “Tierra” o “Sol” en el caso de los textos de la BdPn, se duda en marcarlos, asumiendo que con una ocasión es suficiente o que la

repetición los vuelve obvios y por tanto, sin relevancia para el marcado. En este caso, la dificultad consiste en comprender que el etiquetado está orientado a la construcción de índices, donde es relevante identificar todas las veces que un texto aparece y que, así como tiene funciones de discriminación, también tiene funciones de sistematización.

Así, una vez que se han superado estos problemas iniciales, nos enfrentamos a la cuestión de fondo: ¿cuál es la etiqueta que corresponde a cada palabra?

Para entender mejor esto usaré el siguiente ejemplo: en la *Libra astronómica y filosófica* de Carlos de Sigüenza y Góngora, se alude de manera reiterada a los marquesados y a los marqueses, cosa que no ocurre en otros textos de la biblioteca. Quienes marcaron la *Libra*, tomaron la decisión –sin discutirla previamente con el grupo de trabajo– de marcar los marquesados de la siguiente forma:

Posee esta gran casa el Ducado del
<nametype="inst">Infantado</name>, los Marquesados del
<nametype="place">Cenete</name> de
<nametype="place">Santillana</name> de
<nametype="place" reg="Arguello de Campo">Arguello de
Campo</name>....

Es interesante observar que por una parte separaron el ducado y el marquesado de su denominación completa como Ducado del Infantado o Marquesado del Cenete y después de hacerlo, asignaron a la denominación un sentido: institución o lugar. En una discusión posterior, se les preguntó la razón para hacerlo así. La explicación que dieron fue simple: asumían que un marquesado hace siempre referencia a un lugar, de modo que había que marcar como lugar los nombres que aparecían asociados y lo mismo en el caso del ducado, en el que sin embargo, pensaron que Infantado era una institución. Por supuesto siguió una discusión sobre qué son los títulos nobiliarios, si tenía sentido marcarlos para los fines de la biblioteca y cómo habría que marcarlos en su caso: por ejemplo, como nombre de instituciones, sin separar el título de marquesado del de su denominación. Eso llevó a la discusión de cómo marcar cuando se referían no al título nobiliario sino a la persona que lo ostenta, es

decir, al Marqués de Santillana, en donde había mantenido el mismo marcado que en el caso del marquesado.

Este ejemplo ilustra con claridad la naturaleza de la mayoría de los problemas que emergen en los talleres; los que surgen de la dificultad para una correcta comprensión de lo que se lee para describirlo con una marca que denote su pertenencia a una cierta categoría, como aquellos que tienen su origen en la dificultad para comprender lo que la marca denota como descripción de un término.

Ambas dificultades tienen que ver con el carácter discriminante y reflexivo de la lectura orientada por unas categorías que buscamos ver reflejadas en los textos. Enseñar a marcar es, en realidad, enseñar a leer de una forma específica y a reflexionar sobre las palabras de una cierta forma. En este sentido no podemos sino coincidir con Singer (2012):

“one of the most interesting and important skills TEI can teach is to help our students become reflective about the various kinds of vocabularies they use to describe texts”.

A diferencia de la lectura comprensiva que persigue la captación del sentido general de una lectura, el marcado se enfoca por una parte, a describir la estructura del texto (títulos, párrafos, folios, etcétera) y por otra, a describir palabras, términos o incluso frases que tienen una función específica dentro de él. Pero un rasgo que no enfatiza Singer es que si bien, el marcado de texto fortalece las estrategias de *Close Reading*, también ayuda a la comprensión de lo que implica el procesamiento computacional de los textos y sus problemas; alerta al estudiante sobre el carácter hermenéutico de las marcas y sobre cómo el procesamiento del texto genera un mapa de la lectura que se haga de éste.

7. Discusión final

En México, las Humanidades Digitales y *lae-Research* aún se encuentran en las primeras etapas de práctica y desarrollo. El reto de una formación a gran escala de estudiantes universitarios en el conocimiento de las metodologías y las técnicas computacionales aplicadas a las humanidades y a las ciencias sociales, apenas comienza.

Sin embargo, la experiencia efectiva llevada a cabo en los talleres de marcado TEI muestra hasta qué punto esta metodología puede ser utilizada pedagógicamente para alcanzar dos objetivos muy amplios, adicionales a la habilidad misma de marcar los textos. Por un lado, su aplicación en la enseñanza de ciertas técnicas de lectura, particularmente las asociadas al llamado *Close Reading*, pero sobre todo al ejercicio reflexivo de la forma de describir la estructura del texto y su contenido a partir del uso de ciertas categorías. Por otro, permite enseñar y hacer conciencia sobre las implicaciones del procesamiento computacional de los textos, sus alcances, sus limitaciones. A final de cuentas, como ha escrito Lou Burnard, el marcado de textos “hace explícita la teoría sobre algún aspecto de un documento, el marcado hace un mapa de una interpretación (humana) del texto por medio de un grupo de códigos a través de los cuales el procesamiento de la computadora puede ser realizado.” (Burnard, 1998)

En última instancia, *lae-Research* ofrece metodologías como el marcado TEI que conducen al aprendizaje de una forma de lectura que es útil no sólo para su aplicación computacional, sino para reforzar otros aspectos del aprendizaje como es el análisis detenido de los textos. Esto, con independencia del uso propio de herramientas digitales como la Biblioteca Digital del Pensamiento Novohispano.

Pensando en el futuro de la enseñanza de las humanidades, el marcado sería sólo el primer paso en el aprendizaje de un sistema de lectura que parte de la discriminación y reflexión sobre el texto, al análisis de los mapas de interpretación como índices y visualizaciones, que su procesamiento produce. Planteamos que, en función del crecimiento de *lae-Research* y de la necesidad de la enseñanza de sus metodologías y técnicas, marcar texto, que hoy se puede aplicar como herramienta para enseñar a leer y analizar un texto mediante categorías para reforzar una formación tradicional, será un componente necesario para la incorporación de otras herramientas y métodos para el estudio de las humanidades.

Referencias bibliográficas

Biblioteca Digital de Pensamiento Novohispano (2011). Disponible en: <http://www.bdpn.unam.mx/> (consultado el 4 de octubre 2013).

- Burnard, L. (2006). "On the hermeneutic implications of text encoding". Disponible en: <http://users.ox.ac.uk/~lou/wip/herman.htm> (consultado el 4 de octubre 2013).
- Buzzetti, D. & McGann, J. (2004). "Electronic Textual Editing: Critical Editing in a Digital Horizon". *Electronic Textual Editing*. Ed. Modern language Association. Disponible en: http://www.tei-c.org/About/Archive_new/ETE/Preview/mcgann.xml (consultado 4 de octubre 2013).
- Priani Saisó, E. (2012). "Finding support for disruption: developing a digital humanities project in Mexico". *Aslib Proceedings: New Information Perspectives*, 64(1): 97-103.
- Price, K.M. (2008). "Electronic Scholarly Editions". En Schreibam, S. et al., *Companion to Digital Literary studies*. Oxford: Blackwell. Disponible en: http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405148641/9781405148641.xml&chunk.id=ss1-6-5&toc.depth=1&toc.id=ss1-6-5&brand=9781405148641_brand (consultado el 4 de octubre 2013).
- Schillingsburg, P.L. (1996). *Scholarly Editing in the Computer Age. Theory and Practice*. Michigan: University of Michigan Press.
- Singer, K. (2012). "The MelesinaTrech Project: Markup Vocabularies, Poetics, and Undergraduate Pedagogy". *Digital Humanities 2012, Conference Abstracts*. Hamburg. Hamburg University Press. Disponible en: <http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/the-melesina-trench-project-markup-vocabularies-poetics-and-undergraduate-pedagogy/> (consultado el 4 de octubre de 2013).
- Trabulse, Elías (2001). "Prólogo" a la *Libra Astronómica y Filosófica* de Carlos de Sigüenza y Góngora. México: Sociedad de Bibliófilos Mexicanos.