

ELECCIÓN DE TEXTO, PRESENTACIÓN Y ENFOQUE GENERAL

Al no tratar mi trabajo de TFG directamente con documentación, la elección del texto a editar fue completamente libre. Conociendo Gallica y su amplio y bien digitalizado catálogo, me pareció interesante elegir las *Grandes chroniques de France*¹ y trabajar un texto medieval en un idioma que manejo bien pero no domino del todo. Que no hubiese en esta elección un propósito puramente utilitario condicionó en buena medida el enfoque que se dio a la edición, que planteé desde el primer momento como amplio y lo más completo posible, haciéndola así polivalente. Esto, como explicaré más adelante, pudo haber sido una mala decisión.

Las *Crónicas* fueron encargadas originalmente por Luis IX y continuadas por los monjes de Saint-Denis por orden de Carlos V hasta 1379. En ellas se presenta un relato inédito de la historia de la realeza francesa; texto e ilustración contribuyen a reforzar la idea de la legitimidad y el poder de los Valois en el último cuarto del siglo XIV, trazando una historia de los francos que se remonta a los troyanos y que cubre a las dinastías Merovingia, Carolingia y la de los Capeto y, a las que se añaden posteriormente los reinados contemporáneos de los Valois Juan II el Bueno y de Carlos V el Sabio. Se ofrece así una rara oportunidad de comprender la génesis y la evolución del pensamiento político real y de la propaganda francesa de época. En esta edición se presentan el prefacio de la crónica, la tabla de contenidos que se hace de ella y sus cinco primeros capítulos, que comienzan en la guerra de Troya y terminan con Clodión, de la dinastía Merovingia².

CODIFICACIÓN EN TEI. ELECCIÓN DE ETIQUETAS PRELIMINAR

En el planteamiento de implementación de TEI en este texto tomé ejemplos de varios sitios. Evidentemente, el primero fue el propio proyecto TEI, que puede presumir de tener una excelente documentación³. En ella se puede delinear claramente cuál sería la estructura de etiquetado básica, aunque recurrí a más fuentes, ejemplos concretos de

¹ «Grandes chroniques de France», Français 2813, Bibliothèque nationale de France. Département des Manuscrits, <http://gallica.bnf.fr/ark:/12148/btv1b84472995>.

² Esta misma presentación del texto se encuentra en el apartado «summary» del archivo XML.

³ «The TEI Guidelines», 16 de noviembre de 2023, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>. En las fases preliminares del proyecto fueron de especial ayuda los siguientes tres apartados: «3 Elements Available in All TEI Documents - The TEI Guidelines», 16 de noviembre de 2023, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/CO.html>. «10 Manuscript Description - The TEI Guidelines», 16 de noviembre de 2023, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>. «13 Names, Dates, People, and Places - The TEI Guidelines», 16 de noviembre de 2023, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html>.

codificación de manuscritos medievales realizados con el estándar⁴. Es común en estos proyectos marcar la longitud de línea, columna y página, pero eso hace el texto difícilmente legible. Opté, finalmente, por presentar dos formatos del texto, uno en el que se mantuviesen estos aspectos y otro en el que no. Para simplificar, trabajaría sobre un documento en el que sí se mantienen (adjuntado como *Llamedo Gutiérrez_BnF_Français 2813*) del que eliminaría, una vez hecho, las rupturas (adjuntado como *Llamedo_Gutiérrez_BnF_Français 2813 - sin saltos de línea-columna-página*).

Hacer esto automáticamente usando la función «buscar y reemplazar» de Notepad++ no es particularmente difícil. En el apartado «sustituir», y marcando en «modo de búsqueda» «expresión regular» en vez de «normal», se sustituyen por nada (en efecto, se eliminan) `<l/>`, `<cb/>`, `<pb/>`. Para ello hace falta hacer varias acomodaciones. Como el fichero del proyecto principal no solo dividía el documento en líneas, sino que asignaba a cada una un número, esto se debía tener en cuenta. Así mismo, se pensó inicialmente en utilizar la función «Editar» – «Operaciones con líneas» – «Unir líneas» (Ctrl+J), pero esta generaba espacios indeseados. Así, quedó claro que la operación debía seguir este orden:

- Primero, eliminar genéricamente los `<l/>`, `<cb/>` y `<pb/>`, conteniendo los atributos que tuvieran. Notepad++ permite hacer esto mediante la expresión regular `<l.*?/>` (con sus respectivos para columnas y páginas). De esta manera era fácil eliminar los saltos. Pero ejecutar el código en este estado resultaba en roturas de palabras separadas aún por las tabulaciones y los espacios indeseados.
- Por ello, una vez eliminado, se reemplazan primero los espacios entre líneas mediante la expresión `\R` (dejando el segundo cuadro, de nuevo, vacío). Esto solo deja las tabulaciones, que son eliminadas repitiendo la expresión, pero con `\t` esta vez.

Esta operación elimina los espacios indeseados que el comando «Unir líneas» generaba, pero ello implica que se ha de diferenciar de alguna manera una palabra que es continua a lo largo de distintas líneas de otra que no. Para ello, y con propósito, se utiliza el carácter vacío, el espacio: al final de cada línea sin continuidad se deja un espacio; cuando hay

⁴ Dos proyectos en especial sirvieron en la codificación: Andrew Dunning, «Transcribing Medieval Manuscripts with TEI», Andrew Dunning, 28 de enero de 2019, <https://andrewdunning.ca/transcribing-medieval-manuscripts-tei>. «Normas de transcripción», *7 Partidas Digital* (blog), accedido 9 de mayo de 2024, <https://7partidas.hypotheses.org/normas-de-transcripcion>.

continuidad entre ellas, ese espacio se elimina. De esta manera, unir las líneas y eliminar las tabulaciones mediante sustitución de las palabras deseadas y separa las que no. La expresión general de todo ello, utilizando «|», que funciona como operador booleano «or», resulta así: `<l.*?/>|<cb.*?/>|<pb.*?/>|\\R\\t`.

La cuestión del idioma ha sido una que ha planteado difíciles decisiones a lo largo del trabajo, y en los atributos de las etiquetas no ha sido distinto. He optado finalmente por utilizar el inglés en ellos por dos motivos principales. El primero de ellos es la mera costumbre: etiquetar en inglés es un estándar en programación, un mundo donde nunca se sabe en manos de quién acabará el código (hay quien argumentaría que el código se hace más accesible, aunque tengo mis reservas en este caso concreto). Además de ello, las normas TEI P5 dan valores recomendados a varios atributos, todos ellos en inglés, y seguirlo da mayores visos de coherencia y de sistematicidad. «Inglés salvo que sea absolutamente necesario» (véanse casos como `<region type="provincia">`) es de alguna manera una norma más clara para mí que mezclar en los atributos inglés y español.

Así, las etiquetas utilizadas se pueden dividir en aspectos gráficos y aspectos textuales.

Aspectos gráficos

- Salto de página (`<pb/>`), columna (`<cb/>`) y línea (`<l/>`). Cada uno de ellos utiliza el atributo `@n` para indicar su número. En el caso de que una palabra se corte al final de la línea, `<l/>` incorpora el atributo `break="no"`.
- Mayúsculas (`<hi>`). En caso de encontrar una capital decorada, se utiliza el atributo `rend="init#colour"`, donde el «#» corresponde al número de líneas que la capital abarca y «colour», por el color de la misma.
- Daño (`<damage>`). En caso de encontrar una sección del texto dañada de cualquier forma, se encapsula el texto legible afectado en esta etiqueta, utilizando el atributo `@type` para dar alguna concreción si se puede.

Aspectos textuales

- División lógica del texto (`<div>`). El prólogo, el índice y cada uno de los capítulos está marcado con esta etiqueta, que utiliza los atributos `@type` ("prologue", "table of contents" y "chapter") y `@n` (numeración) para hacer precisiones.
- Abreviaturas desarrolladas con `<ex>`.
- Para adiciones de texto (sea en un margen, en la parte superior o directamente encima de otra palabra) se utiliza `<add>` con el atributo `@place`. En caso de que

la adición sea encima del propio texto, <add> se encapsula dentro de la etiqueta <subst>, y se incorpora el atributo hand="#mano", donde «mano» se sustituye por un elemento identificador del corrector (como, por ejemplo, la tinta más oscura que usa "#black").

- d. Para elementos de la transcripción dudosas se utiliza la etiqueta <nuclear>.
- e. Para elementos considerados errores en el contexto del texto, se utiliza <sic>. Haría falta un conocimiento verdaderamente profundo de la escritura francesa de la época para utilizar esto efectivamente, así que su uso es bastante reducido.
- f. Nombres personales codificados con <persName>, nombres de organizaciones, con <orgName> topónimos con <placeName> y gentilicios con <rs type="demonym">. A todos ellos acompaña un atributo @type para clasificarlo ("king", "city") y el atributo ref="#código", donde «código» es reemplazado por un identificador único que permite singularizar y referenciar ese elemento en concreto.
 - En la singularización que se hace de personajes, organizaciones y lugares antes del texto se utilizan <trait> para indicar vinculaciones. En los personajes, <roleName> con el atributo ref="#organización" (donde «organización» se sustituye por el identificador concreto) se utiliza para vincularlos a organizaciones listadas.
- g. Fechas y duraciones se etiquetan con <date>, añadiendo el atributo @when para una fecha concreta, @notAfter o @notBefore para dataciones *ante quem* y *post quem*, @from y @to para un rango concreto y @dur para indicar duraciones (se utilizan YS para indicar años). Lo mismo aplica para <birth> y <death> en el caso de los personajes. Se añadió el atributo @type para indicar algún tipo de datación específica, como *Ab Urbe Condita* (AUC).

Miscelánea: comentarios

Los comentarios (<!-- comentario -->) se han utilizado a lo largo de todo el texto para marcar elementos que no se consideran merecedores de otro tratamiento pero que sí quiero que queden reseñados, o para singularizar cosas que se suponen descritas a un nivel más general. Consideraciones lingüísticas para la transcripción, símbolos y ornamentos, aclaraciones de la codificación, conjeturas acerca de significados, etc.

ASPECTOS FORMALES DE LA TRANSCRIPCIÓN

La transcripción, por supuesto, se realiza a partir de la reproducción digital del fragmento. En ella, se han intentado adaptar las normas de la Comisión Internationale de Diplomatique, actualizando al francés moderno normas de acentuación y puntuación.

Utilizo la acentuación y la puntuación del francés actual. Para ello, me valgo tanto de mis propios conocimientos, como del diccionario y de la ayuda de MsWord, utilizando el corrector automático en francés. En el caso de las palabras que utilizan la diéresis en francés, que marcan un diptongo en la actualidad, se han mantenido también: «païs». De hecho, esta última es una forma arcaica del actual «pays», al que se tuvo que aplicar la diéresis para marcar el diptongo que con «y» es evidente. El caso del acento circunflejo es curioso. En francés es un acento irregular, que marca vocales a las que antiguamente seguía una «s». Tal es el caso de «maître», que en el texto aparece como «maistre». En los casos en los que conserva la «s», he optado por no añadir el acento circunflejo, siendo inútil su utilización. Sin embargo, en palabras que ya en el texto han perdido la «s», tales como «îlle» —actualmente «île»— mantengo el acento circunflejo.

Mantengo las grafías del texto: «c» en presuntos sonidos [s] («meçonge»), e «i» en [ʒ] («ajouter»). Trazar la evolución de cada «c» y cada «i» consonántica podría complicar demasiado el ejercicio. En el desarrollo de abreviaturas se opta por modernizar la palabra si cabe. Así, una «q» abreviada y seguida de palabra por vocal no se desarrolla como «que», sino como «qu’».

PROCEDIMIENTO

Teniendo todo esto decidido, me lancé a codificar el texto. Los metadatos del mismo quedan recogidos en <teiHeader>, un apartado que preparé —tal vez demasiado— concienzudamente, revisando constantemente las normas TEI P5, haciendo un esquema en árbol con qué podría contener cada elemento (en las normas, esto se indica en el cuadro «*may contain*») y dirimiendo qué de ello podría ser útil. Esto ya presenta un problema: el enfoque general de la codificación, que no está hecha para ningún fin concreto, puede hacer infinita la adición de metadatos y, como veremos más adelante, el etiquetado.

Todo lo presentado en <teiHeader>, excepto <profileDesc>, se hizo en esta primera fase previa a la transcripción, y llevó, entre la escritura, la búsqueda de información y la consulta de documentación, 10 horas de trabajo. La transcripción tenía una longitud orientativa, y se realizó utilizando al mismo tiempo todas las etiquetas excepto las

nombradas en los apartados f) y g) de los aspectos textuales. Esto tenía sentido en tanto en cuanto se ocupaba de los aspectos más formales del texto, dejando para una revisión de la transcripción final cuestiones más relacionadas con el contenido —el etiquetado de personas, lugares, organizaciones, gentilicios, fechas y cantidades—. Transcribí hasta el capítulo quinto, no sin dificultades —transcribir francés antiguo implicó un esfuerzo no insignificante, a pesar de que la letra carolina del manuscrito era inmejorable—. A lo largo de todo ello no advertí realmente que había infravalorado tremendamente el tiempo que me llevaría aplicar las etiquetas restantes a todo el texto y crear para cada una de ellas un elemento en `<profileDesc>`. Este proceso me llevó unas 24 horas (tres sesiones de 8 horas de trabajo).

Las etiquetas de los puntos f) y g) resultaban ser, en la implementación que les di, tremendamente costosas en tiempo. Crear los elementos `<listPerson>`, `<listOrg>` y `<listPlace>`, referenciarlos adecuadamente y dotarlos de contenido, teniendo en mente la voluntad de hacer una codificación «realmente completa» (añadiendo, por ejemplo, referencias a los elementos Wikidata que los tuvieran, geolocalizando todos los lugares, etc.) era demasiado en una crónica, repleta de por sí de nombres, eventos y lugares muy dispares —y ampliamente conocidos, cosa que comporta, paradójicamente, más trabajo de referenciación—. El etiquetado del prólogo, la tabla de contenidos y el primer capítulo —que supondría un 47% del texto transcrito— comportaba ya 44 personajes únicos, 8 organizaciones y 24 lugares. Toda esta parte del proceso se realizó a lo largo de 16 horas de trabajo.

La perspectiva de duplicar esto —o seguramente superar el doble, en el cap. 5, por ejemplo, hay una larguísima retahíla de personajes, pero, sobre todo, de ciudades— resultaba completamente imposible, teniendo en cuenta además que hacía ya tiempo que había superado la cantidad de horas asignadas al trabajo por la guía docente, rondando las 50.

RESULTADO Y VALORACIÓN

Por todo ello, decido dejar el trabajo así, no estando la segunda mitad etiquetado con las etiquetas f) y g). El archivo, tal y como está, es una buena codificación en TEI, que no atiende en su segunda parte tanto al contenido, pero que permite su lectura y trabajo en lo sucesivo y que refleja fidedignamente la fuente original.

El principal error, sin duda, ha sido la falta de planificación. A pesar de que definí previamente etiquetas propias de un trabajo general y completo, decidí transcribir todo antes de etiquetarlo, resultando en una carga de trabajo excesiva que, efectivamente, no he podido cumplir —espero que no en perjuicio de la calificación—.

¿Habría hecho lo mismo, y con el mismo grado de detalle, si tuviese que trabajar con muchas más páginas de texto, o con cientos de documentos? Probablemente no, sería inviable. Aunque aquí se juntan dos factores: primero, se entiende que este, al ser un trabajo de final de curso, tiene que ser una demostración de entendimiento del lenguaje y pulcritud de su presentación; en segundo lugar, la singularidad y riqueza del texto, que merece el grado de atención y detalle recibido por la multiplicidad de intereses que puede suscitar —y que ha suscitado entre historiadores del arte, semiólogos, lingüistas, e historiadores que se han fijado en aspectos muy variados del manuscrito—. Con todo, no me arrepiento del enfoque perfeccionista que he dado a su codificación, que ha resultado, por otro lado, en una edición bastante sólida y que se presta a análisis casi tan variados como a los que se ha sujeto hasta ahora a la obra. Pero esto debería haber ido acompañado, sin duda, de una buena consideración acerca de cuáles son los límites de un trabajo de curso, necesitando esta un proyecto de mayor envergadura.

El proceso de codificación de este manuscrito en TEI ha sido un proyecto que ha requerido una gran cantidad de esfuerzo. El uso del estándar TEI P5 ha permitido una codificación detallada y precisa del texto, lo que es fundamental para su preservación y estudio. Sin embargo, la codificación también ha sido un proceso que ha llevado una gran cantidad de tiempo, especialmente en la implementación de las etiquetas y, sobre todo, en la creación de los elementos de perfil. Aunque el resultado final es satisfactorio, es importante reconocer que el proceso ha sido complejo y ha planteado varios desafíos. Así, los ficheros adjuntos son prueba simultáneamente del potencial de este estándar y de algunas de sus limitaciones. Con todo, hay pocas alternativas que presenten la simpleza, la maleabilidad y la accesibilidad de TEI en la codificación de textos. Es plenamente cierto que ha resultado en una experiencia dura —aunque sin duda la costumbre de utilizarlo limaría muchas de estas asperezas—, pero hay que recordar, para citar a la autoridad, que así son los senderos de la ciencia, escarpados.