

La codificación de la variación textual del paradigma impreso al digital

Encoding Textual Variation from the Printed to the Digital Paradigm

Gustavo FERNÁNDEZ RIVA

Universidad de Buenos Aires-CONICET

gustavo.riva@filo.uba.ar

Dirección

Clara Martínez
Cantón
Gimena del Río
Riande
Ernesto Priani

Secretaría

Romina De León

ABSTRACT

Difference among witnesses is one of the fundamental problems in Textual Criticism. In the 19th century, the field developed a tool specifically designed to convey textual variation in a printed format: the critical apparatus at the bottom of the page. This technique made it possible to store and visualize information efficiently within the typographical paradigm, but in the current digital paradigm features of electronic media should be applied to create new ways of exploring textual variation. This paper evaluates the two most common ways of encoding variants that have been developed in the last decades: the TEI critical apparatus module and variant graphs.

KEYWORDS

Textual variation, Critical Apparatus, Text Encoding Initiative (TEI), Variant Graph.

RESUMEN

La diferencia entre los testimonios es uno de los problemas fundamentales de la Crítica Textual. En el siglo XIX la disciplina terminó de desarrollar una herramienta diseñada para dar cuenta de esta variación textual en ediciones impresas: el aparato crítico a pie de página. Esta técnica permitía almacenar y visualizar la información de una manera eficiente dentro del paradigma tipográfico, pero dentro del paradigma digital es recomendable aprovechar las características de los medios electrónicos para crear nuevas maneras de explorar la variación textual. Este artículo evalúa críticamente las dos formas más comunes de codificar las variantes que han surgido en las últimas décadas: el módulo critical apparatus de la TEI y los grafos de variantes.

PALABRAS CLAVE

Variación textual, aparato crítico, Text Encoding Initiative (TEI), grafo de variantes.

RHD 2 (2018)

ISSN
2531-1786



1. INTRODUCCIÓN

El problema de la variación está en el centro mismo de la Crítica Textual¹. La edición académica de textos antiguos debe lidiar con el hecho de que los testimonios conservados difieren entre sí y, probablemente, con la obra tal como fue compuesta originalmente². La conceptualización y consideración de estas diferencias no se ha mantenido estable a lo largo del tiempo. Es posible establecer dos tipos ideales de actitudes extremas frente a la variación, sin que ninguna se corresponda totalmente con la posición de individuos particulares y siendo conscientes de que existe un enorme rango de posiciones intermedias³. En un extremo estaría la consideración de la variación como un obstáculo para alcanzar el texto crítico que se desea reconstruir. En el extremo opuesto estaría la posición que sostiene que todas las particularidades de cada testimonio son valiosas en sí mismas, que cada manuscrito debe ser considerado como una obra autónoma. Evidentemente, formuladas de esta manera, ambas posiciones son reductivistas y endebles. Por lo general, incluso los editores que con confianza buscan reconstruir un arquetipo reconocen que pueden existir versiones alternativas de un texto y que la variación es importante desde ciertos puntos de vista⁴. De la misma manera, hasta los más firmes representantes de la *New Philology* proponen algún tipo de edición crítica⁵.

En cualquier caso, más allá de la postura teórica, es claro que editar o trabajar con un texto conservado en una serie de testimonios divergentes implica considerar algún método de análisis y registro de las variantes de la tradición textual en cuestión. La importancia de este registro ha sido reconocida desde los primeros ejemplos de la edición académica y se han

¹ Aquí me centraré en los casos de tradiciones textuales con múltiples testimonios divergentes, pero el problema de la variación también está presente en ciertos casos de la crítica genética en los que las versiones del texto se superponen en el mismo testimonio.

² Si bien un fenómeno minoritario, también es necesario tener en cuenta que en algunos casos la variación no surge de copias sucesivas, sino que puede deberse a versiones paralelas creadas por el autor original. Ver, por ejemplo, Bumke (2005).

³ Tomo la idea de esta división polar de Thomas Bein (2000, p. 84) quien se refiere a los dos paradigmas contrapuestos como *originalorientiert* y *überlieferungsorientiert*.

⁴ Varios de los más importantes representantes y formalizadores de lo que solemos denominar el método lachmanniano en el siglo XIX como Paul Lejay, Mario Roque y Gaston Paris se referían muchas a veces a *innovaciones* (*innovations*) o *modificaciones* (*modifications*) en lugar de *errores* para referirse a las lecturas divergentes del arquetipo que permitían emparentar diferentes manuscritos (Trovato, 2014, p. 54). Esto muestra que no tenían una postura despectiva en contra de estas lecturas, sino que simplemente las consideraban una desviación del arquetipo que intentaban reconstruir. Asimismo, incluso Werner Schröder (1994, p. 9), uno de los más acérrimos opositores de la perspectiva cercana al código en la filología alemana, sostiene que puede haber algún valor en las variantes, aunque más no sea en la investigación de la historia de la lengua.

⁵ Bernard Cerquiglini (1989), uno de los referentes de la *New Philology*, sostenía la postura de que la mejor edición posible es la vista sinóptica de los testimonios. A pesar de que los métodos para hacer este tipo de ediciones existen en formato digital, solo en ciertos casos esto constituye la totalidad de la edición ya que, por lo general, los investigadores siguen privilegiando la utilización de textos críticos. Para más detalles sobre estas discusiones en el contexto actual consultar especialmente el volumen de la revista *Florilegium* editado por Markus Stock: *Rethinking Philology: Twenty-Five Years after The New Philology* (2015).

desarrollado diferentes métodos para lograrlo. Un hito por su influencia en la historia posterior fue la edición del Nuevo Testamento griego de Robert Estienne, publicada en Paris en 1550, que registra variantes en sus márgenes (ver figura 1).

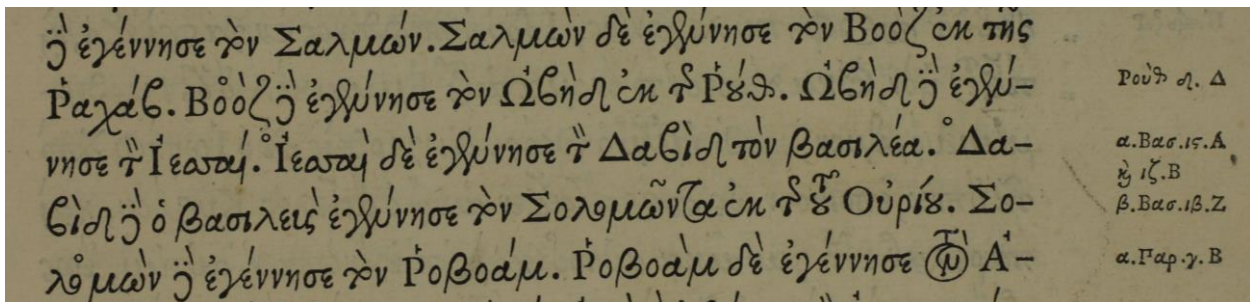


Figura 1

Durante todo el período dominado por la imprenta se desarrollaron diversas formas de registrar la variación textual. Otra edición del Nuevo Testamento griego, hecha por John Mill en 1707, constituye un punto de inflexión en la evolución del aparato crítico, al agregar a pie de página y con una forma estandarizada las referencias a las variantes existentes en otros códices (ver figura 2).

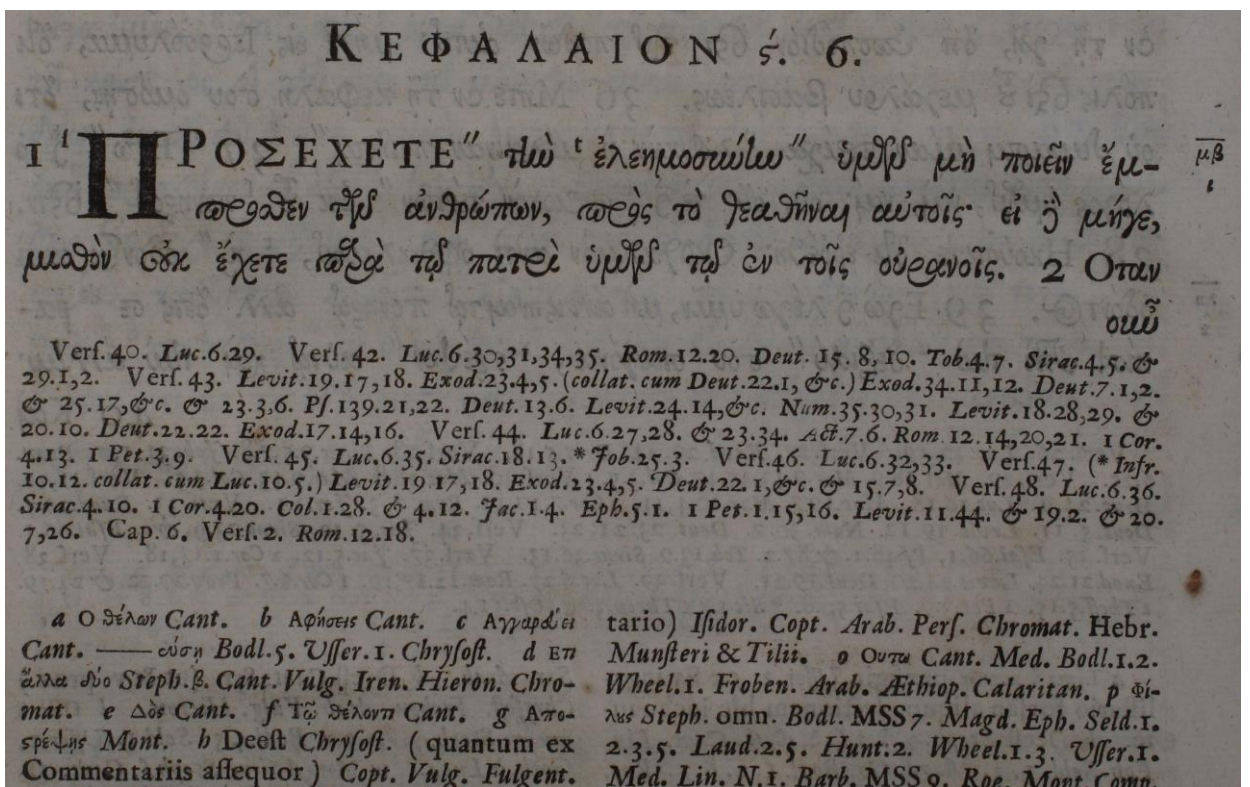


Figura 2

A principios del siglo XIX, ya está desarrollada e instalada la forma moderna del aparato crítico a pie de página, en su variante positiva y negativa. Desde esa época hasta hoy esa forma se ha mantenido como dominante, aunque ha habido intentos de encontrar

posibilidades alternativas, como el *Treppenapparat* de la edición genética de la poesía de Hölderlin hecha por Friedrich Beißner (1943). En este aparato, las diferentes variantes numeradas y copiadas cada una en un renglón en el lugar del texto en que corresponden, como se puede apreciar en la figura 3. Esta forma es relativamente similar al grafo de variantes que explicaré más adelante.

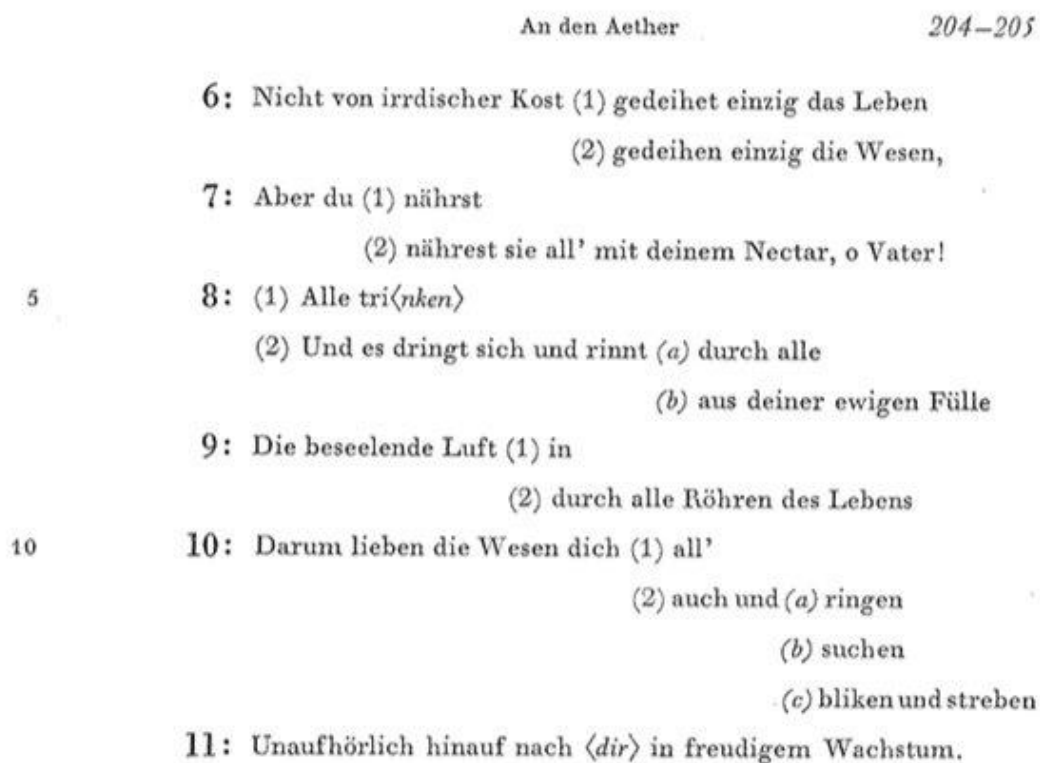


Figura 3

La llegada de la tecnología informática ha posibilitado nuevas formas de registrar y procesar la variación textual imposibles dentro de un paradigma impreso. En este trabajo deseo repasar las cualidades principales del aparato de variantes a pie de página, del módulo *critical apparatus* de la TEI (*Text Encoding Initiative*) y de los grafos de variantes, para evaluar los problemas y ventajas de cada una de estas tecnologías. Una de mis hipótesis principales es que hay dos grandes transformaciones causadas por el paradigma digital en relación a la variación textual: la posibilidad de realizar colaciones automáticas de testimonios y la clara diferenciación entre un formato de almacenamiento y uno de visualización.

2. EL APARATO CRÍTICO A PIE DE PÁGINA

Como sostiene Cynthia Damon (2016, p. 202), el aparato crítico es más que un repositorio de variantes textuales, es un repositorio de todo aquello que los editores juzgan necesario para que los lectores comprendan las decisiones editoriales que dieron como resultado un texto. Sin embargo, es cierto que la mayor parte del aparato crítico de casi

cualquier edición está compuesto principalmente por el registro de la variación textual. Para evitar confusiones, en este trabajo utilizaré el término *aparato de variantes* para referir a la parte sustancial del aparato crítico a pie de página destinada a registrar la variación textual.

El éxito del aparato de variantes a pie de página no es fortuito: se trata de una forma muy eficiente de resolver el problema de registrar las variantes utilizando las posibilidades del código y la imprenta. Particularmente, son dos habilidades técnicas fundamentales las que encauzaron su desarrollo. Por un lado, el refinamiento de la diagramación de página a lo largo de los siglos de uso del código, que permite integrar dos flujos de texto paralelos (el texto y el aparato) sobre la misma página. Por otra parte, el uso conjunto de diferentes tipografías y estilos de letra, especialmente la negrita y la cursiva. Por ejemplo, es común que tanto las siglas de los testimonios luego de la mención de las variantes como las intervenciones del editor se hagan en cursiva. También deben usarse signos de puntuación especiales para separar las diferentes entradas del aparato, como doble barra (||) o punto y coma (;) y ciertas abreviaturas convencionales para las indicaciones editoriales (*om.*, *add.*, *del.*, etc.). En otras palabras, el aparato crítico es el resultado de un desarrollo técnico y de la creación de un sistema de convenciones que, de todas maneras, nunca terminaron de ser completamente estandarizadas.

El aparato crítico es un sistema funcional y efectivo dentro de su contexto por tres razones. Por un lado, es de fácil consulta. Si, por ejemplo, se pusieran las transcripciones completas de todos los testimonios de una obra en un mismo libro, buscar el verso equivalente en todas ellas sería muy difícil. En segundo lugar, ofrece la información ya procesada. Es decir, en lugar de reproducir completos los fragmentos equivalentes, el aparato menciona solo las diferencias y las agrupa. Finalmente, almacena gran cantidad de información en poca cantidad de espacio. Particularmente, el aparato negativo, aquel que menciona el contenido de otro testimonio sin explicitar las palabras del texto crítico a las que refiere (pero que pueden inferirse por contexto), es una gran forma de ahorrar espacio. En resumen, el aparato de variantes permite almacenar y consultar la información sobre la variación presente en un fragmento de una manera rápida y eficaz; ocupando poco del preciado espacio de la página.

De todas maneras, a pesar de sus aspectos positivos, el aparato de variantes tiene grandes restricciones. En primer lugar, la escasez de espacio en la página impresa lo obliga a ser parcial. Por la necesidad de brevedad, el aparato normalmente debe limitarse a registrar cierto tipo de variantes y excluir otras, por ejemplo, normalmente no incluye las variantes ortográficas. Incluso dentro del tipo de variantes no ortográficas, en muchos casos, para ahorrar espacio, los editores no las registran todas sino solo las más importantes. En qué radica la importancia de una variante puede variar según diferentes criterios. Hay variantes que pueden ser importantes para un editor porque contribuyen a crear el estema. En otros casos, variantes importantes pueden ser aquellas que generan un cambio semántico relevante en el texto. No

existe un criterio universal para decidir qué variantes incluir y cuáles excluir, sino que el editor debe decidirlo teniendo en cuenta la tradición manuscrita en cuestión y sus objetivos e incluirlo en su introducción a la edición.

Como consecuencia de estas prácticas y restricciones, el aparato no almacena de modo completo y eficiente las variantes y, generalmente, no permite reconstruir el texto de los testimonios en detalle. Otra característica del aparato crítico es que se trata de una estructura heterogénea, es decir, que la información no está siempre codificada de la misma manera. La mayoría de los fenómenos se marcan refiriendo a una lectura del texto crítico y listando las lecturas alternativas, pero fenómenos como omisiones o versos añadidos se marcan, generalmente, como aclaraciones del editor. Como ya he mencionado, existen ciertas convenciones para marcar cada uno de estos tipos de fenómenos, pero, en cualquier caso, sigue tratándose una creación heterogénea, que no era un problema cuando la información de ese aparato estaba pensada exclusivamente para lectores humanos; pero si se desea procesar la información computacionalmente, esta estructura heterogénea lo hace más difícil.

Desde un punto de vista de la historia de los soportes textuales, resulta claro que el aparato de variantes, con sus fortalezas y limitaciones, es un hijo de las tecnologías que lo posibilitan y condicionan: el códice y la imprenta. Todas sus características (la necesidad de resumir, el lugar sobre la página, la disposición en reglones) se desprenden de la confluencia de estas dos tecnologías. Por lo tanto, es esperable que el cambio de paradigma de lo impreso a lo digital posibilite una nueva aproximación a la variación textual.

Existen, de todas maneras, ediciones electrónicas que siguen utilizando el aparato crítico como un modo de visualización, aunque usualmente con ligeros cambios. Tomemos como ejemplo la edición de las *Epistulae heroidum* de Ovidio en el portal [curculio.org](http://www.curculio.org)⁶. Como se ve en la figura 4, esta edición diagrama sobre la superficie de la pantalla una estructura muy similar al aparato crítico sobre la página impresa, pero, al tratarse de un texto en verso y ya que la pantalla es una superficie predominantemente horizontal (a diferencia de la página), se ha preferido introducir el aparato al costado del texto y no a pie de página. Esto permite a la vista poder consultarlo de manera más sencilla. Sin embargo, la configuración del aparato propiamente dicha no introduce novedades.

25 Argolici rediere duces, altaria fumant,
ponitur ad patrios barbara praeda deos;
grata ferunt nuptae pro salvis dona maritis,
illi victa suis Troica fata canunt.
mirantur iustique senes tropidaeque puellae,

27-28 *delevit Reeve* | 27 *nuptae Heinsius, Planudes (?) : nymphae Ω : sponsae Bentley* ||
28 *Troica EGΩ : Troia Fξ* | *fata EGΩ : facta F¹ξ* ||
29 *iustique] laetique Schenkl, lassique Riese* ||

Figura 4

⁶ Accesible la edición de Hendry (2008) desde <http://www.curculio.org/Ovid/h01.html>.

No hay nada equivocado en ofrecer esta visualización en una edición electrónica, pero es, en cierto sentido, algo anacrónica, pues sigue adaptándose a las limitaciones de la página, en un contexto en donde esto ya no es necesario. En otras palabras, es entendible que algunas ediciones electrónicas sigan representando la variación textual reproduciendo la estructura del aparato crítico del paradigma impreso, porque posee una larga tradición y los expertos se sienten cómodos con ella. Es normal en un período de transición de paradigmas que las formas antiguas continúen durante cierto tiempo. No obstante, ya se han desarrollado alternativas, como el grafo de variantes que explicaré más adelante, que son propias del paradigma digital y que no se ven limitados por las restricciones del aparato crítico que he explicado más arriba.

3. LA VARIACIÓN COMO PROBLEMA INFORMÁTICO

El proceso de comparación y evaluación de las variantes existentes entre distintos testimonios de un texto se conoce como *colación*. En el mundo pre-digital la colación solía realizarse con escritura manuscrita sobre papel. Pueden existir diversas maneras de hacerlo. Alberto Blecua recomienda en su *Manual de crítica textual* (1983, p. 44) seleccionar un texto base, que puede ser tanto un manuscrito como una edición anterior, contra el cual se compararán los demás testimonios tomando notas. Blecua reconoce que, en caso de que no existan ediciones anteriores o que la tradición manuscrita no haya sido estudiada con detenimiento en el pasado, es necesario realizar una colación por calas, es decir, seleccionar ciertos pasajes y comparar todos los testimonios en igualdad de condiciones para tratar de establecer provisoriamente sus relaciones mutuas y seleccionar el más apropiado para servir de base. Una descripción mucho más detallada de la tarea es dada por West (1973, p. 66) en su manual. Allí se dice, por ejemplo, que ciertos editores escriben las variantes de los manuscritos en el margen de ediciones anteriores (a veces con colores diferentes), pero el autor recomienda usar un cuaderno.

Es posible dividir el proceso de registro de la variación textual en cuatro pasos:

- 1) Preparación de las fuentes
- 2) Proceso de colación
- 3) Almacenamiento
- 4) Visualización⁷

⁷ Este modelo es compatible con el *modelo de Gothenburg*, desarrollado como esquema para la colación informática en 2009. El modelo de Gothenburg está pensado exclusivamente para el desarrollo de software de colación automática, mientras que aquí estoy pensando en un modelo de colación más fundamental y abarcativo, que incluya tanto la colación con ordenador como la colación humana. Los pasos 1, 2 y 3 de Gothenburg (tokenización, alineación y análisis) son instancias particulares de la colación computacional, aquí subsumidas en el paso 2. El paso 4 de ambos modelos se corresponde. El

Una novedad de la computación es que la operación de comparación, el proceso de colación en sí, ya no necesariamente requiere un agente humano, sino que es posible realizarla de manera automática con programas como *CollateX*⁸, *Juxta*⁹ o *TRAViz*¹⁰. Estos programas, a diferencia de la colación manual, necesitan necesariamente transcripciones de todos los testimonios. Generalmente estos programas requieren como *input* un formato de texto plano, aunque *CollateX* ofrece la posibilidad de utilizar un formato más complejo, JSON, para mejorar los resultados del alineamiento¹¹. El resultado de la colación de estos programas puede ser almacenado o visualizado en una pluralidad de formas, algunas de las cuales referiré más adelante. Según mis propias experiencias, sus algoritmos dan generalmente buenos resultados, pero no son perfectos, particularmente en tradiciones complejas con transposiciones y muchos testimonios divergentes.

No quiero detenerme en las características de estos programas de colación automática que actúan en el paso 2, porque mi preocupación central en este artículo radica en los últimos dos pasos: almacenamiento y visualización¹². En el mundo del libro impreso estos dos pasos no se distinguen claramente. El aparato de variantes es, al mismo tiempo, el formato de almacenamiento de la información y la manera que tienen los humanos de acceder a ella. En cambio, en el mundo digital la diferencia entre estas dos instancias adquiere una importante relevancia. Existen formatos capaces de almacenar la información que no están diseñados para ser leídos por humanos, sino para ser procesados por un ordenador, ya sea para realizar algún tipo de análisis particular o para ser transformados en alguna forma de visualización. Se han desarrollado básicamente dos formatos de almacenamiento de la variación textual en los últimos años: marcado y grafo. Como mostraré en los siguientes apartados, tanto su estructura de la información como su manera de conceptualizar las variantes son distintas y esto debe tenerse en cuenta en el momento de codificar información relativa a la variación textual en una edición digital.

4. LA VARIACIÓN CODIFICADA EN UN LENGUAJE DE MARCADO: XML-TEI

La manera en que los lenguajes de marcado indican la variación textual está fuertemente influenciada por el aparato de variantes de la época impresa. Su principio consiste en identificar el lugar del texto base sobre el que se produce la variación, marcarlo con un

paso tres de mi modelo está implícito en el modelo de Gothenburg. Para más detalles sobre el modelo de Gothenburg, ver <https://collatex.net/doc/> y https://wiki.tei-c.org/index.php/Textual_Variance.

⁸ Accesible desde: <https://collatex.net/>.

⁹ Accesible desde: <http://juxtacommons.org/>.

¹⁰ Accesible desde: <http://www.traviz.vizcovery.org/>.

¹¹ En *CollateX*, un *input* en JSON permite indicar para cada token una versión original y una regularizada de la palabra, de manera que las diferencias ortográficas no sean tenidas en cuenta al realizar la comparación.

¹² Un resumen más completo de las herramientas de colación asistida por ordenador y sus características puede encontrarse en Bermúdez Sabel (2017).

elemento e indicar las variantes dentro de otro elemento. Esta estructura se sostiene claramente sobre la misma base conceptual que el aparato de variantes a pie de página. El módulo *critical apparatus* de la TEI¹³ y el paquete *reledmac* para LaTeX funcionan según este principio¹⁴. Aquí me detendré en analizar el sistema más complejo, el de la TEI, pero muchos de los problemas abordados también se aplican al sistema *reledmac*.

Dentro de las guías de la TEI se plantean tres métodos posibles para codificar la variación textual: *location-referenced*, *double end-point attachment* y *parallel segmentation*. En el primero, el aparato se crea en alguna parte del documento, separado del lugar donde ocurre la variación. Para esto se utiliza el elemento <app>, dentro del cual cada una de las variantes se anota con el elemento <rdg>. La referencia al texto al que se refieren las variantes es vaga y se incluye en el atributo @loc del elemento <app>, que debe referir al elemento del texto principal dentro del cual se encuentra la variante. Por ejemplo, supongamos que este es un fragmento de un texto codificado según las normas de la TEI:

```
<I n= "1">Aquí me pongo a cantar,</I>
<I n= "2">al compás de la vigüela</I>
```

El aparato crítico según el método *location-referenced* se podría escribir de la siguiente manera:

```
<app loc= "1"><rdg wit= "#B">bailar</rdg></app>
```

El atributo @loc= "1" indica que la variante ocurre en el elemento con el número 1. Podrían usarse, por supuesto, atributos @xml:id como referencia. Como se ve, este método es altamente ambiguo, pues no se explicita a qué palabra o sintagma concreto refiere la variante. Por el contexto, resulta claro que, en este caso, se trata de la palabra *cantar*, pero en otras ocasiones puede resultar ambiguo. Este método está fuertemente influenciado por la estructura del aparato crítico negativo del libro impreso, donde la ambigüedad era tolerada si permitía ahorrar espacio, algo que no tiene sentido en el mundo electrónico. Asimismo, debido a esta ambigüedad inherente, se trata de un tipo de codificación muy difícil de interpretar por el ordenador y que está pensada para lectores humanos. En otras palabras, este método es conveniente solo cuando se quiera codificar en XML una edición impresa con un aparato crítico negativo y no necesariamente en otros casos.

Por su parte, según el método *double end-point attachment* también se utiliza el elemento <app> en una parte del documento distinta a donde ocurre la variación, pero se hace una referencia explícita al fragmento que se refiere, lo que elimina la ambigüedad inherente al método anterior. Esta referencia explícita se logra mencionando a uno o más elementos con un

¹³ Accesible desde: <http://www.tei-c.org/Vault/P5/2.0.1/doc/tei-p5-doc/es/html/TC.html>.

¹⁴ Accesible desde: <https://ctan.org/pkg/reledmac>.

atributo `@xml:id` que delimiten perfectamente el lema en cuestión. Si no hay ya un elemento que marque esta parte del texto, se agregan elementos vacíos (como `<anchor>`) con este propósito:

```
<I n= "1" xml:id= "MF_line_1">Aquí me pongo a <anchor xml:id =
"MF_anc_1"/>cantar</I>
<I n= "2">al compás de la vigüela</I>
[...]
<app from= "MF_anc_1" to= "MF_line_1"><rdg wit= "#B">bailar</rdg></app>
```

Existen muchas alternativas diferentes dentro de este método. Por ejemplo, se puede ubicar el elemento `<app>` dentro del texto luego del lugar donde ocurre la variación, en cuyo caso solo el atributo `@from` es necesario, pues el elemento `<app>` se considera automáticamente como indicador del final del lema. De todos modos, parte del atractivo de utilizar *location-referenced* y *double end-point attachment* radica en que no resulta necesario incluir toda la información de la variación dentro del flujo del texto, de manera que se vuelve más sencillo y legible (para humanos y ordenadores por igual). Una de las mejores maneras de aprovechar este método sin necesidad de incorporar una gran cantidad de elementos `<anchor>` consiste en tokenizar el texto con elementos `<w>` y otorgando a cada uno de ellos un atributo `xml:id`. De esta manera se pueden referir siempre a cada palabra particular.

Finalmente, con *parallel segmentation* se incluyen las variantes en el lugar donde corresponden en el cuerpo del texto. Esta forma de proceder se corresponde casi perfectamente con el funcionamiento del paquete *reledmac* para *LaTeX*:

```
<I n= "1">Aquí me pongo a
  <app>
    <rdg wit= "#A">cantar</rdg>
    <rdg wit= "#B">bailar</rdg>
  </app>
</I>
```

Una característica de este método, ausente en los dos anteriores, es que no implica necesariamente privilegiar una de las lecturas sobre las demás. En aquellos, el aparato indica las variantes refiriendo al texto base, de alguna manera considerado como superior. En *parallel segmentation*, en cambio, todas las variantes tienen la misma jerarquía, se indican en el mismo lugar y utilizando el mismo elemento: `<rdg>`. Si se desea otorgar mayor jerarquía a una variante es posible utilizar el elemento `<lem>` en lugar de `<rdg>` para marcarla. El problema principal de *parallel segmentation* es que, en muchas oportunidades, no resulta posible lograr una codificación satisfactoria cuando las variantes se superponen. Imaginemos el siguiente caso de tres versos:

1. Aquí me pongo a cantar
2. Aquí comienzo a cantar
3. Aquí me viene bien cantar

Si se desea codificar estas variaciones con una estructura anidada y válida de XML utilizando *parallel segmentation* es necesario desglosar las variantes individuales. Una forma de hacerlo sería la siguiente:

```
<I>Aquí
  <app>
    <rdg wit="#A #C">me</rdg>
    <rdg wit="#B"></rdg>
  </app>
  <app>
    <rdg wit="#A">pongo</rdg>
    <rdg wit="#B">comienzo</rdg>
    <rdg wit="#C">viene</rdg>
  </app>
  <app>
    <rdg wit="#A #B">a</rdg>
    <rdg wit="#C">bien</rdg>
  </app>
cantar</I>
```

Como se ve, esta estructura no representa realmente el fenómeno. En esta codificación parece que el testimonio B omite *me* y tiene *pongo* en lugar de *comienzo*, lo que no es absolutamente correcto. La lectura más lógica es considerar que el sintagma *comienzo a* es una variante de los sintagmas *me pongo a* de A o *me viene bien* de C. Sin embargo, no es posible indicar esto sencillamente con este método, sino que es preciso desglosar las variantes de esta manera algo artificial o repetir información (el *me* de los testimonios A y B) de forma redundante, como a continuación:

```
<I>Aquí
  <app>
    <rdg wit="#A">me pongo</rdg>
    <rdg wit="#B">comienzo</rdg>
    <rdg wit="#C">me viene</rdg>
  </app>
  <app>
    <rdg wit="#A #B">a</rdg>
    <rdg wit="#C">bien</rdg>
  </app>
cantar</I>
```

De estos tres métodos, el más usado en la práctica de edición académica es *parallel segmentation*. Según Buzzoni (2016), esto se debe a que los otros son más difíciles de anotar manualmente y porque no implica considerar un texto superior a los demás. Sin embargo, estas ventajas son pobres frente a su falta de precisión y la gran dificultad para corregirlo, ya que en tradiciones complejas termina causando una confusa profusión de etiquetas. No obstante, su preeminencia es tal dentro de la comunidad que muchas aplicaciones como *Juxta*, *StemmaWeb*¹⁵ y el *Critical Apparatus Toolbox*¹⁶ presuponen este método para los archivos TEI que quieran utilizarlas.

Si comparamos críticamente estos tres métodos, *location-referenced* es el menos recomendable de todos, por su inherente ambigüedad y por su estructura no pensada para ser procesada por ordenadores. *Parallel segmentation* parece el más sencillo y adecuado en un primer momento, pero en tradiciones textuales complejas es muy difícil de llevar a cabo y entorpece la estructura del XML. El método *double end-point attachment* es ampliamente superior a los demás porque no genera ambigüedad y no hay fenómenos que no se puedan codificar de manera satisfactoria, aunque si se utilizan elementos `<anchor>`, también puede crear una estructura compleja de elementos en el medio del texto. Por esto, la manera más eficiente de usarlo es, como ya he mencionado anteriormente, tokenizando el texto completo con elementos `<w>` y tomándolos como referencias. Sin embargo, como la mayoría de los archivos TEI no tokenizan el texto, este método también puede ser difícil de codificar. En resumen, no hay ningún método que resulte totalmente satisfactorio.

Más allá del método particular, un problema fundamental de la codificación de la variación textual en TEI es que debe hacerse manualmente. Si bien *CollateX* y *Juxta* ofrecen como *output* archivos de XML con etiquetas de TEI usando el método *parallel segmentation*, esto solo funciona en casos sencillos y falla rotundamente en tradiciones complejas. TEI ofrece un buen formato para codificar aparatos críticos ya existentes o simples con dos o tres testimonios, pero no es un buen formato para estudiar la variación textual de tradiciones complejas. En otras palabras, para codificar dos versiones de una obra impresa moderna (por ejemplo, primera y segunda edición), es posible usar TEI exitosamente. En cambio, para el caso de tradiciones complejas de textos medievales o antiguos no es necesariamente un formato adecuado.

5. LA VARIACIÓN CODIFICADA COMO GRAFO

El grafo de variantes (*variant graph*) fue propuesto por Schmidt y Colomb (2009) como una forma de registrar la variación textual. El concepto de grafo se remonta, por supuesto, a desarrollos de la matemática, particularmente de la topología, así como reformulaciones dentro

¹⁵ Accesible desde: <https://stemmaweb.net/stemmaweb/>.

¹⁶ Accesible desde: <http://teicat.huma-num.fr/>. La aplicación fue desarrollada por Burghart (2015).

de la biología y las ciencias de la computación. Un grafo en este sentido es un conjunto de puntos y conexiones entre ellos. Los grafos proveen una manera sencilla e intuitiva de trabajar con variantes en una secuencia al crear *caminos* alternativos en las regiones donde se producen divergencias (Bjørndal, 2017). Esta propiedad es aprovechada desde hace tiempo por la biología para representar las variaciones en las secuencias de ADN y es la inspiración para el uso en crítica textual. Supongamos los siguientes textos:

- A – Al despertar el dinosaurio ya estaba allí
 B – Al despertar el monstruo ya no estaba allí

Un grafo de variantes que representa esto sería el de la figura 5, generado utilizando *CollateX*:

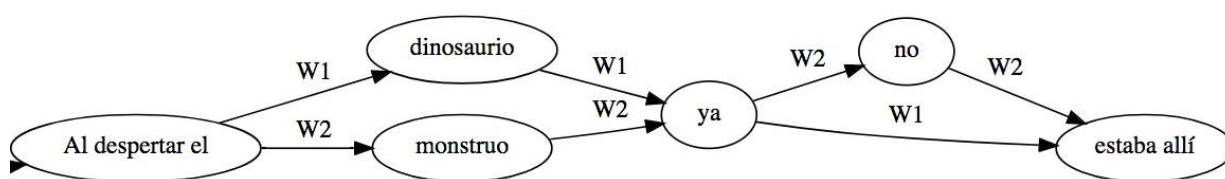


Figura 5

Como se ve, se trata de un grafo dirigido (pues las conexiones tienen una dirección) y acíclico (porque no está permitido volver hacia un nodo ya seleccionado). Cualquier representación visual de un grafo tiene como sustrato algún formato de codificación particular. Lenguajes comunes para representar grafos son *SVG* y *DOT* (Gansner, Koutsofios y North, 2015). El grafo anterior, por ejemplo, se puede escribir en lenguaje *DOT* de la siguiente manera¹⁷:

```
digraph G {
  v0 [label = ""];
  v1 [label = "Cuando despertó el "];
  v2 [label = "monstruo "];
  v3 [label = "dinosaurio "];
  v4 [label = "todavía "];
  v5 [label = "no "];
  v6 [label = "estaba allí"];
  v7 [label = ""];
  v0 -> v1 [label = "W1, W2"];
  v1 -> v2 [label = "W2"];
  v1 -> v3 [label = "W1"];
  v2 -> v4 [label = "W2"];
}
```

¹⁷ Copio el archivo generado utilizando la herramienta en línea provista por *CollateX*: <https://collatex.net/demo/>.

```

v3 -> v4 [label = "W1"];
v4 -> v5 [label = "W2"];
v4 -> v6 [label = "W1"];
v5 -> v6 [label = "W2"];
v6 -> v7 [label = "W1, W2"];
v0 -> v7 [color = "white"];
}

```

Esta estructura define, en primer lugar, los nodos y luego los vértices que los unen, asignando la etiqueta (*label*) del testimonio correspondiente a cada uno de ellos. Por lo tanto, en un grafo podemos diferenciar claramente entre el formato de almacenamiento de la información (*DOT*, *SVG*, entre otros), y la visualización de esa información sobre la pantalla. A diferencia del mundo impreso donde el aparato crítico cumplía ambas funciones, es posible que un mismo archivo sea leído por diferentes programas que generen grafos con diferentes propiedades. Por ejemplo, podemos representar el mismo grafo de variantes en una visualización más compleja como la de la figura 6, generada con el ya mencionado TRAViz:



Figura 6

El grafo generado con TRAViz tiene la ventaja de ofrecer, por un lado, un color diferente para cada testimonio, de manera de identificarlos mejor, y, por otro, de cambiar el tamaño de letra de una palabra según esté presente en más o menos testimonios. Además, TRAViz permite una interacción del usuario con el grafo. Al posar el cursor sobre una palabra, desaparecen todos los vértices correspondientes a testimonios que no la atestiguan. También es posible unir dos nodos para que sean considerados como uno solo (por ejemplo, en el caso de una misma palabra escrita de maneras diferentes).

Cualquier tradición textual puede ser representada exitosamente en forma de grafo. Los programas de colación automática como *CollateX*, que presentaban problemas para generar un archivo TEI correcto, resultan muy eficientes a la hora de crear grafos de variantes. En este sentido, se trata de un formato de codificación mucho más eficiente para tradiciones complejas. Por otra parte, el grafo de variantes no solo es una excelente forma de visualizar la variación, sino que se pueden realizar diferentes operaciones sobre el archivo base en formato *SVG* o *DOT*, tales como recuperar el texto de cada testimonio o calcular el grado de diferencia entre dos de ellos.

En cuanto a su dimensión como visualización, también es importante resaltar que la forma misma del grafo de variantes aporta información importante para el lector, algo que sucedía en mucha menor medida en el aparato de variantes a pie de página. Con esto me refiero a que en el aparato crítico a pie de página era posible saber, viendo su tamaño, si el texto de esa página presentaba más o menos variación en general que el resto de la obra. Pero esta era realmente la única información que surgía de la representación gráfica del aparato crítico, sin considerar su contenido particular. En cambio, la forma del grafo de variantes resume y representa mucha más información. Por ejemplo, en los relatos en verso medievales el grafo de variantes de un verso suele tener predominantemente dos formas: o bien, como un rombo horizontal (figura 7) o como un triángulo que converge hacia la derecha (figura 8)¹⁸:

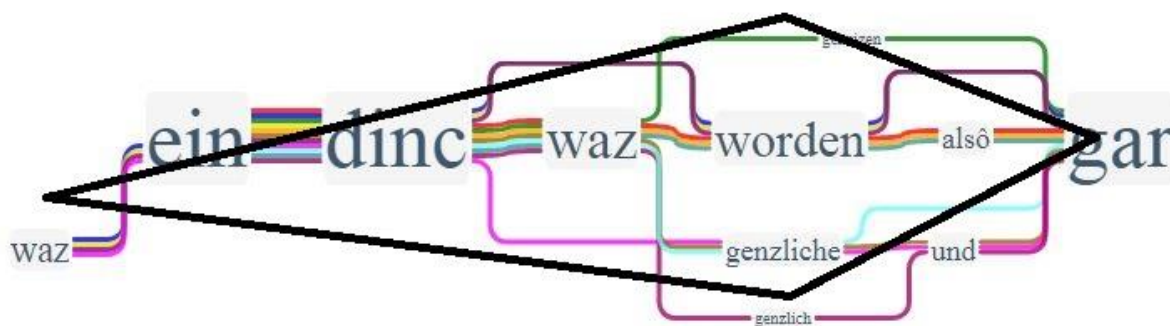


Figura 7

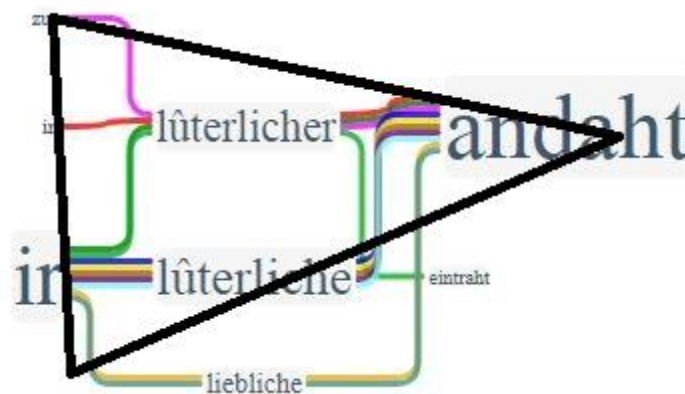


Figura 8

Al observar varios de estos grafos resulta obvio que la tendencia general de los textos es mayor variación hacia la mitad del verso, menor hacia el comienzo y aún menor en la palabra final. La explicación es sencilla: las palabras de rima tienden a sufrir mucha menos variación que el resto del verso porque son, de alguna manera, el *esqueleto* del texto. Un poco más sorprendente puede resultar el hecho de que las primeras palabras del verso tengan más

¹⁸ Los grafos provienen de la edición digital de los relatos cortos en verso de Konrad von Würzburg por Gustavo Fernández Riva (2018).

estabilidad que el medio. Lo que debemos resaltar de este mínimo análisis de estos grafos es que se puede realizar simplemente mirando la forma de los grafos, sin considerar su contenido.

6. CONCLUSIÓN

Al reflexionar sobre este asunto desde la perspectiva de la historia de los medios, resulta claro que no hay ninguna razón esencial a la variación textual que determine que el aparato de variantes deba tener la forma que tradicionalmente ha tenido, es decir, una lista ordenada por número de verso/línea, indicando lemas y variantes. Si bien podría haber ediciones digitales que continúen ofreciendo este formato, se trata, en cierto sentido, de un anacronismo, ya que la creación de esta convención fue un desarrollo tecnológico que se vio guiado por el objetivo de encontrar la mejor manera de representar esta información en una página impresa ocupando el menor espacio posible. El modelo del aparato crítico a pie de página (y en menor medida al final del libro), es una técnica muy eficiente para su contexto mediático, el libro impreso, pero no hay ninguna verdadera razón para seguir utilizándola en un contexto digital, más que la persistencia de la tradición. Pasar del paradigma impreso al paradigma digital implica, tal vez, apostar por las nuevas formas de representar la variación textual que ofrece el nuevo medio.

Como he mostrado, considero que el módulo *critical apparatus* de la TEI es útil para almacenar la información en tradiciones relativamente simples con pocos testimonios, pero muy problemático para tradiciones complejas, como suelen ser los textos medievales o antiguos. En general, considero que el grafo de variantes ofrece, por el momento, la mejor manera de codificar la información relativa a la variación textual en tradiciones complejas.

Finalmente, vale la pena resaltar que, aunque todavía hay mucho trabajo por hacerse en lo que refiere a crear maneras de codificar y procesar la variación textual, hay dos innovaciones del paradigma digital que ya parecen estar establecidas: la separación del registro y la visualización de la variación (algo común a los sistemas de marcado y de grafos) y la posibilidad de colación automática.

REFERENCIAS BIBLIOGRÁFICAS

- Bein, T. (2000). Die Mediävistische Edition und Ihre Methoden. En R. Nuttkofoth, B. Plachte y H. Zwerschina (Eds.), *Text und Edition-Positionen und Perspektiven* (pp. 81-98). Berlin: De Gruyter.
- Beißner, F. (Ed.). (1943). *Hölderlin Sämtliche Werke*. Leipzig: Insel.
- Bermúdez Sabel, H. (2017). Colación asistida por ordenador: estado de la cuestión y retos. *Revista de Humanidades Digitales*, 1, 20-34. Recuperado de <http://revistas.uned.es/index.php/RHD/article/view/16678/16555> el 10/10/2018.

- Bjørndal, M. G. (2017). *Distance Metrics in Variant Graphs*. (Tesis de Maestría). Universidad de Oslo, Oslo.
- Blecua, A. (1983). *Manual de crítica textual*. Madrid: Castalia.
- Bumke, J. (2005). The Fluid Text: Observations on the History of Transmission and Textual Criticism of the Thirteenth-Century Courtly Epic. En K. Starkey y H. Wenzel (Eds.), *Visual Culture and the German Middle Ages* (pp. 99-113). Nueva York: Palgrave Macmillan.
- Burghart, M. (2015). *TEI Critical Apparatus Toolbox*. Recuperado de <http://teicat.huma-num.fr/> el 25/06/2018.
- Buzzoni, M. (2016). A Protocol for Scholarly Digital Editions? The Italian Point of View. En M. Driscoll y E. Pierazzo (Eds.), *Digital Scholarly Editing: Theories and Practices* (pp. 59-82). Cambridge: Open Book Publishers.
- Cerquiglini, B. (1989). *Éloge de la variante: histoire critique de la philologie*. Paris: Seuil.
- Damon, C. (2016). Beyond Variants: Some Digital Desiderata for the Critical Apparatus of Ancient Greek and Latin Texts. En M. Driscoll y E. Pierazzo (Eds.), *Digital Scholarly Editing: Theories and Practices* (pp. 201-218). Cambridge: Open Book Publishers.
- Estienne, R. (1550). *Novum Testamentum Graecum*. Paris: R. Stephanus.
- Gansner, E. R., Koutsofios, E. y North, S. (2015). *Drawing Graphs with Dot*. Recuperado de <http://www.graphviz.org/pdf/dotguide.pdf> el 25/06/2018.
- Hendry, M. (Ed.). (2008). *P. Ovidii Nasonis. Heroides I*. Recuperado de <http://www.curculio.org/Ovid/h01.html> el 09/09/2018
- Interedition Development Group (2013). *CollateX*. Recuperado de <http://collatex.net/> el 25/06/2018.
- Jänicke, S. (2014). *TRAViz*. Recuperado de <http://www.traviz.vizcovery.org/> el 25/06/2018.
- Fernández Riva, G. (Ed.) (2018). *Konrad von Würzburg-Versnovellen*. Recuperado de <http://kvwdigital.000webhostapp.com/> el 25/06/2018.
- Mill, J. (1707). *Novum Testamentum*. London: Guil. Lancaster.
- Performant Software Solutions LLC (2014). *Juxta*. Recuperado de <http://www.juxtasoftware.org/> el 25/06/2018.
- Schmidt, D. y Colomb R. (2009). A data structure for representing multi-version texts online. *International Journal of Human-Computer Studies*, 67(6), 497-514.
- Schröder, W. (1994). *Textüberlieferung und Textkritik. Kleinere Schriften, VI*. Stuttgart: Hirzel.
- Stock, M. (Ed.). (2015). *Rethinking Philology: Twenty-Five Years after the New Philology. Florilegium*, 32. Toronto: University of Toronto Press.
- TEI Consortium (2016). 12. Critical Apparatus. En TEI Consortium, *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Recuperado de <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/TC.html> el 25/06/2018.

Textual Variance Discussion Group (TEI). Recuperado de <https://goo.gl/B2WVFr> el 25/06/2018.

Tree of Texts. Stemmaweb. Recuperado de <https://stemmaweb.net/stemmaweb/> el 25/06/2018.

Trovato, P. (2014). *Everything You Always Wanted to Know about Lachmann's Method: A Non-Standard Handbook of Genealogical Textual Criticism in the Age of Post-Structuralism, Cladistics, and Copy-Text*. Firenze: Libreriauniversitaria.it.

West, Martin L. (1973). *Textual Criticism and Editorial Technique Applicable to Greek and Latin Texts*. Stuttgart: Teubner.

Wilson, P. (2003). *Reledmac*. Recuperado de <https://ctan.org/pkg/reledmac> el 25/06/2018.