

The Estoria de Espanna Digital: collating Medieval prose - challenges... and more challenges

Ward, Aengus

DOI:

[10.1353/dph.2018.0004](https://doi.org/10.1353/dph.2018.0004)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Ward, A 2018, 'The Estoria de Espanna Digital: collating Medieval prose - challenges... and more challenges', *Digital Philology: A Journal of Medieval Cultures*, vol. 7, no. 1, pp. 65-92. <https://doi.org/10.1353/dph.2018.0004>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Publication forthcoming in: *Digital Philology: A Journal of Medieval Cultures* - <https://www.press.jhu.edu/journals/digital-philology-journal-medieval-cultures>

Checked for eligibility: 20/02/2018

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

**"The *Estoria de Espanna Digital*: collating medieval prose - challenges...
and more challenges."**

Aengus Ward

University of Birmingham

Abstract:

Collation as an element in the production of digital critical editions is no longer in its infancy. The current article, based on the experience of editing a lengthy medieval prose text for the *Estoria de Espanna Digital*, addresses the theoretical and practical implications of using digital tools to collate extensive passages of medieval text.

Word Count: 9122

The *Estoria de Espanna* is a thirteenth-century chronicle of Spain, written by, or at the very least under the direction of, Alfonso the 10th of Castile and Leon -known to posterity as el Sabio, 'the Wise' or perhaps 'Learned' King. Although it may appear at first sight that there is an authoritative version of the chronicle composed under the direction of the king, we know that there were three significantly distinct redactions of the chronicle in its first two decades of existence. First written around the year 1270 in a redaction known as the *versión primitiva*, when Alfonso still retained the ambition to be elected Holy Roman Emperor, the *Estoria* was re-written (the redaction known as the *versión crítica*) in 1282 in radically different circumstances - Alfonso had been abandoned by his subjects in the face of a rebellion by his son Sancho-, and again re-cast under the direction of the now King Sancho in 1289 (the

versión enmendada de 1289). None of these versions was ever fully completed, and over the following decades and centuries all three would see multiple further re-writes. The importance of the *Estoria* over the years can be seen in the fact that there are over 40 extant manuscripts representing all three of these redactions in a variety of configurations. Recent philological advances have demonstrated to a large degree the textual relationships at work amongst the extant codices, but despite this recognition of the complexity of its composition and reception the *Estoria* is most widely known and cited in the form of the edition published by Ramón Menéndez Pidal under the title of *Primera crónica general* in 1906 (re-issued and expanded in 1955 and again in 1977).

Menéndez Pidal's edition is based almost exclusively on the two royal manuscripts from the monastery of El Escorial (known as E₁ and E₂, see below), with occasional emendations from other codices. Although the language and text structure are regularized to suit a twentieth-century readership, the editorial practice places a heavy degree of reliance on the base text manuscripts –but these, as we will see below, are not philologically uncomplicated. If Menéndez Pidal's edition is one part of the background to the establishment of the *Estoria Digital*, another is represented by the tradition of editing medieval Iberian text. Critical editing of medieval Peninsular texts has a long history, of course, and not all of it is so conservative in editorial principles as is the edition of Alfonso's chronicle. Thus, in more recent years, and exemplified first by the work of Alberto Blecua and extensively by the SECRET team in Buenos Aires, a more Lachmannian approach, heavily influenced by Italian textual criticism, has come to greater prominence. The chronicles of Pedro López de Ayala are perhaps the most significant outputs similar to the extensive prose of Alfonso's *Estoria*, and the editions of them produced by SECRET are especially

fine examples of the art of neo-Lachmannian editing principles. But the extensive tradition of textual editing amongst ibero-medievalists has, of course, until now taken the form of printed text – and to some extent, the theoretical framework that informed it has been conditioned by the possibilities of its representation in the printed form.

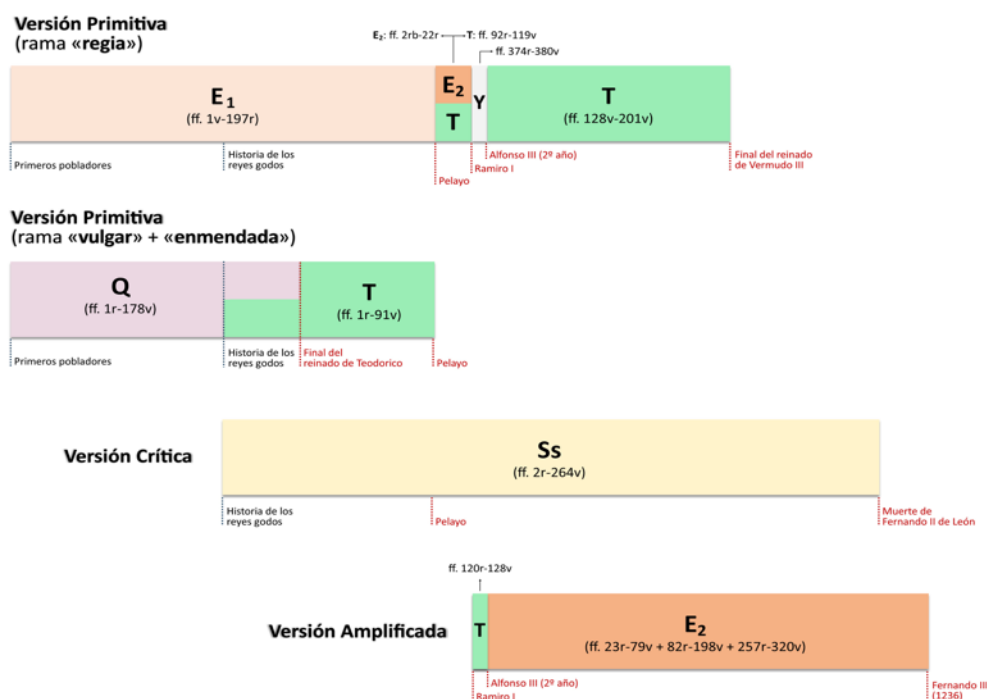
The textual history of the *Estoria de Espanna* is, as has been demonstrated by Catalán and Fernández-Ordóñez in particular, extremely complex.¹ No printed edition can hope to account for this complexity, although the forthcoming edition of the *versión primitiva* by Inés Fernández-Ordóñez will certainly represent the finest of philological accomplishments in print; hence the attempt to edit the *Estoria* digitally. The digital format is not conceived as replacing the philological traditions of ibero-medievalist textual criticism, but rather as building upon and complementing them, since the aims of a digital edition are necessarily different to those of other media.

The *Estoria de Espanna Digital* is the first project to edit digitally (and critically) a large-scale work of medieval Castilian prose.² It consists of the presentation of transcriptions of five of the most significant manuscripts of the *Estoria*, a hypothesis of the *versión primitiva* of Alfonso's chronicle and the presentation of a reader's version of the *primitiva*. There are also ancillary digital tools to aid in the analysis of and access to the *Estoria*. The five manuscripts chosen are:

- E₁: El Escorial Y-I-2, 13th century
- E₂: El Escorial X-I-4, 13th century with 14th century additions
- Q: Biblioteca Nacional de España 5795, 14th century
- Ss: Caja Duero Salamanca 40, 15th century
- T: Biblioteca de Menéndez Pelayo, M-550 14th century.³

In addition, a short fragment of manuscript Y, El Escorial Y-II.11, was employed to cover a gap in the *primitiva* text of T and E₂.

These five were chosen because of all the extant witnesses they best represent (in various proportions) all three of the early redactions of the chronicle. This can be represented graphically as follows:⁴



Nota: Estos diagramas se basan en los trabajos de Catalán (1962, 1997) y Fernández-Ordóñez (2001). Posteriormente, se ha sugerido (Bautista, 2006) que la rama de la *Versión primitiva* conocida como «enmendada» (y quizá también la «vulgar») no sean sino testimonios de la *Versión amplificada* en tramos anteriores al reinado de Ramiro I, lo que modificaría notablemente la concepción de Q y T.

Figure 1. Correspondence between redactions of the *Estoria* and manuscripts transcribed. The *versión primitiva* alone was subject to collation. ([Estoria Digital](#), [About this edition](#))

It should therefore be noted that the *versión primitiva* is found only in four of the five manuscripts. As the object of the current analysis is specifically to examine

the "edited" text of the *versión primitiva* - the product of collation- and not the two other significant redactions of Alfonso's chronicle, I will not deal in great depth with the rest of the edition here. Nonetheless, there are several strategic decisions with regard to the edition as a whole which impacted upon collation policy, so I will briefly deal with these.

The aim of the *Estoria Digital* was to ensure that the manuscript evidence was central to the construction of the edition.⁵ At the same time, it was recognized that the edition also implies other ways of considering medieval texts than just transcriptions and that all of these elements should dovetail in intellectually significant, and useful, ways such that the edition is more than the sum of its parts. Thus, the guide for transcribing TEI5-compliant xml files was designed bearing in mind that the transcriptions were a resource in themselves but also that they would be the raw material for collation.⁶

The edition was constructed on the principle of permanent provisionality.⁷ In the first instance, this edition is provisional since it contains a mere five of the forty extant witnesses of the *Estoria* - no edition can make a claim for exhaustivity if it does not contain all of the evidence and this is demonstrably not (yet) the case for the *Estoria de Espanna Digital*. But the absence of such exhaustivity (in this sense) is not a recognition of failure. On the contrary, the five manuscripts were chosen because they permitted the greatest range of variation to be included. Additionally, at the outset it was decided that fewer manuscripts would initially be transcribed, but at a high level of tagging, thereby sacrificing quantity while increasing the quality of the data. This policy therefore allows the addition of subsequent data at similarly high

levels of tagging in future phases without significant revision of what has already been achieved. The high level of granularity of the transcriptions also permits a more sophisticated collation of manuscript evidence. And the edition is, of course, provisional in another sense in that it is based upon the exercise of critical judgment, both in the transcription phase and (perhaps more crucially) in the collation phase. Revisions of these judgments will naturally give rise to alternative views and outcomes. Although this has always been a characteristic of critical editing, the digital modes of preparation and presentation allow for a more dynamic dialogue between the edition, its users and any future editors; hence the acceptance of provisionality as a strength and not a weakness.

As stated above, the central aim of the edition is to present the text of the *Estoria* in a number of different guises. These are: (i) transcriptions of each of the witnesses concerned by manuscript folio, linked to images of the equivalent manuscript folio (the use of images of the Escorial codices was not possible); (ii) the presentation of a hypothesis of a *versión primitiva* drawn only from the manuscript evidence employed in the edition - in practice this means all of E₁ supplemented with a fragment of E₂ and another of Y (the additional Escorial manuscript) and completed with the direct text of the *primitiva* from T until the point at which T ends (chapter 811, equivalent to chapter 800 of the PCG, the marriage of Fernando I of Castile in 1014 A.D.); and, (iii) a regularized reader's version of this *primitiva* text.

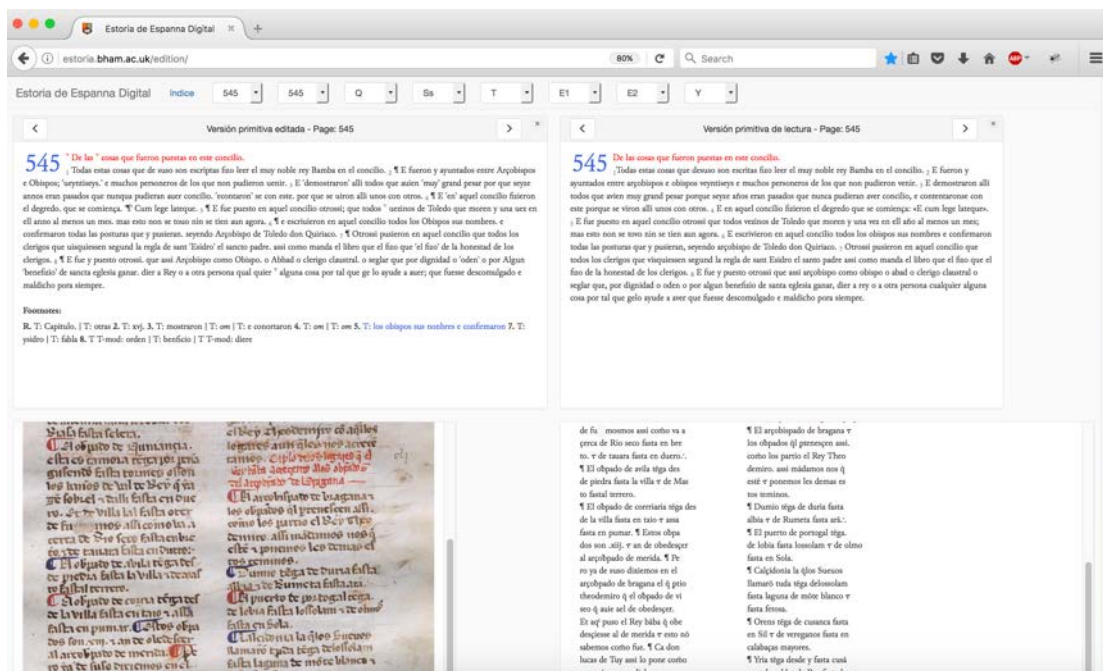


Figure 2. *Estoria de Espanna Digital* showing edited text and reader's text (top); equivalent passage in transcription of manuscript T and corresponding image.

In consequence, the remaining sections of the *Estoria*, which bring the history of the Peninsula up to the mid thirteenth century, are not included in the *edited* section of the *Estoria Digital*, since there is limited direct textual evidence of the *versión primitiva* for such sections, and none at all from the manuscripts we transcribed. The hypothesis of the *primitiva* –the first 811 of the 1146 chapters in total, is therefore the only section for which collation was carried out. The transcriptions are also presented in such a way as to allow the user to access the text of the other redactions (the *versión crítica* and *versión de 1289* respectively), but these also are not subject to collation, not least because there is no comparable text in any of the other manuscripts employed at this stage. The editorial policy after the transcription stage was therefore one of light intervention since neither the edited text nor the regularized version have significant editorial emendation. In the case of the regularized version, the text employed is the direct manuscript evidence available to us and presented in a

graphically regularized form, but no 'defective' readings are replaced with any 'better' substitutes, since even for the general reader the editorial policy is to emphasize the materiality and mobile nature of manuscript culture and not to regard these elements as an impediment to correct or original meaning. A similar, but slightly different, policy is applied in the case of the edited text -the output of the collation editor- since the aim is always to foreground manuscript text and not to emend heavily for any other purpose.

The transcriptions, of course, are both an end in themselves and the raw material for the collation stage. The guide for the preparation of transcriptions is available at the edition home page. The transcriptions are presented in the edition both in a semi-paleographic form (see Figure 2, no attempt is made to mimic the forms of letters, but all abbreviations are respected) and a fully expanded form.⁸

*

An ideal digital edition of the *Estoria de Espanna* would seek to have an edited hypothesis of the *versión primitiva*, the *versión crítica* and the *versión retoricamente amplificada de 1289* in parallel with each other and perhaps also present critically edited texts of the other added elements of the *Estoria*, such as the *Crónica particular de San Fernando*, and indeed later chronicles which draw on the Alfonsine text such as the *Crónica de Castilla* or the *Crónica de 1344*.⁹ Bearing in mind that none of the three principal redactions are extant in complete form, the use of digital tools to present a hypothesis of what they may once have looked like is very tempting. But within the current edition, given the constraints of materials employed and in the

absence of fully tagged transcriptions of all of the witnesses of the *Estoria* (not to mention the subsequent chronicles), such an aim is simply out of reach. The existence of the xml files and collations allows for the broader ambition to be realized in the future if the remaining evidence is transcribed and tagged, but for the moment the reach of the *Estoria Digital* is rather more modest; in respect of collation and edited text, we have confined our ourselves to the collation of those elements of the *versión primitiva* which enter our edition and the presentation of the results in the form of a critical text.

Before dealing in depth with the practical considerations involved in the application of digital collation to the witnesses of the *Estoria*, it is necessary to sketch out the underlying principles which governed the whole process. The first of these concerns the "why" of collation; that is, why is collation necessary or desirable in the first instance? And in what way should it be implemented?

An authoritative view on the importance and desirability of collation in the establishment of critical digital editions is given by David Parker, whose reflections on lengthy experience of editing Gospel texts hold a particular weight. If one accepts that collation is a desirable activity (and this is by no means a given) there are multiple theoretical questions to be asked. One particular issue raised by Parker is that of excluding what he calls "noise" in the printed editions of Biblical texts -understood as the overload of information on the printed page caused by the collation of too many witnesses. Clearly, such an aim is an important one, but before attempting to address how the textual noise of the *Estoria* might be removed, it is worth outlining the specificities of each edition. In the case of the Gospel editions, the removal of such

noise is an admirable aim, but this might not always be the case. The Gospel texts are (relatively) short, but have numbers of witnesses that potentially reach into four figures. By contrast the *Estoria* has but 40 witnesses, and of these only five are employed in the first phase of the edition; on the other hand the two Escorial manuscripts which together constitute the full range of the chronicle from legendary origins to the death of Fernando III (whatever the editorial status of the elements of the codices) comprise no fewer than 556 folios of densely written, double column Gothic script - a total of approximately 103,000 lines of xml transcription.

In the *Estoria* edition, the issue is not therefore one of numbers of witnesses but rather numbers of variants. In parallel to Parker's view that the advantage of digital collation is that of removing such textual noise, one might add that in the case of the *Estoria* the greater advantage of digital collation is the removal of a different type of noise, that is, those variants which are considered trivial with a view to the establishment of a text which is substantively authoritative; thus avoiding cluttering up the textual apparatus with multiple variants from the same manuscript – something which was a central consideration given the level of detail in the tagging and the resulting volume of possible variants. But since it is not always clear what constitutes noise in this sense and what might be a valuable element of the establishment of a critical apparatus, the decision of what is considered trivial and what editorially interesting is one of the most important in the *Estoria* edition.

Even more problematic is the difference in the status of the base texts in the two editions at hand here. In the case of the Gospel texts, there is a relatively stable base text that can serve as a solid basis for collation. In the case of the *Estoria*, there is no ready-made base text (or at least there was no such authoritative base available for this edition) and no lengthy cultural history of textual authority to which to appeal.

The specifics of the establishment of the base text are dealt with below, suffice to say at this point that the status of the base text was a particular difficulty in the establishment of collation principles in the edition of the *Estoria*, and the consequence of this was that a high degree of editorial caution was required in the elimination of trivial variants. Since the edition was not at a stage where evidence could be ruled out easily, editorial conservatism (in this respect at least) followed; in some sense then, in the *Estoria* the noise *is* the edition. For example, as might be expected, the extensive minor orthographic variation was regularized in the collation process (thus "u" and "v" are treated as the same) unless there was another, more pressing reason not to regularize. This is the case of proper nouns, which were left in their manuscript form as variants even if the only difference was a minor orthographic one as it is conceivable that such difference might be stemmatically significant - especially in the case of toponyms or anthroponyms which would have been unfamiliar to the scribe.¹⁰

This notion of stemmatic significance is one of the underlying principles to the exercise of editorial judgment in the collation phase, but it is not the only one, nor even the most significant. The output of the collation is used in the *Estoria Digital* to establish the critical text, that is, to present the hypothesis of the *versión primitiva* - a *primitiva* witness plus variants if you will. It is also available for use in the eventual construction of a *stemma codicum* and to establish the textual relations between the witnesses. But since the philological heavy lifting in this regard has already been done and because the *Estoria Digital* (as yet) contains only the most significant manuscripts but is certainly not exhaustive, the possibility of establishing textual relations using the data of the *Estoria Digital* (whether this be to confirm the previously analyzed relations or to modify them) lies far in the future. Thus, although the exercise of collation was carried out in the knowledge that the resulting data might

in the future have this function, the principal aim of collation was to identify significant variation and (perhaps more immediately relevant) to permit the presentation of the resulting data in the format of a variant edition of the principal redaction of the *Estoria de Espanna*.

Before outlining in depth the practical issues that arose as a result from the collation process in the *Estoria Digital*, I deal here with the methodology employed in the project as it affected the implementation of collation.

Methodology

The transcriptions were initially prepared on the Textual Communities system at the University of Saskatchewan.¹¹ In the early phases of the project it was intended that Textual Communities would be able to provide a full system of transcription, automatic parsing and collation, however it proved more effective to undertake the collation phase independently (for which see below). The transcriptions were prepared according to guidelines specifically designed for the project by Bárbara Bordalejo in conjunction with the rest of the editorial team.¹² As the project was starting from scratch, the initial task was that of compiling a base text. Although this required transcribing xml files from zero, the transcriptions of manuscripts E₁ and E₂ compiled by the Hispanic Seminary of Medieval Studies then at the University of Wisconsin, Madison proved an especially valuable resource in establishing the base text.¹³ The first fundamental difficulty in providing a base text -both for the purposes of collation and, even more immediately, for the purposes of transcribing all of the witnesses- is the lack of an authoritative manuscript text to choose as a base.

Manuscript E₁ is an Alfonsine witness, which therefore could be used for this purpose, but its text covers only the years from the beginning of the chronicle to the reign of the eighth century king Alfonso I, and it is the only such extant Alfonsine codex of the chronicle. Its companion manuscript, E₂, is also from the royal scriptorium, but from the year 1289, after Alfonso's death and it contains a composite text, none of which is the redaction from Alfonso's own scriptorium. On the other hand, the two manuscripts together are the sole royal witness, and together they also form the only complete early version of the *Estoria de Espanna* from start to finish. The resulting dilemma, worthy of Saramago's *History of the Siege of Lisbon*, was that in order to have a base text worthy of use it was necessary to collate the witnesses, but in order to collate the witnesses it was necessary first to have a base text.¹⁴ In the face of this dilemma, it was decided to use the text of E₁ and E₂ together as a base text since, for the reasons outlined above, it was not possible to devise a more methodologically appropriate base text at present. Of course, the ultimate output of the process of collation may give rise to an edited text which could serve as such in future, but once more, that is not yet possible. The second fundamental difficulty concerned the nature of the base text itself. For the purposes of transcription and collation, it was necessary to devise a numbering system to cover the entire text of the chronicle so that direct comparison could be made between all of the witnesses. Again, the devising of a numbering system served the needs of collation, but it also has the virtue of allowing (at least in theory) direct cross reference between the different witnesses of the *Estoria*.¹⁵

The numbering system was devised by Aengus Ward and employed as its most significant elements the <div> tag to represent chapter level divisions and the <ab> tag to segment the text of the chapters semantically.¹⁶ The <div> tags respected the

rubric divisions of the two Escorial manuscripts. This gave rise to a number of difficulties in later stages when as a result of the collation it was realized that some textual divisions (e.g. chapters 741-743 in the edition) were not represented by separate rubrics in the Escorial manuscript although this was clearly an error, since they appear as separate chapters in the rest of the textual tradition. In consequence, after the collation (when these issues arose) the transcriptions and collations had to be revised and the numbering altered. The <ab> tags attempted to respect, where possible, the indicators of semantic division in the text; in practice this meant respecting the punctuation and capitalizing of the manuscript text. In the royal, thirteenth century, sections of the two base manuscripts this was relatively unproblematic, since there was clearly a closely managed form of textual organization, but in the fourteenth century additions to E₂ this was not the case, and the resulting employment of <ab> tags was less guided by the textual organization of the manuscript than might have been wished for. The overall effect, nonetheless, was to provide a system of segmentation of the *Estoria* which permitted collation and which also should provide a new mode of reference to the text of the *Estoria de Espanna*.

Having established both the transcriptions of the five witnesses concerned and a segmentation system fit for purpose, it remained to employ an appropriate digital tool to allow effective collation. The tool in question is CollateX, written by Ronald Dekker at the Huygens Instituut voor Nederlandse Geschiedenis in the Netherlands.¹⁷ CollateX, described by its writer as the "spiritual successor to Peter Robinson's Collate", allows for the collation of multiple witnesses -although in the case of the *Estoria Digital* there were relatively few, indeed, the maximum of four separate witnesses for any one segment of the text was reached only between chapters 397 and

432 as there is no great overlap between most of the witnesses employed. Further development was carried out by Cat Smith at the Institute for Textual Scholarship and Electronic Editing at the University of Birmingham to allow for the use of the editor in the *Estoria Digital* project.¹⁸ The raw xml transcription files were converted to json files using the latest incarnation of the Anastasia software, originally developed by Peter Robinson.¹⁹ The conversion of files for use in the collation editor was greatly aided by support from Peter Robinson at the University of Saskatchewan and Zeth Green at the University of Birmingham. The collation of the 811 chapters of the *Estoria* which represent the text of the *versión primitiva* available to the *Estoria Digital* was carried out between November 2015 and October 2016. A separate section, corresponding to the final chapters of the *Estoria* known as the *Crónica particular de San Fernando* (chapters 1049-1146) is also available for collation - both manuscripts E2 and Ss contain this section- but it is not representative of the *versión primitiva*; in addition, as mentioned above, this section is the subject of ongoing work by Polly Duxfield.

In other aspects of the work on the project (most notably the transcription phase) the labor was carried out collaboratively, that is, team members transcribed separate sections and these were then subject to overall editorial control by the editorial team and specifically the general editor.²⁰ The use of the Textual Communities project facilitated this work greatly, as it enabled all concerned to observe and comment on the work of all, and also allowed the editors to monitor progress. However, in the case of collation, it was decided that it would be more effective to entrust this task to a single individual as (i) there was not a similar tool which would allow monitored collaborative working in the same way and (ii) it was

easier to guarantee a greater degree of coherence of practice this way, and this was especially important given the evolutionary nature of the collation guidelines.²¹

In practice, the collation operation consisted of three fundamental operations. The first is the regularization phase, at which the first editorial decisions are taken over significance of variants.

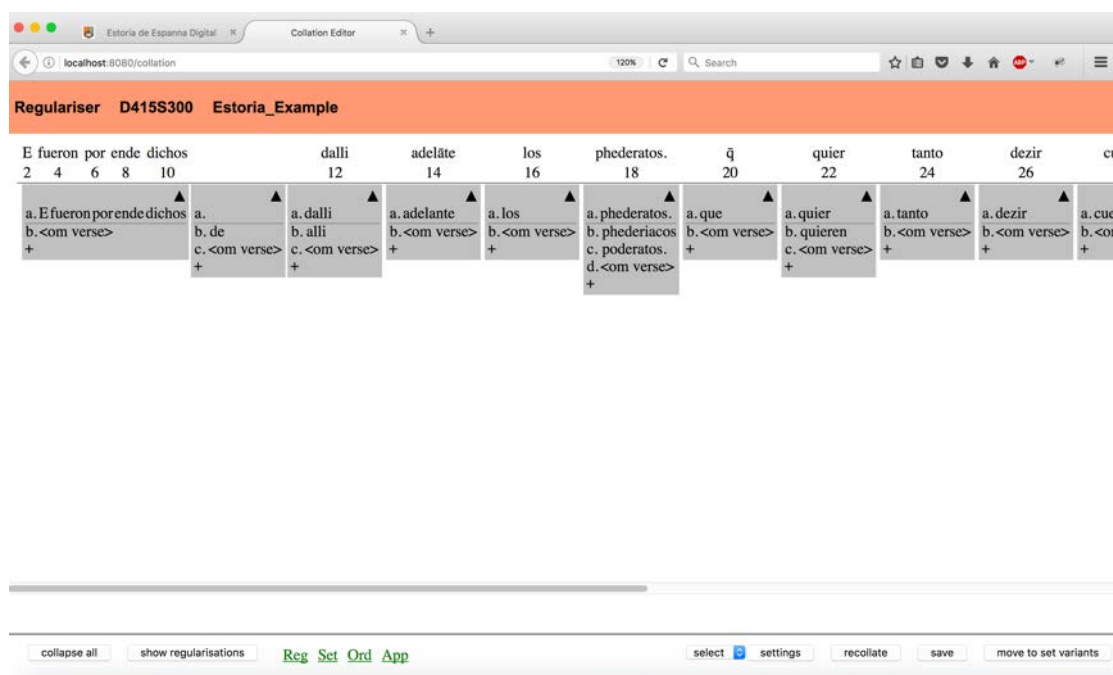


Figure 3. Regularization of <div> 415 <ab> 300

The following stage consisted of the setting of variants.

Set Variants D415S300 Estoria_Example

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26			
a. E fueron por ende dichos Base, E1, Q, Ssr, Tr											a. dalli						a. adelante los b. <om verse>			a. phederatos. b. phederiacos c. poderatos. d. <om verse>		a. que b. <om verse>		a. quier b. quieren c. <om verse>		a. tanto dezir cu b. <om verse>		

collapse all show subreadings Reg Set Ord App select Show Shared Units Save Move to Reorder Variants

Figure 4. Setting of variants for <div> 415 <ab> 300</div>

And the final stage was the ordering of variants, in which the priority of different variants could be altered in the apparatus. In practice, since there were few witnesses, this was rarely used.

Order Readings D415S300 Estoria_Example

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
a. E fueron por ende dichos +b. <om verse>											a. dalli +b. <om verse>						a. adelante los +b. <om verse>			a. phederatos. +b. phederiacos +c. poderatos. +d. <om verse>		a. que +b. <om verse>		a. quier +b. quieren +c. <om verse>		a. tanto dezir cuemo +b. <om verse>		a. auenidos +b. arrenidos +c. <om verse>		a. por amistad +b. <om verse>					

collapse all show non-edition subreadings Reg Set Ord App select Save Approve

Figure 5. Ordering of variants for <div> 415 <ab> 300.

The practical implications of the use of CollateX and the way in which the guidelines for collation were derived are outlined below. The resulting data was then available for processing into the edited form, an example of which can be seen above in Figure 2. The collation editor was employed remotely and all of the results backed up on the project servers at the University of Birmingham, which also now host the edition and all of the data and metadata generated. All of these will be openly available in due course.

Practical considerations and collation guidelines

It should always be borne in mind that the output of the collation editor was to be used primarily as a way of presenting the text of the *versión primitiva* in the first instance, and that the establishment of textual relations was a different, and long term, goal. A conditioning factor in the presentation of the edition of the *primitiva* is that, at present, no alterations can be made to the base text within the collation editor. Thus, the guiding principle of the collation was that the base text was (editorially) prior and immobile and if any variants arose which, in the view of the editor, were better than those offered by the base text, these could not (at least at this stage) be incorporated into the base as a better reading. In another project, this would not perhaps be a major consideration, for if the base text is already understood to be authoritative and the witnesses a source of (implicitly inferior) variants, there is no question of needing to alter the base text (at least, not regularly). However, as outlined above, the base text in the case of the *Estoria de Espanna Digital* was comprised of the manuscript

testimonies which most closely transmitted the chronicle. In the case of the opening 627 chapters, this might not be considered to be a significant issue, as the text is drawn from E₁, the only Alfonsine manuscript extant. Not only for reasons of proximity to Alfonso's scriptorium but also (as it proved) for reasons of textual consistency and accuracy, E₁ is an excellent base text and rarely would require emendation for reasons of defective readings. However, the same is not true of those sections which employ E₂, Y and (rather more extensively) T as the base text. Here the physical manuscripts are rather more removed in time (and in consequence textual coherence) from the Alfonsine exemplar and so there were many occasions on which editorial emendation could have been made. As the collation editor did not allow this, an alternative -outlined below- had to be arrived at, although there were also sound intellectual reasons (aside from the practical ones associated with the collation editor) why this emendation of base text was not in the end employed, as is detailed below.

The process of editing the *Estoria* could thus be split into a series of discrete operations, each of which is separate but related to the rest. It is helpful to distinguish two meta-operations, following the views of Paul Spence, those of preparation of data and presentation of the results.²² In the current analysis I focus solely on the former, although it is important to note that the two do not work completely in isolation from each other, since the principles governing the preparation of data (transcription guidelines etc.) are necessarily inflected by the editorial view on how the final edition should look; although the choices made at the presentation phase should not reduce the flexibility of choice in the ultimate presentation – something which is clearly not the case in print. For the purposes of this article, however, I do not intend to deal in any great depth with presentation of data and confine myself to the first meta-

operation, that of preparation of the data - broadly conceived as the establishment and putting into practical effect of transcription guidelines and, in sequence, the establishment and putting into practice of collation guidelines. With regard to the edited text, for reasons outlined above, the object of the exercise of collation in the first instance was not that of establishing an authoritative edited text of the *Estoria*, but rather allowing the redactions and witnesses of the chronicle to breathe in their diversity. Since the fundamental aim was not that of establishing a taxonomy of scribal error, nor taking a view on the perceived value of any of the variants, the exercise of collation at both the regularization stage and the setting of variants was targeted at the identification of variation which might have some stemmatic significance (whether this was scribal error or not) but which also might also have other contextual value related to the conditions of composition of the witness in question (although this would also most likely have stemmatic significance). In this regard, questions of regularization (or not) of graphic difference or punctuation will inevitably give rise to criticism on the grounds of inconsistency, although the failsafe in this sense is that the detailed transcriptions which contain as accurate a representation of original orthography and textual structure as were deemed possible are provided in direct links from the edited text.

Regularization

The initial phase of collation -that of regularization- is in principle the most straightforward as it implies the elimination of variants regarded as trivial. This is done by means of a drag and drop mechanism in the collation editor; the variant regarded as trivial is dragged to whichever variant (base text or other manuscript) is

considered superior.²³ Because of the manner of operation of the editor, the base variant is not alterable, and since the base text in the *Estoria Digital* is made up of text from particular witnesses the direction of regularization, whether one thinks the base variant is superior or not, is always the same. This may give rise to minor inconsistencies - for example, if the base contains the variant "auia" then any "auja" variant will be regularized to it. But if elsewhere the base contains "auja" and the variant manuscript "auia" (or "avia") then the opposite regularization will take place.²⁴

However, before dealing with the implications of this, we return to question of trivial variants, for arriving at a definition of precisely what constitutes triviality in this sense is not unproblematic. As mentioned above, the first criterion addressed in the regularization phase was that of orthographic variance. Thus, it was assumed that where the sole distinction between two manuscript readings was one of the following:

u/v/b;

i/j and occasionally y;

m/n before bilabial consonants;

c/ç and z

culto consonant clusters (such as ph/f etc.)

the distinction between the two variants was solely graphic and therefore the variant was not preserved for the apparatus but rather regularized out; the same is true of contractions as in the case of “de alli/dalli” in Figure 4. The underlying principle for collation was that any variant has the potential to offer contextual value, so there had to be a positive reason to exclude rather than a negative assumption that particular

classes of variants were of no inherent value. Of course graphic variants have a value for the history of orthography and the evolution of such questions an important element in the diachronic study of written discourse; nonetheless, since the variant forms are available in the transcriptions and the study of the evolution of orthography is not a specific aim of the edited text, it is entirely reasonable to confine the positive products of collation to a different set of values. A similar point might be made with regard to phonetic difference. Once more, the purpose of the collation is not to offer a history of Castilian phonetic change, so one might suggest that the editorial exclusion of those variants considered to have uniquely phonological value would be similarly consistent. And indeed, establishing a clear dividing line between the graphic and the phonological is not unproblematic, especially when it is considered that the evidence employed in the five manuscripts concerned spans the late thirteenth to the mid fifteenth centuries. In any case, the transcription decisions taken already presuppose a certain degree of editorial intervention in this regard. The collation employs the expanded xml text and therefore presents, for example "nombre" in T as a variant for "nombre" in E₁. But in fact, the underlying xml is likely to be "no<am></am><ex>n</ex>bre" and "no<am></am><ex>m</ex>bre" respectively and both would appear in the manuscript as "nõbre" -the prior decision was taken to respect the *usus scribendi* of each manuscript where possible when expanding abbreviations – and in this case the expanded for “nombre” is more common in E₁/E₂ and “nombre” more common in T. In consequence, specifying one as a variant of the other would be absurd. In this light, and that of the variance of scribal practice, the aforementioned limited number of orthographic variants to be regularized was established. But once more, it is emphasized that this list was followed unless there was an over-riding motive to retain the variant. The most pressing of these, as

mentioned above, was the question of toponyms and anthroponyms. With the exception of borderline differences (for example, the length of descenders in i/j etc.) which were regularized, it was decided that other graphic differences would not be regularized out as these were potentially indicative of stemmatic differences in the manuscript tradition; this is particularly true of those proper nouns which would have been unfamiliar to scribes and which might therefore have undergone significant variation at later stages.

In consequence, there may be occasions on which variants have more than one graphic difference but because these are both considered trivial for our purposes the variant will be regularized. It may also be the case that some variants were preserved even though the difference between base and variant was just one graph; these variants are preserved because the difference between them is considered otherwise significant. Such is the case, for example, of tonic preterites. Since there is (potentially) a significant linguistic, rather than solely graphic, difference between the variants, the alternative reading is preserved. Although this may appear to be the product of an overly conservative approach, it was felt that the notion of variational significance should be as broad as possible. Similarly, minor linguistic alterations such as, for example, "o/do" or "so/suyo" should also be preserved as such distinctions at least have the possibility to indicate stemmatic or contextual difference. A rather more borderline case is that of metathesis, but again in this case a more conservative approach was taken and such variation as "peligro/periglo" was preserved in the apparatus. A further case of preservation of variants relates to the question of numerals; although numbers written discursively and those in Roman form clearly mean the same thing it was decided that the different format could be a differentiating element between manuscript traditions. And while it has not yet been

demonstrated that this is actually the case, nonetheless, the preservation of possible lines of comparative enquiry made the retention of different number formats worthwhile. However, in line with the overarching principle, differences of punctuation were regularized. The rationale for this is not that punctuation is trivial, since the disposition of the text on the manuscript page can be of great interest in the establishment of our knowledge of individual practice and the relationships between manuscripts. On the contrary, the physical disposition and textual organization of a manuscript (in particular the range of punctuation employed and the manner of its employment) is a much-underused element in contemporary manuscript studies. However, this must be done on a more coherent, and manuscript specific basis, and for the purposes of this exercise, the analysis of comparative punctuation would not add greatly to the *edited* text.

A final issue with the regularization phase relates back once more to the status of the base text. Once more, in those sections for which E_1 is the base text, there are few occasions on which the variant text is superior, and fewer still in which the base is missing sections that do appear in the alternative manuscripts. However, in the sections in which Y and (especially) T form the base text, it is more frequently the case that alternative readings (in practice from E_2) are clearly better. When this is a single variant, the difficulty is minimal; the variant can be left as such and the editorial judgment on the relative superiority of individual readings left to a later stage. But occasionally that variation is caused by the absence from T of significant passages of text. In some cases, this is caused by substantial lacunae in the manuscript, as between chapter 711₃₆₋₇₀, but it can also be the case that scribal error has given rise to the loss of a line or a whole sentence, as in chapter 739. In this event, it is clear that the base text is defective and is missing elements which truly derive

from the *primitiva*; in consequence, the editorial intervention will be to insert those sections of the alternative manuscript (in practice E₂) which although not directly *versión primitiva* at least have the virtue of filling a clear gap in the base text (at present this has not yet taken place). The aim would ultimately be to indicate the presence of such text in the edited version by the use of a different font, or by the use of alternative witnesses when all 40 manuscripts are incorporated into the edition. It should also be noted that the opposite case did not apply; that is, if the *primitiva* text was clearly complete in the base and Q or Ss had additional text, this was not included as a separate chapter/sentence in the collation. The reason for this is that the edited text is intended to represent the *primitiva*, as closely as possible, and so additional text from alternative versions does not enter into this representation of it. As a result, the use of variants from the *versión crítica* (here represented by the manuscript Ss) is only operable up to chapter 621. At this point, the organizational structure of Ss alters greatly to the extent that it is not collatable with the base text drawn from the *primitiva* so no effort is made after this point to reconcile what are significantly redactions of the *Estoria de Espanna*.

Setting of variants

Having eliminated all those variants deemed trivial, the following stage allowed the editor to undertake another a further operation -the setting of variants- and also to complete the regularization. Since the alignment of the witnesses was undertaken automatically in each verse/sentence, it was inevitably the case that not all of the variants lined up directly with each other. The advantages of dealing with verse are absent here, and the length of the verses/sentences can mean significant variation in

the order of textual elements. The alignment of variants is perhaps the most significant editorial task since its principal function is not just to identify the variants but more crucially to define the nature of variation.

An example of this can be seen in Figure 6 below:

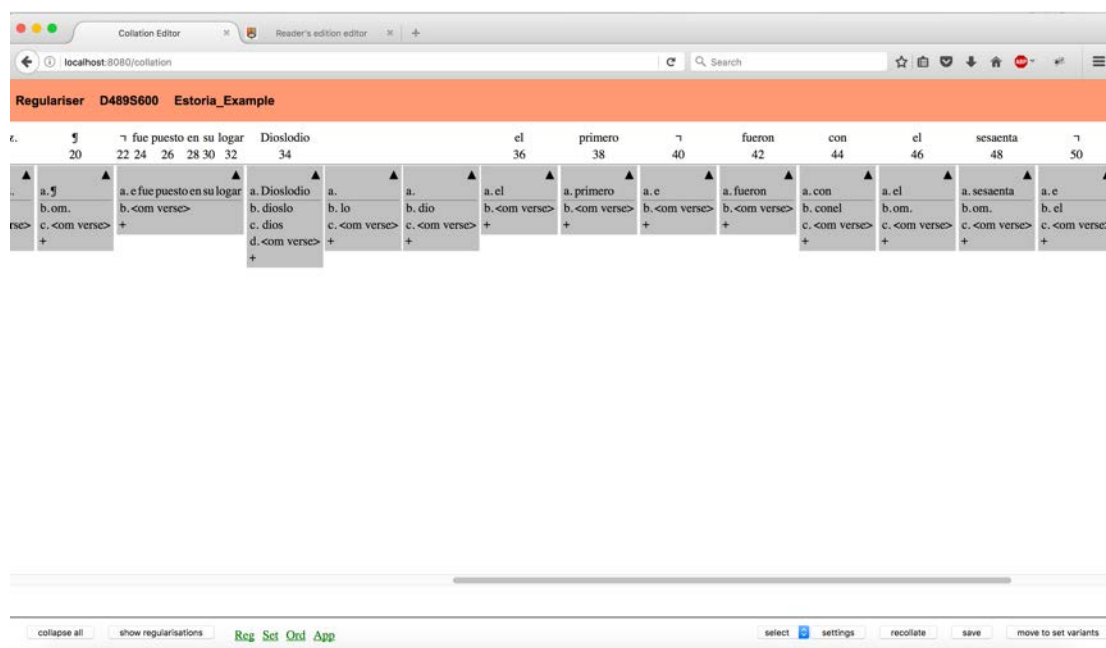


Figure 6.

The base text in segment 34 is the anthroponym "Dioslodio" and precisely the same letters appear as the variant readings, albeit crucially in one case, with word spacing. The grouping together of the following two (unnumbered) segments is logical as the three form a single entity. Human intervention is required at this point as the settings on the collation editor do not permit the editor to recognize these as a single segment. However, this is also not sufficient on its own, as suppressing the spaces would give rise to all three being collapsed into one "Dioslodio". In fact, the operation of the variant setting gives rise to the following:

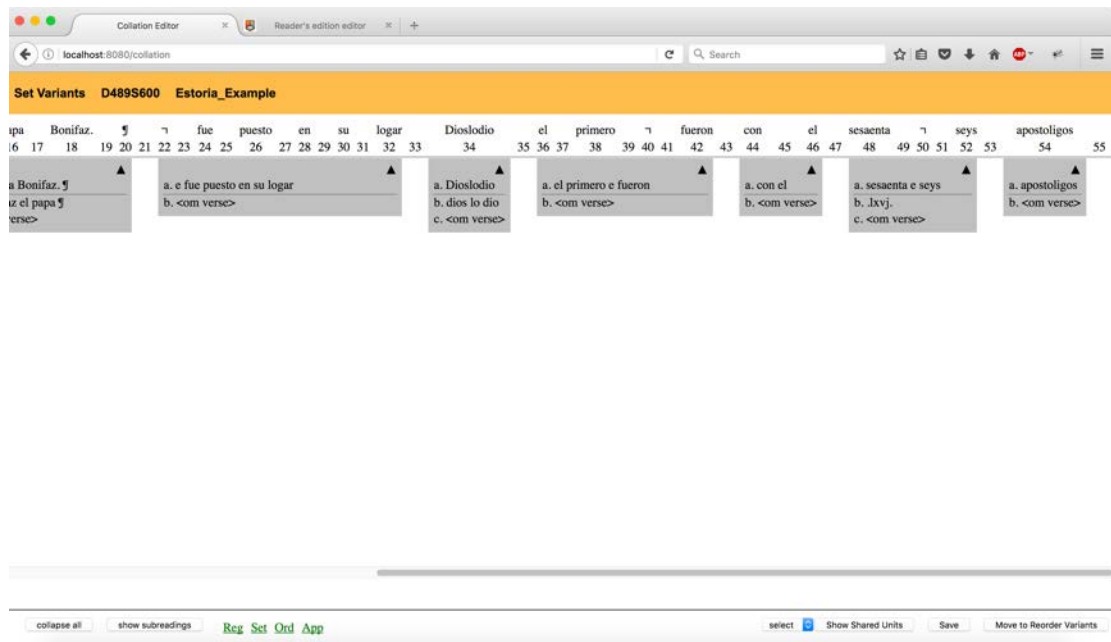


Figure 7.

which recognizes that the scribe of (in this case) Ss has misread the anthroponym as the verbal clause "dios lo dio". Since this is potentially of stemmatic significance it is important that the variant be preserved, and it is only through editorial intervention *at this stage*, that the variant can be correctly identified as such.

Few of the issues arising from the setting of variants were so straightforward, but most concerned the definition of variation. Another such example is illustrated in figure 8 and 9:

Collation Editor

localhost:8080/collation

Regulariser D758S7300 Estoria_Example

2	4	6	8	10	12	14	16	18	20	22
a. [text]	a. Pero	a. quelos	a. moros	a. [text]	a. [text]	a. luego	a. e descabecaron le y.			a. [text]
b. Et	b. la	b. muchedumbre	b. cercaron	b. le	b. le	b. manos	b. <om verse>			b. luego.
c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>			c. <om verse>
+	+	+	+	+	+	+	+			+

collapse all show regularisations Reg Set Ord App select settings recollate save move to set variants

Figure 8

In this case, the base text (T) and the variant readings (E_2) differ rather more substantially. Segments 2 to 8 are relatively unproblematic, since the whole sequence can be considered to be a variant. The issue is rather more with the presence of "luego" in the two witnesses. The resulting set variants screen is as follows:

Collation Editor

localhost:8080/collation

Set Variants D758S7300 Estoria_Example

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
a. [text]	a. Pero	a. quelos	a. moros					a. prisioner le					a. luego				a. e descabecaron le y.					a. [text]
b. Et	b. la	b. muchedumbre	b. cercaron	b. e				b. <om verse>					b. a manos				b. <om verse>					b. luego.
c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>	c. <om verse>				c. <om verse>					c. <om verse>				c. <om verse>					c. <om verse>
+	+	+	+	+				+					+				+					+

collapse all show subreadings Reg Set Ord App select Show Shared Units Save Move to Reorder Variants

Figure 9

The outcome of this is to recognize that the verbal sequence in segments 16-22 is present in the same way in both witnesses, but the question of the place and status of "luego" is more problematic. The variant text will appear in the footnotes as an addition at the end of the sentence and the "a manos" appears as a variant of "luego" in segment 14, neither of which is exactly correct. But recognition of "luego" as part of the verbal string 16-22 would make the entire segment a variant, which is also not correct. The upshot of this is that editors must make whatever editorial decisions they see fit. In the case of the *Estoria de Espanna Digital* it was decided to acknowledge the greatest degree of similarity possible, and the result is to give rise to a greater number of short variants rather than lengthy strings with little variation. One phenomenon of this type that is especially notable is the question of variation within a variant; that is, those occasions on which variant text not present in the base text is itself scribally modified. The only way of dealing with this currently is to set each segment separately, as otherwise the entire sequence of variants will be gathered together; but the result is that an extensive sequence of variants will appear instead of one string with a small modification. Part of this issue can of course be dealt with at the presentation phase, but the definition of the collation units at the set variants stage remains one of the key conditioners for the compilation of the edition.

Presentation

The output of the collation editor, comprising 122,824 individual json verse files, was then available for presentation in the form of an edition. However, as

alluded to above, the collation editor presupposes a fully edited base text against which to measure the value of variants. In the case of the Estoria de Espanna Digital since the base text was the aggregated transcription files of those sections of the manuscripts which contained the *versión primitiva* (and supplemented by the text of E₂ after chapter 811 when there no longer was direct evidence for the *primitiva* in the manuscripts concerned) it was quite possible that non-base variants would be considered to be better readings on occasion. And yet, the output of the collation editor did not allow for this possibility. To allow for a fully edited text presented according to the principles of textual criticism, another stage would be necessary. An emendation editor would also have the virtue of allowing the presentation of the hypothesis of the *primitiva* in a reader's format, that is, with a standardized and regularized orthography in keeping with the mode of presentation of most current printed editions of medieval Iberian texts. To this end, Zeth Green of the University of Birmingham designed a reader edition editor precisely for this purpose.²⁵ An example of one short sentence can be seen in Figure 10 below:

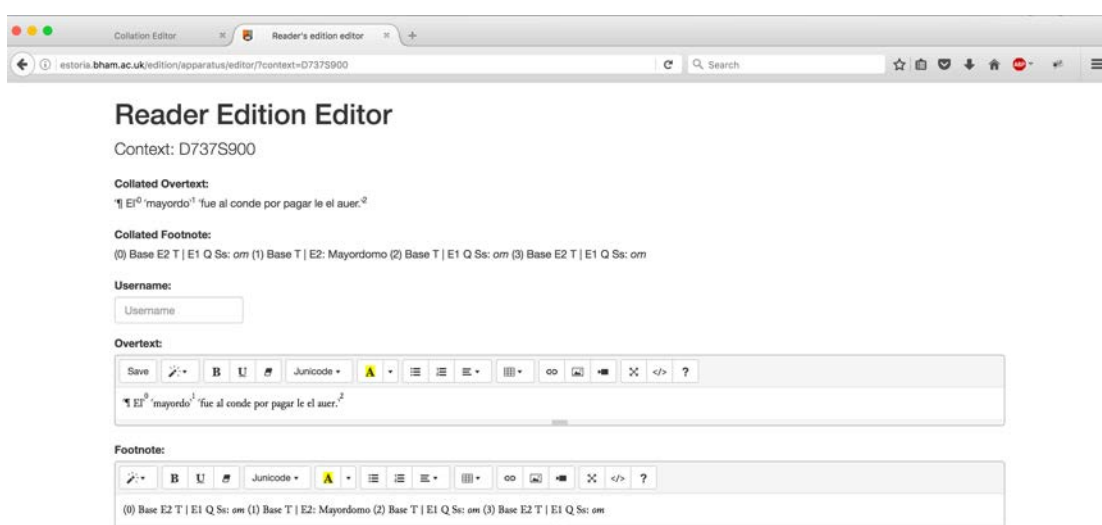


Figure 10

The editor presents the collated overtext, that is, the output of the collation editor, and the resulting variants in separate editing boxes. The overtext can be edited for the purposes of regularization and the footnotes can similarly be edited; in this case, the variant reading "mayordomo" from E2 can be inserted into the edited text and the reading "mayordo" from T (which is the base text here) relegated to the footnotes. In this way, a fully edited text can be compiled according to the appropriate criteria since the editor developed here allows for a hierarchy of variants while at the same time permitting the replacement of base text readings with others and facilitating the regularization of graphs for a reader's edition. However, despite the existence of this facility, as can be seen below in Figure 11, it was decided not to edit the text in this way at this time and to keep the edited text and the reader's text separate.

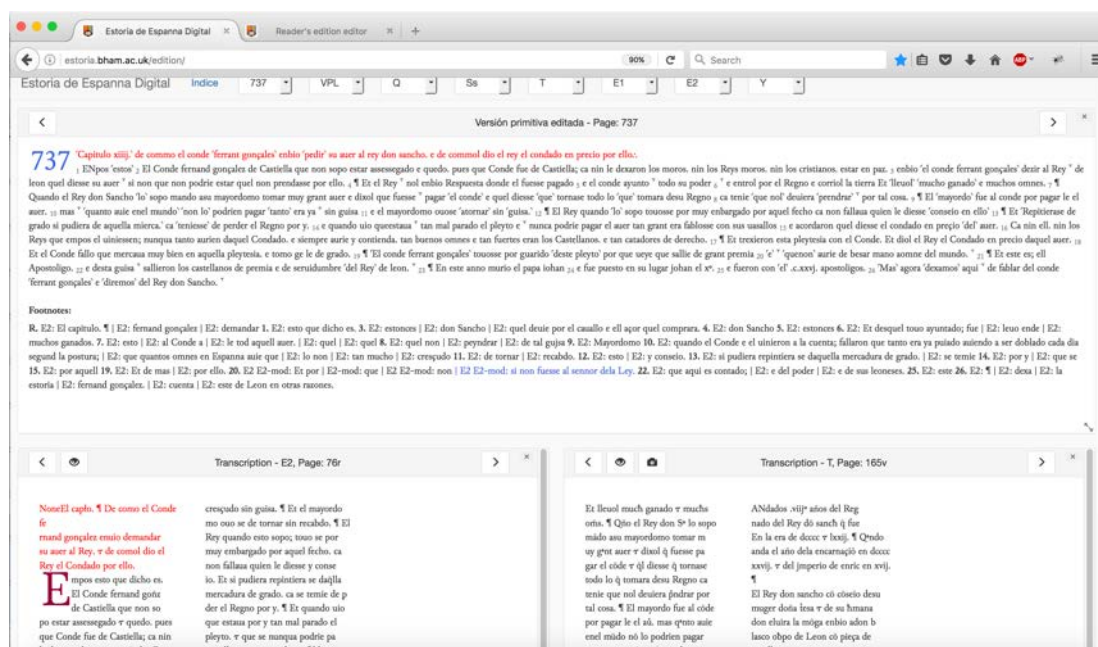


Figure 11.

The presentation of the edition takes the form of an *unregularized* transcription of the base text with variant readings in the footnotes, with no attempt to hierarchize these according to philological principles. Clicking on the variant opens the relevant part of the transcription and clicking on the footnote number opens the base text at that point. A regularized version of the base transcription, again not emended with superior readings, is available in parallel, but not in the same edited text box. The rationale for presenting the text in this way and not seeking to amend is twofold: (i) In the first instance the emphasis in the *Estoria Digital* is resolutely on manuscript text and since the digital means employed permit a range of presentations which aid this, it was decided to foreground the variance in the text and not present a text which has all the appearance of fixed authority. (ii) The evidence employed in the edition is limited to five manuscripts of the 40 extant codices; in this light, any edited text could only be extremely partial, to the point that it was felt that such a presentation would be theoretically and intellectually unjustifiable. The development of a reader's edition editor permits the construction of such an edition, but in the view of the current editor there is too much missing evidence at this stage to warrant an editorial presentation of a type which is, of all the possible presentations, the most easily replicable in print. For these reasons, it was decided to keep the reader's text and the edited text as *separate* until sufficient evidence is available to allow for a fully edited format.

Conclusion

The direct benefits for the study of the *Estoria de Espanna*, both now and in a future realization of the *Estoria Digital* in which all of the manuscript evidence is

included, may be clear, but any digital editing project should also have rather more generalized benefits – what might these be in this case?

Recent debates about the nature of digital scholarly editions have raised the possibility that collation may not be the future of such editions. In such an optic, the principal focus of the digital editor would be on the individual document, with the aim of analyzing in as great a depth as possible the discursive complexities of each document in as wide a range of enmeshed presentations as technology permits. However, the inter-related nature of medieval manuscript traditions (particularly in the case of extensive prose texts like the *Estoria de Espanna*) means that a major form of analysis which reflects that inter-relatedness would be lost to us in the absence of some mode of collation. The experience of collating and editing the *Estoria de Espanna* (the two verbs are related but resolutely not the same) allied to the digital presentation of the outputs of these processes suggests that the collation process can be a very revealing one for the ongoing analysis of medieval texts and textuality. There are, of course, drawbacks. As is the case with any digital editing process, the value of the results is only maintained while the format of these is supported. Indeed, this may be even more the case for collation than for, say, transcription, since there is a globally agreed standard for the latter and it is assumed that any major technological change will account for the fact that the xml/TEI5 standard texts must be incorporated in some way in the future. This is not the case for the outputs of collation, or at least not in the same way.

Those outputs form a central part of the edition itself. At all stages of the process editorial judgement is required, as outlined in the examples above. The creative ways in which such outputs can be presented, in tandem with other presentational and analytical tools are yet to be explored in full but there can be no

doubt of the utility of the contrastive analysis which collation permits. The collation system employed in the *Estoria de Espanna Digital* is not a panacea for all the ills of contrastive analysis, but it does point to ways in which editorial judgement can be exercised in a consistent and overt way to allow for creative and multidimensional presentation and analysis in the future.

The extensible benefits are two-fold: on the one hand, the *Estoria Digital* has pointed to the specific advantages of collation in the first instance, since the edition creates an environment in which the manuscript evidence can be accessed and analyzed both as document, on its own terms and as a function of its context of composition, and as part of a wider related whole – the work *Estoria de Espanna*, if you will. The ways in which the results of these separate, but related, editorial operations can be presented creatively remain to be explored, but the output of the *Estoria Digital* at least demonstrates the advantages of the possibility of collation and the juxtaposition of different ways of reading. It also begins to raise the question of *what* should be collated as well as that of *how* (something which is probably project specific in any case) since it implicitly raises the question of the relative relatedness (conceived in a variety of ways) of different documents. The second extensible element is, as always, that of tools and their development. CollateX has undergone significant alterations in its life-span and the *Estoria Digital* highlights its adaptability to a rather different subject matter to that of its original iteration. Tools, of course, are always both under development and in imminent danger of obsolescence, but the experience of the *Estoria Digital* may go some way towards contributing to extended use of this and related forms of digital textual scholarship. In this sense, the experience of the *Estoria Digital* collation has both theoretical and practical implications for future projects.

Manuscripts cited

- El Escorial, Real Monasterio de San Lorenzo de El Escorial Y-I-2, *E₁*
- El Escorial, Real Monasterio de San Lorenzo de El Escorial X-I-4, *E₂*
- Madrid, Biblioteca Nacional de España, BNE 5795, *Q*. <http://bdh-rd.bne.es/viewer.vm?id=0000134882&page=1> and estoria.bham.ac.uk/edition
- Salamanca, Caja Duero Salamanca 40, *Ss*. estoria.bham.ac.uk/edition
- Santander, Biblioteca de Menéndez Pelayo M-550, *T*. estoria.bham.ac.uk/edition

Works cited

Blecuá, Alberto, *Manual de crítica textual*. Madrid, Castalia, 1983. Print.

Bordalejo, Bárbara. “The Texts We See and the Works We Imagine: The Shift of Focus of Textual Scholarship in the Digital Age.” *Ecdotica* 10 (2013): 64-76. Print.

Catalán, Diego. *La Estoria de Espanna de Alfonso X, creación y evolución*, Fuentes cronísticas de la historia de España, 5, Madrid, Seminario Menéndez Pidal/UCM, 1992. Print.

_____. *De la silva textual al taller historiográfico alfonsí: Códices, crónicas, versiones y cuadernos de trabajo*, Madrid, Fundación Ramón Menéndez Pidal/Universidad Autónoma de Madrid («Fuentes Cronísticas de La Historia de España», 9) 1997. Print.

Estoria de Espanna Digital. v.1.0, General editor Aengus Ward, Birmingham: University of Birmingham, 2016. <estoria.bham.ac.uk> [accessed 31/01/18]

Fernández-Ordóñez, Inés. *Las Estorias de Alfonso el Sabio*, Madrid, Istmo, 1992. Print.

- _____. *Versión crítica de la Estoria de Espanna*, FCHE, 6, Madrid, Seminario Menéndez Pidal/UAM, 1993. Print.
- _____. “La transmisión textual de la “Estoria de Espanna” y de las principales “crónicas” de ella derivadas”, in ed. Inés Fernández-Ordóñez, *Alfonso X El Sabio y Las Crónicas de España*, Valladolid, Secretariado de Publicaciones e Intercambio Editorial/Universidad de Valladolid/Centro para la Edición de los Clásicos Españoles, 2000, 219-264. Print.
- Gabler, Hans Walter. “Theorizing the Digital Scholarly Edition”, *Literature Compass*, 7/2 (2010): 43-56. Print.
- López de Ayala, Pero. *Crónica del rey don Pedro y del rey don Enrique, su hermano, hijos del rey don Alfonso Onceno*. I. ed. Germán Orduna y José Luis Moure, Buenos Aires, SECRET, 1994. Print.
- Parker, David. Web. http://www.tei-c.org/About/Archive_new/ETE/Preview/parker.xml
- Pierazzo, Elena. “A rationale of digital documentary editions.” *Literary and Linguistic Computing* 26.4 (2011): 463-477. Print.
- Robinson, Peter. “Towards a Theory of Digital Editions.” *Variants* 10 (2013): 105-131. Print.
- Sahle, Patrick. “What is a digital scholarly edition?” *Digital Scholarly Editing, Theories and Practices*. eds. Matthew James Driscoll and Elena Pierazzo. Cambridge: Open Book Publishers, 2016: 19-39. Print.
- Saramago, José. *History of the Siege of Lisbon*. London: Harvill, 1996. Print.
- Saracino, Pablo Enrique. *Crónica de Sancho IV*. Buenos Aires: Incipit, 2014. Print.
- Spence, Paul. “Siete retos en edición digital para las fuentes documentales”, *Scriptum Digital*, 3 (2014): 153-181.

Ward, Aengus. “Editar la *Estoria de Espanna*: retos y problemas de la edición digital.” *Incipit XXXVII* (2017), 13-43.

¹ See for example the extensive work carried out by Catalán, *La Estoria* and *La silva*; Fernández-Ordóñez *Las estorias* and *Versión crítica*; and the detailed summary of manuscript witnesses contained in Fernández-Ordóñez “La transmisión”.

² estoria.bham.ac.uk The edition proper can be found at estoria.bham.ac.uk/edition.

³ E₁ and E₂ together contain the full extension of the chronicle and are therefore companion volumes, although with the exception of the first two gatherings, E₂ does not contain material directly from the Alfonsine *scripdtorium* as it was composed in the reign of Sancho IV immediately after Alfonso’s death. Other materials were later added to this codex in the fourteenth century. Despite this composite nature, and partially because it has always been in the royal library, the two together have always been accorded a special status.

⁴ The original version of the diagram and additional visual resources can be found [here](#) along with a full justification of the choice of manuscripts and rationale for the edition.

⁵ For a discussion of recent debates on the centrality of manuscript evidence see, *inter alia*, Bordalejo and Pierazzo.

⁶ xml (extensible markup language) is the standard language used to compile transcriptions for digital editions. TEI5 – the fifth release of the Text Encoding Initiative norms are the standard humanities framework within which xml is used to transcribe and describe the subject matter of digital editions.

⁷ Sahle, 29, describes the digital edition as a process rather than a product.

⁸ The full criteria can be found [here](#).

⁹ The *CPSF* is the subject of the forthcoming thesis and digital edition by Polly Duxfield.

¹⁰ The relationship between the notions of work, text and document, so central to the justification for employing collation in the first instance, is dealt with by, *inter alia*, Gabler, Bordalejo, Robinson and Ward.

¹¹ Textual Communities: <http://www.textualcommunities.usask.ca>

¹² The roles of the team members can be found [here](#). The guidelines for transcription can be found at [the edition website](#).

¹³ Although the HSMS guidelines provide for a significantly lower level of tagging, nonetheless the transcriptions were of no little help in deciding what should, and what should not, be tagged.

¹⁴ Saramago, epigram: “Unless you attain the truth, / you will not be able to amend it. / But if you do not amend it, / you will not attain it. Meanwhile, / do not resign yourself. From *The Book of Exhortations*”.

¹⁵ It should be noted, however, that the numbering system cannot apply to the *versión crítica* in the same way, as the structure of the redaction is significantly different for most of its extension.

¹⁶ xml tags, compliant with the norms of TEI5 are employed to encode the transcriptions. The *Estoria Digital* project used the <div></div> tags to mark chapter

level divisions; thus any text falling between (say) `<div n="50">` and the next `</div>` tag is understood to correspond to Chapter 50. The next level down is the level of sentence, which is marked with the `<ab></ab>` tag.

¹⁷ Details of the development of CollateX can be found at:

<https://www.huygens.knaw.nl/collate-x/?lang=en>

¹⁸ Smith will deal with the question of development in a forthcoming paper. For the software itself, see https://github.com/itsee-birmingham/collation_editor

¹⁹ Details of the original software: <http://anastasia.sourceforge.net>

²⁰ For further details of the management of the editing process see Polly Duxfield's article in the present volume.

²¹ The relationship between transcription and collation is most clearly outlined by Parker: "The fact that Collate works by the making of transcriptions fundamentally affects the way in which editor and user view the edition. In the first place, the user has a check on the way in which the materials are presented in an apparatus. In the second, the editor's attention is shifted away from textual variation as a series of short differences, to textual variation as multiple copies. In the third, the manuscript as an artefact comes back into its own. We have adopted tagging procedures (building on what had already been done with Collate and working in partnership with the INTF) which allow a digital reconstruction of the lay-out of a manuscript. In the days of collating, one tended to focus only on the clearest way of presenting variations from the collating base. In transcribing, the textual variation is seen within the framework of recreating the scribe's procedure of copying the manuscript." http://www.tei-c.org/About/Archive_new/ETE/Preview/parker.xml

²² Spence, "Siete retos", "...dos procesos distintos en el ciclo de la edición: la *preparación* de una edición y su *presentación* posterior, que antes debían ir fundidas por las limitaciones del papel impreso, pero ahora pueden separarse." (156)

²³ Priority in this sense was given to whichever manuscript was closest to the *versión primitiva*. In practice, when E1/E2 is the base text, this meant that variants were preferred in the following order: E₁/E₂ – T – Q – Ss.

²⁴ This phenomenon is especially notable when the base text changes from the 13th century E₁ and E₂ to the fourteenth century T (and briefly Y). Such differences are regularized out in the reader's text, but not in the edited text which is the product of the collation output.

²⁵ Details to go here, especially link to github. (not yet available, April 2017)