Visualization of Variants in Textual Collations to Analyze the Evolution of Literary Works in The Cervantes Project

Carlos Monroy, Rajiv Kochumman, Richard Furuta, Eduardo Urbina,
Eréndira Melgoza, and Arpita Goenka

TEES Center for the Study of Digital Libraries
Texas A&M University
College Station, TX 77843-3112, USA
{cmonroy, rajiv, furuta, e-urbina, ere, arpita}@csdl.tamu.edu

Abstract. As part of the Cervantes Project digital library, we are developing an Electronic Variorum Edition (EVE) of *Don Quixote de la Mancha*. Multiple editors can create an EVE with our Multi Variant Editor for Documents (MVED), which allows collation of one base text against several comparison texts to identify, link and edit all existing variants among them. In this context we are investigating the use of visualizations to depict graphically variants in order to validate the accuracy of the textual transcriptions and to understand the similarities and differences among different printings and editions. Our broader goal is to enable users to analyze the collation's results and to discover facts about the evolution of the Quixote textual history, and to provide evidence to eliminate printing and compositor's errors and thus to produce a more correct edition closer to Cervantes' original manuscript. This paper describes the visualization tool, and presents the initial results of its use.

1 Introduction

The Cervantes Project was initiated in 1995 and is hosted at the Center for the Study of Digital Libraries at Texas A&M University. As stated in [9], its main goal is to provide "a comprehensive on-line reference and research site on the life and works of Miguel de Cervantes (1547-1616), the author of the classic *Don Quixote de la Mancha*."

The main components of the project are: the Cervantes Digital Library (CDL), which provides an electronic repository of novels, plays, and other writings in different formats and versions; the Cervantes Digital Archive of Images (CDAI), an archive of photographic images documenting Cervantes' life and time; the Cervantes International Bibliography Online (CIBO), which contains a cumulative annotated bibliography of studies, editions, and translations of Cervantes' works.

The cornerstone of literary studies is the availability of critical editions that provide scholarly annotated texts. A key aspect of this process for Cervantes scholars is the access to documentary texts accurately transcribed as well as to the early printings of his works. To this end, one of the activities of our project has been to obtain multiple microfilmed copies of all the textually significant editions of Cervantes' best-known work, *Don Quixote*. Initially, the novel was published in two parts. The first edition (the *princeps*) of part one was published in Madrid in 1605. The *princeps* of part two dates from 1615, again published in Madrid. We have already obtained over thirty copies of the nine key early editions—indeed we believe that our collection is the most comprehensive ever put together—and our acquisition efforts are continuing. For example, only approximately 18 copies of the 1605 *princeps* are believe to have survived, and of these only 12 are currently accessible to scholars for digitizing and electronic editing. We have acquired already microfilms of 8 copies of the *princeps* of both parts. The work reported in this paper uses 6 of the first part *princeps* copies; the 2 remaining were acquired more recently and are still being incorporated. Additionally we report here on work using the first part princeps and copies of five additional of the significant early editions published between 1605 and 1637.

We seek to make this corpus easily-available to the Cervantes scholar. Until now, even the best of editions are based on limited knowledge of and access to textual sources, and even more limited forms of collation of such sources. Furthermore, traditionally, access to source materials has been through "filtered" sources, such as critical editions, often offering only a modernized eclectic text, and thus reflecting conjectural choices and preferences of the editors. Financial considerations associated with print have restricted the ability to include original source material. Given the current digital environment, we seek to create an *Electronic Variorum Edition* (EVE) of the *Don Quixote*, providing both the complete scholarly apparatus associated with a critical edition, but also access to the original sources in facsimile and transcribed formats. Towards that end, we have implemented tools for building, accessing, and viewing EVEs. One of the central tools, the MVED, described earlier in [8], provides the means for editing and creating an EVE.

The collation module of the MVED (Multi Variant Editor for Documents) enables an editor to collate a base text against a set of comparison texts. The output of this process is a hyperlinked set of variants, which in turn can be edited and/or annotated to provide further information and construct the critical apparatus   To assist the editor in this process, we sought to find a method to highlight anomalies and discover textual patterns in the collation as basis for investigating the nature of the variants as authorial, compositorial or printing errors, and as points of departure to establish textual correlations and editorial filiation. Furthermore, we recognize a practical need in our project to validate the accuracy of the transcriptions of the texts, since those are created manually from an electronic template.

Previously, we have described briefly [11] our use of visualizations to depict the variants, and its utility in presenting anomalies that require further investigation. In this paper, we report further on our application of visualizations in detecting patterns of interest within collections of copies or editions of a text. Our earlier visualization

tool presented its information as a static display. Beginning with the representations developed for static display, we have extended the visualizations to permit interactive exploration of the relationships depicted. In implementation, we have achieved this purpose by representing the information space in ItLv, a tool originally intended for presentation of timeline-based information spaces.

ItLv (Interactive TimeLine Viewer) [13] is based on earlier work in our Center by Kumar [12] regarding the use of interactive timeline viewers in representing relationships among items in an information space. Kumar investigated manipulations of the timeline representation and presentation to allow the flexible display of the attributes of a collection's information, as well as the synthesis of new attributes. Kumar noted that any attribute with ordered values could be mapped to the x or y axis of the display, and it is this observation that we use as the basis for our current visualization. ItLv represents a re-implementation and expansion of Kumar's prototype, allowing stable use of larger datasets. ItLv incorporates features and options such as searching, filtering, splitting, superimposing, and highlighting, to enable users to manipulate the results of the collation flexibly. As we will see in later sections, in our visualization we map copies/editions to the y axis and offset from the beginning of the text to the x axis. Initially, one could think that the editons should be placed on the x axis since they evolved over time; however, we extended the metaphor of the timeline in such a way that the "events" that take place in this space are the variants among the texts. A variant is positioned according to its starting point (i.e., offset from the beginning). Attributes of the variance (e.g., its length) are reflected in the visual appearance of its corresponding marker.

We are currently using the ItLv visualizations to accomplish two goals; 1) to refine the results of the collation, identifying false variants or errors in the textual transcriptions, and 2) using the refined collations to analyze the results and make targeted inferences, validating or invalidating hypotheses, and advancing interpretions about the similarities and differences among the texts. This process can lead to further discoveries such as patterns across the texts, which in turn could serve as indicators of the presence of several compositors working on different sections of the text and their individual tendencies regarding the introduction of textual variants. Therefore, the visualization provides a higher level of abstraction of the relationships among different copies of the text. In this paper, we describe the interactive options of ItLv and report the results of its use in refining the results of textual collation, as well in discovering facts about the nature of the texts themselves, and the relationships among them, to carry forward with a high degree of authority and confidence the needed editorial emendations.

2 Digital Libraries and Visualization Tools

Presently there are several initiatives and projects dedicated to the creation and use of digital libraries about ancient manuscripts, books, maps, and a variety of artifacts. The idea is not only to provide an online repository, but also to help scholars to perform analyses about the relationships among the components of the collections in the hope of understanding facts, data, and information in their historical and cultural contexts, and then to make those discoveries available to a wider audience. The Perseus Digital Library is one such example. Hosted at Tufts University, the Perseus Project [4] conceived as an evolving digital library in the humanities, was initially focused on the classical Greek world. However, it has evolved to include other topical areas in the humanities.

More focused on the analysis of textual history and variants, the Canterbury Tales Project [18] provides a computer-based repository of the transcriptions of all manuscripts and early editions of the *Canterbury Tales*. The manuscripts are then compared to generate a collection of agreements and disagreements, a process called collation. The results of the collation are used as input for computerized tools that help in reconstructing the history of the text. Finally, the materials and results of the analyses are published electronically. However, due to the fact that every word is categorized grammatically, the Canterbury Tales Project goes beyond being a collation-only project, making it interesting for areas such as; history of the language, dialects, and orthographic and morphological change.

In the process of analyzing patterns, similarities, and differences in large collection of texts, scholars often face challenges in visualizing different abstractions and datasets of those collections. Therefore, visualization tools have an important role as means to reduce the complexity of such analyses. The Compus visualization system [6] is an example of a visualization tool that depicts information about the structure of ancient textual manuscripts. Compus enables users to analyze the structure of a collection of texts by depicting TEI-enabled texts components relative to their position in the text, the tool also supports dynamic queries and structural transformations.

Visualizing elements in textual transcriptions can be accomplished in different ways based on the nature of the data and the needs of the users. LifeLines [14] illustrate the use of a visualization tool for personal histories applied to medical and court records, professional histories and other types of biographical records. In fact, regarding the benefits of the tool, they conclude, "[LifeLines] provide a complete visualization environment offering overview, zooming, filtering and details on demand. Color coding, filtering and dynamic highlighting unveiled relationships between events that may otherwise be difficult to see in paper records."

Finally, the power of visualizations using point representations in two-dimensional spaces for representing information collections and their relationships were pioneered in StarFields [1]. Ahlberg and Shneiderman focus particularly on the relevance of these displays in visual information seeking; discovery of the important attributes of large information spaces through browsing.

3 Visualizations in the Cervantes Project
3.1 Visualizations for validating the results of the collation

Once the collation has generated the index of variants and prior to beginning the analysis of patterns and differences among them, it is necessary to validate the results of the collation process. The traditional approach involves the manual re-collation of the transcriptions against the digitized images in order to determine which variants are real and which are not. Obviously this constitutes a very cumbersome and time consuming process, especially for a collation with a high number of variants. Another approach is to use a graphical representation of the results of the collation, and to focus on patterns that seem unexpected or anomalous.

As an example, figure 1 shows the result of collating six copies of the *Quixote* 1605 *princeps*. One of the copies of the *princeps* is defined to be the "base", and is not represented directly in the display. The rows of the display show the result of comparing the "base" to each of the copies, in turn. The display shows the first chapter which contains four pages, delimited by the vertical bars in the visualization. Ideally, one would expect that all six copies would be the same and that any differences would flag errors. As it turns out, this is not the case for this edition. Following Knowles' early textual studies and partial collations of several copies of the *princeps*, Casasayas [3] believes that of the approximately 18 known copies, no two are completely alike. Some differences are likely due to changes made during the print run, but some may result from deliberate attempts at deception in later times [10]. Our examination of the visualization pointed out examples of transcription errors (due to unclear patches in the original microfilms), one "unexpected" action taken by the collation algorithm (represented by the unusually long length of one of the variances), and differences of the form identified by others previously (including those in the first two pages). See [11] for further explanation of this display.

Fig. 1 Depicting variants in six copies of chapter one of
the princeps

Our static visualization only depicted the results in the screen, but did not provide further information such as the content of the variant, its length in characters, and the offset in the text; nor does allow interaction. Switching to the ItLv allowed us to provide a more interactive visualization. In figure 2, variants between five editions of *Don Quixote* and the *princeps*, which is the base text in this case, are shown. The five editions are depicted in the Y-axis, the offset in the text is depicted in the X-axis, and the vertical bars represent page divisions in the *princeps*, as in the previous example. As would be expected, many more variances have been identified in this collection, since the different editions were printed in separate locations between 1605 and 1637 (more detail will be given later). In this display, the mouse cursor has been positioned over one of the variances. A pop-up window is displayed which contains information about the variant, such as its content, its offset in the text, its length, and the image of original text. This is connected to the axes by blue lines, one horizontal and the other vertical, which show the context of a variant. All variants with some similarities are displayed in a different color, in this case, are those containing the same string, possibly as a substring. In this particular case the variant is between the word "que" and its abbreviation q̃ , which we encode in our transcriptions as " <q>".

Fig. 2 The timeline viewer depicting variants
among six editions of *Don Quixote*

3.2 Visualizing the results of collating several copies of one edition

Figure 3 shows the results of collating the first five chapters of the six copies of the Quixote 1605 *princeps*. In this figure, as before, the Y-axis from bottom to top corresponds to copies one to five respectively. The X-axis from left to right corresponds to the offset in the text. Unlike the earlier example, the vertical lines here depict the beginning of each chapter.

As before, a relatively small number of variants have been identified; this is not a surprise since copies of the same edition are presumably identical. Three uncommon patterns stand out. The first is a high number of variants in chapter one of copy five– these are the rectangles depicted on the top leftmost part of figure 3. Second are a high number of variants in chapter five of copy two– these are the rectangles depicted at the bottom rightmost part of figure 3. And third is a variant containing 49 characters in chapter three of copy one, the large rectangle depicted at the center bottom in figure 3.

After reviewing the original images and the textual transcriptions we determined that the first pattern represented true variants such as capitalization, missing punctuations, abbreviations, and accentual differences. The second pattern in chapter five of the second copy was a problem in the conversion of the text because the text editor we used did not encode accented letters or special characters correctly. The third pattern was a miscalculation due to transcription of page numbering and did not represent a true variance. We noted that variances two and three required correction to the textual transcriptions and requested that the transcribers make the corrections. We then returned to our analysis of the text.
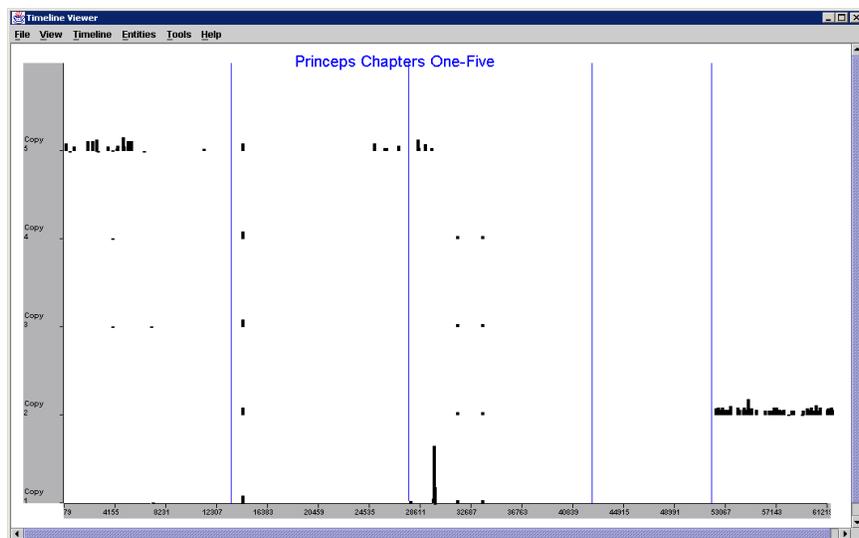
Fig. 3 Results of the collation of chapters one through five of five
copies of Don Quixote's 1605 princeps before filtering false variants
and transcription errors

Having determined that the variances in groups two and three were not significant to our analysis, we specified that ItLv should suppress their display. Figure 4 depicts the results of the collation after removing those variants. Note that the cluster of false variants in chapter five of copy two and the long variant in chapter three of copy one, have been removed. Further analysis could then be carried out, suggesting, for

example, that copy 5 was "aberrant" in some fashion, and consequently would not be a good choice for base text when we subsequently compared multiple editions.
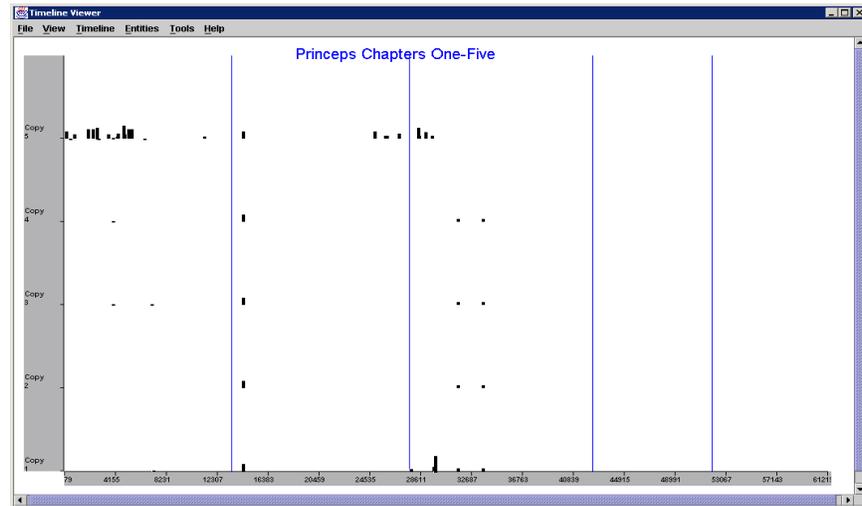
Fig. 4 Results of the collation of chapters one through five of five copies of Don Quixote's edition of Madrid 1605 called the Princeps after filtering false variants and transcription errors

In the analysis of copies of the same edition, the visualization tool can be used to detect "states" produced, as defined by Carter, [2] "When alterations, corrections, additions or excisions are effected in a book during the manufacture, so that copies exhibiting variations go on sale. These variants copies are conveniently classified as belonging to different 'states' of the edition." We believe this kind of analysis could be extremely useful for determining the existence of states of a given edition with a high degree of authority and confidence, particularly when more copies become available in the future.

3.3 Visualizations for detecting patterns

After false variants are detected and removed, the resulting dataset constitutes a more reliable source for textual analysis. For instance, regarding the evolution of a literary work, it is crucial to analyze the results of the variants to discover transmission and composition patterns that may show evidence regarding what happened to a given text and when. Moreover, a more ambitious goal would be to recreate the original manuscript or at least an idealized, as intended by Cervantes. This corresponds to a stated goal of the Canterbury Tales project [18]: "Our first aim is as old as textual scholarship: we want to find out, as nearly as we can, what Chaucer actually wrote."

The second case analyzed compares the results of collating chapters one and two of six editions of the first part of *Don Quixote*. Figures 5 and 6 depict the results of the

collation; the Y-axis from bottom to top corresponds to five of the editions in the following order, Valencia 1605, Madrid 1605 (Second Edition), Brussels 1607, Madrid 1608, and Madrid 1637. The sixth edition is the *princeps*, Madrid 1605, which was used as a base text for the collation. Figures 5a and 6a depict the original results of the collation for chapters one and two respectively.
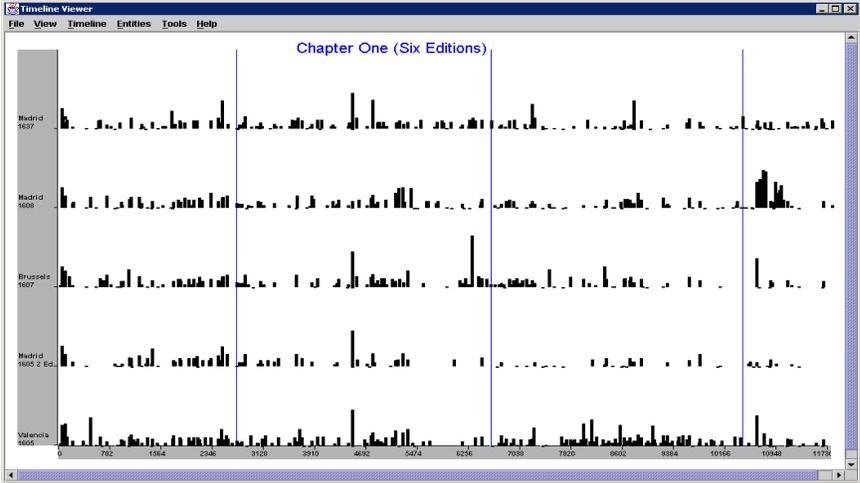

Fig. 5a Results of the collation of chapter one
before applying the filter
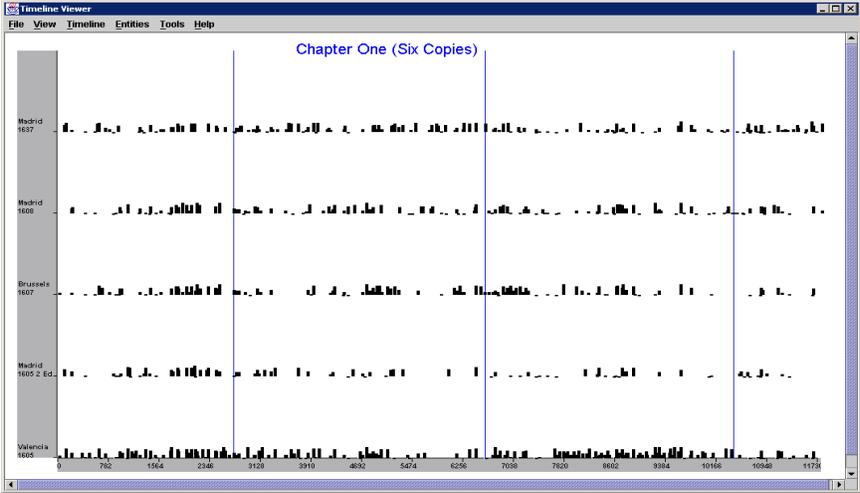
Fig. 5b Results of the collation of chapter one
after applying the filter

After reviewing the results, we found that several of the variants of length greater than ten characters were due to errors in the collation process rather than true variants. Therefore, these false variants need to be removed, otherwise the user can make false conclusions. Removing false variants as a previous step in the analysis of the collation can be done in two different ways. The first way is by modifying the textual transcriptions, rerunning the collation process, and finally displaying the results in the timeline viewer. A better and more efficient approach is having a mechanism to remove false variants directly in the visualization tool without generating the collation again; this is accomplished with the ItLv filtering option. Figures 5b and 6b depict the result of the collation for chapters one and two respectively after applying the filtering option.
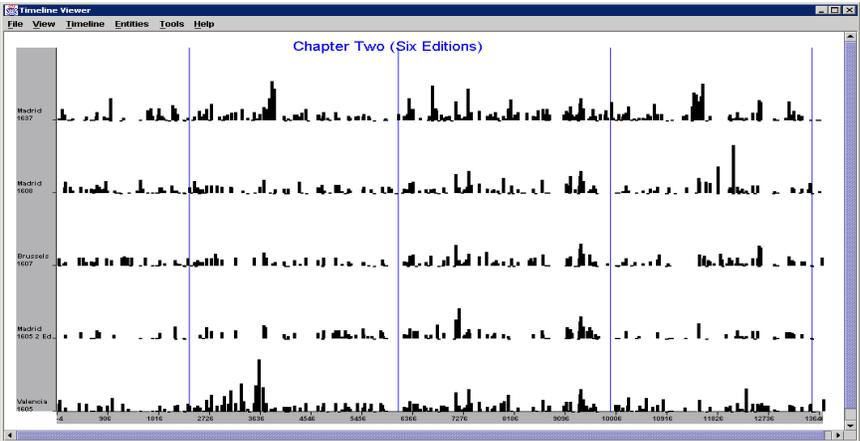
Fig. 6a Results of the collation of chapter two before applying the filter
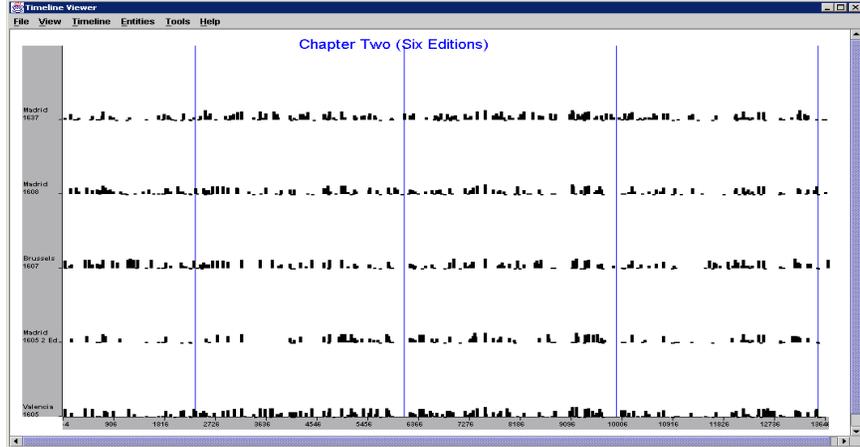
Fig. 6b Results of the collation of chapter two after applying the filter

The filter option is constrained-based; therefore the user is presented with all the attributes of the variants, as well with the following conditions: a) equal, b) not equal, c) less than, d) greater than, e) less or equal to, f) greater or equal to, and g) containing. The last condition enables users to filter the dataset based on a substring, e.g. the user wants to filter those variants containing the string "que". Once the filter has been applied, only those variants that satisfy the condition will be depicted. The filtering option can be applied again to this subset of variants, using either the same or a different attribute and condition. And this process can be applied again and again. The filter can also be removed, this will restore the original dataset.

The filtering process do not remove false variants from the data source, they are only removed from the screen. However, if the user wants to create a clean data source, he or she can use the export option. In the export option the user applies a condition to the dataset (similarly as in the filtering option) and then only those variants that satisfy that condition are exported to a new file.

We started by analyzing variants that include abbreviations, since the presence or absence of abbreviations was the result of compositors' decisions during page setting of the pages and not necessarily present in Cervantes's manuscript. Figure 7 depicts the use of one particular abbreviation in the princeps edition—that of the common Spanish word "que". On the left, the string *que* has been replaced by the abbreviation q̃, which we encode in our transcription as "<q>". The second instance represents the word "aquel". On the right, two instances of the string *que* are shown, one corresponds to a single string and the other one to a string inside another string, in this case the hyphenated word "aunque". The top windows of figures 8 and 9 simply show all occurrences of the string "que" in the base text, emphasizing the frequency with which the word is found in chapters one and two respectively.
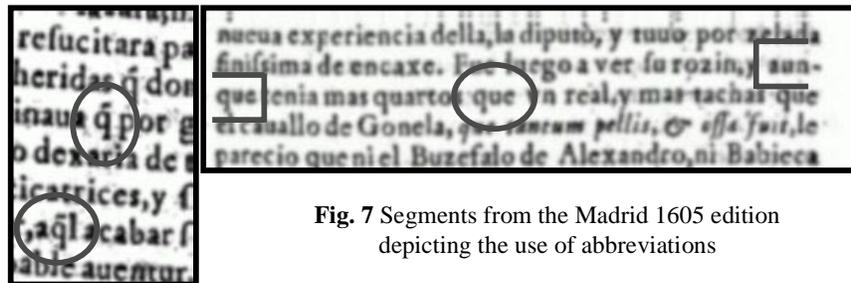
Fig. 7 Segments from the Madrid 1605 edition
depicting the use of abbreviations

Because q̃ was inserted during page composition, many of the identified variants differ only in its inclusion or its representation as "que". Consequently, we filtered the results of the collation removing those variants that did not include the abbreviation q̃ . This is shown in the lower windows of figures 8 and 9.

For chapter one, with the exception of two variants at the end of page four and none in page three, we found that most variants appear from the second half of page one up to the end of page two. Initially we thought that those variants would be replacing all

the "que" strings in the base text, therefore, we depicted all the ocurrences of the string "que" in the base text (window at the top of figure 8). But since the string "que" appears regularly across the whole chapter, we can therefore raise the hypothesis that the pattern of the abbreviation $\tilde{q}$ as a compositorial variant could reflect the preference of an individual compositor in charge of those sections of the book in the Madrid 1605 edition. Significantly, that abbreviation was passed on to the next editions as we can see a similar pattern from bottom to top (window at the bottom of figure 8).

For chapter two, variants were more scattered. However, with the exception of two variants in page one and one more in the second half of page two, most of the variants are concentrated on the first half of page three and across page four (window at the bottom of figure 9). Ocurrences of the string "que" in the base text are depicted in the window at the top of figure 9. In this case, it is more difficult to conclude about the intervention of compositors in certain parts of the text.
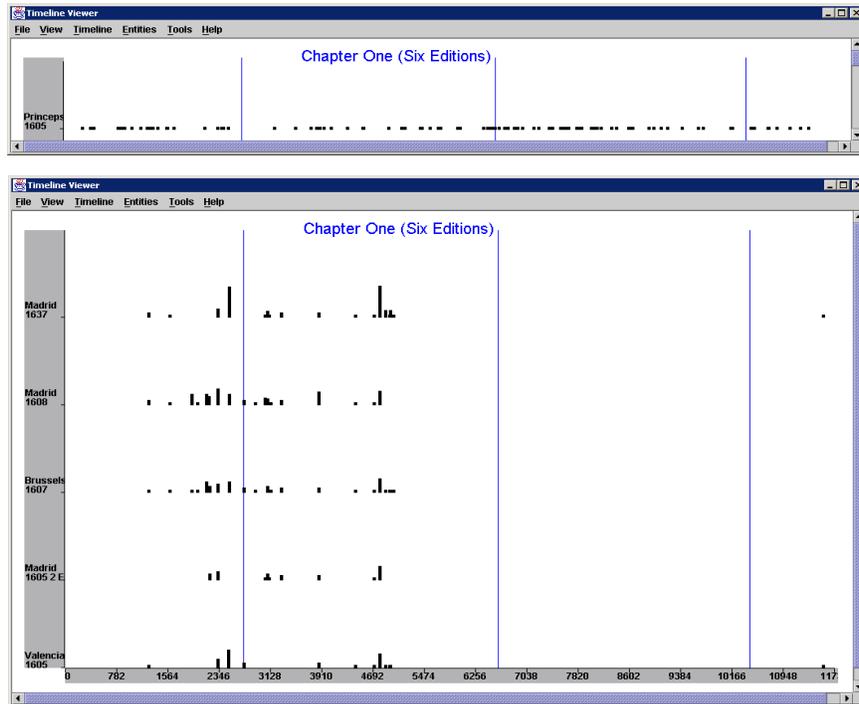
Fig. 8 The abbreviation $\tilde{q}$ in chapter one across six editions of Don Quixote

Fig. 9 The abbreviation $\tilde{q}$ in chapter two across six editions of Don Quixote

At this point it is worth pointing out that the six textually significant editions of the 1605 *Quixote* included in the EVE were published in three different cities during a 32 year span of time. Through the inclusion of at least three copies of each edition, multiple collations, and the analysis of variants afforded by the visualization tool we hope to be able, among other matters, to identify precisely patterns of transmission and textual filiation among them. In particular, we intend to investigate current theories regarding the presence of patterns of variants that can be attributed to the compositors [16], and those supposedly that resulted from Cervantes' own intervention correcting the leaves during the printing process of the 2nd corrected edition of 1605, and possibly the 3rd edition of 1608, all produced at the Juan de la Cuesta printing house in Madrid [15,16]. Based on partial collations of the texts, it is believed that all the editions under consideration were based on the Madrid 1605 2nd edition [17]. We intend to use the MVED collation module and the visualization tool to test and evaluate this and other similar hypothesis. For example, it appears from the limited context visualized in figures 8 and 9 that there are relatively few changes between the Madrid 1605 *princeps* and the Madrid 1605 second edition (the next to bottom row in the lower windows). This is consistent with the knowledge that both editions were produced at the same printing house. In contrast, more differences are shown between the *princeps* and the 1605 Valencia edition (the top row in the figures). Continuing this analysis across larger segments of the corpus and with

selection of different editions to serve as base text for collations may provide further evidence about the relationships.

4 Conclusions and future work

ItLv more than an improvement with respect to state of the art visualization systems, is a tool that exploits visualization to allow the exploration of variants among different versions of texts. ItLv takes the concepts of previous visualization systems and tailors it to enable the exploration of similarities and differences among several texts, by means of its flexibility and functionality.

Despite the limited number of existing copies of the *princeps* available for digitization and textual analysis, the sample editions and chapters included in the EVE at this time in which we have tested the ItLv visualization tool shows its usefulness for validating the results of textual collation and the reliability of the transcriptions, and its functionality as a heuristic device to help to understand the relationships, similarities, and differences among different editions and/or copies of the same edition of the *Quixote*.

Nevertheless, our testing also indicates the need to further refine and expand its capabilities. For instance, in the current implementation, there is no provision to keep track of the filters that have been applied to the dataset in any given session. We think that a useful feature would be for ItLv to record the filter conditions as rules, allowing users to recreate that session and make comparisons between the application of a series or combination of filters.

As of now the collation module of the MVED generates a set of linked variants; the process of categorizing those variants is done separately during the editing process by the MVED. We believe that the interactive features of ItLv will increase the editor's ability to analyze differences and similarities among the texts, and that this will be further enhanced by including the variance categorizations identified during editing (e.g., authorial, compositorial or typographical errors) into the set of attributes available for visualization.

By graphically encoding variant attributes such as puntuation, abbreviation, typographical and spelling errors, as well as lexical changes; ItLv can support ortographical, morphological, typographical, and lexical analysis of the texts. This can be accomplished by using the filtering option, since each of the categories of variants can be isolated and visualized separately.

Edition attributes such as formal divisions in parts or chapters, presence of extratextual materials, annotations, and illustratations, can also be encoded as a separate data set. Superimposing the edition attributes on the variants dataset, we expect that will expand scholar's perspective and understanding of context in variant analysis.

Finally, we believe that ItLv can be used to better understand the textual history of a literary works such as the *Quixote* by identifying, connecting and visualizing relationships among the different components of an edition such as illustrations, extra textual materials, variants and notes, along a period of time. Therefore, it will contribute significantly in the preparation of a critical variorum edition tracing the evolution of the contributions of a particular editor or commentator, as well as in the comparative study of intertextual borrowings and influences.

5 Acknowledgements

6 References

[1] Ahlberg, C., and Shneiderman, B., "Visual information seeking: tight coupling of dynamic query filters with Starfield displays", Conference proceedings on Human factors in computing systems: celebrating interdependence, Boston, MA, 1994, pp. 313-317.

[2] Carter, J., "ABC for Book Collectors." 7th edition. Oak Knoll Press, 1995.

[3] Casasayas, J. M., "Ensayo de una guía de bibliografía cervantina," Tomo V. Mallorca, 1995.

[4] Crane, G., "The Perseus Project and beyond: How building a digital library challenges the humanities and technology", *D-Lib Magazine*, January 1998.

[5] Crane, G., Smith, D., Wulfman, C., "Building a Hypertextual Digital Library in the Humanities: A Case Study on London.". Proceedings of the first ACM/IEEE-CS joint Conference on Digital Libraries, Pittsburgh, Pennsylvania, May 1998, pp. 426-434.

[6] Fekete, J., Dufournaud, N., "Compus: Visualization and Analysis of Structured Documents for Understanding Social Life in the 16th Century", Proceedings of the fifth ACM Conference on Digital Libraries, San Antonio, Texas, 2000.

[7] Flores, R. M., "The Compositors of the First and Second Madrid Editions of Don Quixote Part I." London: Modern Humanities Research Association, 1975.

[8] Furuta, R., Kalasapur, S., Kochumman, R., Urbina, E., Vivancos-Pérez, R., "The Cervantes Project: Steps to a Customizable and Interlinked On-line Electronic Variorum Edition Supporting Scholarship", Research and Advanced Technology

for Digital Libraries: 5th European Conference, ECDL 2001, Darmstadt, Germany, September 2001, pp. 71-82.

[9] Furuta, R., Hu, S., Kalasapur, S., Kochumman, R, Urbina, E., Vivancos-Pérez, R., "Towards an Electronic Variorum Edition of Don Quixote", Proceedings of the first ACM/IEEE-CS joint conference on Digital Libraries, Roanoke, Virginia, 2001, pp. 444-445.

[10] Knowles, E. B., Jr., "Notes on the Madrid, 1605, editions of Don Quijote." *Hispanic Review* 14 (1946): 47-58.

[11] Kochumman, R., Monroy, C., Furuta, R., Goenka, A., Urbina, E., and Melgoza, E., "Towards an Electronic Variorum Edition of Cervantes' Don Quixote: Visualizations that support preparation.", to appear in the ACM Joint Conference on Digital Libraries, Portland Oregon, July 2002.

[12] Kumar, V., Furuta, R., Allen, R., "Metadata Visualization for Digital Libraries: Interactive Timeline Editing and Review", in Proceedings of the third ACM conference on Digital Libraries, Pittsburgh, Pennsylvania, May 1998, pp. 126-123.

[13] Monroy, C., "Augmenting Cognition through Information Visualization Using a Timeline Viewer," Masters of Computer Science Project Report, Department of Computer Science, Texas A&M University, 2002.

[14] Plaisant, C., Milash, B., Rose, A., Widoff, S., Shneiderman, B. "LifeLines: visualizing personal histories", in Proceedings of CHI'96, Vancouver, BC, Canada, April 14-18, 1996, pp. 221-227.

[15] Rico, F., "Historia del texto." *Don Quijote de la Mancha.* Barcelona: Instituto Cervantes-Crítica, 1998.

[16] Rico, F., "El 'original' del *Quijote*: del borrador a la imprenta," *Quimera* 173 (Oct 1998): 8-11.

[17] Urbina, E., Furuta, R., Goenka, A., Kochumman, R. , Melgoza, E., Monroy, C. **"**Texto, contextos e hipertexto: la crítica textual en la era digital y la *Edición electrónica variorum del* Quijote-*IV Centenario*." To appear in *Quaderni di Letterature Iberiche e Iberoameicane*, Issue #27, Dipartimento di Scienze del Linguaggio e delle Letterature Comparate dell'Università degli Studi di Milano, 2002.

[18] "The Canterbury Tales Project", De Montfort University, Leicester, England. http://www.cta.dmu.ac.uk/projects/ctp/index.html. Accessed on May 25, 2002.