The Poetess Archive at Miami University includes a database of electronic documents encoded using the TEI (Text Encoding Initiative) schema extended with Poetess Archive tag-set, and derived from widely used terms in literary analysis and criticism. The extended TEI schemas based on the Xml standards hold elements of interest to a literary scholar, and are spread across multiple encodings. The Xml representation of literary texts is suitable for machine processing and electronic exchange of information, but does little to promote adoption and intuitive use of these resources by scholars. We are developing a visualization tool that seeks to integrate multiple encodings while allowing comparative analysis of multiple poems encoded using the extended TEI tag-set. The proposed solution is an interactive visual representation of differently encoded versions of text that can enhance cognition, and aid in uncovering of new knowledge. This approach will facilitate identification of frequently changing hotspots in encoded text and aid in the process of close reading.
A Thesis

Submitted to the
Faculty of Miami University
in partial fulfillment of
the requirements for the degree of
Master of Computer Science

Department of Computer Science and Software Engineering
by
Manish Chaturvedi
Miami University
Oxford, Ohio
2011

Advisor
Dr. Gerald Gannod

Committee Member
Dr. Laura Mandell

Committee Member
Dr. Alton Sanders
Contents

1 Introduction 1
  1.1 Contributions ................................................. 4
  1.2 Organization .................................................. 4

2 Background and Related Work 6
  2.1 Background ..................................................... 6
    2.1.1 The Text Encoding Initiative ............................ 7
    2.1.2 TEI Encoding and Prosody in English Poetry .......... 7
    2.1.3 Information Visualization and Cognition ................ 8
  2.2 Related Work .................................................. 9
    2.2.1 Information Visualization Tools ......................... 9
    2.2.2 Content-Based Visualization Tools ....................... 10
    2.2.3 Visual Differencing Tools ................................ 12
    2.2.4 Stylometric Tools and Authorship Analysis ............ 13

3 Approach 14
  3.1 Text Visualization ............................................. 16
  3.2 Content-Based Visualization ................................. 19
  3.3 Visualization Pipeline ....................................... 21
  3.4 The Visual Thinking Process .................................. 22
    3.4.1 Depth Perception and 3D ................................. 23
# Appendix

A User Guide 67

B User Study Handout 88

C Myopia Tool Installation 107
List of Tables

3.1 Pipeline and Poetry Visualization Event Mapping . . . . . . . . . . . . 22
3.2 Myopia Tool - Bundled Categorized Encoded Poems . . . . . . . . . . . 35
3.3 Visualization Pipeline: Implementation . . . . . . . . . . . . . . . . 36

4.1 User Study Research Questions . . . . . . . . . . . . . . . . . . . . . 38
4.2 Pretest Questions and Purpose . . . . . . . . . . . . . . . . . . . . . 39
4.3 Pretest Questionnaire: Familiarity with TEI . . . . . . . . . . . . . . 44
4.4 Pretest Section: TEI Visualization Activities . . . . . . . . . . . . . . 45
4.5 Pretest Section: Value Of TEI Encoding . . . . . . . . . . . . . . . . 46
4.6 Participant Task Performance . . . . . . . . . . . . . . . . . . . . . . 47
4.7 Posttest Section: Usefulness of Myopia tool in Textual Analysis . . . 48
4.8 Participant Task Experience . . . . . . . . . . . . . . . . . . . . . . . 48
4.9 Expectations and Perceptions: Pre and Post-Test Statements . . . . . 51
4.10 Pretest Section: Expectation from a Visualization Tool . . . . . . . . 52
4.11 PostTest Section: Perception of the Myopia Tool . . . . . . . . . . . . 54
4.12 PostTest Section: Usability Of Myopia Tool . . . . . . . . . . . . . . 56
4.13 PostTest Section: Experience With the User Study . . . . . . . . . . 57
List of Figures

2.1 Visualizing n-grams in Homeric Poetry ................. 13

3.1 Myopia Tool - Text Overlayed With Graphics .............. 15
3.2 Myopia Tool - Visual Comparison Mode .................. 16
3.3 Sound Densities In the Odyssey [20] ....................... 17
3.4 Myopia Tool - Context Specific Tool Tips ................ 19
3.5 Myopia Tool - Content Based Features ................... 20
3.6 The Visualization Pipeline .............................. 20
3.7 Myopia Tool - Information Foraging Options .............. 23
3.8 Myopia Tool - Structuring 2.5 Dimensional Space ......... 24
3.9 Myopia Tool - GUI widgets with wxPython ................ 25
3.10 Myopia Tool - Loading Bundled Encodings ............... 26
3.11 Myopia Tool - Active Encoding and Graphical Elements .... 28
3.12 Myopia Tool - Navigational Controls ................... 29
3.13 TEI Encoding - Poem with Meter Encoding ............... 29
3.14 Myopia Tool - Hotspot Abstract View ................... 30
3.15 Visualization Mapping - Graphics With Text .............. 32
3.16 Visualization Mapping - Graphics Without Text ........... 33

4.1 Pre and Post Test Expectations and Perceptions .......... 55
To my parents.
I would like to thank my advisor, Dr. Gerald Gannod, for making this thesis possible through his guidance and dedication. I am grateful for his help and direction during the duration of my studies for the Master’s Degree and throughout this thesis.

I am very thankful to Dr. Laura Mandell for being the primary driver for this work and her enthusiastic and timely help from inception to completion of this thesis. I also acknowledge Dr. Mandell for providing the encodings utilized in the Myopia tool. I owe much gratitude to Professor Helen Armstrong who has been instrumental with her graphical design expertise that forms the basis for all the visualizations developed as part of the thesis. I would like to thank Dr. Eric Hodgson for his guidance and for providing resources of the AIMS visualization labs which has played an invaluable part in development of the poetry visualization tool. Thank you to Dr. Alton Sanders who I first knew as the director of the graduate program in the CSA department for his help and guidance over the years.

Thank you to Dr. Mandell, and Dr. Sanders for serving on my committee.

Thank you to the faculty of the Computer Science and Software Engineering department for the contributions to my education and research.

Finally, thank you to my children, Avishi and Arya, for giving me so much joy and love over the years. Most of all, thank you to my wife, Tapasya, for loving and supporting me always.
Chapter 1

Introduction

The goal of critical analysis of poetry or any other literary text is to develop an appreciation and a deeper understanding of such work. The reader employs rhetorical devices to derive meaning from the physical text and images contained in her object of study [16]. Such meaning may be plainly deduced from the words themselves. More often however, literary critics seek to uncover hidden connotations. These connotations may manifest themselves through structure of the narrative, the organization of the poem, through sound of the language, and the metaphors employed. Close Reading involves focus on the connotation of the text, not simply what the words literally mean [4]. The critic looks for textual hotspots, where connotations converge in the text. The hotspots might point to a unifying idea behind the poem, and help in uncovering new ideas about the work. Such hotspots are rich in meaning, have complicated undertones, and thus are discovered in a slow and deliberate process.

The purpose of Information Visualization is to develop a model or image and apply it to abstract, often nominal data and relations to gain insight into said data. Tremendous amounts of data can be analyzed with appropriate visualization tools. This is in contrast to scientific visualization where the model operates on physical or
quantitative data, and there is a strong correlation between the dataset and applied imagery. Information visualization covers areas such as visual reasoning, visual data modeling, visual programming, visual information retrieval and browsing, visualization of program execution and visual languages [28].

Within Information Visualization, text visualizations are a function of the type of the visualized document (prose, software, etc.). The intent of the visualization is to derive an overview, analyze different versions, or answer domain specific questions. Content-based visualization is the most common and intuitive representation of textual documents. This visualization is characterized by simple 2D layout, in conjunction with mapped thumbnails, and thus leverages a context-and-focus technique [28]. Text visualizations can be effective in obtaining useful insights into large repositories of textual information in an efficient manner, as compared to direct reading or viewing of content. However, effective and efficient visualization of textual and relational information remains a challenge due to inherent abstract nature of data, non-interoperable set of attributes and need for user-friendly, easily understandable visualizations that keep in mind the endusers - who are subject matter experts, and often non-engineers.

An Information Visualization tool is characterized by visual representation of data, its suitable presentation, and facilitates user interaction. Graphics in documents and other visualization tools have been used since ancient times to supplement innate perception and most often have led to discovery of new knowledge [10]. In our current environment, digital technology has matured and presents a new medium for development of visual aids and tools. In turn, such tools have the potential for use in new research methodologies that manage cognition and expansion of knowledge [12].

The Poetess Archive at Miami University is a digitized resource for the study of poetry and literary history of late eighteenth and nineteenth century British and
American poets. The collection includes a database of electronic documents encoded using the TEI (Text Encoding Initiative) guidelines. Scholars study elements and structures in poetic works as they relate to features such as rhythm and sound, and encode their evaluations in language of TEI.

The XML representation of literary texts is suitable for machine processing and electronic exchange of information, but does little to promote adoption and intuitive use of these resources by scholars. Our method utilizes TEI encoded texts in the Poetess Archive to generate interactive visualizations in an multi-dimensional environment.

There has been recurrent efforts to develop text visualization and visual statistical tools to better understand the underlying structures of chosen resources. Some of these tool suites provide opportunities for exploration and collaboration in area of textual analysis, [11, 14, 19, 21], others are more narrowly focussed on comparison of plain or encoded text, across versions [1, 18]. Despite attempts to do so these digital tools have not resulted in significantly easing the process of close reading [6]. The side-by-side method of highlighting differences is not conducive to spot connotations and analyses, and thus does not help in fully exploiting potential of digital tools in aid of close reading of literature. XML by its very nature is hierarchical, and these hierarchies cannot overlap within a XML document. The elements and prosodic structure of poetry however do overlap - a metaphor can extend out of a metrical foot. Different features thus need to be encoded in separate XML files making a composite analysis difficult. A tool which fuses TEI encodings spread over separate files into a single visually integrated document could help overcome the limitations of encoding literary texts in XML.

We are presenting a method to support close reading of literary texts using multi-dimensional composite visualizations. An interactive visual representation of differ-
ently encoded versions of text can enhance cognition, and aid in uncovering of new knowledge. The Poetry Visualization tool allows for multidimensional representation of TEI encoded Poetry and text. Presenting an interactive visual representation of differently encoded versions of text, the tool seeks to amplify understanding and uncover new knowledge. This tool will facilitate identification of frequently changing hotspots in the encoded text and will visualize the attendant changes across various TEI documents. The goal is to help scholars analyze the original documents in new ways and explore information visualization as a rhetorical device in literary criticism.

**Thesis Statement**

*Close reading of literature is a time consuming activity. Often combining the results of close reading can be difficult. By using information Visualization, new knowledge about text can be synthesized.*

### 1.1 Contributions

The contributions of this research are in the areas of Information Visualization and Digital Humanities. The purpose of the research is to support close reading of literature using information visualization. This will facilitate abstract analysis of literary texts by fusing multiple perspectives.

### 1.2 Organization

The rest of the thesis consists of four chapters. Chapter 2 presents the background and related work of this thesis and the contributions. In Chapter 3 our approach, methodology and implementation are presented. Chapter 4 describes the evaluation of our approach and provides the results of the user study we conducted to answer
the research questions used to test our approach. Finally, Chapter 5 summarizes the thesis and provides conclusions and future work for our research.
Chapter 2

Background and Related Work

2.1 Background

Digital and online archives provide a viable outlet for libraries, museums and other repositories of literary artifacts to propagate information. Scholars now have access to tremendous resources to further explore and research disparate knowledge. The TEI guidelines are a lynchpin in transcription and digital representation of textual material. The XML based TEI encodings of texts is however not conducive to intuitive use by the subject experts and scholars who are its target audience. Text visualizations can be effective in obtaining useful insights into large repositories of textual information in an efficient manner, as compared to direct reading or viewing of encoded content. We are interested in developing an Information Visualization tool for visual representation of differently encoded poetry, developing a suitable presentation environment, and facilitating real time user interaction. The goal is to see how differently marked up versions of an encoded poem from the Archive compare to each other. We seek to identify hotspots in the poem, which always change in each of the representations. This will help scholars analyze the original documents in new
ways and explore information visualization as a heuristic device in literary criticism.

2.1.1 The Text Encoding Initiative

The TEI consortium is an international organization charged with development of guidelines for digital encoding of literary and linguistic texts. Libraries, museums, publishers, and individuals for digital representation of textual material use these guidelines. Such material spans manuscripts, research papers, historical archives, early printed books, linguistic corpora, anthologies, critical editions, ancient inscriptions, and other literary, historical, and cultural material. P5 is the latest version of TEI guidelines [27]. TEI provides guidelines to encode digital texts include digital facsimiles, which are composed of digital images of the original resource, and preparation of digital texts [15]. TEI format can help transcribe or convert existing text to character form and mark them up in XML format. These recommendations provide a starting point and can be elaborated towards transcriptional needs of scholars. capturing information about alterations, corrections or errors in the source text. For example physical damage and alterations may make the primary source ilegible, with the transcriber resorting to conjecture. Addition, deletions and substitutions, layout peculiarities of text are frequently encountered in literary works. TEI encoding guidelines have also been utilized for capturing the physical structure of books[10]. The text corpus in the Poetess Archive contains literary manuscripts, which are suitable to be transcribed utilizing the TEI guidelines to be preserved and accessed digitally.

2.1.2 TEI Encoding and Prosody in English Poetry

The English language has a long-standing tradition and methodology in place to study elements and structures of poetic works. Prosody is the term used for study of ele-
ments and structures in rhythmic speech and literature. Metrical organization is an important focus of prosody. Meter is derived from the Greek term for measure, with four metrical systems in vogue - Syllabic based on number of syllables per line, Accentual measures number of accents in a line of poetry, third system is Accentual-Syllabic which relies on a unit called Foot. It is the base for Standard English meters. Lastly, Quantitative meter focus on the durational feet. Foot is a measurable, patterned unit of poetic rhythm imported into the Accentual-Syllabic metrical system. Common feet types are - Iamb (short long), Anapest (short short long), Trochee (long short), Dactyl (long short short), Spondee (long long) and Pyrrhic (short short). Iambic and Anapestic feet are ascending feet, Trochee and Dactyl descending, Spondee and Pyrrhic are generally substitute feet [24].

Sounds in Poems and their phonological structure is brought out by use of Alliteration (repetition of same sounds or syllable in two or more words in a line), Assonance (vocalic rhyme, repetition of middle vowel sound), Consonance (pleasing combination of sound, counterpart of assonance), Rhyme (metrical rhetorical device based on sound-identities of words).

The prosodic structures described above have been utilized to encode poems for specific features in the Poetess Archive. The encoding process starts with applying Text Encoding Initiative (TEI) schema to the basic poem. Further attributes and elements are added to yield encoding descriptions for Meter, Sound, Tropology, and Syntax semantic structures [17].

2.1.3 Information Visualization and Cognition

It is generally acknowledged that the capacity of the human comprehension is greatly enhanced by utilizing linkages between external perception and internal mental pro-
cesses [12]. Cognitive processes are frequently described in terms of visual metaphors - we see, when we understand and bring our ideas into focus. Visualization is not simply presence of illustration or graphics to supplement textual information. The role of visualization is to augment cognition, and channel innate visual perception towards knowledge discovery and creation of new knowledge. This knowledge however is product of an interactive process.

2.2 Related Work

2.2.1 Information Visualization Tools

Methods developed for visualization of textual data are often derived from quantitative processes used initially for scientific visualizations [12, 28]. Text visualization tools sometimes draw on statistical visualizations in form of graphs and charts [13] based on word counts, frequency distributions and occurrences [11, 19]. Projects like TAPoR display not just results of analysis, but include visualizations of source text as well. Such visualization tools seek to allow interpretive analysis of literary texts, while introducing an element of artistry in the rendered images [31]. Certain visualization tools have been developed with focus on comparing multiple variations of a single source text. Juxta [1] and TEI Comparator [18] are projects which contain tools devoted to visually differentiating versions of source texts. Multi-dimensional (3D) visualization tools have been used primarily to render architectural artifacts and buildings, not textual material. These visualizations cater to historians, archaeologists, or any scholar interested in interplay of extant, modified or now inaccessible buildings on their occupants over time [12].
2.2.2 Content-Based Visualization Tools

Zillig has helped create a text visualization tool called TokenX [19]. It is a multi-document text visualization, analysis, and play tool developed with help of Center for Digital Research in the Humanities at University of Nebraska-Lincoln. It is designed for use on text represented as a XML tree - similar to TEI encoded documents. The tool can operate on pre-selected text, or user specified link to texts in XML format. The visualization aspects of the tool are the Word Cloud, highlighting non-words and substituting words with images. TokenX can also highlight text based on word patterns, and replace words with blocks. The tool can help with text analysis with word concordance presented in a sorted fashion, count keywords in context, word replacement and other word usage statistics.

Many Eyes is a project of Collaborative User Experience (CUE) research groups Visual Communication Lab [11]. The main idea is to present user generated visualizations over the Internet, parallelize data analysis, and leverage human visual intelligence to discern patterns. In some ways Many Eyes attempts to bring social networking to bear on massive data sets and allow online community to generate and discuss derived visualizations. Of the various visualization categories in Many Eyes, the following relate to analyzing text - Word Tree, Tag Cloud, Phrase Net, and Word Cloud Generator. A Word Tree is a visual search tool useful for nominal data like a book, article or a poem. A user picks a search word or phrase, and the tool displays all the different contexts in which it occurs in the text. The output is arranged in a tree like structure, which reveals recurrent themes and patterns. Many Eyes allows any unstructured text with up to a million words as input to the Word Tree visualization tool. A Word Cloud generator allows visualizing frequency of words in given input. Phrase Net is an analytical tool which looks at word pairs for specified
patterns, and creates a network diagram of the matched words. The size of a word is proportional to its frequency of matches, and size of connecting arrows depends on frequency of occurrence of nodal words in a match.

The ArchiveZ project aimed to create a searchable visualization tool for aggregated metadata about archival collections [13]. The data source used for this visualization are over 500 Xml encodings of finding aids, called Encoded Archival description (EAD). ArchiveZ is meant to be used by archivists towards enhanced understanding of archival collections, researchers for identification of holding institutions, and promote use of local archives by students. The tool puts the scope and availability of archival records across multiple collections in context. The principal visualization is a dual-sided histogram combined with traditional histograms to support temporal exploration of multiple subjects assigned to each collection.

National Institute for Technology and Liberal Education (NITLE) has developed SVG based information visualization tools to help correlate large, unstructured documents in a corpus [25]. The underlying methodologies range from probabilistic models, computational linguistics and graph theory. The NITLE Semantic Engine (NSE) tool suite is comprised of cluster-based visualizations, sunburst dendrogram, content stars and graph based character diagrams. The content star with its color coded, keyed clusters is a useful visualization towards comparing large number of documents, and discover anomalies. Proposed updates such as animations would greatly help with time series or temporal analysis of documents.

Text Arc is a summarized visual representation of selected works of text developed by W. Bradford Paley [22]. The tool uses a combination of indexing, concordance, and summary of the text to render visualizations for the viewer to explore. Visualizations rely more on human intuition and perception than pre-defined algorithms to potentially uncover hidden patterns in the underlying text. Words with higher levels
of occurrence are arranged closer to their place of usage. Around the highlighted words is the overall textual arc arranged in a spiral comprised of the entire text. A text view is available to show Keyword In Context. TextArc can also be utilized against collections of documents, news articles and email archives.

2.2.3 Visual Differencing Tools

NINES is a scholarly organization devoted to forging links between the material archive of the nineteenth century and the current digital research environment. A suite of Open Source tools developed under Nines aegis is comprised of IVANHOE, Collex and Juxta [2]. Juxta Collation software is an open source, cross platform tool, which is designed to compare and collate multiple editions of textual works. It includes various analytic visualization tools, such as split frame comparison of base and variant texts, and Heat Maps based on all variations. A collation view is useful in visualizing the degree of differences between variants [1]. TEI-Comparator is a tool developed for the Holinshed project at University of Oxford, England. It is designed as a comparison engine to visualize any two versions of TEI encoded Holinshed texts. TEI-Comparator can be used to follow a trail of re-used terms and points of editing across multiple versions. In contrast to other purely visualization tools, the TEI-Comparator employs a Bespoke fuzzy n-gram algorithm called Shingle cloud to pre-process the pair of texts being compared. In this respect TEI-Comparator is more of a diff engine with a visual interface. The main limitation of the tool is the intensive pre and post processing which needs to be done for it to work effectively. Perhaps more importantly the tool is restricted to comparing just two versions of the source text at any given time, relies to a lesser degree on human perception and is algorithmic in nature [18].
2.2.4 Stylometric Tools and Authorship Analysis

The Word Spectrum developed by Chris Harrison at Carnegie Mellon utilizes Google’s bi-gram dataset as basis of simple statistical comparison and its visualization [17]. The use frequency of words following the components of the selected bi-gram is analyzed; word placement is a function of the frequency of related term. The process iterates for several other word combinations creating a Word Spectrum of these combinations. The font size is an inverse power function set uniquely for each visualization. The various sample spectrums on the website demonstrate some striking correlations and trends mined from the bi-grams word associations.

![Figure 2.1: Visualizing n-grams in Homeric Poetry](image)

Forstall and Scheirer have described a work in progress that attempts to examine internal homogeneity in Homer’s works and its stylistic evolution over time [7] using character, word and metrical levels. In a poster presentation at the Digital Humanities 2011 the authors explain their goal to bridge the gap between statistical analysis and subjective interpretation of poetry. The poem is divided into samples of various sizes, and then n-gram frequencies for most common features is calculated. The resulting visualization is an abstract representation of Iliad and Odyssey in 35-line samples composed of bi-gram character frequencies.

The stylometric techniques described above are based on the assumption that word frequencies are not wholly within the conscious control of the poet. Instead each author generates an identifiable word pattern due to force of habit and prior use.
Chapter 3

Approach

Visualizations based on statistical analysis of textual content have been attempted in the past. The literary critic is expected to utilize the graphical, and sometimes abstract output in pursuit of subjective analysis of the source text. Attempts have been made to employ statistical criteria in analyzing sound patterns [20]. Packard tabulated the frequency of various sounds in Homers Iliad and Odyssey. The statistics in Figure 3.3 show how many verses in Homer do not contain a particular sound, has the sound pattern occurring once, twice and so on. However Packard does caution against relying inordinately on such statistical exercises to unearth distinctive patterns in the analyzed text.

“We cannot expect to identify expressive passages merely by counting letters.”

In our work the text based visualizations are in aid of the process of close reading. Close reading means reading a poem word by word, repeat the process line by line, explore sound, syntax and rhythm of the work. The reader might eventually succeed in uncovering hidden connotations. Multiple iterations are necessary to understand
the structure, unearth meaning and tenor of a poem. Only then the reader is in a position to address underlying prosodic elements in literary text. A non-abstract, enhanced visualization of syntax and sound structures of a poem can greatly ease the laborious process of close reading.

The Poetry Visualization tool *Myopia* developed as part of this thesis employs a direct representation of textual content overlayed with graphical metaphors to aid in literary analysis of poetry, Figure 3.1.

![Figure 3.1: Myopia Tool - Text Overlayered With Graphics](image-url)
3.1 Text Visualization

As we seek to provide a visualization tool to bring out the prosodic structures of poems in the Poetess Archive, a pertinent question is *How to Visualize Text?* [30]. The information contained in a text document can be structured into three main categories:

- **Content**
- **Structure**, and
- **Metadata**

The *Content* is description of the information contained in the document itself, in other words content can be directly read and understood. On the next level is *Structure*, a hierarchical layout which in case of literary text can be verses, stanzas and
<table>
<thead>
<tr>
<th>Table 1. Sound Densities in the Iliad*</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
<tr>
<td>d</td>
</tr>
<tr>
<td>e</td>
</tr>
<tr>
<td>f</td>
</tr>
<tr>
<td>g</td>
</tr>
<tr>
<td>h</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>j</td>
</tr>
<tr>
<td>k</td>
</tr>
<tr>
<td>l</td>
</tr>
<tr>
<td>m</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>o</td>
</tr>
<tr>
<td>p</td>
</tr>
<tr>
<td>q</td>
</tr>
<tr>
<td>r</td>
</tr>
<tr>
<td>s</td>
</tr>
<tr>
<td>t</td>
</tr>
<tr>
<td>u</td>
</tr>
<tr>
<td>v</td>
</tr>
<tr>
<td>w</td>
</tr>
<tr>
<td>x</td>
</tr>
<tr>
<td>y</td>
</tr>
<tr>
<td>z</td>
</tr>
<tr>
<td>α</td>
</tr>
<tr>
<td>β</td>
</tr>
<tr>
<td>γ</td>
</tr>
<tr>
<td>δ</td>
</tr>
<tr>
<td>ε</td>
</tr>
<tr>
<td>η</td>
</tr>
<tr>
<td>θ</td>
</tr>
<tr>
<td>ι</td>
</tr>
<tr>
<td>κ</td>
</tr>
<tr>
<td>λ</td>
</tr>
</tbody>
</table>

Figure 3.3: Sound Densities In the Odyssey [20]
lines. A table of contents of a text is an example of a visualization based on structure of the document. The Metadata describes various information about the document not contained in the text itself. Metadata includes cross-references, keywords, indexes, as well as information on the author, and publication details. Metadata visualization needs to provide insight as well as allow the user to correlate this information with the content and structure in the source document.

The aforementioned elements (content, structure, and metadata) might all be present in a document or be generated by the process of Text Analysis. The techniques employed in text analysis include statistical, lexical, syntactic and semantic analysis. An example of such analysis is in document-retrieval systems in Google search engine where content is analyzed and resultant information leads to classification and indexing. A desirable characteristic of text visualization applications is that they should be scalable. The visualization technique should be efficient in helping the user quickly understand the structure and content of the source text. The visualization should allow quick access to detailed information, and use of the visualization must result in faster comprehension as compared to going through the text line by line. In the Myopia tool the visual metaphors of Syllabic line-lengths are augmented with context sensitive tool tips to access detailed information in an efficient manner, Figure 3.4.

Text visualization methods vary in methodology depending on the type of the document being analyzed (narrative text, software code), or the type of task at hand (obtaining a structural overview, comparing different documents, or comparing different versions of the same document, or answering domain specific questions). Myopia allows a visual, structural overview of selected poems, see Figure 3.2. The content-based visualization methods are most relevant to our work as plain text forms the corpus of the Poetess Archive, and understanding underlying structure and syntax is
the goal of the process of close reading.

### 3.2 Content-Based Visualization

Document contents can be intuitively visualized using the documents default layout, as we do with numerous word processing and document viewing systems such as Adobe Acrobat. Such content-based visualizations also incorporate additional structural and metadata information. However since the main focus of the visualization happens to be the content in the document, such systems are referred to as content-based visualizations [30].

The main utility of textual content-based visualizations is in their simplicity. All views are familiar two-dimensional layout and mapping techniques such as rendering of document pages with full or reduced size. The Key legend is an important tool to convey an intuitive mapping of color semantics. The display might constitute of plain textual content, or include user provided or inherent annotations. In combination with simple to use navigational aids and interface, these techniques lend themselves to quick adoption and wide usage. These hallmarks of content-based visualization
systems are being integrated in the *Myopia* tool as we seek to enhance ease of use and supplement cognition, Figure 3.5.

![Myopia Tool - Content Based Features](image1)

**Figure 3.5: Myopia Tool - Content Based Features**

![Visualization Pipeline](image2)

**Figure 3.6: The Visualization Pipeline**
3.3 Visualization Pipeline

The purpose of a visualization application or tool is to provide insight into raw data underlying the images rendered by the visualization. The process of visualization conceptually is constituted of a series of steps that manipulate the data and produce the desired image. This modular process can thus be represented as a pipeline of discrete steps, each step representing a sub-process involved in transformation of data and information. The sequence of data transformations that take place during the visualization process are called the visualization pipeline [29]. The visualization pipeline typically consists of four stages: data importing, data filtering and enrichment, data mapping and data rendering. The data importing, filtering and enrichment steps can be difficult to deduce in actual applications. The main separation point in the visualization pipeline is when abstract data becomes actually visible during the mapping process. The implementation of visualizations can thus have all or some of the separate steps outlined in the conceptual outline. The mapping of the various stages of the conceptual visualization pipeline and it’s implementation in the Poetry Visualization tool is displayed in Table 3.1: A schematic outline of the visualization pipeline processes is shown in Figure 3.6.

The users can interact with the image to obtain fresh insights into the underlying dataset. Applications which allow the users to interactively modify the input parameters to generate new renderings implement a process called Computational Steering. This interactive process is an integral part of the Myopia tool, as the user controls the presentation of various encoding elements, derived from the underlying Xml metadata.
### Table 3.1: Pipeline and Poetry Visualization Event Mapping

<table>
<thead>
<tr>
<th>Poetry Visualization Process</th>
<th>Visualization Pipeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read In Xml Encoded Text</td>
<td>Importing Data and Enrichment</td>
</tr>
<tr>
<td>Parse Xml Elements</td>
<td></td>
</tr>
<tr>
<td>Map TEI encoded features to</td>
<td>Mapping Data</td>
</tr>
<tr>
<td>Visual Features</td>
<td>Rendering</td>
</tr>
<tr>
<td>Compose Mapped 3D scenes</td>
<td></td>
</tr>
</tbody>
</table>

#### 3.4 The Visual Thinking Process

The human and the computer can be thought of as a single cognitive identity, with the computer acting as a *cognitive co-processor* to the brain [26]. The human transmits low-bandwidth information to manipulate the user interface via keyboard and mouse, while the computer handles high-bandwidth information processing, making it available for subsequent pattern discovery. The brain’s capability to rapidly process visual information, and discern visual patterns forms the basis for any visualization tool, such as we are developing. Pirolli and Card developed a theory of information access to help in design of interfaces to visual cognitive tools. People follow *information scent* as they *forage* for information, making just in time decisions about which path to follow [23]. An ideal cognitive loop between a computer and human would require the computer providing exactly the information needed at any given time. Thus only relevant information should be on screen, allowing user to *forage* as they explore the interface. In our Poetry Visualization tool the cost of visual queries is reduced by minimal and user controlled layering of visual information, see Figure 3.7 where user has turned off a particular visual element to focus on other aspects.
3.4.1 Depth Perception and 3D

We live in a three dimensional world, however these dimensions are not perceived by the human eye in equal measure. There are millions of rays reflected from surfaces in the up-down and sideways dimension which are processed by the brain at any given instant from cone receptors in the eye. Thus we have ample visual detail about objects in the picture plane. The perception of depth is provided by the away direction. Here only non-occluded surfaces will be visible, and only one point of color is available for each ray entering the eye. Amongst the millions of pieces of information about color and texture, the distance information is very limited. This results in the visual space being referred to as having 2.5 dimensions, with the .5 referring to the away dimension [30].

We want to display syntax and structural elements parsed from the encoded poems in 2.5 dimensional environment. Content-based visualizations primarily render text, which is more easily read in a 2 dimensional layout. Relief and perception of depth however allows for an effective use of screen space to layer information [32]. The visualization of the stress patterns in our tool is based on the 2.5D design idea. Subtle but discernible projection of stressed syllables provides visual cues to the user.
without interfering with readability or overdoing the 3 dimensional aspects, Figure 3.8

**Rhyme for a Child Viewing a Naked Venus in a Painting of "The Judgment of Paris"**

He gazed and gazed and gazed and gazed,

Amazed, amazed, amazed, amazed.

Figure 3.8: Myopia Tool - Structuring 2.5 Dimensional Space

### 3.5 Implementation

The Poetry Visualization tool, *Myopia* has been developed in Python and utilizes open source Python libraries for rendering multidimensional graphics integrated with an intuitive user interface. The framework used for multidimensional graphics is called Panda3D [9], which was developed, at the Carnegie Mellon University.

From their site:

“Panda3D is a game engine, a framework for 3D rendering and game development for Python and C++ programs. Panda3D is Open Source and free for any purpose.”

It is useful to mix two-dimensional interface design elements with higher dimension graphical design elements. Navigational controls should always be visible and accessible to the user [32]. In the Poetry Visualization tool the two-dimensional Graphical User Interface (GUI) elements are created using the wxPython framework [33]. The
GUI Panel of the Poetry Visualization tool is an example of the types of widgets developed using wxPython, Figure 3.9.

From the wxPython site [33]:

“wxPython is a GUI toolkit for the Python programming language. It allows Python programmers to create programs with a robust, highly functional graphical user interface, simply and easily. It is implemented as a Python extension module (native code) that wraps the popular wxWidgets cross platform GUI library, which is written in C++.”

![Figure 3.9: Myopia Tool - GUI widgets with wxPython](image)

3.5.1 Loading Resource Bundles and Encodings

The User Interface of the Visualization tool is split into three distinct visual regions...
• The Main Visualization Area to display and interact with the rendered text

• A Key/Legend Panel at bottom of the visualization panel which holds a tabbed display of the Key associated with the loaded visualization

• GUI Panel at the right of the Visualization panel. This allows loading and manipulating source TEI documents and the resulting 3D visual elements.

Figure 3.5 shows the layout of the visualization tool.

![GUI Panel](image)

Figure 3.10: Myopia Tool - Loading Bundled Encodings

User starts by selecting an entry from the Hierarchical set of Categories in Load Meter Encoding section. The TEI encoded poems have been categorized into 12 types. Each category holds one or more poems with Metrical encodings following the TEI schema. Clicking on name of a poem Loads the underlying Xml document in the Poetry Visualization tool. The Figure 3.10 is a screenshot of the Hierarchical Tree widget used to load encoding in the visualization tool.
The digitized poems in the Poetess Archive at Miami University are XML documents, encoded with extended TEI schemas. The extended tag set is comprised of four separate schemas for meter of a poem, its syntax, literary figures in the poem, and its sound structures. Figure 3.13 shows a Coleridge sestet encoded with the meter schema. The default encoding loaded by the tool is Meter as the archive has a preponderance of meter encodings. These XML encodings are parsed into the Poetry Visualization tool utilizing Python libraries, and mapped into runtime structures for rendering.

The poems visualized by the tool are included with the source code distribution as bundled resources. As the encodings follow a standard XML schema, additional poems can easily be appended to the list of visualized poems. These bundled encodings have been organized into pre-defined Categories utilizing the element tag "type" in the XML source document. The table 3.2 provides a list of Categorized poems which are presently included with the Myopia tool.

3.5.2 Interacting With Encoding Layers

Meter is default encoding supported by the tool as it is most widely available TEI encoding in the Poetess Archive. Loading a particular poem always brings in the Metrical version initially, and this is the default visualization of the tool as well. The Set Encoding section of the GUI panel is comprised of two areas:-

- Meter Encoding
- Tropological Encoding

The Meter Encoding Checkbox is initially checked allowing the Meter visualization to be displayed.
User can affect the main Visualization panel elements utilizing this section of the GUI panel. Figure 3.11 shows the GUI options available to the user as they switch the currently active encoding of the loaded poem, as well as visual metaphors employed for that encoding.

The Poetry Visualization tool allows up to three distinct poems to be visualized at any given time. If the content does not fit the viewable area then Vertical and Horizontal scroll bars can be used as in any other GUI based tool. Zoom In/Out feature is available as well as shown in Figure 3.12.

A direct method of navigating to a loaded visualization can be availed using the Loaded Poems drop box. By selecting the desired poem the user moves the active visual pane to that visualization.

Figure 3.11: Myopia Tool - Active Encoding and Graphical Elements
Figure 3.12: Myopia Tool - Navigational Controls

Figure 3.13: TEI Encoding - Poem with Meter Encoding
3.5.3 Abstract Representation - The Hotspot Viewer

The hotspot viewer is an abstraction of combined TEI tags used in Meter and Tropological encoding. This visualization shown in Figure 3.14, is derived from the underlying encoded text as a matrix, with each character represented initially as number 0. For meter encoding each occurrence of Metrical feet is marked in the matrix by incrementing the corresponding cell value. Similarly each tag location in trope encoding results in the matrix abstraction being updated. This abstraction can be helpful to get a bird’s eye view of the most heavily encoded areas of the poem. These heavily marked up sections can potentially point to areas in the poetry where connotation converge.

Figure 3.14: Myopia Tool - Hotspot Abstract View
3.5.4 The Implementation Pipeline

The conceptual visualization pipeline described in Section 3.3 is modular, and has discrete, well defined segments. Actual implementation of such a conceptual pipeline often does not find such clearly demarcated functions. The stages of data importing, filtering and enrichment may often coincide or transition seamlessly into one another. In our Poetry Visualization tool, the input data is in form of Xml files constrained by enhanced TEI based schemas. The process of parsing these Xml elements, and bringing them into Python structures for rendering is a single composite step encompassing data import and filtering (unwanted Xml tags are simply ignored). Table 3.3 highlights the difference with the conceptual visualization pipeline outlined in table 3.1. The enrichment of parsed data (converting Arabic numbers into Roman numerals for display for instance), and the rendering stage where abstract data becomes tangible are often clearly identifiable. This is the case with our visualization pipeline, as all data mapping, assignment of visual metaphors to Xml elements, and final rendering in multidimensional space is handled by modules based on Panda and wxPython libraries.

The effectiveness of a given visualization is directly related to the properties of the mapping between the raw data and its rendering [29]. In general the mapping should be invertible, so that the user is assured that the visualization actually relates to the source data being rendered. In our visualization tool the textual content of the encoded poetry being visualized is always available to the user, see Figures 3.15 and 3.16. The ability to revert to the source text at any stage of the visualization from abstract graphical mappings is a conscious design choice in our tool, and enhances its effectiveness.
God's Grandeur

The world is charged with the grandeur of God. 
It will flame out, like shooting star! 
It gathers in a greatness, like the noon of day! 
Crushed. Why do men then new not reck his rod! 
Generations have trod, have trod, have trod. 
And all is seared with trade, bleared, smeared with toil. 
And wears man's smudge and shares man's smell the soil. 
Is bare now, nor can feel, being shed. 
And for all this I natural am Iber upon.
There lives the deepness deep down things. 
And thought the last lights off the black West wind. 
Oh, morning, all the brown brink eastward, springs. 
Because the Holy Ghost over the bank. 
World broods with warm breast and with a hill bright wings.

Figure 3.15: Visualization Mapping - Graphics With Text
3.5.5 Tools

We implemented the visualization tool using the Eclipse Interactive Development Environment (IDE). Eclipse is ubiquitous in java language based software development, and is increasingly becoming popular for other languages as well. The python components were integrated into the Eclipse IDE with the Pydev plugin [3]. Pydev is a Python IDE for Eclipse, which can be used for Python, Jython or IronPython development. PyDev greatly simplifies Python development with features like code completion with auto import, syntax highlighting and refactoring help.
### Table 3.2: Myopia Tool - Bundled Categorized Encoded Poems

<table>
<thead>
<tr>
<th>Category</th>
<th>Poem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sonnet</strong></td>
<td>Holy Sonnet 1 by John Donne</td>
</tr>
<tr>
<td></td>
<td>Holy Sonnet 14 by John Donne</td>
</tr>
<tr>
<td></td>
<td>Mowing by Robert Frost</td>
</tr>
<tr>
<td></td>
<td>Bright Star by John Keats</td>
</tr>
<tr>
<td></td>
<td>On His Blindness by John Milton</td>
</tr>
<tr>
<td></td>
<td>Anthem for Doomed Youth by Wilfred Owen</td>
</tr>
<tr>
<td></td>
<td>The Sonnet William by Wordsworth</td>
</tr>
<tr>
<td><strong>Roundel</strong></td>
<td>The Roundel by Algernon Charles Swinburne</td>
</tr>
<tr>
<td><strong>Couplet</strong></td>
<td>Rhyme for a Child Viewing a Naked Venus in a Painting of ”The Judgment of Paris” by Robert Browning</td>
</tr>
<tr>
<td></td>
<td>Epigram Engraved on the Collar of a Dog Which I Gave to His Royal Highness by Alexander Pope</td>
</tr>
<tr>
<td><strong>Quintet</strong></td>
<td>Ah Sun-flower by William Blake</td>
</tr>
<tr>
<td></td>
<td>Jordan by George Herbert</td>
</tr>
<tr>
<td></td>
<td>Song by Edmund Waller</td>
</tr>
<tr>
<td><strong>Octave</strong></td>
<td>Gods Grandeur by Hopkins</td>
</tr>
<tr>
<td></td>
<td>Though I am Young and Cannot Tell by Ben Jonson</td>
</tr>
<tr>
<td></td>
<td>Even Such Is Time by Walter Ralegh</td>
</tr>
<tr>
<td></td>
<td>He Wishes For the Clothes Of Heaven by W.B. Yeats</td>
</tr>
<tr>
<td><strong>Spenserian Stanza</strong></td>
<td>Adonais (stanzas 54-55) by Shelley</td>
</tr>
<tr>
<td></td>
<td>Despayre in Praise of Suicide (Faerie Queene 1.9.39-40) Edmund Spenser</td>
</tr>
<tr>
<td><strong>Stanza</strong></td>
<td>Jabberwocky by Lewis Caroll</td>
</tr>
<tr>
<td></td>
<td>Ode On a Grecian Urn by Keats</td>
</tr>
<tr>
<td><strong>Quartrain</strong></td>
<td>Westron Wynde by Anonymous</td>
</tr>
<tr>
<td></td>
<td>The Sick Rose by William Blake</td>
</tr>
<tr>
<td></td>
<td>The Brain Is Wider than the Sky by Emily Dickinson</td>
</tr>
<tr>
<td></td>
<td>The Lowest Place by Christina Rossetti</td>
</tr>
<tr>
<td><strong>Tercet</strong></td>
<td>Upon Julias Clothes by Robert Herrick</td>
</tr>
<tr>
<td>Poetry Visualization Process</td>
<td>Visualization Pipeline</td>
</tr>
<tr>
<td>-------------------------------------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Parse Xml Data and Map to Objects</td>
<td>Importing Data and Enrichment</td>
</tr>
<tr>
<td>Map TEI encoded features to Visual Features</td>
<td>Mapping Data Rendering</td>
</tr>
<tr>
<td>Compose Mapped 3D scenes</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4

Evaluation

4.1 User Study Design

We conducted a user study to evaluate the effectiveness of our approach in helping the process of close reading. To achieve this objective we utilize the Pre-Experimental Design methodology known as Pretest-Posttest Experiment design explored by Campbell, Stanley et. al. [8]. The experiment seeks to measure the effectiveness, adequacy and usability of the Poetry Visualization tool. Since this experimental design does not include randomized assignments to experimental and control groups, our study does not conform to scientific standards of experimental design [5]. However it does capture behavior of real-users in controlled circumstances.

The study is composed of an experimental run consisting of a Pre-test questionnaire, followed by three assigned tasks (performed with help of installed Visualization tool, or screen-casts, slides and screen-shots of the software, and a comprehensive User’s Guide. The participants also provide response to a Post-test questionnaire which records their impressions of the tool and the assigned tasks themselves. The experiment and survey document is included as Appendix B.
4.1.1 Research Questions

The three aspects of the tool, *effectiveness, adequacy and usability* that we wanted to evaluate are addressed by defining the research questions outlined in Table 4.1. The main research question (RQ1) is concerned with measuring the effectiveness of the visualization tool. This examines whether participants are able to efficiently close read given text. Three sub-research questions are intended to measure the effectiveness of the more specific aspects of the tool. The goal of our other research questions is to collect responses about the design of the *Myopia* Visualization tool in comparison to what the participant expected (Adequacy) and to get feedback about how easy the tool is to use and navigate (Usability).

<table>
<thead>
<tr>
<th>RQ1</th>
<th>Can a Visualization tool based on TEI help in close reading?</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1.1</td>
<td>Can the visualization tool be used to identify metrical similarities?</td>
</tr>
<tr>
<td>RQ1.2</td>
<td>Is the tool an efficient medium to quickly deduce sound patterns and rhythms?</td>
</tr>
<tr>
<td>RQ1.3</td>
<td>Does the tool provide flexible ways to perform close reading activities?</td>
</tr>
<tr>
<td>RQ2</td>
<td>Can Myopia meet the expectations of a visualization tool in aid of close reading?</td>
</tr>
<tr>
<td>RQ3</td>
<td>Do the users find the tool practical enough to use frequently?</td>
</tr>
</tbody>
</table>

Table 4.1: User Study Research Questions

4.1.2 PreTest-PostTest Design

Participants are first given a survey (questionnaire) to evaluate their initial responses before the experiment is conducted. The participant is then asked to complete the experiment (i.e. Tasks). A second questionnaire is taken after the participant has completed the tasks in the experiment. Other than general background questions, the questions in the questionnaires were written as closed-ended matrix questions. These closed-end questions ask the participants to rate their opinion using a five point Likert
scale (i.e. 1 = strongly disagree, 2 = disagree, 3 = neither disagree nor agree, 4 = agree, 5 = strongly agree).

**PreTest Design**

The pre-test questionnaire consists of six sections, and aim to discern participants background, attitude towards TEI, visualized encodings and visualization tools.

Table 4.2 lists the sections in the Pretest questionnaire and their purpose.

<table>
<thead>
<tr>
<th>Section</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Background</td>
</tr>
<tr>
<td>2</td>
<td>TEI Knowledge</td>
</tr>
<tr>
<td>3</td>
<td>TEI or Xml Visualizations</td>
</tr>
<tr>
<td>4, 5</td>
<td>Value of TEI Visualizations</td>
</tr>
<tr>
<td>6</td>
<td>Expectations from a Visualization Tool</td>
</tr>
</tbody>
</table>

First two sections in the pre-test are dedicated to understanding the background of the participants and their familiarity with TEI which underlies the visualized encodings in the *Myopia* tool.

In Section 3 the participants are asked to rate statements about their familiarity with TEI and Xml visualizations. If a participant has used other visualization tools, this could influence their perception of what a TEI encoding visualization tool should be like.

Section 4 and 5 seeks to gauge participants attitude towards TEI encoding activities, visualizations, and its possible uses.

Section 6 starts with a short description of the process of close reading. After this
definition the participant is given a range of questions (listed in Table 4.10), to gauge their expectation of a TEI based visualization tool in the process of close reading. The participant response to the statements in this section are subsequently tied to their perception of the *Myopia* tool after completing the assigned tasks.

**Post-Test Design**

The Post-Test questionnaire is given to participants after the experiment and consists of five sections. The goal of distributing this questionnaire is to measure the participants attitudes and opinions after they have performed the assigned tasks. These tasks provide the users the opportunity to interact with the Myopia tool, and evaluate it’s features. The Post-Test feedback will help establish if Myopia is effective as a TEI based visualization tool in aid of close reading.

Section 1 of the Post-Test questionnaire pertains to the user experience with the study while performing the assigned tasks. The questions want the user to: judge the reasonableness of the tasks, whether each task had sufficient time allocation, comment on the ease of use of the related resources such as slides and screenshots, and finally if they found the tasks to have a realistic purpose.

Statements in Section 2 deal with the experience of the users with the *Myopia* tool as part of the experiment. The idea is to judge the effectiveness of the tool in performing each of the assigned task. This section hence addresses the main research question **RQ1**.

Section 3 statements are directly related to the last section in the Pre-Test segment, which related to user expectations from a visualization tool. The questions in this section however mention Myopia explicitly, to address Research Question **RQ2**.

Section 4 consists of statements to gauge the usefulness of individual features in the Myopia tool. The participants rated the tool from 1 (very useless) to 5 (very useful).
An additional rating of 6 (Not Applicable) was also provided for those participants who might not be able to respond to TEI specific statements in this section. The post-test questionnaire finishes with a set of statements which relate to the usability of the Myopia tool. The participant responses help address the research question RQ3.

The final segment in the post-test questionnaire allows the participant to provide any comments or suggestions they might have about the tool. The comment segment finishes by soliciting participant comments about the user study itself.

### 4.1.3 Tasks

Our tasks for the user study have been designed to elicit the effectiveness of the visualization tool we have developed. The tasks had participants look at slides and screen casts of the tool and answer questions dealing with literary analysis of the visualized poems. The tasks seek to (1) Find Metrical Features and Identify Encoding Patterns (2a) Visualize and Validate TEI encodings (2b) Compare and Contrast encodings, and (4) Identify Stress Patterns from visualizations. Tasks 1, 2b, and 3 are meant to address the sub objectives of RQ1. Task 2a is specific to statement 4 in the Post-test questionnaire, and geared towards measuring the usefulness of the tool.

**Task 1** addresses the requirements of **RQ1.1**, (Can the visualization tool be used to identify metrical similarities?). The effectiveness of the Poetry visualization tool can be gauged by the ease with which similar poetic texts can be located and analyzed. The metric of similarity in this context is the underlying TEI encoding and the resultant visualizations in the Myopia tool. The participant needs to identify the poem with a constant meter, and identify that meter.

**Task 2a** is designed to find the participant’s impressions of the differently encoded versions of the same poem, and record any variations they find in resulting
visualizations.

Task 2b addresses the research question RQ1.2, comparing two differently encoded versions of the same poem. This comparison gauges the efficiency of the tool in deducing sound and rhythm patterns in a poem. Task 3 is designed keeping RQ1.3 in mind. Two different approaches to identify stress patterns are put forward in this task. This showcases the flexibility of the tool in performing close reading activities. The complete user study is included in the Appendix B of this thesis.

4.1.4 Pilot Study

The initial batch of four participants in the user study were comprised of graduate students from the English department at Miami University. Their feedback from the study pointed out problems with accessing the hand-outs and audio visual material associated with the tasks. Feedback also suggested that the Task descriptions needed to be well explained in order to obtain full participation and ensure more complete responses. We also modified the Pre-test and Post-Test questionnaire in light of the Pilot Study to better align the two sets of surveys. These changes increased the number of questions in the Pre-test survey, and made them more in line with the Post-test questions.

4.2 Results

The results obtained from the user study and experiment are presented in this section. The background of the participants, and their prior experience with TEI, XML and Visualization tools is first described. This information is obtained by the responses to the pre-test segment of the questionnaire. We evaluate the performance of the experiment participants by collating their task performance and responses to the
post-test questions. The Post-Test responses also allow us to answer our research questions.

4.2.1 Participant Characteristics

7 graduate students from the Department of English at Miami University were participants in our user study. These students mostly specialized in English Comprehension and Rhetoric, and their participation was co-ordinated with help of the English Department. A smaller and separate group of students had earlier participated in the Pilot Study, which helped fine tune the experiment undertaken by the larger group. None of the participants had prior knowledge of TEI and related encoding activities.

4.2.2 TEI Knowledge and Comprehension

The Pre-Test sections 2 to 5 intend to find participant familiarity and attitude towards TEI encoding, textual visualization, and close reading. The results of these questions are described in rest of this section.

Familiarity with TEI

Overwhelmingly the participants described their knowledge and familiarity of TEI encoding as None. For most respondents this user study was their first exposure to the encoding standards and its use in marking up textual material.

Attitude towards TEI and Xml based Visualizations

The purpose of the statements put forward in Table 4.4 is to gauge participant exposure to other visualization aids in the past. With limited knowledge of TEI and encoding in general, most participants responded in the negative or were noncommit-
Table 4.3: Pretest Questionnaire: Familiarity with TEI

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning TEI Guidelines</td>
<td>71.4%</td>
<td>0%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Learning Enabling technologies of TEI</td>
<td>71.4%</td>
<td>0%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Using TEI support resources (e.g. wiki, tools, software)</td>
<td>71.4%</td>
<td>0%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Developing custom TEI schemas</td>
<td>71.4%</td>
<td>0%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Encoding text using TEI</td>
<td>71.4%</td>
<td>0%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
</tbody>
</table>

Table about prior usage of visualization tools and software to explore encoded text. 60% of the participants were also neutral on the usefulness of learning TEI. However a significant number of respondents were favorable towards learning TEI, and agreed that visualizing TEI based encodings is a good idea. These positive responses mitigate to a certain extent threat to internal validity of the study.

4.2.3 Task Performance

For the experiment, the participants were provided access to a website with screen-shots and screen-cast movies to help complete the outlined tasks. Task1 is setup to find Metrical features and to identify encoding patterns in supplied extract of a poem. This task seeks to ascertain the effectiveness of the tool in analyzing and lo-
Table 4.4: Pretest Section: TEI Visualization Activities

Some statements regarding TEI Visualization activities are shown below. Please, rate each statement on a scale from 1 to 5 to indicate to what extent they apply to you according to the following scale: 1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5 Strongly Agree

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>When viewing an unfamiliar encoded text, I:</td>
<td>16.6%</td>
<td>16.6%</td>
<td><strong>50%</strong></td>
<td>16.6%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(1)</td>
<td>(3)</td>
<td>(1)</td>
<td>(0)</td>
<td>6</td>
</tr>
<tr>
<td>Use tools or software to try to visualize or describe the encoded text</td>
<td>16.6%</td>
<td><strong>33.3%</strong></td>
<td><strong>33.33%</strong></td>
<td>16.6%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
<td>(1)</td>
<td>(0)</td>
<td>6</td>
</tr>
<tr>
<td>Use XML based technologies such as XSLT</td>
<td>16.6%</td>
<td><strong>33.3%</strong></td>
<td><strong>33.3%</strong></td>
<td>16.6%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
<td>(1)</td>
<td>(0)</td>
<td>6</td>
</tr>
<tr>
<td>Use software tools to visually validate TEI encoding</td>
<td>16.6%</td>
<td><strong>33.3%</strong></td>
<td><strong>33.3%</strong></td>
<td>16.6%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
<td>(1)</td>
<td>(0)</td>
<td>6</td>
</tr>
<tr>
<td>Rely on XML schema editors to view and validate encoding of unfamiliar text</td>
<td>16.6%</td>
<td><strong>33.3%</strong></td>
<td><strong>33.3%</strong></td>
<td>16.6%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
<td>(1)</td>
<td>(0)</td>
<td>6</td>
</tr>
<tr>
<td>Use visual tools to spot differences in encodings</td>
<td>16.6%</td>
<td><strong>33.3%</strong></td>
<td><strong>50%</strong></td>
<td>0%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(0)</td>
<td>(0)</td>
<td>6</td>
</tr>
<tr>
<td>Consult with other TEI user for help and guidance</td>
<td>16.6%</td>
<td><strong>33.3%</strong></td>
<td><strong>33.3%</strong></td>
<td>16.6%</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(2)</td>
<td>(1)</td>
<td>(0)</td>
<td>6</td>
</tr>
</tbody>
</table>

cating similarly encoded texts. Task1 asked the participant to identify the poem with a constant meter, and identify the line lengths in this text, and was sub-divided into three parts. All 7 participants successfully completed the first sub-task of identifying poem with a same meter throughout. There were two correct options in the resources for this sub-task, however only one was applicable to the followup questions. This caused some participants to respond to sub-task 2 and 3 incorrectly Overall 67% of respondents were able to correctly perform Task1.
Table 4.5: Pretest Section: Value Of TEI Encoding

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of TEI is a valuable asset</td>
<td>0%</td>
<td>0%</td>
<td>60%</td>
<td>20%</td>
<td>20%</td>
<td>5</td>
</tr>
<tr>
<td>Learning to encode texts using TEI is a worthwhile exercise</td>
<td>0%</td>
<td>0%</td>
<td>60%</td>
<td>0%</td>
<td>40%</td>
<td>5</td>
</tr>
<tr>
<td>Visualizing TEI encodings is helpful in gaining additional insight</td>
<td>0%</td>
<td>0%</td>
<td>60%</td>
<td>20%</td>
<td>20%</td>
<td>5</td>
</tr>
</tbody>
</table>

Task2a provides the participant opportunity to encode a stanza from a Walter Landor poem *Upon Julia’s Clothes*. This encoding is then compared by the participants to the screenshots of the same text visualized with the *Myopia* tool. Task2a concludes by recording the participant’s impressions of the these versions of encodings. Though 64% of responses to this section can be deemed satisfactory, the participants had difficulty with this task, largely because of their inexperience in encoding and the inherently subjective nature of encoding.

Task2b continues with the theme of comparing two versions of encodings, this time encodings produced by the *Myopia* tool. This task demonstrates the tool’s usage as a visual diff tool, and as an aid in prosody. The first sub-task asked the participants to identify visually any encoding differences from screen-shots. Over 90% responded correctly to this sub-task. The second sub-task had a lower success rate, as it required participants to have some prior knowledge of poetry and criticism. This comparative aspect of *Myopia* has been validated with 64% respondents completing the assigned task correctly.
Table 4.6: Participant Task Performance

<table>
<thead>
<tr>
<th>Task</th>
<th>Sub Task</th>
<th>Correct</th>
<th>Incorrect /Unsure</th>
<th>Total</th>
<th>%Correct</th>
<th>Task Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>42.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>57.1%</td>
<td>66.6%</td>
</tr>
<tr>
<td>2a</td>
<td>a</td>
<td>4.5</td>
<td>2.5</td>
<td>7</td>
<td>64.3%</td>
<td>64.3%</td>
</tr>
<tr>
<td>2b</td>
<td>a</td>
<td>6.5</td>
<td>.5</td>
<td>7</td>
<td>92.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>2.5</td>
<td>4.5</td>
<td>7</td>
<td>35.7%</td>
<td>64.3%</td>
</tr>
<tr>
<td>3</td>
<td>a</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Task3** has the objective of identifying Stress Patterns from provided visualizations. Meter is a recurring pattern of stressed (long) and unstressed (unaccented, or short) syllables in lines of a set length. Identifying Stress patterns in a poem is an important part of literary analysis. The participant is asked to comment on different ways available in the tool to achieve the same objective, highlighting the flexibility of the tool. 100% of the participants were successful in achieving the objective of this task.

Overall the tasks were not considered by the participants to serve any realistic purpose, with a majority of the respondents unsure about how to complete some or all of the tasks. A fuller introduction to the tool and tasks might have alleviated some of the issues faced by the participants in completing the tasks.
Table 4.7: Posttest Section: Usefulness of Myopia tool in Textual Analysis

Please, rate each statement on a scale from 1 to 6 according to the following scale:
1-very useless, 2-useless, 3-Neither useful nor useless, 4-useful, 5-very useful, 6 N/A

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare encodings</td>
<td>0%</td>
<td>0%</td>
<td>28.5%</td>
<td>57.1%</td>
<td>14.3%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Validate Encodings</td>
<td>0%</td>
<td>0%</td>
<td>42.3%</td>
<td>42.8%</td>
<td>14.3%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Identify quickly the rhythm and syntax ele-</td>
<td>0%</td>
<td>0%</td>
<td>71.4%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>ments of Poetry</td>
<td></td>
<td></td>
<td>(0)</td>
<td>(2)</td>
<td>(0)</td>
<td>(0)</td>
<td></td>
</tr>
</tbody>
</table>

The Visualization tool is useful for this task

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2a</th>
<th>2b</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Visualization tool is useful for this task</td>
<td>33.3%</td>
<td>100%</td>
<td>66.6%</td>
<td>66.6%</td>
</tr>
<tr>
<td>During this task I validated my encoding for a poem</td>
<td>16.6%</td>
<td>66.6%</td>
<td>33.3%</td>
<td>33.3%</td>
</tr>
<tr>
<td>The Stress Patterns can be identified in this task</td>
<td>66.6%</td>
<td>83.3%</td>
<td>66.6%</td>
<td>83.3%</td>
</tr>
</tbody>
</table>

Table 4.8: Participant Task Experience

4.2.4 Tool Effectiveness and Flexibility

The effectiveness of the Myopia tool, and the related answer to the main research question RQ1 is explored in this section. RQ1 addresses if a Visualization tool such as Myopia based on TEI helps in close reading. This research question has three sub questions described below.
RQ1.1: asks *Can the visualization tool be used to identify metrical similarities?* The participants in the user study were asked to identify metrical similarities as part of task Task2b. 65% of the participants successfully completed the assigned task. The majority of respondents reported that the tool was useful in finding visualizations depicting metrical similar structures. We conclude the *Myopia* tool can be used to identify metrical similarities, and research question RQ1.1 can be answered positively.

RQ1.2: asks *Is the tool an efficient medium to deduce sound patterns and rhythms?* Task1 was designed to test if participants could identify rhythmic structures, as well as efficiently locate any anomalies, or change in sound patterns. Over 60% of the participants correctly responded to this exercise. The users were also successful in deciphering deviation from expected encoding from the visualizations. Only 33% of participants found the tool useful for the task which is a matter of concern. However, a majority were able to identify stress patterns as part of this task, and considering the high success rate we conclude that the *Myopia* tool is an efficient medium to quickly deduce sound patterns and rhythms.

RQ1.3: *Does the tool provide flexible ways to perform close reading activities?* Meter is a recurring pattern of stressed (long) and unstressed (unaccented, or short) syllables in lines of a set length. Identifying Stress patterns in a poem is an important part of literary analysis. Task3 allows the participants to complete an analytical activity using the *Myopia* tool in two different ways. A large majority of participants were able to correctly identify stress patterns using different visual metaphors employed by the tool. As the responses were accurate despite different methods employed to perform the task, we can conclude that the *Myopia* tool adequately addresses the research question RQ1.3.
**Tool Usefulness**

A small number of respondents (30%) found the *Myopia* tool useful in identifying rhythm elements quickly. Over 70% were noncommittal. The response was largely positive when asked about use of the tool to compare and validate encodings. On balance the participants found the tool to be useful, however large number of neutral responses suggests further work needs to be done to increase the effectiveness of the tool.

The positive outcome of sub questions related to close reading activities described above leads us to conclude that we have addressed **RQ1**, and the *Myopia* tool is helpful in close reading.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S1</strong></td>
<td>A Visualization tool can aid in close reading.</td>
</tr>
<tr>
<td><strong>S1’</strong></td>
<td>The Myopia Visualization tool is not useful for purpose of Close reading.</td>
</tr>
<tr>
<td><strong>S2</strong></td>
<td>Visualizations of encodings can be subjective and abstract limiting usefulness of such a tool.</td>
</tr>
<tr>
<td><strong>S2’</strong></td>
<td>Abstract or Graphical representation of meter and feet in Myopia was helpful.</td>
</tr>
<tr>
<td><strong>S3</strong></td>
<td>Encodings without visual metaphors and transformations are difficult to validate and learn.</td>
</tr>
<tr>
<td><strong>S3’</strong></td>
<td>Encodings without visual transformation by the Myopia tool are difficult to learn and validate.</td>
</tr>
<tr>
<td><strong>S4</strong></td>
<td>Encodings are difficult to compare across versions. A visualization tool would help with this task.</td>
</tr>
<tr>
<td><strong>S4’</strong></td>
<td>Encodings are difficult to compare across versions. The Myopia tool helped in this task.</td>
</tr>
<tr>
<td><strong>S5</strong></td>
<td>Close reading is only worthwhile if source text is directly utilized in the process.</td>
</tr>
<tr>
<td><strong>S5’</strong></td>
<td>Close reading is possible since Myopia displays source text as well as visual metaphors.</td>
</tr>
</tbody>
</table>

Table 4.9: Expectations and Perceptions: Pre and Post-Test Statements

### 4.2.5 Tool Expectations and Adequacy

Statements in section three in the Pre-Test and section six from the Post-Test segments are co-related, as the responses help determine if the tool adequately fulfills participant expectations. This is first done with respect to a generic visualization tool, and then their perception after using the *Myopia* tool. The responses to the...
Pre-Test expectation, in contrast to the Post-Test impressions of the the *Myopia* will help answer the research question **RQ2**.

**Table 4.10: Pretest Section: Expectation from a Visualization Tool**

<table>
<thead>
<tr>
<th>Some statements regarding Visualization and Close Reading are shown below.</th>
<th>1%</th>
<th>2%</th>
<th>3%</th>
<th>4%</th>
<th>5%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please, rate each statement on a scale from 1 to 5 to 1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5 Strongly Agree</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Total</td>
</tr>
<tr>
<td>A Visualization tool can aid in close reading</td>
<td>0%</td>
<td>14.3%</td>
<td>14.3%</td>
<td><strong>42.8%</strong></td>
<td>28.5%</td>
<td>7</td>
</tr>
<tr>
<td>Visualizations of encodings can be subjective and abstract limiting usefulness of such a tool.</td>
<td>0%</td>
<td>28.5%</td>
<td><strong>57.1%</strong></td>
<td>14.3%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Encodings without visual metaphors and transformations are difficult to validate and learn</td>
<td><strong>14.3%</strong></td>
<td>0%</td>
<td><strong>57.1%</strong></td>
<td>28.5%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Encodings are difficult to compare across versions. A visualization tool would help with this task.</td>
<td>0%</td>
<td>14.3%</td>
<td><strong>71.4%</strong></td>
<td>14.3%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Close reading is only worthwhile if source text is directly utilized in the process</td>
<td>0%</td>
<td>0%</td>
<td><strong>42.8%</strong></td>
<td>28.5%</td>
<td>28.5%</td>
<td>7</td>
</tr>
</tbody>
</table>

The responses to the co-related statements in the Pretest and Posttest segments are plotted on a Radar or Star chart. This chart plots the values of each category along a separate axis that starts in the center of the chart and ends on the outer ring. The source data for the chart in Figure 4.1 is derived from tables 4.10, and 4.11. In this Star or Radar chart the category comprised of co-related statements are
displayed along the periphery. The Value axis ranges from the center to the outer ring, with its domain comprising of number of distinct responses for the rating levels. The rating levels themselves start at 1 (Strongly Disagree) to 5 (Strongly Agree) and are displayed in the chart legend. This radar chart thus allows us to view data along more than three variables. We have placed co-related statements next to each other, so as to better highlight any change in participant responses in Pre Test to Post Test segments. Each set of co-related statements are separated by a blank section (shown as Sepr-1 to 5) to enhance readability of the graph.

Statement S1 regarding expectations from a generic visualization tool in aid of close reading is responded to by participants positively. The corresponding statement S1’ is framed negatively, and 60% participants disagreed that Myopia was not useful for close reading. The chart visualizes the correspondence with the number of people agreeing with the S1, and disagreeing with S1’ remaining the same (Purple and Red lines respectively).

A minority of respondents expected visualized encodings to have limited use while answering statement S2. After interaction with the tool the perception of the participants became significantly positive in relation to graphical metaphors and their use in Myopia. This shift can be noticed in the chart with a significant decrease in neutral responses (Green line signifying Neither Agree nor Disagree).

S3 states encodings are difficult to learn and validate without visual transformations. Only 14% participants expected visual transformations to be effective, however post test positive responses increased significantly suggesting the participants found Myopia meet or exceed their expectations. Participants similarly changed their responses with respect to statement S4, after interaction with the tool in a positive fashion.

Statement S5 theorizes that close reading is possible only in presence of the source
text. Participants response similarly support the statement S5’ (Close reading is possible since Myopia displays source text as well as visual metaphors).

Analyzing the pre test expectations and post test perceptions, we can conclude that Myopia adequately meets expectations, and RQ2 can be answered positively.

Table 4.11: PostTest Section: Perception of the Myopia Tool

<table>
<thead>
<tr>
<th>Some statements regarding Myopia Tool and Close Reading are shown below.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Myopia Visualization tool is not useful for purpose of Close reading.</td>
<td>14.3%</td>
<td>42.8%</td>
<td>14.3%</td>
<td>28.5%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>Abstract or Graphical representation of meter and feet in Myopia was helpful.</td>
<td>0%</td>
<td>14.3%</td>
<td>28.5%</td>
<td>42.8%</td>
<td>14.3%</td>
<td>7</td>
</tr>
<tr>
<td>Encodings without visual transformation by the Myopia tool are difficult to learn and validate</td>
<td>0%</td>
<td>28.5%</td>
<td>28.5%</td>
<td>14.3%</td>
<td>28.5%</td>
<td>7</td>
</tr>
<tr>
<td>Encodings are difficult to compare across versions. The Myopia tool helped in this task.</td>
<td>14.3%</td>
<td>14.3%</td>
<td>14.3%</td>
<td>42.8%</td>
<td>14.3%</td>
<td>7</td>
</tr>
<tr>
<td>Close reading is possible since Myopia displays source text as well as visual metaphors</td>
<td>0%</td>
<td>14.3%</td>
<td>28.5%</td>
<td>57.1%</td>
<td>0%</td>
<td>7</td>
</tr>
</tbody>
</table>
4.2.6 Myopia User Feedback

Tool Usability

The participants were equally divided when asked about the visual layout of the *Myopia* tool. This response can be attributed to the participants not interacting with the tool directly, but through slides and screen-shots. Over 50% participants however felt the tool could be potentially used as an aid to close-reading and analysis of poetry. Majority of the responses to the statement about the tool as an encoding help were positive. One participant commented “I really liked the tool, particularly the color coding and images”. Another comment about the tool was “It seems to be very helpful and intuitive already”. Couple of responses mentioned limited usefulness of the tool, with one participant expressing concern that it was too narrowly focussed on analyzing or visualizing metered poetry. Such responses however did not dispute the utility of *Myopia* in visualizing encoded poems, and its help in close reading.
Overall the responses supported the tool as satisfying the need to visualize encodings. From this we can conclude that *Myopia* tool is practically useful and positively answers research question **RQ3**.

<table>
<thead>
<tr>
<th>Statements in this part relate to the usability of the Myopia Visualization tool</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I found the visual layout of the tool easy to navigate.</td>
<td>28.5%</td>
<td>14.28%</td>
<td>14.28%</td>
<td><strong>42.8%</strong></td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>I could potentially use the tool as an aid to close reading and analysis of Poetry.</td>
<td>0%</td>
<td>14.2%</td>
<td>28.5%</td>
<td><strong>42.8%</strong></td>
<td>14.2%</td>
<td>7</td>
</tr>
<tr>
<td>The visualization tool is a encoding help, not meant for literary critics.</td>
<td>0%</td>
<td>28.5%</td>
<td><strong>42.8%</strong></td>
<td>14.3%</td>
<td>14.3%</td>
<td>7</td>
</tr>
<tr>
<td>Visualizing encodings is a great idea.</td>
<td>0%</td>
<td>0%</td>
<td>28.5%</td>
<td><strong>71.4%</strong></td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>This tool does not satisfy the need to visualize encodings.</td>
<td>14.3%</td>
<td><strong>28.5%</strong></td>
<td>14.3%</td>
<td>14.3%</td>
<td>14.3%</td>
<td>7</td>
</tr>
</tbody>
</table>

### 4.3 Threats to Validity

Various factors like differential selection, experimental mortality, reactive or interaction effect of testing jeopardize the validity of experimental design [8]. *Internal Validity* is the basic minimum without which any experiment is uninterpretable. Are the results affected by the process of experiment itself, or is there inherent bias in the
participant pool? *External Validity* on the other hand relates to generalizability of the results.

Table 4.13: PostTest Section: Experience With the User Study

<table>
<thead>
<tr>
<th>Some statements regarding Your Overall Experience with the Study</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall, the tasks were reasonable to perform</td>
<td>0%</td>
<td>28.5%</td>
<td>14.3%</td>
<td>57.1%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>I felt there wasn’t enough time for each task</td>
<td>14.28%</td>
<td>57.1%</td>
<td>28.5%</td>
<td>0%</td>
<td>0%</td>
<td>7</td>
</tr>
<tr>
<td>I felt unsure about how to complete some or all of the task</td>
<td>0%</td>
<td>0%</td>
<td>28.5%</td>
<td>57.1%</td>
<td>14.28%</td>
<td>7</td>
</tr>
<tr>
<td>The Poetry Visualization slides were hard to follow.</td>
<td>0%</td>
<td>14.2%</td>
<td>28.6%</td>
<td>42.8%</td>
<td>14.2%</td>
<td>7</td>
</tr>
<tr>
<td>I felt the tasks did not resemble a realistic purpose</td>
<td>0%</td>
<td>28.5%</td>
<td>42.8%</td>
<td>28.5%</td>
<td>0%</td>
<td>7</td>
</tr>
</tbody>
</table>

4.3.1 Internal Validity

Participants

The participant’s familiarity with TEI encoding and desirability of visualizing these Xml based document was evaluated in the pre-test segment of the study. Attitude towards visualization tools and their expected use in close reading was another factor we tried to capture in the pre-test questionnaire. The majority of the participants were not aware of TEI encoding and visualization tools based on these encodings.
before taking part in this user study. The participants were offered the choice to skip TEI specific sections in the pre-test questionnaire, and respond directly to the visualization tool expectation statements. However only one participant explicitly skipped TEI related questions, which suggests the participants were fully engaged in completing the study and assigned tasks to the best of their ability.

**Tasks and Evaluation Methodology**

The participants in the study were provided access to slides and screen-casts of the *Myopia* tool through a website. The screen-casts provided overview of the tool and its functionality. As the tasks in the study are based on the resources contained in the tool, some familiarity of the tool layout and functioning is desirable. This approach was adopted as individual installation and access to the *Myopia* software could have been technically challenging for some participants for purpose of this study. Moreover the pre-experimental study design utilizes tasks performed in a controlled environment. The overview slides and screen-casts in conjunction with the defined tasks provides the user with enough resources to help evaluate the effectiveness, adequacy and usability of the *Myopia* visualization tool. The sliding Lickert scale used to capture user feedback suffers from the potential problem of a pattern of similar questions eliciting responses without much thought from the respondent. We have attempted to avoid this problem by mixing the tone and tenor of questions in a section, and alternating negative and positive questions.

**Pilot Study**

The initial batch of participants in the study pointed out problems with accessing the hand-outs and audio visual material associated with the tasks. Feedback also suggested that the Task descriptions needed to be well explained in order to obtain
full participation and ensure more complete responses. We also modified the Pre-test and Post-test questionnaire in light of the Pilot Study to better align the two sets of surveys. These changes increased the number of questions in the Pre-test survey, and made them more in line with the Post-test questions.

4.3.2 External Validity

Participants

The participants in our pre-experimental study were from the university student body, primarily graduate students in the English Department, not specializing in Poetry. The motivation of participants could be affected due to them being unfamiliar with TEI encodings and the usefulness of getting a composite visual from these encodings. Also, if participants have had no prior exposure to guidelines or other tools based on TEI, they might not realize the benefits of the Poetry Visualization tool in comparison with other tools. However, the feedback from the Post-test questionnaire and Post Experiment comments suggest that the participants did see value in developing such a tool. A participant noted,

“This program is useful for metered poetry."

While two-thirds of the respondents thought visualizing encodings was a good idea.

The study results are generalizable despite lack of participants background in TEI and Poetry criticism. We also note that majority of the participants were able to successfully complete allocated tasks in reasonable amount of time.
Chapter 5

Conclusion

5.1 Discussion

In our work we developed a multi-dimensional visualization tool, which helps in the process of close reading. The Myopia tool utilizes TEI encoded texts and renders visual metaphors based on rhythmic and structural elements in these encodings. The encodings are spread across multiple files due to the non-interleaving and hierarchical nature of XML. Our tool fuses multiple perspectives which helps in literary analysis of the underlying source text. In developing this Information Visualization tool to represent differently encoded poetry, we have developed a suitable presentation environment, and facilitated real time user interaction. This approach has helped towards the goal of visually differentiating how marked up versions of an encoded poem compare to each other. We also seek to identify hotspots in the poem, which always change in each of the representations. Our hypothesis is that an interactive, multi-dimensional visualization tool based on TEI encoding is effective and practical in close reading of the source text. Additionally the visual representation can help validate and highlight problem areas in such encodings. We conducted a user study
to evaluate the effectiveness of our approach in helping the process of close reading. To achieve this objective we utilize the Pre-Experimental Design methodology. The experiment seeks to measure the effectiveness, adequacy and usability of the Myopia Poetry Visualization tool. The study is composed of an experimental run consisting of a Pre-test questionnaire, followed by three assigned tasks. The participants also provide response to a Post-test questionnaire which records their impressions of the tool and the assigned tasks themselves. The data collected from these questionnaires and the performance in the assigned tasks suggests that the participants were able to utilize the tool to help in close reading activities. The three research questions we had framed to test our hypotheses was answered positively, and we can conclude that participant expectations from a visualization tool developed to help in close reading was met by Myopia.

## 5.2 Future Work

Currently the TEI encodings in the Poetess Archive, and by extension the Myopia tool are confined primarily to metrical structures. An immediate addition to the tool is to support encodings based on Syntax and Sound structures for all the visualized texts. This would greatly enhance the usefulness and visual impact of the tool.

The feedback from the user study participants suggests that they had limited success with using the tool for completing Task 1. As Task 1 is critical to address portions of research question RQ1, future work needs to focus on rectifying this reported shortcoming in the tool. Myopia is also designed to support 3D visualizations from the outset, however the required hardware and environment might not be readily available for some users. This limitation can be overcome by deploying the tool as a web-based application, in addition to porting to mobile platforms.
The *Myopia* tool utilizes XML resources bundled within the tool itself. These encodings are static, any encoding changes have to be uploaded separately. A potential enhancement is to allow users to edit XML encodings from within the tool, and immediately visualize and validate the changes.

Feedback from the User Study is a good source of enhancements for future releases of the tool. The responses to the Post-test questionnaire will be utilized to prioritize features users would like to see in the *Myopia* tool.
Bibliography


Appendix A

User Guide
The Poetess Archive at Miami University collection includes a database of electronic documents encoded using the TEI (Text Encoding Initiative) schema and an extended tag-set derived from widely used terms in literary analysis and criticism. The Poetry visualization tool seeks to integrate multiple encodings while allowing comparative analysis of encoded text.
Contents

Overview........................................................................................................................................3
  Loading Poetry Encodings........................................................................................................4
Exiting the Tool................................................................................................................................4
Set Active Encoding.....................................................................................................................6
  Meter Encoding........................................................................................................................6
  Working With Meter Elements....................................................................................................7
  Meter Elements – Additional Information....................................................................................9
  3D Syllable Visualizations........................................................................................................9
  Additional Tropology Attributes...............................................................................................12
Popup/Hover Tropological Information.......................................................................................12
Alternate Views............................................................................................................................12
Navigation.....................................................................................................................................14
Navigate With Loaded Poems Drop Box...................................................................................15
Load and Play Audio Resources.................................................................................................16
  Bundled Poetry Resources..........................................................................................................19
Overview

The Poetry Visualization tool allows for multidimensional representation of TEI encoded Poetry and text. Presenting an interactive visual representation of differently encoded versions of text, the tool seeks to amplify understanding and uncover new knowledge.

The User Interface is split into three distinct visual regions

- The **Main Visualization Area** to display and interact with the rendered text
- A **Key/Legend Panel** at bottom of the visualization panel which holds a tabbed display of key associated with the loaded visualization
- **GUI Panel** at the right of the Visualization panel. This allows loading and manipulating source TEI documents and the resulting 3D visual elements.
Loading Poetry Encodings

User starts by selecting an entry from the Hierarchical set of Categories in “Load Meter Encoding” section. The TEI encoded poems have been categorized into 12 types (see Appendix). Each category holds one or more poems with Metrical encodings following the TEI schema. Clicking on name of a poem “Loads” the underlying Xml document in the Poetry Visualization tool.

Remove/Reset Loaded Encodings

A maximum of three (3) poems can be loaded in one panel of the visualization tool. In order to change the loaded encodings or to simply restart the visualization without exiting the tool, use the “Reset Encodings” button in the GUI Panel.
Exiting the Tool

User can exit from the visualization tool at any time

- Make GUI Panel as the active panel by clicking on it
- Selecting **File->Quit** from the main Window
**Set Active Encoding**

Meter is default encoding supported by the tool as it is most widely available TEI encoding in the Poetess Archive. Loading a particular poem always brings in the Metrical version initially, and this is the default visualization of the tool as well.

**Meter Encoding**

The “Set Encoding” section of the GUI panel is comprised of two areas:-

- Meter Encoding
- Tropological Encoding

The Meter Encoding Checkbox is initially checked allowing the Meter visualization to be displayed.

User can affect the main Visualization panel elements utilizing this section of the GUI panel.

- Display Only Tropes of a particular poem
  - Uncheck “Meter Encoding”, Check “Tropological Encoding”
- Display Both Encodings
  - Check “Meter Encoding”, Check “Tropological Encoding”
- Text Only
  - Uncheck “Meter Encoding”, UnCheck “Tropological Encoding”
Working With Meter Elements

The *Graphical Elements* section allows the user to selectively turn on and off Metrical Feet, Syllable representation as well as Text. These visual elements can be manipulated while the user is in the Meter Encoding visualization.

![Anthem for Doomed Youth](image)

**Meter Feet Not Displayed**
Anthem for Doomed Youth

1

What passing bells for these who die as cattle?

Only the monstrous anger of the guns

Only the stuttering rifles’ rapid rattle

Can patter out their hasty orisons.

No mockeries now for them: no prayers nor bells,

No darklingalleys save the choirs.

The shrill, demoted choirs of waiting shells.

Syllables Turned Off

Display Graphical Elements Only
Meter Elements – Additional Information

Feet and Syllable representation are the dominant elements in display of Meter encoding in the tool. Hovering over the text segments in the visualization, the user can access additional information based on underlying metrical structure.

3D Syllable Visualizations

A visually intuitive representation of Long and Short syllables in the loaded encoding is available in the Meter Visualization. In order to turn on enhanced depth perception and 3D font capability click the “Stressed Syllable” checkbox. The Long text segments appear raised as compared to the Short, the effect is more pronounced if the user is accessing the tool in a 3D environment.

The “Stressed Syllable” must be turned on before loading the desired meter encoded poem.
Trope Encoding

Visualization of Tropology features (Literary Figures, Literary Image, Ambiguity, Negation, Connotation, Emotion and Agency) is possible using this tool.

From the Set Encoding section click on the “Tropological Encoding” to overlay the text with graphical Tropology symbols.

Currently only tropology encoding for the John Keats poem “Ode To a Grecian Urn” is included in the Visualization tool. Selecting the “Tropological Encoding” option has no effect if poems other than the Ode are loaded.
Additional Tropology Attributes

Tropology features in a poem, such as Theme are encoded as a singular tag in Xml. However the theme itself varies even though there is just one underlying type.

Such informational complexity is presented with help of Popup text (tool tips) as user hovers over the concerned tropological element.

Alternate Views

User might be interested in viewing plain source text for the loaded encodings within the tool itself. “Load Text and Xml” section of the GUI panel allows this functionality.

Checking the “Display Text” box overlays the current visualization with a text editor view containing all the plain text poems currently being visualized.

A user might desire to see the underlying Xml or TEI encoding of the poem being visualized. Checking the “Display Xml” box in the “Load Xml Section” allows this functionality. The Xml editor does not currently allow in place editing of the code.

De-selecting the respective check boxes can dismiss Text and Xml viewers.
Ah Sun-flower

Ah Sun-flower, weary of time,
Who countest the steps of the Sun,
Seeking after that sweet golden shine,
Where the traveller’s journey is done,
Where the Youth and world both die,
And the pale Virgin shraddled in snow,
Are from their graves and aspire,
Where my Sun-flower wishes to go.

Dying Speech of an Old Philosopher
Navigation

The Poetry Visualization tool allows up to three distinct poems to be visualized at any given time. If the content does not fit the viewable area then Vertical and Horizontal scroll bars can be used as in any other GUI based tool. Zoom In/Out feature is available as well.

A direct method of navigating to a loaded visualization can be availed using the “Loaded Poems” drop box. By selecting the desired poem the user moves the active visual pane to that visualization.
Multiple Poems in Overview Mode

Navigate With Loaded Poems Drop Box
Load and Play Audio Resources

The “Loaded Poems” section and drop box performs a navigational function outlined above – It also allows loading audio resource pertinent to the selected Poem in the drop box.

The audio controls in this section are only enabled if an audio file for the selected poem is available within the Visualization tool’s resource folder. The audio controls feature the familiar Play, Pause and Stop functions.

Audio Controls
Overview Mode

The “Overview” section in the GUI panel allows the user to get a birds eye view of the loaded poems and associated layers of encodings. This abstract view can be helpful in analysis of the underlying textual content, in this case - poetry.

Multiple Encodings Overview Turned On
Hot Spot Viewer

The hotspot viewer is an abstraction of combined TEI tags used in Meter and Tropological encoding. This visualization is derived from rendering the underlying encoded text as a matrix, with each character represented initially as number 0. For meter encoding each occurrence of Metrical feet is marked in the matrix by incrementing the corresponding cell value. Similarly each tag location in trope encoding results in the matrix abstraction being updated.
Appendix

Bundled Poetry Resources (A Partial List)

- **Sonnet**
  - *Holy Sonnet 1* John Donne
  - *Holy Sonnet 14* John Donne
  - *Mowing* Robert Frost
  - *Bright Star* John Keats
  - *Renouncement* Alice Meynell
  - *On His Blindness* John Milton
  - *Anthem for Doomed Youth* Wilfred Owen
  - *Sonnet 29* William Shakespeare
  - *Ozymandias* Percy Bysshe Shelley
  - *The Sonnet* William Wordsworth

- **Roundel**
  - *The Roundel* Algernon Charles Swinburne

- **Couplet**
  - *Epigram Engraved on the Collar of a Dog Which I Gave to His Royal Highness* Alexander Pope

- **Quintet**
- Ah Sun-flower William Blake
- Jordan George Herbert
- Song Edmund Waller

- Sicilian Quatrain
  - Dying Speech of an Old Philosopher Walter Savage Landor

- Sestet
  - Meeting at Night Robert Browning
  - My Sweetest Lesbia Thomas Campion
  - Life of Life (from Prometheus Unbound II.v.48-71) PB Shelley

- Octave
  - God’s Grandeur Hopkins
  - Though I am Young and Cannot Tell Ben Jonson
  - Even Such Is Time Walter Raleigh
  - He Wishes For the Clothes Of Heaven W.B. Yeats

- Spenserian Stanza
  - Adonais (stanzas 54-55) Shelley
  - Despayre in Praise of Suicide (Faerie Queene 1.9.39-40) Edmund Spenser

- Stanza
  - Jabberwocky Lewis Caroll
  - Ode On a Grecian Urn Keats

- Blank Verse
  - Hyperion I.1-14 Keats
  - Chorus (from Henry the Fifth III.i.1-17) Shakespeare

- Quatrain
  - Westron Wynde Anonymous
  - The Sick Rose William Blake
  - When a Man Hath No Freedom George Gordon, Lord Byron
  - The Brain Is Wider than the Sky Emily Dickinson
  - Since There's No Help Michael Drayton
  - A Slumber Did My Spirit Seal William Wordsworth
  - from In Memoriam; section 7 Alfred Lord Tennyson
  - Sonnet 73 William Shakespeare
  - Sonnet 18 William Shakespeare
  - The Lowest Place Christina Rossetti

- Tercet
  - Upon Julia’s Clothes Robert Herrick – 2 Encodings
Appendix B

User Study Handout
The Poetess Archive at Miami University is a collection of late 18th and 19th century’s popular poetry. The collection includes a database of electronic documents encoded using the TEI (Text Encoding Initiative) schema derived from widely used terms in literary analysis and criticism.

The Myopia Poetry Visualization tool allows a user to view different structural features within poems based on concepts of close reading. Myopia seeks to integrate these TEI based encodings while allowing comparative analysis of encoded poems. The tool presents literary critics and encoders with an easy-to-use environment to aid in the close reading of such texts.

What the Study is about?

We have designed a set of questions and specific tasks to aid in judging the effectiveness and utility of the Myopia tool. The experiment concerns a student or scholar using the Poetry visualization tool to visualize and compare various encodings utilizing an interactive interface.

What you are asked to do?

You are being asked to participate in a research study to judge effectiveness of the Myopia Visualization tool. You will interact with the Visualization through screen-casts and screen-shots (the .png files, called “slides”). You will be asked to perform 3 tasks. You should take about 20 to 30 minutes to perform each of the three tasks. There is a questionnaire before and after the experiment.

Your Answers will be confidential:

The records of this study will be kept private. In any sort of report we make public, we will not include any information that will make it possible to identify you. The experiment does not collect any information, which can potentially identify the participants.

Taking part is voluntary:

Taking part in this study is completely voluntary. You may skip any questions that you do not want to answer. If you decide to take part, you are free to withdraw at any time.

If you have questions:

The researchers conducting this study are Manish Chaturvedi and Gerald Gannod. Any questions regarding this experiment may be directed at chaturm@muohio.edu

Statement of Consent: I have read the above information, and have received answers to any questions I asked. I consent to take part in the study.

Your Signature ______________________ Date ____________
Your Name (printed) __________________________

This consent form will be kept by the researcher(s) for at least two years beyond the end of the study.
Poetry Visualization Experiment

Introduction

Standards for encoding text and other data types were created to structure the data from various domains into machine-readable representations. These standards are flexible to accommodate the range of data possibilities but also adhere to strict guidelines in order to establish machine-readable content. While using standards has many advantages, those wanting to analyze the text encoded by the standard are often faced with a complex and time-consuming task.

Our approach to address these barriers to close reading and literary analysis of encoded text is to present an interactive visualization tool with an intuitive interface. The Myopia Poetry Visualization tool utilizes a subset of digitized resources from the Poetess Archive and presents literary critics and encoders with an easy-to-use environment to aid in the close reading of such texts.

The experiment concerns a student or scholar using the Myopia tool to visualize and compare various encodings utilizing an interactive interface.

The underlying standard in this experiment is the Text Encoding Initiative (TEI).

From the TEI site:

*The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics.*

You will interact with the Visualization tool either through screen-casts and screen-shots (the .png files, called “slides” below). You will be asked to perform 3 tasks.

You are expected to take about 20 to 30 minutes to perform each of the three tasks.

We kindly ask you:

* to write down your answers in a legible way;
* not to consult any other participant during the experiment;
* to perform the tasks in the specified order;
* to write down the current time before starting each task and after completing all the tasks;
* not to return to earlier tasks because it affects the experiment;
* to fill in the required information for each task and give additional information, if requested.

There is a questionnaire before and after the experiment. Thank you for your participation in this experiment!

Manish Chaturvedi and Gerald Gannod
Pre-Test Questionnaire

General Survey

In this short survey a number of questions regarding your experience and attitude towards TEI comprehension and Close Reading will be asked to get an impression of your skills and expectations.

1. Please answer these questions about your personal background. These answers will be kept private and only serve to put other responses in context.

What is your education background (e.g. field of study)?

What is your current job/education level?

Note: Please skip to Question 5 in the Pre-Test questionnaire if you wish to not respond to TEI specific questions.

2. This part of the questionnaire contains questions about your knowledge of TEI. Please rate each statement about your personal experience according to the following scale:
1-None, 2-Beginner, 3-Knowledgeable, 4-Advanced, 5 Expert

<table>
<thead>
<tr>
<th>Learning TEI Guidelines</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Enabling technologies of TEI (e.g. XSD, XML, XSLT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using TEI support resources (e.g. wiki, tools, software)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developing custom TEI schemas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encoding text using TEI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Pre Test Questionnaire

Some statements regarding TEI Visualization activities are shown below.

Please, rate each statement on a scale from 1 to 5 to indicate to what extent they apply to you. Mark an “X” to indicate your choice under the appropriate rating.

1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5 Strongly Agree

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>When viewing an unfamiliar encoded text, I:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use tools or software to try to visualize or describe the encoded text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use XML based technologies such as XSLT to transform unfamiliar text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use software tools to visually validate TEI encoding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rely on XML schema editors to view and validate encoding of unfamiliar text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use visual tools to spot differences in encodings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consult with other TEI user for help and guidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TEI is an XML based standard that allows exchange of machine-readable documents. The TEI encoded documents can be enhanced and transformed into more complex data formats.

Visualizations based on the TEI encodings leverage these strengths.

If you are familiar with TEI, rate each statement on a scale from 1 to 5 to indicate to what extent they apply to you. Mark an “X” to indicate your choice under the appropriate rating.

1-Strongly Disagree, 2-Disagree, 3-Neither Agree or Disagree, 4-Agree, 5 Strongly Agree

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of TEI is a valuable asset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning to encode texts using TEI is a worthwhile exercise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visualizing TEI encodings is helpful in gaining additional insight into source texts like Poetry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pre Test Questionnaire

5

Please rate the following statements from
1 (Strongly Disagree to 5 (Strongly Agree)

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have no knowledge of TEI and its uses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encoding and Visualizing texts is a good idea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The concluding section of this pre-test survey is about your expectations for using a visualization tool in aid of close reading.

“Close Reading involves focus on the connotation of the text, not simply what the words literally mean. The critic looks for textual hotspots, where connotations converge in the text. The hotspots might point to a unifying idea behind the poem, and help in uncovering new ideas about the work”.

Please rate the following statements from 1 (Strongly Disagree to 5 (Strongly Agree)

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Visualization tool can aid in close reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visualizations of encodings can be subjective and abstract limiting usefulness of such a tool</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encodings without visual metaphors and transformations are difficult to validate and learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encodings are difficult to compare across versions. A visualization tool would help with this task.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close reading is only worthwhile if source text is directly utilized in the process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Which features do you think a text visualization tool should have, that will simplify or ease the process of close reading?
Experiment

To help you get familiar with the Poetry Visualization tool and options, and complete the tasks outlined in this experiment we have provided a set of screen cast movies and screen shots of the visualization tool.

The Screen-cast movies required to complete the tasks are as follows-

1. VizToolOverview
2. EncodeCompare
3. StressPatterns

Following Screen Capture images are required to complete the tasks in this experiment-

1. Herrick 1
2. Herrick 2
3. Stress Pattern 1
4. Stress Pattern 2
5. Emily MindIsWider

These resources are to be retrieved from the webpage at

http://www.users.muohio.edu/chaturm/MovieHost.html

Please access these screen-cast movies and images AFTER completing the Pretest Questionnaire section.

Once you have an idea of the visual elements and layout of the tool, please continue with the given tasks.
Tasks

Starting Time: ________________

Task 1 – Find Metrical features and Identify Encoding Patterns

Slides used in Task1:

EmilyMindIsWider.png
Herrick1.png
Herrick2.png
StressPattern1.png
StressPattern2.png
The effectiveness of the Poetry visualization tool can be gauged by the ease with which similar poetic texts can be located and analyzed. The metric of similarity in this context is the underlying TEI encoding and the resultant visualizations in the *Myopia* tool.

**Use the Screenshot slides:**

*EmilyMindIsWider.png*

*Herrick1.png*

*Herrick2.png*

*StressPattern1.png*

*StressPattern2.png*

**Please perform the following task:**

Looking over all the pictures, identify the poem, which employs a constant meter throughout and displays an exact convergence of rhythm and meter. Please identify the meter. Utilize the Key legend panel (The KEY) for the last step.

…………………………………………………………………………………………………

……………………………………………………………………………………………

………………………………………………………………………………………………

Can you identify the two distinct line lengths used in this poem (Monometer, DiMeter etc.) From the Visualization panel?

Name the line lengths used in the poem.

…………………………………………………………………………………………………

……………………………………………………………………………………………

………………………………………………………………………………………………

Please record the first line of the poem, which is NOT a Tetrameter.

…………………………………………………………………………………………………

……………………………………………………………………………………………

………………………………………………………………………………………………

……………….

**Current Time:** ________________

8
Task 2a – Visualize and Validate TEI Encodings

Starting Time: ________________

Slides used in Task 2a&b:

Herrick Encoding 1.png
Herrick Encoding 2.png
Metrical and Rhythmic elements in a poem can be encoded with any set of pre identified symbols.

The following extract is from a Sicilian Quatrain by Landor:

The above visualization might have been encoded using — for short syllables and ′ for long in the following fashion:

_ ........ ___ ........ ___ ........ ___ ........ ___ ........
I strove with none. for none was worth my strife:

Task:
Encode the text of Herrick’s Poem Upon Julia's Clothes given below.

This encoding can be done by hand using pencil and paper and consists of marking stressed/unstressed syllables and meter using — for short syllables and ′ for long:

Upon Julia's Clothes

Whenas in silks my Julia goes,

Then, then (methinks) how sweetly flows

That liquefaction of her clothes.

Next, when I cast mine eyes and see

That brave vibration each way free,

O how that glittering taketh me!
Task 2b Compare and Contrast TEI Encodings

Compare the two versions of the poem “Upon Julia’s Clothes” depicted in screenshots Herrick Encoding 1 and Herrick Encoding 2.

Complete the following section based on the meter encodings of “Upon Julia’s Clothes”:

Locate and report metrical feet dissimilarities between the two versions (line number, meter and text)

The encoding for the two versions of the poem are similar in that they replace a regular Iambic Foot with a 3 syllable foot to convey a sense of sparkle and twinkling.

Using the visualization slides for Herrick’s poem identify and record below, the line where you think this replacement occurs.

Ending Time: _________________
Task 3 Identify Stress Patterns from Visualizations

Slides used in Task 3:
- StressPattern1.png
- StressPattern2.png

Meter is a recurring pattern of stressed (long) and unstressed (unaccented, or short) syllables in lines of a set length. Identifying Stress patterns in a poem is an important part of literary analysis.

With help of the screen shots and screen casts provided, Identify and List the Stressed Syllables in “Rhyme for a Child Viewing a Naked Venus in a Painting….."

Perform this task using:

a. The Graphical Element corresponding to Syllables (refer to Key Legend in the picture StressPattern1.png)

b. The “Stressed Syllable” view with 3d text elements in StressPattern2.png

Which method between a & b appears to be more intuitive and preferable to analyze unfamiliar verse and poetry?
# Post-Test Questionnaire

## Experiment Evaluation

Thanks for completing the tasks! To get a summary of your experiences with the Poetry Visualization tool and to allow you to give your comments, please respond to the following questions.

The first part is about your overall experience in performing the experiment. Please rate each statement on a scale from 1 to 5 to indicate to what extent they apply to you.

1-strongly disagree, 2-disagree, 3-neither agree or disagree, 4-agree, 5-strongly agree.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall, the tasks were reasonable to perform</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I felt there wasn’t enough time for each task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I felt unsure about how to complete some or all of the task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Poetry Visualization slides were hard to follow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I felt the tasks did not resemble a realistic purpose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This part is about your experience with the Visualization tool for each of the tasks.

Please indicate every task for which the following statements apply.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2a</th>
<th>2b</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Visualization tool is useful for this task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During this task I validated my encoding for a poem</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Stress Patterns can be identified in this task

Before the experiment you were asked about your expectations for a Visualization tool.

Please rate each statement on a scale from 1 (strongly disagree) to 5 (strongly agree) to indicate to what extent they apply to you.

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Myopia Visualization tool is not useful for purpose of Close reading.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstract or Graphical representation of meter and feet in Myopia was helpful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encodings without visual transformation by the Myopia tool are difficult to learn and validate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encodings are difficult to compare across versions. The Myopia tool helped in this task.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close reading is possible since Myopia displays source text as well as visual metaphors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This part is about the usefulness of individual features of the Myopia tool. Please rate how useful the features were on a scale from 1 to 5:

1 - very useless, 2 - useless, 3 - neither useless nor useful, 4 - useful, 5 - very useful, 6 – Not Applicable.

<table>
<thead>
<tr>
<th>Feature</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare encodings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validate Encodings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identify quickly the rhythm and syntax elements of Poetry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Statements in this part relate to the usability of the *Myopia* Visualization tool.

Please rate the following statements on a scale from 1 (strongly disagree) to 5 (strongly agree).

<table>
<thead>
<tr>
<th>Statement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>I found the visual layout of the tool easy to navigate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I could potentially use the tool as an aid to close reading and analysis of Poetry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The visualization tool is a encoding help, not meant for literary critics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visualizing encodings is a great idea.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This tool does not satisfy the need to visualize encodings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Post-test Questionnaire

Enter comments and/or suggestions you may have about the **Myopia Visualization tool**. These will be useful for improvements.

Enter comments and/or suggestions you may have about the **experiment**. These will be useful for future studies.

THANK YOU!
Appendix C

Myopia Tool Installation
Myopia Tool Installation Instructions

Mac OS X
1. Download and install Panda 1.7.2 from http://www.panda3d.org/download.php?sdk&version=1.7.2

Nvidia needs an account to be setup before they allow the download. Above steps involve running the downloaded dmg, and accepting the default settings.

Download and Unzip the Myopia source file. Change Directory to the base directory (MyopiaPoetryVizToolv1.2) and execute ./runPViz.sh

Windows
- Download Panda
- Download wxPython (win32 bit, unicode version even if on 64-bit windows)
- Download Myopia Source Code
- Edit Run Myopia Poetry Tool.lnk in Myopia root directory
- Update Panda install Path
- Update Start In directory to Myopia install directory