

## *Abstract*

This paper describes a prototype system that allows readers to view an electronic text in multiple simultaneous views, providing insight at several different levels of granularity, including a reading view. This prospect display is combined with a number of tools for manipulating the text, for example by highlighting sections of interest for a particular task. The result is a powerful approach to working with electronic text for various purposes: sample scenarios are outlined involving directors reading scripts, students studying novels, and second-language learners familiarizing themselves with grammatical constructions.

# *M u l t i - l e v e l D o c u m e n t V i s u a l i z a t i o n*

Stan Ruecker, Eric Homich & Stéfan Sinclair

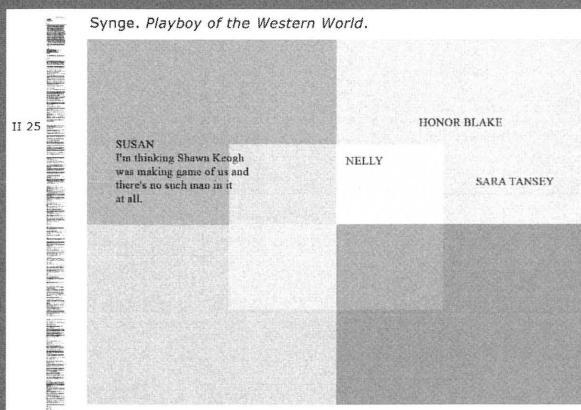
## *Introduction*

Digital text offers software developers and designers the opportunity to provide readers with a variety of new perceptual experiences and possibilities for action that have simply not been available through printed texts (Bork, 1983). An obvious example is the widespread adoption of digital texts connected by hyperlinks and identified by many theorists as a significant change in the way people are able to interact with the written word (Bolter, 1991; Landow, 1994, etc.). However, many other new affordances of digital text remain to be identified, developed and studied. One of these possible new affordances is the ability to have text or layout features change over time (Chang et al, 1988; Ford et al, 1997). In kinetic text research, traditionally static design elements such as font, size, leading, color and placement can all be used dynamically to achieve layout effects that were previously available only in non-interactive media such as film (Lee et al, 2002).

This project extends research in hypertext and kinetic text theory to provide readers with a text document display that combines simultaneous prospect — an overview of the entire text — and detail views, with related tools. Much as architectural blueprints allow the person reading them to get a sense of an entire building or some key feature, such as the wiring or the ventilation, allowing readers to see an entire text at once (that is, providing text prospect) has perceptual advantages. These advantages, which we will explore in this paper, are not available in cases where the text can only be accessed sequentially. The system also includes related tools that allow the reader to carry out new kinds of actions that would not otherwise be available.

From hypertext theory comes the concept of associated text elements, where interaction with one text moves the reader into a related text. However, zooming through prospect views differs from a hypertextual implementation in that there are no pre-defined links between views. Hypertext is also predicated on the concept of connecting lexia or individual documents, so that following a link has the effect of visually replacing the source text with the destination text. In this project the text is treated as a stable whole and presented so as to minimize interruptions to the reader's literary engagement with the text (Miall, 1999).

Kinetic text theory contributes the notion of a system where text characteristics change as a way of responding to reader interests. In this case, the reader has the ability to identify the portion of the whole text that will display in the reading view. There is also the capacity to highlight specific passages in the entire text, by selecting the features from a set of choices that derive from the tagging available in the document. Finally, in cases where this system has been integrated with related digital reading tools, additional kinetic features may be possible, as in the Watching the Script prototype (Ruecker et al., 2004), where the reader views the script by watching it scroll at various character positions on stage. *figure1*

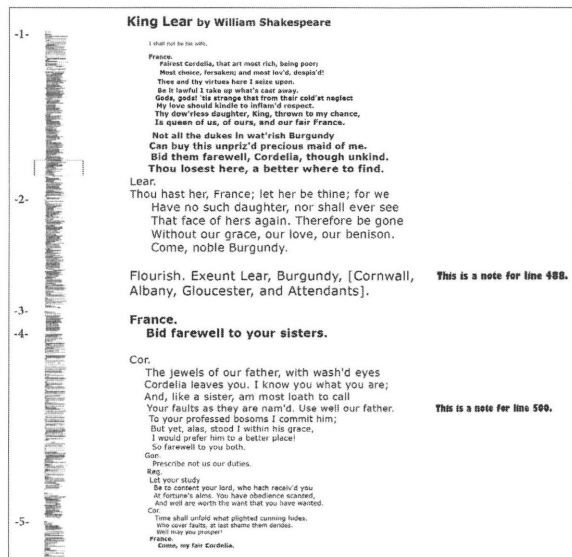


**Figure 1.** Readers using the Watching the Script prototype are able to see a microtext representation of the entire script that is associated with multiple reading views automatically scrolling at character positions on a stylized stage. This screen shot shows four characters on stage; Susan is currently speaking.

## The Multi-level Document Visualization Prototype

In the Multi-level Document Visualization prototype that we have developed, the prospect view indexes a fisheye reading view, where a segment of text of about a dozen lines is shown at full size, while adjacent text is displayed as increasingly smaller lines of microtext (Small, 1996; Furnas, 1986; Bederson, 2000). Prospect on an entire document has been a traditional component of print design, with books, for example, often containing apparatus such as the table of contents, indexes and chapter headings. However, there were inherent limitations, since the static form of the text could not be the basis for any new opportunities for action derived from tools associated with prospect. However, with a digital text there are several advantages that can be made available. These in some ways parallel the advantages gained by people using zoomable electronic maps. By zooming out, the reader is able to gain some insight into the larger terrain; by zooming in, the reader can examine details within context. In the case of digital text, however, there are further advantages that relate to the linearity of the document, as outlined below.

Firstly, the prospect view can be used as an index to the document, allowing the reader to gauge the total amount of text against the current insertion or focus point. <sup>figure2</sup> This feature is similar to the ability to gauge location in a printed book by physically judging the total number of pages against the current page. However, since in this case the text is digital, the gauge can also be used as an access method, where the reader can accurately change the current insertion point by choosing a new point on the prospect view. This capacity resembles to some extent the scrollbar and sliding thumb, the size and position of which correlate to the length of the document and the current viewing position. However, the prospect view provides additional cues to the reader through the visual presence of lines of microtext, which in some documents can help differentiate section breaks or other textual characteristics. The explicit use of the scrollbar as an analog for the entire document has also been explored by projects such as Hill and Hollan (1992), where marks were superimposed on the scroll bar as a form of interaction history, to indicate locations of reading and editing.



**Figure 2.** This version of the multi-level document visualization prototype shows three views: one of the entire document for prospect, another for reading and a third with notes. The prospect view includes a position indicator that allows the reader to quickly determine where the reading view sits within the larger document. Speeches by France are currently highlighted by bold face.

Secondly, a prospect view can be used to gain insight into the overall structure of the document and some of its characteristics. For example, a prospect view with an associated search function might allow the reader to find a particular word or phrase and see at a glance all the points where it occurs in a particular novel. If the novel has been encoded in XML, the search might also reveal segments of text that match an XPath Query (where certain markup tag names and/or encoded attributes are located). By extension of this idea, a prospect view on a play might allow the reader to select two or more characters from the cast list and see all the locations where those characters interact on stage (Johnson, 1994). Since the prospect view and reading views are connected, selecting each of the character interactions in turn provides a quick means of seeing how the interactions progress through the course of the play, without losing the larger context of the scenes in which the characters do not appear.

Thirdly, and perhaps most importantly from the perspective of the reader, the tools associated with the prospect view provide a set of new opportunities for action or affordances. The concept of affordances, developed by Gibson (1979), suggests that people learn to directly perceive what they are able to accomplish in a given environment. Designing interface affordances can therefore direct the designer in ways that are somewhat more generalized than designing functions, since one purpose can be to maximize the opportunities for action, rather than attempting to constrain them to the maximum efficiency for a single task. From this perspective, the prospect view serves as the basis for the design of new affordances. Some excellent text visualization systems have included prospect views (e.g., Small, 1996); the current prototype generalizes the capacity of such visualizations through additional affordances, largely derived from the reader's opportunity to choose any available digital text and to apply tools for visual selection and extraction that rely on characteristics of the texts, such as XML markup. Some additional affordances are also being developed, such as the provision of annotation, interaction histories and the mapping of text on stage.

The concept of working with a prospect view can be generalized by including additional levels of display. These levels could conceivably display an arbitrary number of prospect views for increasingly smaller structural sections, beginning at the level of the document collection and descending in a cascading manner into the reading details of a particular document. The levels might include, for example, in descending order of size, the text collection, author collection, genre collection, play, act, scene and currently selected lines. Aside from the full prospect and reading views, the display of the various levels depends on the use of markup in the documents as a means of expressing an ordered hierarchy of content objects (OHCO), where someone has tagged each document according to standard divisions and subdivisions (Renear et al, 1996).

Where an OHCO form of tagging has been applied, a corresponding multi-level representation has been implemented in the prototype. The reader can navigate the document by clicking in any of the displays, and the current insertion point in the text visibly changes in all of them. Features that remain to be implemented in addition to this cascading form of prospect display include any number of related tools that draw on the new perceptual opportunities at each level. For example, at the level of the document collection, the system might allow the user to sort the items alphabetically by author's last name. At the level of the author collection, the system might provide a tool for sorting the documents by date or genre. Within the genre display, an appropriate tool might allow the user to group the items by publisher, sort within publisher groups by date of publication and display the results as a set of multiple timelines.

Associated with these multiple displays are a number of tools and related features, each of which provides a new affordance. The annotation tool, for example, allows users to create and insert comments at any point in the text, which appear as marks on all the displays. As different readers each access and annotate the text, an interaction history in the form of previous readers' annotations becomes available. Some documents may have been previously encoded using a textual markup system (or "tagset") defined in Standard Generalized Markup Language (SGML) or eXtensible Markup Language (XML). One tagset that has been widely used for text collections in the humanities is the one defined by the Text Encoding Initiative (TEI), which specifies that information be provided for structural elements such as chapter breaks for novels, or divisions into act, scene and line for plays. For documents that have been encoded with a TEI-style tagset, an additional index appears attached to the prospect view, which displays these structural elements.

Through the combination of multiple simultaneous document views at different scales and a set of related tools, the Multi-level Document Visualization screen provides a dynamic reading environment that begins to demonstrate some of the unexplored promise inherent in digital text. The application of this system opens a variety of possibilities for reading and studying electronic texts, including any situation in which an advantage can be obtained by viewing a document at multiple structural levels simultaneously.

Three different scenarios will be outlined below, including one where the system would be useful to directors or dramaturges adapting a script for a specific production, another where it would benefit literary students studying novels and a third where it would be used by students attempting to acquire a second language. Each of these scenarios is intended to represent a situated application of the principle of prospect display, as it would be provided through a system like the prototype under discussion. A second phase of research will involve observing participants working with the prototype in each of the ways discussed. These scenarios are not intended to be exhaustive, but to suggest possibilities for different kinds of users with different needs. Further application of the system in other areas will also be considered in future research.

## **SCENARIO 1    READING SCRIPTS**

For a director or dramaturge working on the staging of a play, or for certain academic readers researching drama, one of the common requirements is to determine which of the characters can be played by the same actor or double-cast. Typically, three kinds of information are required. First, which characters never appear together? Second, which characters appear together seldomly (such that it might be possible to delete some of those appearances)? Third, of the characters who rarely or never appear together, which ones appear in quick succession, making costume changes difficult or impossible? These three kinds of information can be obtained by scanning carefully through the script and taking appropriate notes, but the process is time-consuming, and in complicated plots with many characters, it can be prone to error.

Using the Multi-level Document Visualization prototype, a visual analysis of a play for possible double-casting could be carried out in one of two ways. If the system included a tool that allowed the user to identify lines by character, and more than one character could be selected at the same time, the director or dramaturge would be able to run through the list of likely permutations and identify cases of all three kinds. For the third kind, where the task is to identify distance between appearances on stage of two characters who are candidates for double-casting, the user could then select instances of proximal appearance for display in the reading view, in order to determine exactly how close they are.

On the other hand, if the system were tailored for this function, it would also be possible to provide the user with a dedicated tool that would filter characters by simultaneous appearance and show only those characters who never appear together on stage or appear seldomly together according to some pre-determined threshold. Dedicated applications have been developed for this purpose (e.g., Johnson, 1994), and augmenting them with a prospect view once again allows the specific function to be readily generalized into the larger affordance of working directly in the script, rather than moving from output tables to text and back again.

As in the first method, these instances of non-simultaneous appearance could then quickly be scanned in the reading display to determine how closely together they do appear, since time is required for costume changes. It is important to exploit the reading display for this purpose, since the prospect display will in many cases provide only a logarithmic representation of the entire script, where one line of pixels represents several actual lines of text. This conversion is necessitated by screen resolutions that do not provide enough lines for a true representation.



## **SCENARIO 2    STUDYING NOVELS**

For a student of literature who is interested in obtaining an overview of a novel, the Multi-level Document Visualization system can provide structural insights based on the ability to select significant features for highlighting. For example, the student might be looking at an electronic copy of *The Great Gatsby* that has been tagged using an XML tagset with tags for metaphoric material. By choosing various keywords from the list of available tag values, the student can begin to see at which points in the novel different metaphors come into play. Extracting those sections into the reading view will allow the student to analyze the interactions between characters at these significant points in the text.

For example, one of the recurring metaphors in *The Great Gatsby* is the watching eyes, as symbolized by the signboard by the road that advertises the local optometrist. By looking at occurrences of optical references in the text, the student can quickly develop a sense of the structural and stylistic presence of this feature of the novel. If the task for the student included the need to write a paper about the significance of scrutiny in this novel, the ability to efficiently select these passages for extraction into the reading view would also provide the opportunity to place them in a file for subsequent selection of citations.

## **SCENARIO 3    LEARNING A SECOND LANGUAGE**

One of the tasks faced by people who are learning to read a new language is the ready identification of grammatical constructions of various kinds, as actually practiced by people who are proficient in the language. When working with conventional print materials, this intermediate literacy skill relies either on the ability of the student to correctly identify instances of the right type or on texts that have been prepared in advance by sophisticated language users, which can limit the student in terms of text choice.

By using the Multi-level Document Visualization system, a student of a new language would be able to apply the different views to any electronic text that had been encoded with grammatical markup. For example, an ESL (English as a Second Language) student with a philosophical inclination might wish to find instances in Smith's *On the Origin of Objects* of sentences in the perfect past. If an appropriate tool were available to identify and highlight these sentences, the student could then read quickly through the list and discover by example some of the variations that are possible in English.

Furthermore, if a tool is also available for annotating the text and storing the annotations, either the professor or the tutor could choose particular examples for discussion, then pass the text with its associated annotations to all of the students who are interested in learning through the use of that particular document.

## Conclusion

The advantages of a Multi-level Document Visualization tool include the ability for the user to obtain an overview of a text that is keyed to a reading view, and in the general case, to any number of related views at different levels of granularity. In the prototype discussed here, several levels of detail are provided. These different views are combined with tools that allow opportunities for action in relation to electronic texts that are not normally available. Potential uses of this kind of electronic reading system will be limited only by the imaginations of the users and the capacity of the designers to provide appropriate tools, but examples of the kinds of possible uses include script analyses for double casting, thematic study of novels and adoption of literature as a source of grammatical models for people attempting to acquire literacy in a new language. Future research directions will include the experimental observation of participants in each of these areas using the prototype system and the identification of potential new tools to facilitate additional user tasks.

## Author Notes

*Stan Ruecker, Ph.D. is an Assistant Professor in the Humanities Computing program at the University of Alberta. His research involves the use of prospect in interfaces for browsing tasks, text visualization and online reading. He is a founding member of the Experimental Reading Workshop and the Health Information Design Network. He has presented and published on SGML theory, affective human factors, interaction histories, the design of the electronic book and eighteenth-century literature.*

*Eric Homich recently completed his M.A. in Humanities Computing at the University of Alberta. He has a degree in Computer Science and spent several years as a programmer, database administrator and systems developer before returning to school full time. He is interested in visualizations of information, particularly non-numeric information. In September 2004, he started his PhD in the Faculty of Information Studies at the University of Toronto. For more specific information (and some fun Java applets), see his site at [www.cobble.org](http://www.cobble.org).*

*Stéfan Sinclair, Ph.D. is an Assistant Professor in the School of the Arts at McMaster University. His areas of interest include 20th Century French literature (especially Oulipo), computer-assisted text-analysis, literary databases and educational technologies. He is the creator of online Humanities Computing tools such as HyperPo <<http://huco.ualberta.ca/HyperPo/>> and SatorBase <<http://www.satorbase.org/>>.*



## References

- Bederson, B.B. 2000. "Fisheye Menus." *Proceedings of the 13th annual ACM symposium on user interface software and technology*, 217-225.
- Bolter, Jay David. 1991. *Writing Space: The Computer, Hypertext, and the History of Writing*. NY: Lawrence Erlbaum.
- Bork, A. 1983. "A Preliminary Taxonomy of Ways of Displaying Text on Screens." *Information Design Journal*, 3.3, 206-214.
- Chang, B.W., J. Mackinlay, P.T. Zelleweger and T. Igarashi. 1988. "A Negotiation Architecture for Fluid Documents." *UIST98 Conference Proceedings*, 123-133.
- Ford, S. and J. Forlizzi, S. Ishizaki. 1997. "Kinetic Typography: Issues in time-based presentation of text." *CHI97 Conference Extended Abstracts*, 269-270.
- Furnas, G.W. 1986. Generalized fisheye views. *ACM SIGCHI Bulletin, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 17.4. <http://doi.acm.org/10.1145/22627.22342>
- Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton-Mifflin.
- Hill, W. C., J. D. Hollan, D. Wroblewski, T. McCandless. 1992. "Edit Wear and Read Wear." *Conference Proceedings on Human Factors in Computing Systems*, 3-9. <http://doi.acm.org/10.1145/142750.142751>
- Johnson, Eric. 1994. Project Report: ACTORS: Computing Dramatic Characters that are on Stage Simultaneously. *Computers and the Humanities*, 28.6 December, 1994-1995, 393-400. <http://www.unix.dsu.edu/~johnsone/actors.html>.
- Juola, J.F. 1988. "The use of computer displays to improve reading comprehension." *Applied Cognitive Psychology*, 2, 87-95.
- Landow, George P. 1994. *HyperText/Theory*. Baltimore: The Johns Hopkins University Press.
- Lee, Johnny C., Jodi Forlizzi, and Scott E. Hudson. 2002. The kinetic typography engine: an extensible system for animating expressive text. *Proceedings of the 15th annual ACM Symposium on User Interface Software and Technology*. <http://doi.acm.org/10.1145/571985.571997>
- Miall, David. 1999. "Trivializing or Liberating? The Limitations of Hypertext Theory." *Mosaic* 32.2, 157-171.
- Mills, C.B. and L.J. Weldon. 1987. "Reading text from computer screens." *ACM Computing Surveys* 19.4, 329-358.
- Owen, G. Scott. 2002. "Hypergraph." <http://www.siggraph.org/education/materials/HyperGraph/hypergraph.htm>.
- Purcell, Chris. 2004. "Text Visualization." <http://www.usemod.com/cgi-bin/mb.pl?TextVisualization>.
- Raymond, D. R. 1991. "Visualizing texts." *Making sense of words: Proceedings of the Ninth Annual Conference of the UW Centre for the New OED and Text Research*. Waterloo, Ontario: UW Centre for the New OED.
- Renear, A. and E. Mylonas, D. Durand. 1996. "Refining our Notion of What Text Really Is: The Problem of Overlapping Hierarchies." *Research in Humanities Computing*.
- Rodgers, Deborah. 2001. "A grammar for zooming interfaces: Using interaction design strategies to improve user's navigation and spatial awareness." *Information Design Journal* 10.3, 250-7.
- Rockwell, Geoffrey and John Bradley. 1996. "Watching scepticism: computer assisted visualization and Hume's Dialogues." *Research in Humanities Computing* 5, 32-47.
- Ruecker, Stan, Eric Homich, and Stéfan Sinclair. 2004. "Watching the Script of Synge's Playboy of the Western World." *COCH/COSH*, (need detail) Winnipeg: University of Manitoba.
- Ruecker, Stan. 2003. Affordances of Prospect for Academic Users of Interpretively-tagged Text Collections. Ph.D. Dissertation. Edmonton: University of Alberta.
- Sinclair, Stéfan. 2003. "Computer-Assisted Reading: Reconceiving Text Analysis." *Literary and Linguistic Computing* 18.2, 175-184.
- Small, David. 1996. "Navigating large bodies of text." *IBM Systems Journal* 35, 3-4.