Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

Seeing the Text Through the Trees: Visualization and Interactivity in the Text Applications

Dr. Geoffrey Rockwell
Humanities Computing Centre
McMaster University
1280 Main St. W.
Hamilton, Ontario
Canada, L8S 4M2

John Bradley
Centre for Computing in the Humanities,
King's College London
Strand
London WC2R 2LS England

Patricia Monger
Arthur Bourns Bldg Room 131K
McMaster University
1280 Main St. W.
Hamilton, Ontario
Canada, L8S 4M2

Abstract

In this paper we discuss two interactive text visualization systems and then discuss the rhetorical effects of interactivity. The first model is SIMWeb, a data visualization system with connections to TACTweb for full-text searching. SIMWeb provides a graphical representation of the results of statistical processes that can be used to explore a text. The second experiment, Eye-ConTact is a prototype for a process visualization environment for research applications in the study of electronic texts. The paper then discusses the effects of visualization with particular attention to the contribution of interactivity to the process of textual research. We argue that this allows for pragmatic experimentation with processes and information and conclude by discussing some of the dangers of interactive visualization systems.

Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

1. Introduction

In the fourth book of Plato's Republic a story is told about the struggle of seeing.

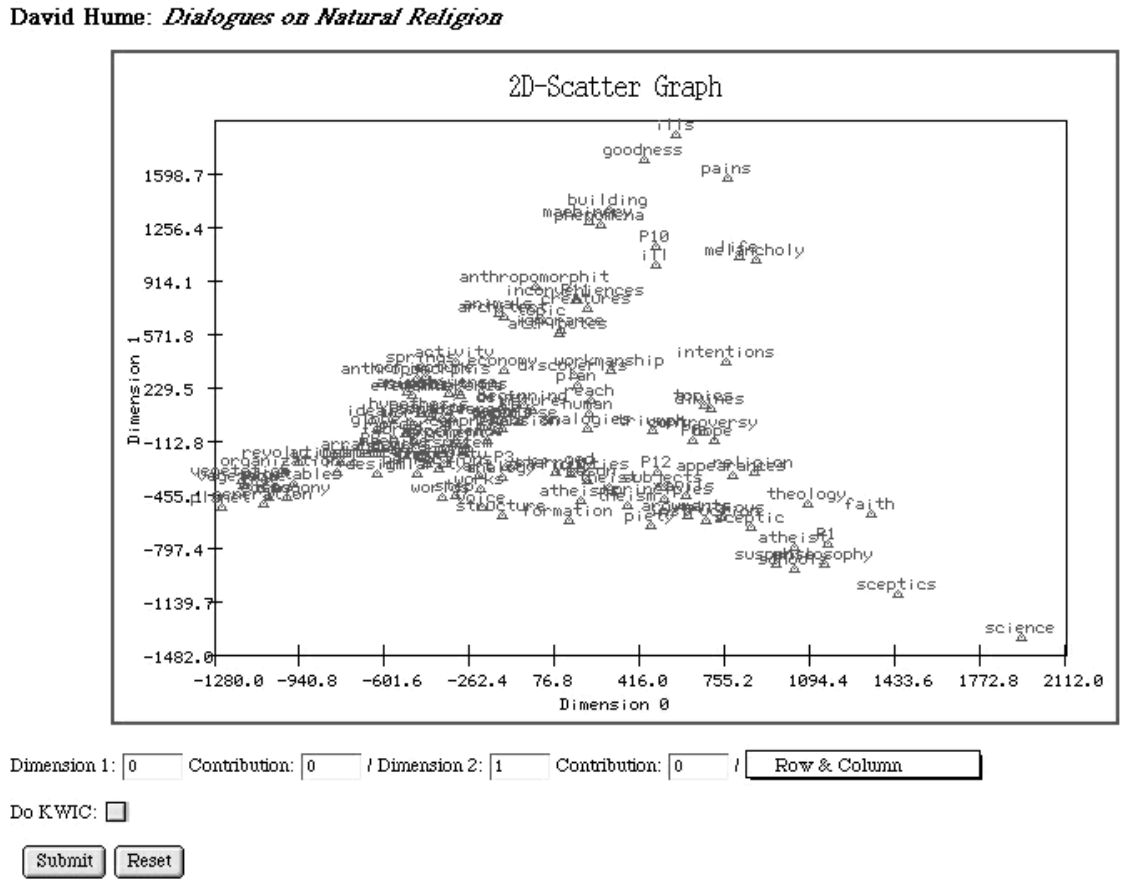
> Leontius, the son of Aglaion, was going up from the Piraeus along the outside of the North Wall when he saw some corpses lying at the executioner's feet. He had an appetite to look at them but at the same time he was disgusted and turned away. For a time he struggled with himself and covered his face, but, finally, overpowered by appetite, he pushed his eyes wide open and rushed towards the corpses, saying, "Look for yourselves, you evil wretches, take your fill of the beautiful sight!" (Plato, *Republic*, Bk. IV, 439e-440a)

We all know about the temptations of seeing. Who is not fascinated by deadly sights while feeling guilty for being tempted? This paper is about the sight and feel of a different type of corpus - the electronic text. We will start with the gruesome sights and then move to theoretical speculations about the rhetorical effects of seeing and touching such sights, a discussion which we hope will relieve you of any guilt. What is new about this paper is a discussion of interactivity, something we appreciate when it is there, but something we rarely discuss in humanities computing. We want with this paper to show two types of interactivity with an electronic text and ask what interactivity contributes to the interpretation of text. We are interested in what it means to be able to touch what you see.

2. SIMWeb

The first type of interactive visualization we will discuss, SIMWeb, was designed to test the feasibility of interactive diagrams derived from Correspondence Analysis on a text that are connected to a full text search engine.[1] This is an example of data visualization that represents the results of a statistical process so that the user can interact with it. Readers of our previous papers on this subject will recognize the images we are flogging to death. The difference with SIMWeb is that the diagram, rather than being the published end of text-analysis, is now presented as a starting point - as an interactive system that combines a text engine (TACTweb) with a visualization environment for the exploration of a text.[2] The

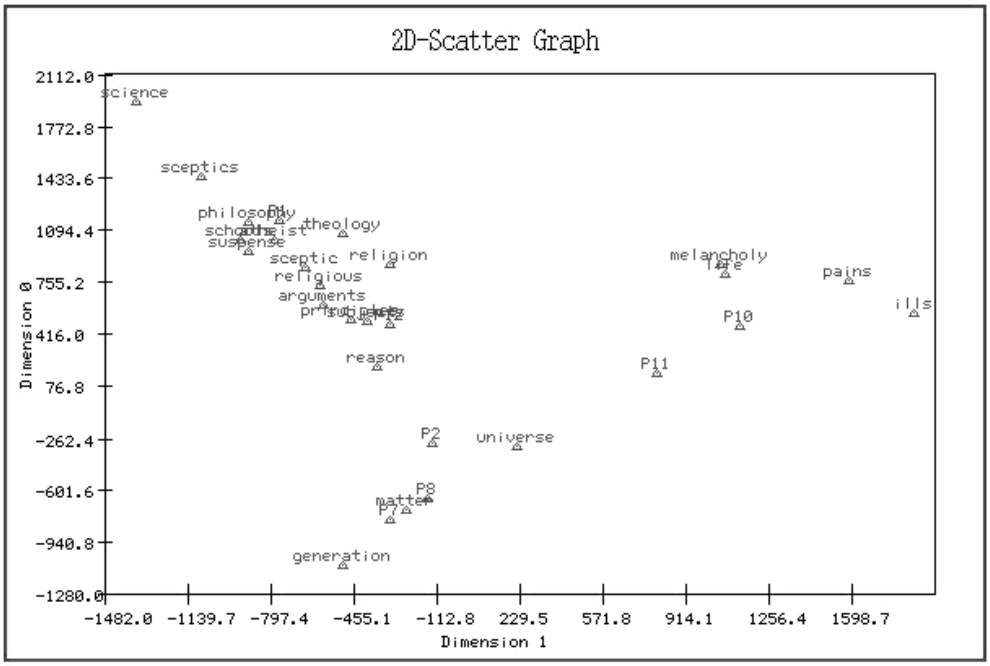
integrated environment is accessed through the WWW so that users can study the text by changing the parameters of the visualization, zooming in to parts of the visualization, and clicking on labels in the visualization to launch a text engine. In lieu of a demonstration we offer an illustrated walkthrough.



<Figure 1: SIMWeb Opening Screen - File = SIMWeb1.jpg>

Figure 1 shows the opening visualization that you see when you connect to the WWW site. You can interact with it in two ways, a) by changing the parameters that define the visualization, or b) second by direct manipulation.

David Hume: ***Dialogues on Natural Religion***



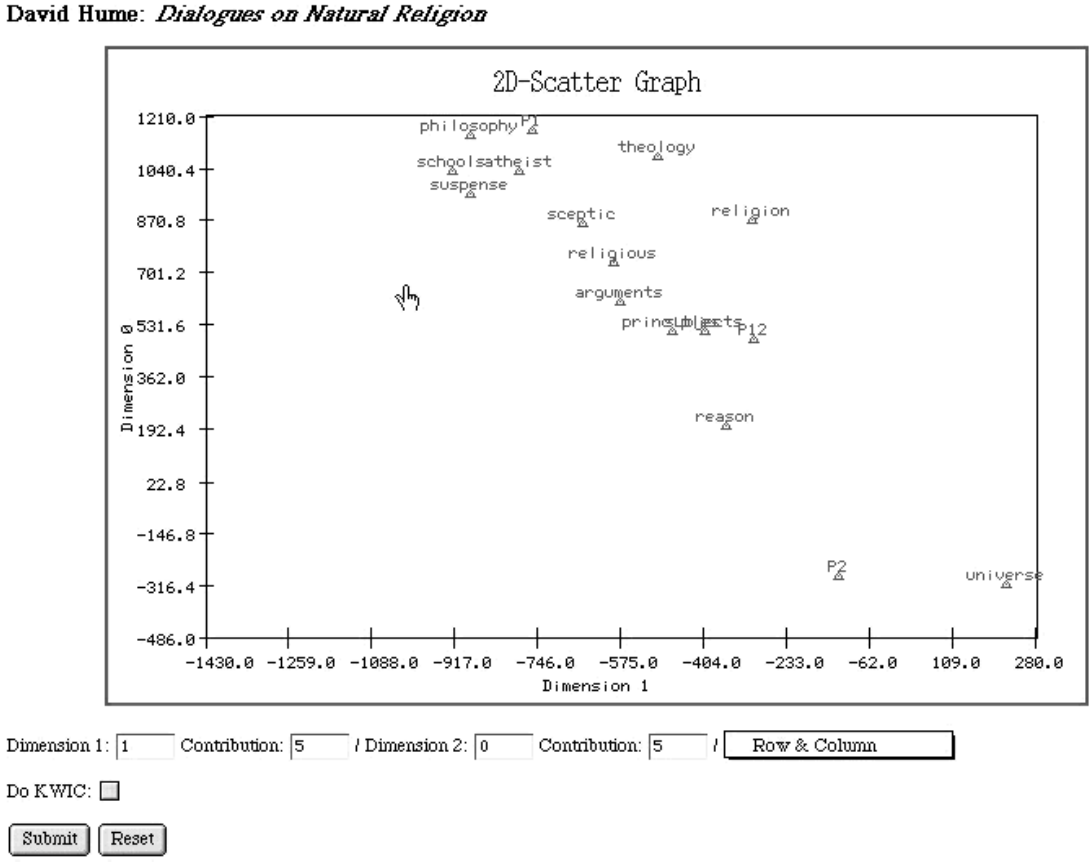
Dimension 1: 1 Contribution: 5 / Dimension 2: 0 Contribution: 5 / Row & Column

Do KWIC:

Submit Reset

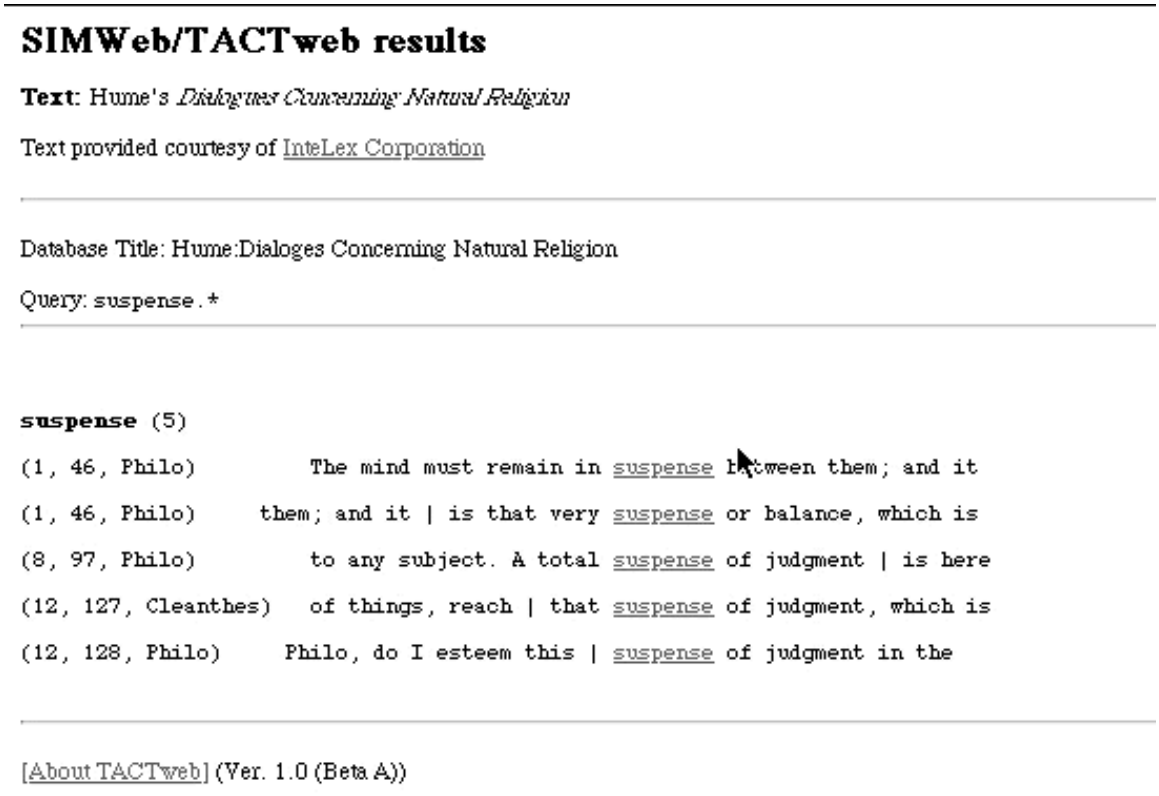
<Figure 2: SIMWeb Screen 2 - File = SIMWeb2.jpg>

By changing the parameters you can choose how to graph the dimensions generated by Correspondence Analysis and you can hide the words that do not contribute significantly to these dimensions. Figure 2 shows the same two dimensions rotated with the Contribution Levels set to 5 so that only the words that contribute significantly appear.



<Figure 3: SIMWeb Screen 3 - File = SIMWeb3.jpg>

Direct Manipulation is when you interact directly with the graphical representation. In this visualization there are two ways you can manipulate the graphic directly. The first is to zoom in by clicking where you want to the center of the resulting zoomed image to be. Figure 3 shows the result of clicking in the upper left-hand quadrant.



<Figure 4: SIMWeb TACTweb Screen - File = SIMWeb4.jpg>

A second type of direct manipulation is available when you turn on the **KWIC** setting. At this point clicking on a word (the label of one of the points graphed) in the visualization launches a text query. SIMWeb passes the word clicked to TACTweb which then searches in a text database of the original text for that word. The results of the query are presented as a Key Word in Context display where the key words are hot and can be clicked on to see the full text of the passage in question.

What SIMWeb models is the possibility for links between interactive data visualizations and traditional interactive text-analysis tools like TACTweb, which is a WWW accessible version of a subset of the functionality of TACT. This environment provides the researcher with a visualization of a statistical analysis of a particular text, in this case Hume's *Dialogues Concerning Natural Religion* which can be explored in a fashion that the tables of numbers output from a statistical analysis could not. The point of this

demonstration is not to defend the statistical analysis employed, but to demonstrate the value of interactive exploration of data and the connection between interactive visualizations and traditional text tools. In particular such visualizations offer the following features.

1. Because this interactive visualization displays the relevant words as labels in a two-dimensional space as opposed to graphing only points, the user can easily make the connection between the words used in the statistical analysis and resulting data. Many graphing tools do not easily provide for the labeling of points with the words whose position they represent.
2. Because you can change the parameters of the representation and zoom in on the results you can easily explore the data for anomalies and patterns. The researcher can thus identify the perspectives that suggest further interpretation and from there launch traditional text tools. Thus the visualization rather than being used to present results to others through publication, is used to explore the information for further study.
3. In our case you can launch a TACTweb search from any display, which means you can call up the full text so that you can easily make the connection between the graphical representation and the text represented. This allows you to confirm possible patterns or decide that what you see is uninteresting. The visualization thus leads to a rereading as it should, instead of distancing the reading.
4. Finally the ability to play with the visualization allows the researcher to explore the process that was used to generate the graph in a way that a static image would not. It allows us to test statistical methods pragmatically – do they produce results that are interesting or provocative of further study.

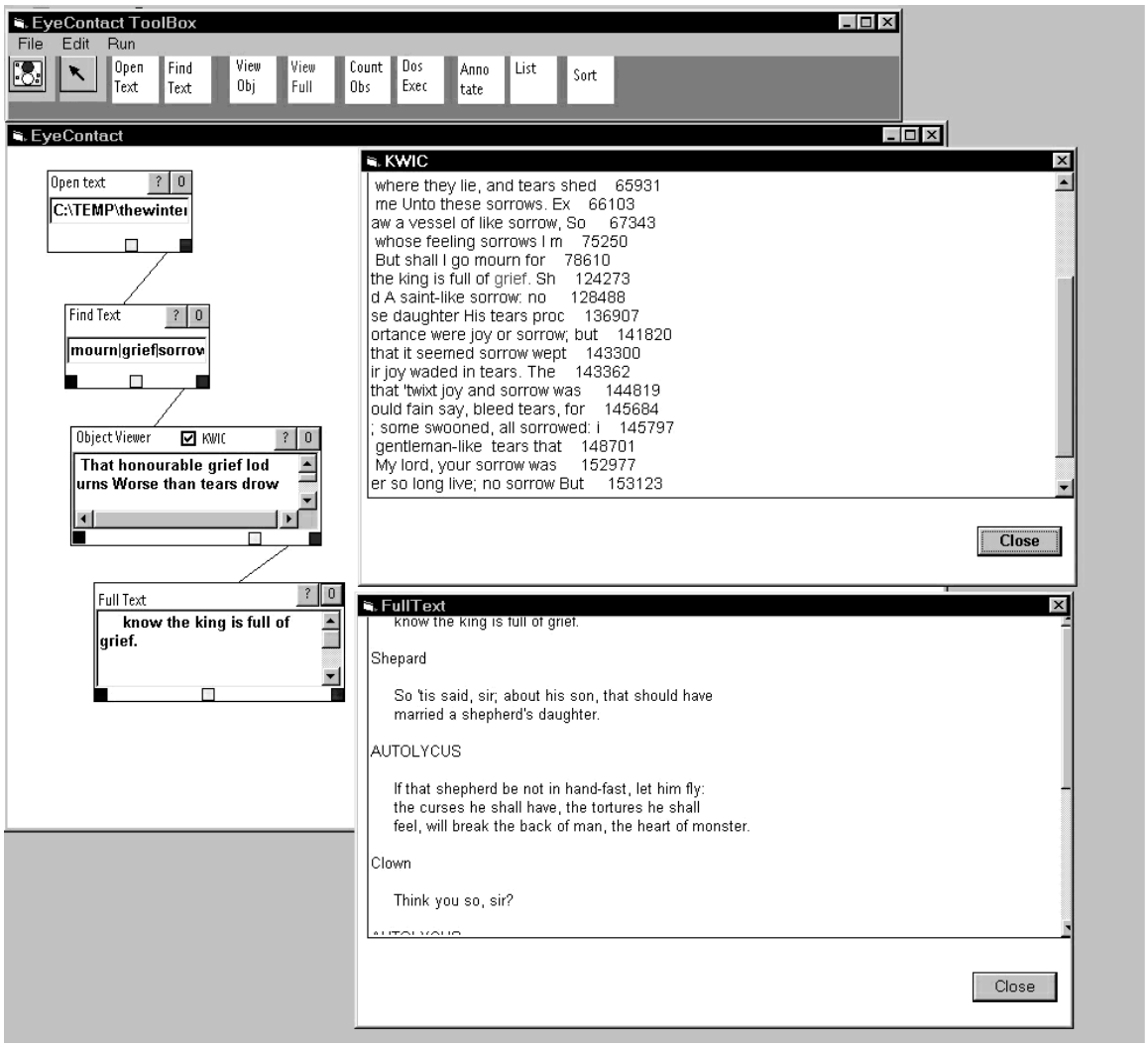
One frequent question viewers have regarding this visualization is how did we generate the graph in the first place. What exactly is it showing? On the one hand we could provide a detailed discussion of the statistics behind the visualization.[3] Such descriptions rarely seem to satisfy viewers; the problem is a deeper one of how the visualization hides its history, an

answer about which more will be said later. For the moment let us suggest that it is irrelevant to this paper exactly how this graph was generated as the point of this paper is to discuss interactive visualizations in general. Even if this graph were shown to be deceptive or useless, graphing texts might still be prove interesting.

3. Eye-ConTact

The second model of an interactive visualization discussed here is of Eye-ConTact, a visual programming environment for researchers who are studying texts. It is a prototype designed to test possible designs for modular visual development. It was built in Visual Basic for the Wintel platform with Perl scripts that do the actual text processing. It was designed to try different paradigms for visual programming in the humanities.

We should mention that this is not the only possible design for a visual text analysis environment. One can imagine other paradigms, such as a spreadsheet model with large cells that hold text and output of processes or a model like Venn diagrams for defining queries. The BNC software Sara, for example provides a visualization of structured queries.



<Figure 5: Eye-ConTact Screen - File = EyeConTact1.jpg>

In Eye-ConTact the user creates a flow chart or Map of how they want their *experiment* on a text to be conducted by placing modules on a screen and connecting them. Some of the modules if double-clicked will open a more complete view. The Map of this flow of data from a complete text to refined displays allows the researcher to see the logic of his or her study of a text and then reproduce such studies with other texts. What is being represented in a visualization of a text application. After looking at other models we believe that what should be displayed is the flow of transformations to a text in order to produce a useful research result.

The visualization or map of the inquiry has been redesigned from an earlier version to be a more rigorous description of the transformations to the original electronic text such that the results could be recapitulated from the Map. A visual programming environment should encourage the researcher to think about the transformations they subject a text to and it should allow the researcher to keep an accurate record of how their analysis produced the results they find significant.

In this model one can place process modules on the map, connect them into a larger process, and run the process. Many of the modules as they appear on the Map have a larger panel that can be opened which has more features. In Figure 5 above you see two such opened panels - one is a KWIC display and one is a Full Text display. Ultimately this type of process visualization will subsume the first (data visualization), since the idea is to include the statistical techniques and data visualization modules that are in SIMWeb in Eye-ConTact. Thus a process visualization can be used to create processes that include data visualizations

What are the limitations of this prototype? This prototype while it works, was not designed to be a distributable product. It is being used to try out designs and is thus not efficient or particularly good at text processing. If you want to learn more about this there is a paper discussing the rationale behind the design of Eye-ConTact along with earlier screen dumps of the program at the Computing and Humanities Working Papers site at www.chass.utoronto.ca/epc/chwp/rockwell/ .

4. Graphical Representation

Now it is time to turn away from the sights, step back, and ask what is the point of such of visualization. Rest your wretched eyes and read.We will start by recapitulating some of our earlier research. In an earlier paper by John Bradley and Geoffrey Rockwell, entitled "Watching Scepticism" (Bradley et al., 1996), we presented a topology of representations of texts that organized the various types of graphical representations of texts.

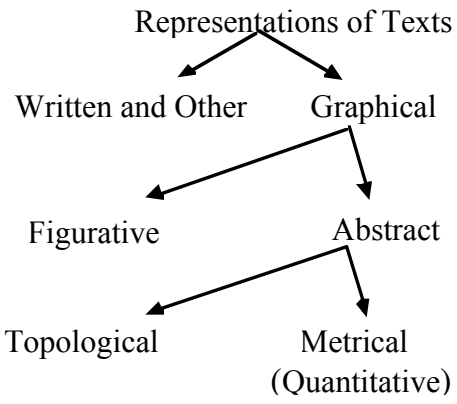


Figure 6: Topology of Textual Representations

In that paper we distinguished between **metrical** representations that are based on measurements of a text and those that are not based on measurements which we called **topological**. The topological graphs are abstract, like the metrical, and often work because we know about metrical graphs, but they are drawn by the human hand rather than based on numbers generated through measurement of the text. We went on to distinguish **visualizations** from graphical representations. A visualization is interactive - you can both change what you see by changing settings and you can also use the graph to do other things, like launch a full text search.

We need now to further refine our topology of graphical representations to include a further distinction based on what is being represented.

Figure 7: Visualization Chart

	Data Representation	Process Representation
Presentation (Representation)	Distribution Chart	Flow Chart of Program
Interactive (Visualization)	*SIMWeb* ←	→ *Eve-ConTact*

Until now we have talked about representations of texts, but in the field of text-analysis there is a second, related type of representation, which is the representation of process. This is the representation of the process by which you got the results in question. If SIMWeb is a visualization of the object of study, Eye-ConTact is a visualization of the process of study.

While it is no longer common, if one goes back in the literature, even in humanities computing, you will find flow-charts of processes from the days when all computer science students were forced to flow-chart their programs. This type visual notation shows the overview of the logic of the code using a graphical notation made up of typed boxes and lines. While flow-charting has, by and large, disappeared from the curriculum, along with the special plastic templates for drawing them, the idea of graphically describing a process has not. The flow-chart has been merged with the code it outlines to become a visual programming language, which is what Eye-ConTact is – a visual programming language for a specific domain – text-analysis.

Visual programming is not a new field. In computer science there has been some research done in the area of visual programming.[4] Visual programming environments provide the user with a graphical way of representing the logic of their application as an alternative to the way most of us program by writing code in a programming script. While there is much excitement about the potential for visual programming there have been few

successes. It would seem that visual programming works best in specific communities, like ours, where there is a consensus about the types of operations and the relationships between operations. Textual scholars offer a coherent community - we have the experience with computer assisted text-analysis to have built a consensus about the basic operations likely to be needed. In addition the promise of domain specific visual programming is that it should allow the humanist to concentrate on the tasks of their discipline, not on the difficulties of programming. A visual programming environment tailored to text-analysis promises to provide accessible programming suitable to researchers in the humanities. The Eye-ConTact project discussed earlier in this paper is a prototype of what such text-analysis visual programming might look like.

As important as the ease of programming is the potential for the rigorous description of the logic of a research project. Too often when using interactive text-analysis tools like TACT the researcher might find useful results but have not kept track of how the results were derived. This becomes a particular problem when one is moving data from one application to another. The design principle behind Eye-ConTact is that the user be encouraged to describe the process by which they get results so that the program can record that process as a map which can then reviewed at a glance (or altered as the research matures). This allows the researcher to recapitulate the process with a different text or to experiment with different processes with the same text. Given that such visual programming environments are usually modular, this also allows the researcher to build larger projects out of smaller programs if they can be repurposed. From the user's point of view, by implementing a text-analysis environment as a visual programming environment, the results always come through the conscious arrangement of operations on information thus preserving the logic of the experiment. Visual programming tools like Eye-ConTact are therefore more difficult to use out of the box, but they encourage a research discipline on

the user which we feel is important to include in the next generation of research text-analysis tools.[5]

5. Rhetorical Effects of Graphical Representations

So why bother with graphical representations or visualizations? Both in "Watching Scepticism" and more thoroughly in a forthcoming paper entitled “Empreintes dans le sable: Visualisation scientifique et analyse de texte” we discussed the rhetorical effect of graphical representations.[6] As William Playfair pointed out 200 years ago,

> Information, that is imperfectly acquired, is generally as imperfectly retained; and a man who has carefully investigated a printed table, finds, when done, that he has only a very faint and partial idea of what he has read; and that like a figure imprinted on sand, is soon totally erased and defaced. ... On inspecting any one of these Charts attentively, a sufficiently distinct impression will be made, to remain unimpaired for a considerable time, and the idea which does remain will be simple and complete, at once including the duration and the amount.[7]

A graphical representation of information has the following advantages:

1. It is easier to remember an image than the numbers or ideas it represents. This is Playfair's argument - that information, especially quantitative information, is difficult to remember unless represented graphically. If it is easier to remember it is easier to hold in your mind and think about.

   As an aside it is worth noting that this fits with the art of mnemonics or mnemotechnics which Frances Yates documents in *The Art of Memory* (Yates, 1966). The traditional art of memory, of which the poet Simonides is supposed to be the father, made use of artificial spaces to help memorize things. The idea is that if one puts things to remember in locations (*topoi)* that are in an imaginary space one can more easily recall them by then wandering through this virtual space. This technique was explained by the not unwarranted belief that the mind is better suited to remembering sensible things in space than abstractions. It is also worth noting

that this art of memory developed to support orators who had to remember large quantities of textual information.

2. Second, it is easier to get an overview of the whole through an image. Images are viewed differently than texts. A text is read in time, while an image is grasped as a whole and then examined for details. An image encourages the viewer to zoom in and out from the whole to the details, while a text is supposed to be read sequentially. Thus an image can provide an overview of a body of information that shows the relationships between the parts.

   The view of the whole is part of the interpretative process. As Northrop Frye pointed out in *Words with Power* (Frye, 1992), criticism starts with the formation of a mental picture of what one has read that shows the unity of work.[8] Textual visualizations while they are not such mental images, can approximate them for those disposed to feast their wretched eyes.

3. Third, certain patterns and anomalies can be more easily recognized in a graphical representation than in a table. Patterns that are difficult to find in large tables of words and numbers will often stand out immediately when graphed properly. In addition, the anomalies that stand out, stand out in relation to the whole. If one finds interesting patterns or anomalies one finds them in the context of an overview. It is how they stand out against the whole that make something an anomaly.

4. Finally, graphical representations being atemporally read are immediate. Images are not read over time the way a text is. They are seen as a whole and then explored the way a hypertext might be. It takes no more than a moment to view an image, though one might choose to linger. When you see an image you see it all in one glance which gives the viewing a feeling of immediacy. A table of numbers or KWIC, by contrast takes time to read. This immediacy is tied to the capacity of a graphical representation to show the unity of the whole. A well formed graphical

> representation can, to paraphrase Tufte (Tufte, 1983), show you the whole in the least time with the least ink and in the least space. It is this efficiency that makes visualizations so tempting and potentially misleading. They appear to be a direct view of the original body, a view which if accepted without thought can crowd out the truth. Socrates is right to warn us about feasting our eyes on dead bodies.

Graphical representations also have problems which in the early thrill of visualization tend to get forgotten:

1. It is not always clear what graphical features mean as attractive as they might be. A picture may be worth a thousand words, but it is not always clear *which* thousand words. This is because graphical features are ambiguous. What does it mean in an image if one object is *above* another or connected to the first by a line. It almost always is the case that for a graphical representation to be understood there has to be a textual key to the graphical features employed.[9] Thus visual programming languages that held such promise in the 1980s have foundered on the problem of the precision of graphical features.
2. A second problem is that the process by which a graph is generated is usually hidden from the viewer. This is the problem mentioned earlier with SIMWeb, the visualization does not explain itself. Any showing is also a hiding. In order to achieve immediacy and a holistic overview a visualization hides what is non-essential - its history. Texts tells stories, and the story of how a graph has been generated is usually supplied by the viewers experience or accompanying notes, but not by the graph itself, unless that graph is connected to one that represents the logic of the result. Only a system like Eye-ConTact keeps its history. (Which is why we settled on process visualization over the other models shown.)

   This means that graphical representations to work depend on either elaborate descriptions of the processes used to generate them, or on a shared graphical

> literacy. Most of us today do not need to be told how a bar graph in a newspaper was generated - we have learned to read such graphs and know what details to pay attention to if we distrust the graph. Textual visualizations, however, are less common and we have to ask whether a shared graphical literacy would develop in the community to make sense of them. This is particularly the case with graphs generated by sophisticated statistical techniques infrequently used in the humanities. Such graphs while attractive are impenetrable without an explanation.[10]

6. Interactivity

Now we turn to the final issue, which is the interactivity in visualization. There is not, to the best of my knowledge, a decent discussion of what interactivity adds to a rhetorical artifact. There is a literature around interface, but that concentrates around interface design. There is a literature around hypertext, but this literature is concerned with a limited set of types of interactivity – in particular the hypertextual link. There is a very small literature and lots of web sites about games, but this literature rarely deals with the rhetorical impact of interactivity. Games enthusiasts talk about *flow* and the rapidity of response, but I have not seen a deeper reflection on what interactivity really contributes.

Brenda Laurel in *Computers as Theatre* (Laurel, 1991) is one of the best starting spots as she deals with interactivity in terms of interface and drama. She suggests we should look at three parameters, the frequency of interaction, the range of choices, and the significance to the user. My favorite definition, however, is from a dialogue between Stewart Brand and Andy Lippman found in (Brand, 1987). In that dialogue, and it is worth noting that it is a dialogue full of interruptions, Lippman defines interactivity as, 'mutual and simultaneous activity on the part of both participants, usually working toward some goal, but not necessarily.' He then provides five corollaries which it is useful to apply these to the two experiments demonstrated here.

1. **Interruptibility** - Either participant (the user or the computer) should be able be able to interrupt the other. Lippman comments that interactivity in this regard is like a conversation rather than a lecture.

   Both SIMWeb and Eye-ConTact are not as responsive as they should be. SIMWeb suffers from its WWW interface. Once a query is submitted you have to wait, though you can stop the query. Eye-ConTact, should, but does not now, have a way to interrupt and walk through a process once you start running it. There are, however, a number of ways one can zoom in and out which is a form of dialogue with the whole - a form of interruption.

2. **Graceful Degradation** - An interactive system should be capable of responding gracefully to requests that are beyond what it was programmed to handle. Again, SIMWeb and Eye-ConTact are weak in this regard as they do not have very informative error messages. The open and modular aspect of Eye-ConTact does, however, provide a way of dealing with degradation through the creation of new processes.

3. **Limited Look-ahead** - An interactive system should allow the user who is working toward a goal to anticipate what is the best way to achieve that goal. Ideally visualizations should provide this through overview. The visualization is in effect the look-ahead that you use to navigate the information beneath. This principle has been specifically applied to Eye-ConTact where we changed the icons that made of the map to be miniature panels that could show some of the information within, thus providing limited look-ahead.

4. **No-default** - There should be no default path through the information or default way of interacting. This is not true of many hypertext systems, that actually have a default path that the user cannot avoid if they want to make sense of the system. Both SIMWeb and Eye-ConTact are open and do not have a hidden default. They

are not like so many axial hypertexts that are essentially books with a few annotations tacked on.

5. **Impression of an Infinite Database** - The user should feel that they can go in any direction and interact forever. The connection to a full-text database in SIMWeb and Eye-ConTact goes a long way to giving the sense that there is, if not an infinite amount of information, at least all of the original text. More importantly, the intention in Eye-ConTact is to allow new modules to be used through a special interface thus extending the potential processes that one can run.

In concrete terms what interactivity adds to visualization is the ability to not just see, but to also manipulate or touch the information. SIMWeb and Eye-ConTact represent two paradigms for manipulation, 1) the manipulation of the interpretative process and 2) the manipulation of the results. They also demonstrate two types of manipulation - the direct manipulation of clicking, dragging, and snapping, and the indirect manipulation through parameters typed or selected from a menu. Touching, by invoking another sense, reinforces the virtues of seeing mentioned above. It makes it more sensible, more immediate, and thus easier to remember or hold in ones mind as a mental construct for reflection. Manipulation also allows one to learn about the processes by experimentation. When manipulating a visualization one is not only exploring the object visualized, but also learning about the constraints of the process of visualization. One can, so to speak, kick the tires of the computing that generated the visualization to see where it breaks down. We believe this pragmatic exploration of process should make difficult statistical techniques more accessible to humanists uninterested in the math.

Interactivity, as hypertext theorists like Landow (Landow, 1992) have pointed out, changes the relationship between author and reader. We believe that providing hypertext links is a limited type of control to transfer to the reader with about as much effect as a good index or table of contents; the challenge is to imagine a deeper transfer of control. If you

want to see what control feels like play the current crop of games. When the reader is given control over the trajectory of the content the differences between author and reader begin to disappear and what you have are different types of author/editors. This is especially the case with process visualization, where the user is creating a larger process out of smaller ones. In Eye-ConTact you have control comparable to that of a programmer who is an author of sorts.

While this resolves one of the problems mentioned above regarding visualization - i.e. that it hides its history – it is a more dangerous thing to bring into sight. Given sufficient control the visualization becomes a simulation of the interpretation of the text. While in SIMWeb you are looking at a representation of the text, in Eye-ConTact you are playing with a representation of scholarly technique, a shift from manipulating the book to manipulating the reading. This is the danger with visualization - if sufficiently rich and interactive the user forgets that not only the original text but also that the interpretative possibilities are mediated by the design of the visualization environment. It can be deceptive when the control of the user increases to the point where the interpretative presence of the simulator disappears so that the simulation (or as Baudrillard would put it the simulacrum) is confused with the real thing and this is what we are often accused of doing with these tools – replacing the book and scholar with its machined image. We should not confuse the corpse for the person or the text for the trees, though in our case it is not always clear what the original text would be..

This raises the question of the status of the visualization. Is it an interpretation of the original the way an essay about a text is, or is it an simulated edition of the original. I have heard editors say that all editions are interpretations, but if that is true an edition is a special type of interpretation, one that aims for transparency through which the illusion of an original can be seen. Likewise with a visualization, the more interactive it is, the more

transparent the interface is, to the point where the user feels they are working with the original, whatever that may be.

To return to Plato's image of Leontius's desire to see and abhorrence of sight, can we conclude by tackling this fear? We trace it back to an interpretative position articulated among others by Heidegger to the effect that one should interpret a text according to how it presents itself. In other words one should respect the way a work chooses to reveal itself. If it is a book you should read it as a book on paper in your hands. This position, despite the impossibility of actually getting a text to tell you anything at all about how it should be read, warns us about the dangers of interpretative violence to a text. In *The Nature of Language* (Heidegger, 1982, p. 60) Heidegger 'converses' in an unscientific fashion with a poem by Stephan George entitled "The Word". We concatenate a few lines of the poem Heidegger quotes to show the connections,

> Wonder or dream from distant land/ I carried to my country's strand/ And waited till the twilit norn/ Had found the name within her bourn-/ Then I could grasp it close and strong.../ And straight it vanished from my hand, .../ Where word breaks off no thing can be.

Carrying, grasping, exposing to light, and naming for Heidegger hide the wonders we should care about. For Heidegger like many others, certain methods do such violence to that which is sought that it vanishes. If simply naming and grasping something is a violence to it, what would visualizing be. To shine the light of quantification on a text and then throttle it with interactivity all using modern technology (gasp) is surely an even greater violence than the metaphysical tendencies Heidegger wished to avoid. It introduces the distraction of the tool. To this we have no answer except to tempt you to feast your eyes on what has been shown. After all seeing is believing.

Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

7. Notes

1. This visualization can be tried at http://tactweb.humanities.mcmaster.ca/cgi-dos/simweb/simweb.bat .
2. TACTweb can be tried and downloaded from http://tactweb.humanities.mcmaster.ca .
3. Such a discussion is available at, http://tactweb.humanities.mcmaster.ca/cgi-dos/simweb/what.htm .
4. For those interested a good place to start is an article by Myers entitled, "Taxonomies of Visual Programming and Program Visualization" (Myers, 1990) or an article by Price, Baecker and Small entitled "A Principled Taxonomy of Software Visualization" (Price et al., 1993).
5. In a paper available on the web we went into the reasons for the recording of research in greater detail. See Bradley, J. and Rockwell, G., "Towards new Research Tools in Computer-Assisted Text Analysis" at http://www.humanities.mcmaster.ca/~grockwel/ictpaper/learneds.htm .
6. “Empreintes dans le sable: Visualisation scientifique et analyse de texte”, written by Geoffrey Rockwell and John Bradley has been accepted for publication in a forthcoming collection of papers edited by Michel LeNoble and Alain Vuillemin entitled *Litterature, informatique, lecture: De la lecture assistee par ordinateur a la lecture interactive*.
7. This quote from Playfair, *The Commercial and Political Atlas*, pp. 3-4, comes from Tufte, 1983, *The Visual Display of Quantitative Information*, p. 32.
8. See *Words with Power* (Frye, 1992) pp. 69-76. Frye writes that the reading of a text is *pre-critical* in the sense that criticism doesn't start until a mental image of the whole has been formed.

9. Marian Petre in "Why Looking Isn't Always Seeing: Readership Skills and Graphical Programming" (Petre, 1995) goes into this problem of understanding what she calls the 'secondary notation' of a graphic.
10. While there seems to be a general acceptance of statistical techniques in France thanks to introductory texts like (Salem et al., 1994); that does not seem true in the English speaking world, perhaps because of the separation of the humanities and social sciences.

Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

8. References

Arnheim, R. (1969). *Visual Thinking*. University of California Press, Berkeley, California.

Bertin, J. (1983). *Semiology of Graphics*. W. J. Berg (tr), University of Wisconsin Press, Madison, Wisconsin.

Bradley, J. and Rockwell, G. (1996). "Watching Scepticism: Computer Assisted Visualization and Hume's Dialogues". In G. Perissinotto (ed), *Research in Humanities Computing 5*, Clarendon Press, Oxford, pp. 32-47.

Brand, S. (1987). *The Media Lab; Inventing the Future at MIT*. Viking, New York.

Earnshaw, R. A., and Wiseman, N. (1992). *An Introductory Guide to Scientific Visualization*. Springer-Verlag, Berlin.

Foulser, D. (1995). "IRIS Explorer: A Framework for Investigation", *Computer Graphics*, 29.2: 13-16.

Frye, N. (1992). *Words with Power*. Penguin, Toronto.

Heidegger, M. (1982). On the Way to Language. P. D. Hertz (tr), Harper & Row, New York.

Hume, D. (1991) *Dialogues Concerning Natural Religion*. S. Tweyman (ed), Routledge, New York.

Lancashire, I., Bradley, J., McCarty, W., Stairs, M., and Wooldridge T. R. (1996). *Using TACT with Electronic Texts*. The Modern Language Association of America, New York.

Landow, G. (1992). *Hypertext: The Convergence of Contemporary Critical Theory and Technology*. John Hopkins, Baltimore.

Laurel, B. (1991). *Computers as Theatre*, Addison-Wesley, New York.

McKinnon, A. (1989). "Mapping the Dimensions of a Literary Corpus", *Literary and Linguistic Computing*, 4.2: 73-84.

Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

Myers, B. A. (1990). "Taxonomies of Visual Programming and Program Visualization", *Journal of Visual Languages and Computing*, 1: 97-123.

Petre, M. (1995). "Why Looking Isn't Always Seeing: Readership Skills and Graphical Programming", Communications of the ACM, 38.6: 33-44.

Petre, M. and Green, T. R. G. (1993). "Learning to Read Graphics: Some Evidence that 'Seeing' an Information Display is an Acquired Skill", *Journal of Visual Languages and Computing*, 4: 55-70.

Plato (1997). *Plato; Complete Works*. John M. Cooper (ed) Hackett, Indianapolis, Indiana. The translation of the *Republic* is by G. M. A. Grube with revisions by C. D. C. Reeve.

Potter, R. G. (1988) "Literary Criticism and Literary Computing: The Difficulties of a Synthesis", *Computers and the Humanities*, 22: 91-97.

Price, B. A., Baecker R. M. and Small, I. S. (1993). "A Principled Taxonomy of Software Visualization", *Journal of Visual Languages and Computing*, 4.1: 211-266.

Raymond, D. R. (1991). "Characterizing Visual Languages": *Proceedings of the IEEE Workshop on Visual Languages*, Kobe, Japan, October 1991, pp. 176-182.

Raymond, D. R. (1993). "Visualizing Texts": *Making Sense of Words: Proceedings of the Ninth Annual Conference of the UW Centre for the New OED and Text Research*, Oxford, England, September 1993, pp. 19-32.

Repenning, A., and Summer, T. (1995). "Agentsheets: A Medium for Creating Domain-Oriented Visual Languages", *IEEE Computer*, March 1995: 17-25.

Rockwell, G. and Bradley, J. (1998). "Eye-ConTact: Towards a New Design for Research Text Tools", *Humanities Working Papers*, URL: http://www.chass.utoronto.ca:8080/epc/chwp/.

Salem A. and Lebart L. (1994). *Statistique Textuelle*. Dunod, Paris.

Preprint of Rockwell, Geoffrey, "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications", (Primary Author) written with John Bradley and Patricia Monger, *Literary and Linguistic Computing*, vol. 14, no. 1, 1999, p. 115-130.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Graphics Press:

Chesire, Connecticut.

Yates, F. (1966). *The Art of Memory*. University of Chicago Press, Chicago.

<Figure 1: SIMWeb Opening Screen - File = SIMWeb1.jpg>

<Figure 2: SIMWeb Screen 2 - File = SIMWeb2.jpg>

<Figure 3: SIMWeb Screen 3 - File = SIMWeb3.jpg>

<Figure 4: SIMWeb TACTweb Screen - File = SIMWeb4.jpg>

<Figure 5: Eye-ConTact Screen - File = EyeConTact1.jpg>

Figure 6: Topology of Textual Representations

Figure 7: Visualization Chart