

What is text analysis, really?

Geoffrey Rockwell
TSH 309A, School of the Arts
McMaster University
1280 Main St. W.
Hamilton, ON
Canada, L8S 4M2

Abstract

In which the author revisits the question of what text analysis could be. He traces the tools from their origin in the concordance. He argues that text analysis tools produce new texts generated from queries through processes implemented on the computer. These new texts come from the decomposition of original texts and recomposition into hybrid new works for interpretation. The author ends the article by presenting a portal model for how text analysis tools can be made available to the community.

Introduction

to analyze is nothing but an operation that results from the conjunction of the preceding operations. It merely consists in composing and decomposing our ideas to create new combinations and to discover, by this means, their mutual relations and the new ideas they can produce. (Condillac, *Essay on the Origin of Human Knowledge*, p. 48)

In a mock confrontation between Allen Renear and Jerome McGann at the ACH/ALLC in 1999 at the University of Virginia, two views as to what a text really is were put forward. Renear put forward, for the sake of the confrontation, the OHCO (ordered hierarchy of content objects) perspective while McGann practiced a view of text as performance.¹ In the context of a humanities computing conference this confrontation was designed to highlight the relationship between theories of text and ways of representing texts digitally. Renear's Platonic view of the text as a real abstract object fits nicely with the dominant practice for the digital representation of texts, as represented by the guidelines of the TEI. McGann instead gave us an example of a reading that was both a performance itself and pointed to the

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

combinatorial possibilities within and around the text. McGann's challenge to Renear was to show how a playful reading of a text was both a new text and that this potential could not be captured easily by an OCHO. The confrontation succinctly opened again the question of the relationship between how we represent texts, how we use them, and our theories of textuality.²

What does this have to do with computer-assisted text analysis?

What was not made clear in the confrontation was the role of the tools we use for accessing and manipulating digital texts; tools which I will call text analysis tools. If we are to take McGann's public performance of a reading as an analogue for what we wish to achieve with these tools, we have to think not only about how we represent the text but also about the performance of analysis and the tools that are used to perform this analysis with a computer. The logic of the tools, despite (or because of) their tendency to become transparent in use, can enhance or constrain different types of reading which in turn makes them a better or worse fit for practices of literary criticism including the performance of criticism.

Another way of saying this is that we have a model of computer-assisted literary text analysis that is guided by a view of what a text is and how we should use it that does not match the practice of many contemporary literary critics. (It should be noted that this is not true in the field of computational linguistics and may not be true in literary criticism in the future.) Consequently, as others have pointed out, text analysis tools and the practices of literary computer analysis have not had the anticipated impact on the research community. This is often blamed on the absence of easy-to-use tools, especially tools that take advantage of OCHO, but there are two other issues that have to be taken into account.

First, the tools we have (and even those we anticipate) have emerged out of a particular tradition that I will call an "editorial" tradition that goes back to tools for editors of concordances starting with Roberto Busa. Second, I believe that the moment when humanities computing could have an impact on literary criticism through the provision of critical tools (accompanied by relevant methodologies and theories that backstop the tools) is passing as industry server based text tools emerge instead. These industry tools provide access to licensed digital archives and satisfy our colleagues while we keep on imagining personal research text analysis tools. The community we hoped to provide with text analysis research tools has found them elsewhere while we fiddle.

Text Tools and ConCORDING

To understand the current state of text analysis tools and their logic we can briefly review their history in terms of the practices they complement and the theories of textual practice they augment.

Text analysis tools have their roots in the print concordance. The concordance, is a standard research tool in the humanities that goes back to the 13th century. Concordances are examples of the sorts of "augmentation" tools that extend our scholarly reach and therefore assist in intellectual work of the sort that Vannevar Bush and Douglas Engelbart imagined.³

The first computer-based text-analysis tools were designed to assist in the production of print concordances. Father Roberto Busa in the late 1940s was one of the first to use of information technology in the production of a concordance, his *Index Thomisticus*, (a remarkable concordance and more to the works of Thomas Aquinas). His project began by using index cards, moved onto analogue information

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

technology in the 50s and migrated to electronic computers as they became available.

The published results were finally delivered in the 1970s with a CD released in 1992.⁴

The technology he used was developed *ad hoc* as he went along rethinking how information technology could facilitate his project.

In the 1960s and 70s the first generation of tools created to be used by others became available. These were tools for mainframes that were batch tools, and they were designed, like Busa's tools, to assist in the production of paper concordances.

The paper concordances would still be the tool that the rest of us used, the computing tools were for the editors of these concordances. It is interesting to review the names of some of these early tools. COCOA stands for Count and Concordance generation on the Atlas.⁵ The Oxford University Computing Service took over COCOA in 1978 and produced OCP or the Oxford Concordance Program.

With the availability and increasing power of microcomputers in the 1980s, text analysis tools migrated from mainframes to personal computers. OCP evolved into Micro-OCP and new programs came out for the personal computer like the Brigham Young Concordance program (BYC) later renamed and commercialized under the name WordCruncher and the TACT environment developed at the University of Toronto and released in 1989 at the ACH/ALLC conference that year. When these tools became available to researchers on their personal workstation they changed how we use tools in three ways.

The scholar could now use tools whenever and wherever they wanted on a personal computer instead of having to wait for mainframe time or having to connect over a tethered terminal. In effect this meant that the humanist was no longer dependent on the paper concordance when doing research in their study, but could use electronic tools **instead** of print. This change in the time and place of computer-

assisted text-analysis, along with interface developments, led developers away from a batch concordancing model towards interactive tools that took advantage of the fact that the scholar would have personal access to tools and e-texts for study in their time and place of study.

Secondly, with interactive tools and a more mature community of users we began to realize we could ask new types of questions that print concordances could not support. As we experimented with new questions we realized that one of the things that was important was this intellectual process of iteratively trying questions and adapting tools to help us ask new questions. We can do so much more now than find words in a string. We can ask about surrounding words, search for complex patterns, count things, compare vocabulary between characters, visualize texts and so on.

Thirdly, as personal tools became available, we began to re-imagine the electronic text, which went from being something created by (and exclusively for) a concordance project to an electronic edition meant to be used by anyone with whatever tools they might have for unanticipated future research. Our models for tools and e-texts began leapfrogging each other as advances in tools triggered the need for improvements in texts. Now advances in text models and markup have surpassed the personal tools.

The Hermeneutics of Text Analysis

Let us pause now to consider the hermeneutical principles behind the concordance and tools that extend it.⁶ As Willard McCarty puts it, "The early history of the concordance suggests that it was invented essentially for the same job to which we apply it today, 750 years later: to discover patterns of coherence in a text or textual

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

corpus. ... the concordance very likely grew out of a habit of mind conditioned by a *typological* or *figural* view of the Bible, i.e. the intratextual notion that the meaning of the biblical text is derived by putting together normally disjunct passages into a *concordantia*, a concord of senses."⁷

The *Encyclopedia Britannica Online* warns in its discussion of "Parallelism" as a form of Scriptural interpretation, against the naive use of concordances.

Parallelism, the interpretation of Scripture by means of Scripture, is a corollary of the belief in the unity of Scripture. But as a hermeneutical principle it must be employed sparingly, since the unity of Scripture should be based on comprehensive exegetical study, rather than itself provide a basis. ... One naive form of parallelism is the 'concordant' method, in which it is axiomatic that a Hebrew or Greek word will always (or nearly always) have the same force wherever it occurs in the Bible, no matter who uses it."⁸

The hermeneutical principles underlying the use of the concordance and the text-analysis tools that evolved from it can be summarized thus:

- First, the use of a concordance for interpreting a text presumes that there is some sort of unity to the text and a consistent use of words.
- Second the concordance is a new text that is assembled out of passages that agree or concord. The concorded hybrid provides a new combination of the parts of the original work. The concordance is a monster new text patched out of the old.
- Third, a concordance is generated according to some procedure, be it a manual procedure or process implemented in software. The procedure that generates the concordance takes as its input a query about a word or pattern. The particular concordance one looks up or generates is a text in response to a choice by the reader that is generated by the software or editor according to established procedure.

In "Seeing the Text Through the Trees" I cautiously suggested that there are problems with these principles. I now want assert what I before left tentative:

1. The hybrid texts generated by computers are new texts that can be called interpretations in that they belong to the class of texts which have a special relationship to an existing text. If we want to distinguish them from human interpretation we can call them interpretative aides.
2. They are analytic in that, following Condillac, they are generated by processes of taking apart and putting back together information into new configurations for the purposes of discovery and reflection.
3. Such hybrid interpretative works can take advantage of the multimedia and transmedia capabilities of the computer. Such interpretations can involve other media, especially measurements (quantifications) and visualizations of these quantifications – visualizations which interpret a text from one media to another.
4. And finally, there is no a-priori privilege to certain processes of decomposition and recomposition like traditional concordancing. The assumptions behind concordancing are as suspect as any, despite the long tradition of using concordances. Familiarity with concordances should not breed contempt for other monsters. The challenge before us is to question our procedural habits and presuppositions as to what are legitimate recombinations – to forget the concordance and ask anew how we can analyze a text with a computer and whether such computer-assisted interpretations are interesting in and of themselves. We need to play again and make playpens available to our colleagues rather than re-implement a limited set of procedures grounded in one hermeneutic.

I therefore want to propose a very different image of what a concordance is.

Following the lines of Lucian of Samosata I call it, and have been calling it, a hybrid

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

text created by choices of the user from the original text. I call it a hybrid (or monster) because it is authored not just by the original author, but also by the user's choices and the procedures used to generate it. It is neither afoot or ahorseback like the centaur Cheiron, the tutor of Achilles.⁹ It is neither the work of the original author nor that entirely of the provoker of the concordance. Its unity comes from the intentions of both in a way that can be recapitulated by others. It is one text in a larger dialogue between authors, readers, and users. Lucian had it right, we are in dialogue with Dialogue recreating monsters out of the old.

Rather than redeveloping tools based on principles of unity and coherence we should rethink our tools on a principle of research as disciplined play. Disciplined play privileges experimentation and modeling over hypothesis testing or concordance publishing. Play is a pragmatic approach of trying something, seeing if you get interesting results, and if you do, then trying to theorize why those results are interesting rather than starting from articulated principles. The astute critic will counter that it is impossible to experiment without some pre-theoretical intuition, but that is my point – we need to imagine environments that allow a much broader set of intuitions to be played out so that they can be theorized and confronted. That means aggregating a variety of tools from the traditional to the bizarre in ways that allow them to be recombined in unanticipated ways we don't approve of. A hermeneutic of play postpones questions of principle indefinitely which another way of keeping them in play.

TAPoR and Text Analysis Portals

Such calls for playful environments are fine, but what would they look like - another digression. In May of 1996 Susan Hockey, who at the time was the director of the

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

Center for Electronic Texts in the Humanities at Princeton and Rutgers, invited a number of people together to discuss what Michael Sperberg-McQueen described as a crisis of confidence in our software tools. The problem was that we had developed a rich model for representing a text, as described by the Text Encoding Initiative Guidelines, a model which was (is) being used to produce magnificent scholarly electronic editions with woven into them enriching information so that new and interesting questions can be asked. Unfortunately, we don't have accessible tools that can take advantage of that information and allow us to pose a new generation of questions. In the leapfrogging game of tools and texts we now have analytical texts that are too rich for most tools, just as, before the Web, we seemed to have tools but lacked the critical mass of electronic texts.

The situation, of course, is not that bad. Along came the Web and we found that not only did we now have a common infrastructure for accessing textual information, but we also found that the larger computing community became interested in text manipulation and XML tools for the Web. There actually **are** tools that can make use of XML or SGML tagging, but these are typically server based tools, programming libraries, or commercial information management tools. These tools have emerged out of the incredible energy around the Web and new standards like XML. The tools have emerged from the private sector and from the open source community; they just haven't been designed for us and need to be adapted to fit into our research practices. They can only be deployed on more sophisticated (and expensive systems) by people with a certain level of technical proficiency. Thus only well funded projects can deploy them and they therefore tend to be used to publish scholarly corpora by well- funded projects.

The current situation can be summarized thus:

- 1 The personal computer tools that are accessible to textual scholars are limited and have difficulty taking advantage of the emerging texts.
- 2 Tools that have the capacity to take advantage of the rich encoding woven into the electronic texts are available but they are meant for the deployment of services on well-supported infrastructure that is way beyond what the average humanist is willing to manage.
- 3 Thus the available tools have typically only been used by large electronic text projects to publish their electronic editions. The tools are deployed not for general use but to make available the research of a specific project in the ways imagined by that project. But original research consists of asking new and unanticipated questions and that means giving researchers access to the evidence and a breadth of tools with which to study the evidence.

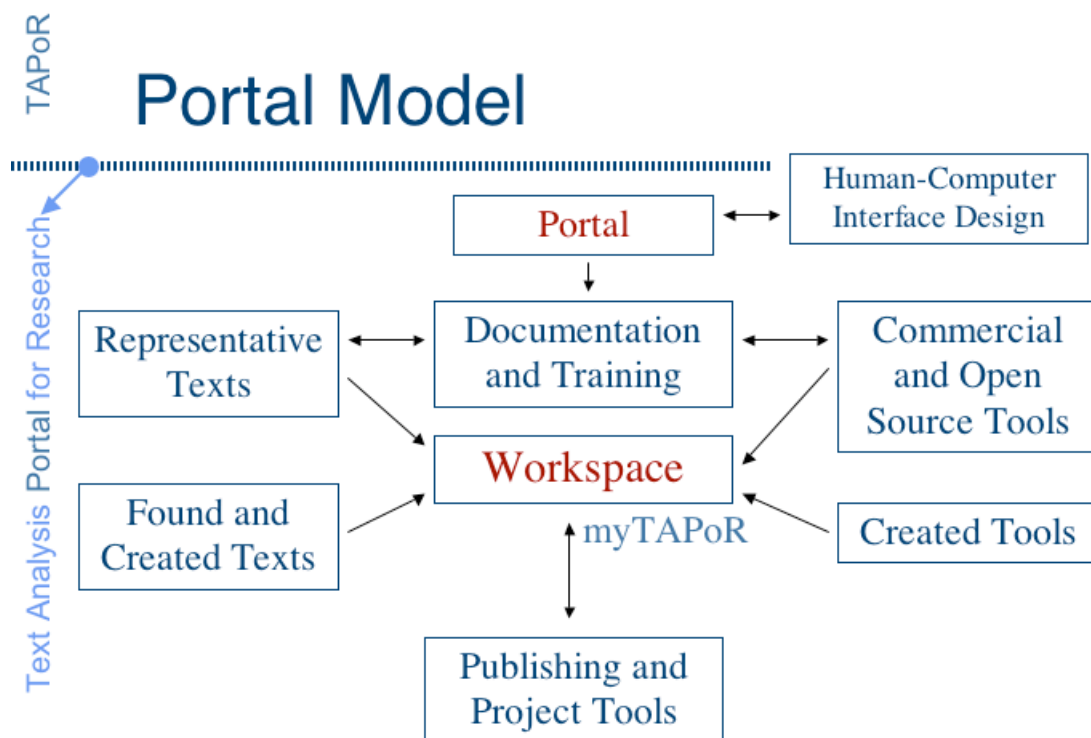


Fig. 1. TAPoR Portal Model

While I doubt we can resist the commercial forces that lead to the bundling of limited tools and texts, we can understand this process in terms of its relevance to the practices of our colleagues and imagine an alternative that is relevant to contemporary literary criticism. This paper will therefore conclude with (yet another) proposal for a model for text analysis tools, a portal model – a portal that is something like TAPoR, where we set up a virtual laboratory which makes available a variety of server based tools properly supported, documented, and adapted for use in the study of electronic texts.¹⁰ The idea is that you should be able to bring your texts to this laboratory and try not just one, but many different tools as your research matures. A portal is a laboratory or workshop where tools that none but the richest of us could afford individually have been gathered and adapted for general use. In other words it is a playpen that can:

- 1 Provide an **introduction** to the field
- 2 Support **casual** use
- 3 Provide a **single point of entry** to common resources
- 4 Be **customizable** for advanced users
- 5 Support **individual** researchers and **projects**
- 6 Support the study of **computing methods** in textual studies
- 7 Support the study of **human-computer interaction**

Users of a portal typically can either dip their toes in and casually use tools without installing them or can create individual accounts (which is one of the features that distinguishes a portal from other types of Web sites) where they can customize sets of texts and tools they wish to use on the texts.

The trick to online tools is that the aggregated tools in the playpen can operate on texts that are not in the portal – texts elsewhere on the Internet, whether mounted

by the user on a personal site or found elsewhere. For example, in the TAPoR Prototype Tool below the user enters a URL for the text to be processed, a key word for which they want co-occurrences, and a distance in lines for the tool to use.¹¹ The tool gets the text from the remote site, processes it and returns a list of word that co-occur with the key word sorted by frequency. Such a model is feasible now that most researchers have reliable high-speed Internet access and because of the processing capabilities of modern Web servers.

The screenshot shows the 'TAPoR Prototype Tools' interface. At the top, there is a blue header with the text 'TAPoR Prototype Tools' and 'Geoffrey Rockwell, ACH/ ALLC 2002'. Below this is a green header with the text 'Co-occurrences (t.Cocur)'. The main content area contains the following elements: a horizontal line, a checkbox labeled 'Strip tags before concording. Note. Resulting HTML is difficult to predict without stripping tags.' which is checked, another checkbox labeled 'Sort by frequency.' which is also checked, a text input field for 'Number of lines of context (0=only the found line, max of 10):' with the value '0', a text input field for 'Word to find:' containing the text 'test', a label 'Enter URL in field below:' followed by an empty text input field, and a label 'Or place text in this field. (If there is text the URL will be ignored.)' followed by a large empty text area.

Fig. 2. Prototype Co-occurrence Tool

The portal model provides us a way of taking advantage of the trend away from personal tools towards community tools while also engaging a different critical practice of playful criticism. TAPoR is in a hermeneutical tradition that incorporates play in method and which is best expressed in the theoretical work of Gadamer and

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

Huizinga. A portal for text analysis can finesse the problems of ease-of-use while also providing a virtual playpen for contemporary critics to try computer-assisted techniques beyond those provided by the commercial publishers of e-texts. The portal ironically could be the backdoor through which our colleagues could be introduced to the playful work of humanities computing.

That said, we should be honest and admit that much of our discourse around tools is for our own sake. It is our humanities computing to play with tools and texts. Does it matter if anyone else ever uses these tools as long as they help us understand the practice of creating and reading digital representations? The portal prototype developed, while it may have practical applications, is for humanities computing an attempt to illustrate a particular relationship between a theory of texts and analysis on the one hand and an interface for text analysis that implements that theory on the other hand.

End

Appendix: The TAPoR Project

TAPoR¹², or the Text Analysis Portal for Research, is a project that brings together a number electronic text projects and researchers across Canada. We gathered in order to apply to the Canada Foundation for Innovation and other agencies for support to set up infrastructure to help us with common research needs. Our proposal was successful in the last round of CFI and we are now in the first phase of implementing this infrastructure. We are that scary moment of starting to do what we said we would do.

Specifically we proposed a network of nodes at universities across Canada; nodes which would have servers and local labs where we could aggregate and make available the best text tools be they from industry or other sources. These would be supplemented by representative texts and special infrastructure like an interaction lab that will go into Faculty of Information Studies at the University of Toronto.

Our aim is to support not only the researchers and existing projects at the nodes, but also to provide a portal to appropriately configured tools for researchers across Canada and abroad. We hope to significantly improve the research infrastructure in this field in Canada not just for those who are part of the project but for our colleagues and collaborators in what is really important, research in the humanities and other disciplines that make heavy use of textual evidence. In the process of doing this we hope to significantly advance the understanding of the place of computer assisted text analysis in the interpretation of texts. We hope to trigger a reexamination of the presuppositions, the types of questions, and the interpretative theories that form our practices.

TAPoR is, however not just computing infrastructure, it is also a collaboration of people at universities across Canada. The universities are, going roughly from

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

West to East, the University of Victoria, the University of Alberta, McMaster University, the University of Toronto, Université de Montréal and the University of New Brunswick. Researchers at other universities like Malaspina UC and the University of Guelph are also involved, and it is our hope that we can, at the appropriate moment welcome as many other organizations into the collaboration as possible. Part of the project is to design a portal that can be extended and mirrored by interested parties.

Notes

1. Susan Hockey, Chair, with Allen Renear and Jerome J. McGann, Panel: "What is text? A debate on the philosophical and epistemological nature of text in the light of humanities computing research", ACH-ALLC, June 9-13 1999, at the University of Virginia.
2. This paper was presented as part of a panel on *Reconceiving Text Analysis* at the ALLC/ACH 2002, University of Tübingen, Tübingen, Germany. Parts of it were also presented in a paper on "TAPoR: Building a Portal for Text Analysis" at the COCH/COSH and SSHRCC Joint Session on "Mind Technologies" which was part of the *Congress of the Social Sciences and Humanities* at the University of Toronto in May, 2002.
3. Vannevar Bush in "As We May Think" (found in Nyce, *From Memex to Hypertext*) and later Douglas Engelbart in "Augmenting Human Intellect: A Conceptual Framework," developed the idea that information technology should augment or extend our capabilities rather than replace us. As Engelbart put it, "By 'augmenting human intellect' we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems." (p. 1)
4. For an account of the project see Busa, "The Annals of Humanities Computing: The Index Thomisticus." For a survey of the early history of computing in the humanities see Joseph Raben, "Humanities Computing 25 Years Later."
5. See the appendix on COCOA written by Robert L. Oakman in Howard-Hill, T. H., *Literary Concordances*. The Atlas was a computer developed in the early 1960s by the University of Manchester with Ferranti Ltd. See David Aspinall, "The ATLAS Computer: The Technology", <http://www.ukuug.org/events/linux2001/papers/html/DAspinall.html>, Accessed, March, 2003.
6. This discussion of the hermeneutics of text analysis is a summary of a longer discussion by the author, John Bradley and Patricia Monger in "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications."

7. Willard McCarty, "Introduction to ConCORDing and Text-Analysis: History, Theory, and Methodology." (p. 2).
8. *Encyclopedia Britannica Online*, in the section on "biblical literature", <<http://search.eb.com/bol/topic?eu=119719&sctn=14&pm=1>> [Accessed May 29, 1999]
9. In Lucian's dialogue "The Double Indictment" a character called Dialogue complains that he is "neither afoot nor ahorseback, neither prose nor verse, but seem to my hearers a strange phenomenon made up of different elements, like a Centaur." (p. 147) This dialogue includes Lucian's dialogical defense of his reuse of the dialogue form. There is hermeneutical tradition stretching from Lucian to Bakhtin pays attention to the dialogical mixing of elements into hybrids or *satura* as literature. The argument here is that the concordance is likewise a dialogical hybrid of other voices.
10. For more information on the TAPoR project see URL: <http://www.tapor.ca>.
11. A first prototype of the TAPoR tools is available at URL: <http://www.iath.virginia.edu/~gmr3f/tapor/main.html>, Accessed March, 2003. These prototype tools were programmed in Ruby by the author with the help of Stephen Ramsay to test the viability of the model. Another example of this model is HyperPo (See the article by Stéfan Sinclair in this collection.)
12. "Tapor", is the Old English for "taper" or candle (not "tapir" the odd-toed ungulate mammal.)

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

Bibliography

- Aspinall, D.** (2001). "The ATLAS Computer: The Technology", URL: <http://www.ukuug.org/events/linux2001/papers/html/DAspinall.html>, Accessed March, 2003.
- J. Bradley and G. Rockwell.** (1996). "Watching Scepticism: Computer Assisted Visualization and Hume's Dialogues ", *Research in Humanities Computing* 5, Oxford: Clarendon Press: 32-47.
- Busa, R.** (1980). "The Annals of Humanities Computing: The Index Thomisticus." *Computers and the Humanities*. 14.2: 83-90.
- Condillac, Etienne Bonnot de.** (2001). *Essay on the Origin of Human Knowledge*. Trans. **Aarsleff, H.** Cambridge, UK: Cambridge University Press.
- Engelbart, D. C.** (1962). *Augmenting Human Intellect: A Conceptual Framework*. Menlo Part, CA: Stanford Research Institute, on Contract AF 49(638)-1024, October 1962. Also available at <http://www.bootstrap.org/augment/AUGMENT/133182-0.html>, Accessed March, 2003.
- Engelbart, D. C.** (1963). "A Conceptual Framework for the Augmentation of Man's Intellect." *Vistas in Information Handling*. Ed. **P. H. Howerton and D. C. Weeks**. Washington D.C.: Spartan Books: 1-29.
- Gadamer, H-G.** (1985). *Truth and Method*. Trans. Glen-Doepel, W. 2nd ed. New York: Crossroad.
- Howard-Hill, T. H.** (1979). *Literary Concordances*. Oxford: Pergamon.
- Huizinga, J.** (1950). *Homo Ludens: A Study of the Play-Element in Culture*. Boston: Beacon Press.
- I. Lancashire, J. Bradley, W. McCarty, M. Stairs, and T. R. Wooldridge.** (1996). *Using TACT with Electronic Texts*, The Modern Language Association of America: New York.
- Lucian.** (1967). *Lucian: With an English Translation*. Trans. **Harmon, A. M.** 8 vols. New York: Macmillan.
- McCarty, W.** (1996). "Introduction to ConCORDing and Text-Analysis: History, Theory, and Methodology." *CETH Summer Seminar*. Ed. **S. Hockey and W. McCarty**. Princeton, New Jersey: CETH, Section 5.
- J. M. Nyce and P. Kahn,** Ed. (1991). *From Memex to Hypertext: Vannevar Bush*

Preprint of Rockwell, Geoffrey, "What is Text Analysis, Really?", *Literary and Linguistic Computing*, Vol. 18, No. 2, 2003, p. 209-219.

and the Mind's Machine. San Diego, CA: Academic Press.

Potter, R. G. (1988). "Literary Criticism and Literary Computing: The Difficulties of a Synthesis", *Computers and the Humanities*, 22.2: 91-97.

Raben, J. (1991). "Humanities Computing 25 Years Later." *Computers and the Humanities*, 25.6: 341-350.

G. Rockwell and J. Bradley. (1998). "Eye-ConTact: Towards a New Design for Research Text Tools", *Computing in the Humanities Working Papers*, URL: <http://www.chass.utoronto.ca:8080/epc/chwp/>. Also at URL: <http://www.humanities.mcmaster.ca/~grockwel/ictpaper/ictintro.htm>. Accessed March, 2003.

G. Rockwell and J. Bradley. (1999). "Empreintes dans le sable: Visualisation scientifique et analyse de texte", in *Litterature, informatique, lecture*, edited by **Vuillemin and LeNoble**, Paris: Pulim: 130-160.

G. Rockwell, P. Monger and J. Bradley. (1999). "Seeing the Text Through the Trees: Visualization and Interactivity in Textual Applications," *Literary and Linguistic Computing*, 14.1: 115-130.

Rockwell, G. (2001). "The Visual Concordance: The Design of Eye-ConTact", *Text Technology*, 10.1: 73-86.

Sperberg-McQueen, C. M. (1996). "Text Analysis Software Planning Meeting, Princeton, 17-19, May 1996, Trip Report". See <http://tigger.uic.edu/~cmsmcq/trips/ceth9505.html>, Accessed, June, 2002.