Afterword

READING THROUGH THE MACHINE

Finally, let's dream for a moment – a literary scholar's dream. Most of us who regularly turn to a computer to support our work might venture to have such a dream. In it, I am sitting comfortably reading, the computer next to me readily accessible. My thoughts about the text as I read can be recorded by the computer as they occur to me (perhaps I speak aloud while the computer's voice recognition system transcribes my words). Perhaps I begin to wonder about a prominent image in a novel I am reading: I can search for this image across the rest of the text or the whole corpus of the novelist's other writings. Or I may recall a thought I had while reading a particular poem last year and now wish to revisit it, perhaps add to it in the light of my present understanding. All the literature I have read (and perhaps much that I propose to read) is available through the computer; all the thoughts I have had during my reading are similarly available. As I read on the screen (sitting with my notebook computer, that is, in a comfortable armchair, with its wireless network connection), the system that I am accessing is an online portal personalized to access my selection of texts and the records of my reading, together with a powerful set of Internet tools for organizing and searching these texts (cf. the portal technology described by Siemens, 2009). Through the portal I

can also semi-automatically create personal web pages, the outcome of my searches and the connections I build between my responses and the texts I read. These are also stored online, available only to me, and for as long as I wish to keep them. The Internet also provides access to the reference works I might need: a dictionary, an encyclopedia, a chronology; each can be activated merely by highlighting a word or phrase on screen and requesting the relevant resource. This system is available to me whether I am an occasional, ordinary reader, a student, or a professional scholar of literature – it will adjust to my needs and offer me the resources appropriate to my interests. It will accommodate both the absorbed, experiential mode of the ordinary reader reading for pleasure, and the analytical, focused mode of the student. Moreover, a part or whole of this virtual structure could be designated as available for public access or access by a limited set of users, thus making it a central tool for discussions in an educational context.

This is just a preliminary sketch of the system I am dreaming about. Many of the elements needed to create it already exist, scattered across different applications, including several described in this book, but not yet brought together on the Internet in the way I have envisaged. As Susan Hockey (2000) remarked, introducing her survey of research with digital texts, these are "tools and techniques which ought to be available via the Internet, but at present are not […] The expectation is that these tools will be available in future versions of the Internet" (p. v). In this context, my primary focus will be on text analysis: the systematic analysis by computer of language, style, narrative forms, and other features. I will suggest that it is here, rather than in hypertext or virtual reality, that the power of the computer can provide an appropriate basis for literary reading. This will require developing and testing text analysis methods rather more systematically than hitherto. In addition, we will interface text analysis methods with another dimension rarely considered: the responses of actual readers.

Thus the computer will offer a facility for registering and mapping the responses of readers to the literary texts they read. Since the computer screen seems destined to become a medium on which readers of the future will experience literature, the scenario I sketch here – albeit a somewhat speculative one – will enable the computational power of the computer to be brought into play in support of reading and understanding literary texts. In particular, the power of the computer will support the significance of the individual's acts of reading. As Lev Manovich (2001) points out, the computer allows for individual variability as never before: "new media technology acts as the most perfect realization of the utopia of an ideal society composed of unique individuals. New media objects assure users that their choices – and therefore, their underlying thoughts and desires – are unique, rather than preprogrammed and shared with others" (p. 42). In this respect, the computer

will provide the ideal medium for exploring the individuality of our responses to the texts we read and then, if desired, sharing them with others.

Why would anyone choose to read a literary work on computer screen? Because the computer can help us to co-create and explore the imaginative world of the text. Our computer tools so far are not adept at this, but in this commentary I mention a little of what has been done, and what we might try to do.

The research I will draw upon in support of this model, then, comes both from the text analysis tradition, and from empirical studies of literary response (Miall, 2006). Both fields have largely been disregarded by mainstream literary scholarship. In combination, however, they offer, first, an opportunity to rethink our expectations of the computer in literary studies, and subsequently, with appropriate research, to help model and support what ordinary readers experience when reading. As I suggest, we can develop a system for the individual reader, whether engaged in a book for pleasure or a student of literature, drawing both on our knowledge of texts and on the systematic recording and analysis of the responses of groups of readers. Individuals will benefit in the long run from being able to engage at any given moment with their own image as readers, that is, with the encoded history of their responses, with the forms and processes embodied in each text they have read, and with the concerns they express through reading (i.e., the issues they are engaged with in their own lives). This will depend on building systems that do not yet exist, but many of the central pieces are already evident, certainly sufficient for a sketch of what the first stage of the building might include. My proposal might seem a Quixotic endeavour, but it is, to cite Jerome McGann (1998), an exercise in "imagining what we don't know" (p. 617).

So far the computer as a literary medium has been developed in three ways. First, as an expressive medium, it has been used to present hypertexts and, more recently, to open the prospect of immersive virtual reality narratives – a problematic development, as Ryan (2001) has suggested. Second, as a delivery medium, the computer is being used as an alternative to the book: editions of existing literary texts are being digitized, or "repurposed," in order to be read on screen or, at least, distributed digitally for printing out. Neither of these applications provides any knowledge about the texts they transmit, nor are they attentive to the reader on the other side of the screen beyond the basic executive functions of interaction, that is, clicking links in hypertext (virtual reality is more complex, but lacks the self-reflective role of literature). Here, then, we must turn to the third way, that of text analysis, which not only presents texts but can enrich our experience of them with all the computational power that is available through this machine. The computer can become a tool for thinking, for enlarging the conversations we have with ourselves while reading, and for focusing and systematizing our experiences. As McGann (1998) observes, "Computerized environments implicitly argue for dialectical

models" (p. 617). The machine will, in this respect, become our symbiotic partner as a reader, simultaneously "reading" both the literary text and our responses to it in order to pattern our identity.

This may also restore to readers the authority they have supposedly lost in post-modern (i.e., Foucauldian) theory, as sites acted upon rather than acting, appropriated by the discursive formations that in literature as in other discourses are said to define what is known. This, however, we can consider a research question that the reading machine may enable us to explore. The machine leaves open the question of how we are to understand the literary system: whether it can be considered text-based, driven by the distinctiveness of literature as a medium (as we have argued in previous studies: e.g., Miall and Kuiken, 1999), or is the effect of certain conventions in classifying and processing texts that came into place primarily in the eighteenth century, as the social constructivism of recent schools of literary theory now argue. The machine will offer us opportunities to assess these alternative views, since it will enable us to collect and analyze data about texts and about readers, as well as their interactions. It will allow us to consider whether the formal structures of a literary text influence reading regardless of the disposition or experience of the reader, or whether it is only the reader's knowledge of the conventions of literary reading that promotes formal features to become agents of influence.

This last issue, of course, echoes a debate of twenty years ago in which critics such as Stanley Fish (1980) and Barbara Herrnstein Smith (1988) attacked stylistics for its formalist premises. I mention this debate, since my proposed reading machine might be considered vulnerable to the same attack. Before I do so, I should clarify three other issues which might otherwise distract from the positions I present later.

First, I am not attempting to describe a machine that reads. Although research on natural language processing has been taking place for several decades, the building of a computer that can read any text it is offered with understanding (let alone aesthetic pleasure) is still a long way off. My aim is not a computer similar to the one imagined by Richard Powers (1995) in his fiction *Galatea 2.2* which is taught to interpret literary texts (and becomes too self-aware for her own good). Reading, especially literary reading, is one of the most complex activities we perform, and seems dependent not only on cognitive functions, but on feeling, and kinaesthetic and other bodily responses. We have recently begun to recognize the embodied nature of thinking, as scholars as diverse as Hayles (1999) and Damasio (1999) have shown. Thus, for a computer to read as we do would require something equivalent to the porting of human subjectivity into the computer. While in 1988 Hans Moravec announced this as a possibility within a few decades, Hayles's (1999) discussion provides strong grounds for considering it improbable, a roboticist's fantasy. By the shorthand phrase "reading machine" I mean a support system for reading, a knowledge-base, a program with some similarity to an expert system, able to model

information, to reason about it using Bayesian logic or a similar stochastic process, and to reflect its findings in an intelligible way to the reader.

Second, the attempt to model literary reading might be considered misguided. Isn't reading too fugitive and idiosyncratic to be amenable to systematic analysis? This certainly seems to be the view implied by current scholarship, which continues to elaborate theories and readings of texts with no regard for the empirical support that might be obtained from actual readers. While reader response theory has formed a significant part of this work, it has been the tradition here as in the rest of literary theory to ignore the actual reader. Jonathan Culler (1975), for instance, declared his indifference to such investigations, remarking that the important question was "what an ideal reader must know implicitly in order to read and interpret works in ways which we consider acceptable, in accordance with the institution of literature" (p. 123; see also Culler, 1981, p. 129). This was a question for theoretical reflection, not data collection. At the same time, complex theoretical arguments continue to be advanced about the interpretive moves that an ideal reader must make. Yet the question of idiosyncrasy is answered in numerous empirical studies of reading: almost all show that in various ways reading is systematic, that different readings, although they may differ interpretively, do so on grounds that are partly recoverable; and on examination, such readings are found to be determined in part by rules that we can trace either to texts or to the psychological processes of reading (including the cultural influences that mediate the process). It is this systematicity that will form an essential core of the reading machine and that will account for its fruitfulness.

Third, my reference to producing data on the formal structures of literature may remind some of my readers of an earlier debate about stylistics, focused in particular on the analysis of Baudelaire's "Les Chats" by Roman Jakobson and Claude Levi-Strauss (1972). This, for Stanley Fish (1980) was an example of the "monumental aridity" (p. 94) of Jakobson's stylistic analyses in general, in which every formal feature that could be analysed was included. As Fish (1980) puts it, in this approach there is nothing in place to govern the field of description, thus "there is no way of deciding either where to begin or where to stop, because there is no way of deciding what counts" (p. 94). While Fish (1980, pp. 322-326) appeals to the experience of the reader as the domain in which formal features are recognized, he is one more theorist for whom actual readers remain of no interest (that is, beyond his bizarre and much-cited "experiment" in asking his students to read a list of names left on a blackboard as a poem). Thus, he fails to suggest the appropriate response to Jakobson's work: that in collecting formal features the task should, first, be driven by hypotheses about reading; and second, the results should be tested against the responses of actual readers. In other words, in the light of a hypothesis about what is at issue in reading, we devise some measure of the reading process. Then we use

an array of formal features we believe to be implicated in reading as predictors of the reading measure – measures range from collecting brain scans during reading, to the analysis of talk-aloud protocols.

Fish's (1980) objection to Jakobson is a double-edged sword. It could apply equally well to Fish's (1980, pp. 21-67) own postulated data about the reading process, his "affective stylistics" which was based on a model of supposed expectations and disappointments as readers construe, or misconstrue, each line of a text. We could, indeed, consider a good deal of literary theory and interpretation in the same light, as unchecked fields of description with no principles in place to define the limits of what might count. While Fish (1980) or Culler (1981) would answer that the conventions of literary reading control what counts, I suggest that this is also an empirical question. What counts should, where possible, be investigated by examining what actual readers are doing. If Fish's (1980) affective model is correct, then it should be possible to design a study to validate it with the experiences of actual readers. The elaboration of such theories, proposing specific types of interaction between text features and readers, is the essential first step in outlining what an analysis of reading should be taking into account. But invariably this is as far as literary scholarship ever gets (including hypertext theory): it makes "implicit premises" on texts and on readers, as Cees van Rees (1985, p. 445) expressed it, in his critique of Gerard Genette's (1980) narrative theory. So many categories of narrative mood, tense, time, etc., are elaborated by the theorist: but, the question should then be, are readers actually influenced by them? Almost certainly they are, at particular times and in specifiable ways, but Genette (1980) is typical in his indifference to this question (although late in his book he acknowledges the importance of Proust's remark, that as readers we are reading ourselves).

Here, then, is the first step towards the computer system we wish to build. On the text side, we take a poem such as Coleridge's (1924) "The Rime of the Ancient Mariner" (186-209) and we not only encode the poem for presentation on screen (so that readers know which part they are reading, which line numbers), but also we analyze the narrative and stylistic aspects of the poem: the turning points of the narrative, the most ambiguous parts, the lines with the most striking poetic features (metaphors, alliteration, etc.) that also seem to typify the tone of the poem overall (but keeping this information off-screen until readers request it). We will initially glean much of the information we need in this respect from systematic empirical studies with readers of the poem. When a reader pauses at a passage that seems particularly interesting or surprising (a passage that the readers in our prior empirical studies will probably also have found striking), the system can be queried: what is known about this passage in terms of the narrative, or its stylistic qualities? With additional facilities, that I will describe later, the system can also sketch its relation to other work by the poet or other poets (a menu of such choices

will be offered on screen). In this way, readers can explore their understanding of the poem, whether to reflect on it, to extend it, or perhaps help to correct it (if, say, a reader has misunderstood a phrase).

"The Mariner" has been interpreted in widely different ways. While critics have diverged from each other in their view of the poem, so have the readers we have studied. While critical differences can be imputed primarily to the theoretical position held by the professional reader (intent on expounding a historicist reading, say, or an ecological one), our volunteer readers more typically depended on their own experience and understanding of life. The reading machine will also accommodate this. Here is one brief example: the responses of two readers to a passage late in the poem. As his ship travels home, the Mariner describes himself as looking out and seeing nothing but the ocean, yet he is:

Like one, that on a lonesome road
Doth walk in fear and dread,
And having once turned round walks on,
And turns no more his head;
Because he knows, a frightful fiend
Doth close behind him tread. (Coleridge 1924: 203)

This passage is undoubtedly disturbing and represents a challenge for interpretation. Both of the readers I cite below made quite extensive commentaries after choosing this passage, but I pick out what seems to be a central concern in each. One of them said:

It feels like, this, there's this certain point in your life that you can't turn back. You know that if you try and turn back or stop what you're doing, it's over, nothing good will come out of it. You have to continue on. It's like riding a bike downhill, you know you can't stop it because if you stop, you know you'll crash… It could be a metaphor for life, there's no turning back once you get started.

In contrast to this sense of the imperative forward movement of life, another reader spoke of guilt over missed opportunities:

This passage also describes someone who lacks a high self-confidence, I guess. It's like the person seems to see his abilities. And so he's trying to avoid things. I know I've done that at times, in situations when I don't feel like I mingle to do things the way I want to. I sometimes turn away. But afterwards, the frightful fiend that follows me, it's not really anything that's external, it's more internal. It's my, well,

guilt is a strong word, sort of regret that I didn't have the courage to deal with it better. It's something that haunts me sometimes.

Whether the concern of either of these readers is more than a passing response to the poem is not clear, although each seems to touch on some long-standing issue in his or her lives. In both cases this was the last of five passages considered by the reader, and in response to previous passages both protocols show indirect anticipations of the concerns reported here, but we cannot tell how important these might be. The reader may not be clear what importance the concern has either, but given a facility for recording such comments, the place of the concern in the reader's life might become evident in the longer term. Other readings may evoke the same issue, allowing it to be reconsidered or developed – or resolved. Just as the features of the poem can be indexed by the computer, so the comments of the reader can be coded, sorted, and ordered for later retrieval, for the reader's reconsideration, whether in another context or in rereading the same text.

A reader might also wish to turn back to the poem itself and probe its language a little further. For instance, given the mysterious nature of the "fiend" in this verse, we can ask if the word occurs elsewhere in Coleridge's writing, in case some similarity in use casts a light on its occurrence here in "The Mariner." The "fiend" appears only once in this poem, and it is the more problematic for being a figurative fiend, not a literal one (the Mariner describes himself as "Like one, that on a lonesome road"). The question can be explored by means of a text analysis program. Here we envisage a facility for surveying all of Coleridge's poetry and picking out passages where the word "fiend" (or "fiends") occurs. While such concordance programs have existed for some time as free-standing computer applications, only very recently has the same activity become possible through the Internet, although only basic look-up facilities are available so far, e.g., *Open Source Shakespeare* (http://www.opensourceshakespeare.org), and these are tied to a specific text, except for Sinclair's *HyperPo* (http://hyperpo.org/). In the Coleridge example, we wish to know, of course, where other occurrences of "fiend" are located, so the text of the poetry is encoded with the titles of the poems, at a minimum, and preferably line numbers as well to help us find our place in longer poems.

A concordance of Coleridge's poetry for "fiend" and "fiends" yields 19 instances. Several of the fiends turn out to be rather conventional: they are mostly from hell (Milton seems responsible for some of these), or they are figurative representations of evil. But two of the examples seem more relevant to the mysterious fiend of "The Mariner": in "Religious Musings" (Coleridge 1924, pp. 108-125) and, especially, "Pantisocracy" (Coleridge 1924, pp. 68-69). These suggest that the Mariner's fiend may, like those in "Pantisocracy," be internal, an agent beneath consciousness known only from dream states or other unconscious promptings. This would help elucidate

the obscure lines that introduce the simile, where the Mariner says he "looked far forth, yet little saw / Of what had else been seen." "Pantisocracy," in other words, helps to gloss the Mariner's comment: it suggests that what he would have seen if he had looked in the right place (within instead of "far forth") would have been like a pursuing fiend. The fiend, then, may be a part of the self.

As a reader of "The Mariner," making use of the reading machine, I can thus quickly put in place a verbal environment drawn from the rest of Coleridge's poetry that will help enrich my understanding of the verse before me. This depends both on facilities which are commonplace in concordance programs and several other facilities that could readily be developed to support the reader. At the same time, my own responses to the verse I am reading are being recorded, if I wish, on the computer. Since we can envisage using the reading machine over many years, a large body of comments will accumulate over time. These too, however, will be searchable through the same concordance and collocation facilities that gave instant access to Coleridge's poetry. Thus, if I wish to review what I have thought about a particular concern I enter a word, or a combination of words for a collocation search, and bring up all the recorded instances. At any time during my reading of "The Mariner" (once I have performed the inquiry into the "fiend," for example), or after assessing a particular concern, I can output the screen pages that record the results of my inquiry as a web page, and have this permanently linked to the contents of my record. In other words, while engaged with the text, I can produce and interlink a hypertextual representation of my image as a reader. When I wish, I can switch to this representation (sidelining the text I am currently looking at, perhaps) and examine, by clicking from one linked view to the next, what has moved me as a reader, how I explored it, when a given insight occurred and in what context, and, in summary, who I am as a reader (information that will be kept secure, private to my Internet portal). Then, if appropriate, designate part of my response record available for viewing by others (if working within an educational context, or a reading group).

I could, of course, simply keep a written diary of my reading, which I index from time to time. What reasons are there to believe that a computer-based system, offering encoded texts and comments, will provide a significantly more powerful context for reading to make it worth building? Here is where the design of a knowledge-base containing literary texts and readers' comments about them is critical: it will take its shape, at least initially, from what we know about the structures of literary texts and their determining effect on readers. This presupposes that literary response is far from idiosyncratic, that it is impelled in systematic ways by the structural and linguistic features of texts. Put another way, even though individual interpretations may diverge, certain kinds of attention are promoted by the underlying field of forces that each reading encounters. The interactive processes that are set in play can then

be analyzed and modeled, and a support system put in place through the computer that includes information about both the text and the individual's responses.

What is currently little appreciated is that a number of examples of empirical research now show that readers' responses to literature are far from idiosyncratic, although they may vary one from another on interpretive grounds (e.g., Martindale & Dailey, 1995). The source of the commonalities we find in response appears to be located to a significant degree in the literary text itself, in its phonetic, figurative, narrative, and other features, which compel a certain kind of attention from readers. This makes possible a range of text analysis techniques which we can apply to map texts and to provide a support system for readers, since we know that readers are likely to be addressing the same features of texts as they read, even if the questions differ that they bring to the system. While further studies of readers' responses are required to test and validate such computer-based techniques, this approach alone will not be enough. The main limitation of text analysis methods on their own is that they cannot address the fundamental feature of reading literary texts: their modifying power over the reader. Text analysis methods typically produce static maps of features. It is the recording and patterning of readers' responses that overcome this limitation.

The system I have described comprises the first stage in building the reading machine. Next we envisage a set of programs for recording and analyzing a group of readings, in which the probabilities and rules of interpretation are inferred from the text elements noticed by readers and from their verbal or written responses to such elements. This would contribute to the building of the subsequent program, where this information is made available to the current or a later reader. This would provide all the facilities we have already sketched, drawn from text analysis and from representations of the reader's own responses. But now this can be enhanced by information from others. Once a reader had entered some data about the response, the system would match the new data to its existing data, and make predictions about the likely path of the reading and about its overall direction. If the user requested it, the program would thus be able to provide advice specific to the point the reader had reached. Each time more data was entered, the program would update its probabilities. A system built on these principles would thus allow for a range of different (and possibly incompatible) readings of the same text: more interestingly, it would also give access to some of the underlying causes for differences in readings and provide a logical basis for analyzing such differences.

The reading machine will thus not only accommodate such divergences of interpretation, it will facilitate them. It will enable readers to engage in a dialogue with the computer about their understanding of a text, and through the computer to relate their developing understandings to what other readers have thought, and what other evidence supports such readings. "Computer technology," as William

Winder (1996) puts it, "like the codex itself, is the basis of a new incarnation of dialogue, and the source of renewed collaboration. The *reincarnation of dialogue* is the ultimate object of humanities research and the source of humanists' fascination." For the same reasons, the reading machine will reinvigorate reading, and may one day become the preferred medium for our experience of literary texts.

David S. Miall
University of Alberta, Canada

REFERENCES

Coleridge, S. T. (1924). *The poems* (E. H. Coleridge, Ed.). Oxford: Oxford University Press.

Culler, J. (1975). *Structuralist poetics: Structuralism, linguistics, and the study of literature*. London: Routledge & Kegan Paul.

Culler, J. (1981). *The pursuit of signs: Semiotics, literature, deconstruction*. London: Routledge & Kegan Paul.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace.

Fish, S. (1980). *Is there a text in this class?* Cambridge, MA: Harvard University Press.

Genette, G. (1980). *Narrative discourse*. (J. E. Lewin, Trans.). Ithaca: Cornell University Press.

Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. Chicago: University of Chicago Press.

Hockey, S. (2000). *Electronic texts in the humanities: Principles and practice*. Oxford: Oxford University Press.

Jakobson, R., & Lévi-Strauss, C. (1972). Charles Baudelaire's 'Les Chats.' In R. T. De George & F. M. De George (Eds.), *The structuralists: From Marx to Lévi-Strauss* (pp. 123-146). New York: Anchor.

Manovich, L. (2001). *The language of new media*. Cambridge, MA: MIT Press.

Martindale, C., & Dailey, A. (1995). I. A. Richards revisited: Do people agree in their interpretations of literature? *Poetics*, *23*, 299-314.

McGann, J. (1998). Textual scholarship, textual theory, and the uses of electronic tools: A brief report on current undertakings. *Victorian Studies, 41*, 609-619.

Miall, D. S. (2006). *Literary reading: Empirical and theoretical studies*. New York: Peter Lang.

Miall, D. S., & Kuiken, D. (1999). What is literariness? Three components of literary reading. *Discourse Processes, 28*, 121-138.

Powers, R. (1995). *Galatea 2.2*. New York: Farrar, Straus, Giroux.

Ryan, M.-L (2001). *Narrative as virtual reality: Immersion and interactivity in literature and electronic media*. Baltimore: Johns Hopkins University Press.

Siemens, R. (2005). Text analysis and the dynamic edition? A working paper, briefly articulating some concerns with an algorithmic approach to the electronic scholarly edition. *Text Technology*, *14*, 91-98. Retrieved from http://texttechnology.mcmaster.ca/pdf/vol14_1_09.pdf

Smith, B. H. (1988). *Contingencies of value: Alternative perspectives for critical theory*. Cambridge, MA: Harvard University Press.

Van Rees, C. J. (1985). Implicit premises on text and reader in Genette's study of narrative mood. *Poetics*, *14*, 445-464.

Winder, W. (1996). Texpert systems. *Computing in the Humanities Working Papers* B.35 (1996). Retrieved from http://www.chass.utoronto.ca/epc/chwp/winder2/