THE HISTORY, DISCOVERIES, AND AIMS OF THE CANTERBURY TALES PROJECT

by Peter Robinson

In memoriam: John Manly and Edith Rickert

Every year countless people throughout the world encounter the *Canterbury Tales* in editions, translations, and adaptations. Ultimately, all these many different forms of Chaucer's work derive from a single source: the text he composed sometime between 1385 and his death in 1400. We have no direct knowledge of this text. We have no authorial manuscript of the *Tales*, nor indeed any single manuscript explicitly authorized by Chaucer, or whose copying was unambiguously supervised by Chaucer. Like Shakespeare, and unlike such contemporaries as Gower, Petrarch, and Boccaccio, we have no evidence of any systematic attempt by Chaucer to regulate the publication of his work. Notoriously, too, Chaucer left the *Tales* unfinished, adding another layer of uncertainty to our ignorance of the first states of the text.

What we actually have are some eighty-four manuscripts of the *Tales* and four early printed editions dating before 1500.[1] This evidence, then, sets two classic scholarly problems. First, what can we deduce from this mass of manuscripts? Second, how should we present what we find to the reader? Since Tyrwhitt, the history of *Tales* textual scholarship is a record of the attempts by various scholars to grapple with these problems.[2] The most ambitious of these attempts was that by John Manly and Edith Rickert at the University of Chicago from 1920 on and resulting in their massive eight

volumes *The Text of the Canterbury Tales Studied on the Basis of All Known Manuscripts*.[3] Before Manly and Rickert, there had been several careful but necessarily limited explorations of the manuscripts of the *Tales*.[4] It is their distinction that they saw that a full account of the whole tradition of the *Tales* must rest on analysis of all the available evidence in all the manuscripts. Logically, this meant discovery of all the manuscripts (itself a massive task), exploration of all that could be learned from the manuscripts as to their origins and later provenance which might cast light on the text they contain, word-by-word collation of the whole text of every manuscript, and analysis of this collation. This is a breathtaking program. One measure of their achievement is that in the years since, only one more manuscript of significant portions of the *Tales* has been discovered;[5] another measure is the extent to which techniques they pioneered and developed of palaeography and codicology are now commonplace in Middle English manuscript scholarship. Throughout their volumes, one is awed by the acuity of their insight, their energy, and the depth of their knowledge: they knew the text of the *Tales*, in every manuscript, at every word, as no one ever did before and perhaps as no one ever will again. At this distance, however, it is clear that they did not achieve their main aim: the establishment on the basis of their analysis of a text that would be accepted by later scholars. In part this was because the vast amount of data they accumulated overwhelmed their methods of analysis; in part because the methods both of data gathering and of analysis themselves were flawed. However, perhaps their greatest failure was their response the second general editorial problem given above: how should what is discovered be presented to the reader? The presentation of variants in their volumes 5 to 8 is obscure, to say the least. Even more damagingly, the connections between the

collation data, their analysis, and the text they present are unclear. While scholars since have paid tribute to the scope of their endeavour and their knowledge, no later edition has followed the text they gave in their volumes 3 and 4.

It is possible to argue that Manly and Rickert's first failure – to present a convincing and coherent analysis of the history of the tradition – derived from the inability of the manual methods of collation and analysis that they had to use to cope with the millions of words in these eighty-plus manuscripts. By the late 1980s, pioneering work by several scholars on methods of computer-assisted collation and analysis offered promising new approaches to this problem. A grant from the Leverhulme Trust in 1989 to Susan Hockey, with myself as the researcher, gave the opportunity to explore this with two manuscript traditions, one of them that of a Chaucer text. The beginnings of the Canterbury Tales Project lay in a series of experimental transcriptions, collations, and analyses carried out in this project in Oxford by myself and Elizabeth Solopova from 1991 to 1992, on the *Wife of Bath's Prologue*. These experiments gave encouraging results, and Solopova and I joined with Norman Blake in 1992 to found the Canterbury Tales Project, with Blake as the first director. The project aimed to carry forward the work done on the manuscripts of the *Wife of Bath's Prologue* across all the texts of all the *Tales*. At this point, we set out the first aim of the project: to explore the textual history of the *Tales* by transcribing, collating, and analyzing the manuscripts of the *Tales* using computer methods. This aim corresponds to the first scholarly editing task outlined above, to deduce what we could from the mass of witnesses. Grants from the British Academy, University of Sheffield, and Oxford University enabled the project to start on

this basis, with part-time researchers in Sheffield (Estelle Stubbs and Michael Pidd) joining myself and Solopova.

Over time, the project expanded. Doctoral students, first at Sheffield and later at De Montfort University, wrote dissertations on individual manuscripts and aspects of the tradition, both developing and making use of the project's resources.[6] There are now some dozen researchers, students, and associates in five universities working under the broad umbrella of the project, and we expect to have transcribed, researched, and published over forty percent of all the texts of the more than eighty fifteenth-century witnesses of the *Tales* by mid-2004, when our current five-year funding from the Arts and Humanities Research Board concludes.[7] Naturally, with so many different people, over such a period, and in the context of so vast a set of questions, there have been many different approaches and many differing conclusions drawn. In what follows, I will try to present what seem to me the least contestable results of our work – but others in the project would present a different summary.

Concerning our work on the analysis of the textual tradition of the *Tales* (the first of the two major questions we must address), we can divide the results into three sections: conclusions about relationships among all the texts in the tradition; conclusions about the earliest state of the text; conclusions about the value of individual manuscripts. We have followed two different directions in our research. The first is analysis of all the witnesses to particular sections of text, where we transcribe and collate all the witnesses and then use sophisticated computer programs – some drawn from evolutionary biology, some which we have developed ourselves – to create a view of the relations of these witnesses from the record of agreements and disagreements provided by the collation.

The theoretical and practical base of this work has been greatly strengthened in the last three years by our collaboration with experts in phylogenetic methods in the Department of Molecular Biology in Cambridge.[8] Three sections of the *Tales* have now (June 2003) been analyzed in full by these methods; a further three will be analyzed by May 2004.[9] The second is analysis of individual witnesses, in terms of their history and codicology: this analysis has been the work of doctoral students mostly under the supervision of Norman Blake, and some nine witnesses have been or are being studied in this manner.[10]

Conclusions about relations among all the text in the tradition

Our work has confirmed Manly and Rickert's suggestion that many of the witnesses of the *Tales* may be grouped into four large groups: **a b c d**. We have clarified the relationships between these: it appears that these are actually two groupings **ab** and **cd**, each descending from a common exemplar (at times, Manly and Rickert suggest this, somewhat hesitantly; see for example 2:42). We have been able to advance beyond Manly and Rickert in analysis of the witnesses that cannot be allocated to these groupings, witnesses we term loosely the **O** witnesses: in many cases these represent independent lines of descent from the originals.[11]

At various places, Manly and Rickert appear to suggest that some, at least, of the differences between these manuscripts are due to "part publication": that is, sections of the *Tales* circulated independently and so have distinct textual histories (e.g., 2:133, 2:193, 2:489). However, our research points in a different direction. The detailed analyses of the relations between all witnesses in three separate sections of the *Tales* suggest that the differing family relations in these sections can be explained by shifts within the

tradition, after this first generation of copies from a single set of originals, and not by the existence of multiple sets of originals from which multiple traditions might arise. That is, we do not have evidence that Chaucer produced separate versions of distinct tales, which circulated independently of each other. This is not to say that this did not happen: just that no trace of any such independent copies has survived.

There are two major arguments in support of this hypothesis of a single textual history for the whole *Tales*. The first is the persistence of the **ab cd** groupings (and, of pairs and other groupings within the **O** witnesses) across the whole text. It is difficult to explain why the different sections should have the same witnesses in the same relations with each other if they were originally "published" separately. The second is the discovery by Barbara Bordalejo that the same pattern of relationships is found in grouping manuscripts by their tale ordering as in grouping the manuscripts by their word-by-word textual variation.[12] This suggests that the same history explains the tale orders as explains the text itself. This discovery removes one of the most attractive arguments for the "part-publication" model: that some of the tale-order fragments might represent such original parts. But if this were so, then how could the history of the text within each fragment be the same as for other fragments? And, how could this be so if the whole history of the arrangement of the fragments into more or less connected sequences also has the same history as the text within each fragment? So far, we have found the same textual history in each section; and we have found that the arrangement of the sections themselves follows the same history again. The easiest explanation for this is that the whole text descended from a single set of originals, copied as a whole at the earliest stages of the extant tradition, and that the arrangement of the parts also evolved with the

text itself, through the same stages of "descent with modification" that created the tradition as we have it. An advantage of this view is that it simplifies understanding of what happened. We do not need to create a new set of family relations for each stretch of text; we can use what we learn of one part of the text to illuminate another.

From this, we suggest that all extant copies descend from a small number of manuscripts produced by a group of scribes within a few years before and after Chaucer's death. Several of these scribes knew each other (three of them copied stints of the Trinity Gower manuscript), and they seem to have worked closely together in making the first copies of the *Tales*.[13] Surviving manuscripts from this period appear to be Aberystwyth, National Library of Wales Hengwrt Peniarth 392D (Hengwrt); San Marino, Huntington Library 26.C.9 (Ellesmere); Oxford, Corpus Christi MS 198 (Cp); Cambridge, CUL Dd.4.24 (Dd); and London, British Library MS Harley 7334 (Ha4). All of these seem to be independent copies derived from a single set of originals. Further, other manuscripts and witnesses (some of these much later) appear to descend independently from other distinct but now lost copies made in the same period. Together with these putative first-generation copies, all these form what I have called the **O** witnesses: manuscripts which cannot be shown to be related within the subfamilies (notably, the **ab**, and **cd** groupings) and so must be considered as evidence of the existence of several other independent and early copies no longer extant. Key examples of these later **O** witnesses are London British Library Additional MS 35286, Oxford Christ Church MS 152, and Cambridge UL Gg.4.27; there are several others that await detailed study. Establishing as far as we can the exact relations among all these, and between these and all the other manuscripts and the presumed originals, remains our major task.[14]

Conclusions about the earliest state of the text

Our work suggests that all extant copies descended from a single set of originals. What can we determine about this "single set of originals"? It appears that this was not throughout an ordered "fair copy," with all sections existing in polished text with a settled and well-defined sequence of prologues, tales, and links from start to end. At least one part, the whole of "Fragment I" from the *General Prologue* to the end of the *Reeve's Tale*, does seem to have existed in such a state: hence the unanimity in the manuscripts in this sequence.[15] But in other sections there is evidence that the originals contained material in something more like a "working draft" form, with no clear indication as to how the separate sections should be connected. In a few cases, we have evidence of Chaucer reorganizing material – assigning tales to different tellers, for example – but not completing the revision. In others, it appears that the originals contained separate versions of linking passages, or drafts of links that were not patched into the surrounding texts. It appears too that some pages may have had lines, or whole passages, either first written within the text but marked for deletion, or written elsewhere on the page but marked as additions to the text. This meant that at each such point, the scribe would have the option of deleting or adding the passages in question. Each of the first generation of copyists from these originals seems to have made a slightly different set of decisions: for the choices of the Hengwrt/Ellesmere scribe, see below.[16]

A complicating factor would have been the physical form itself of these papers. In some cases, a single tale and associated links might form a discrete booklet of one or more quires (compare, the *Parson's Tale*, and possibly the *Nun's Tale* [=SNT] in

Hengwrt, both of which Stubbs suggests might have been copied separately from the rest of Hengwrt).[17] It is notable that several tale segments have around 650 lines, about the right amount to fit neatly into a regular quire of eight leaves (accommodating between 624 and 672 lines, at 16 pages with between 39 and 42 lines per page).[18] But many others exceed this, yet do not have enough to fill out a second regular quire without being linked to another tale; others fall short of filling a quire on their own without linking to another tale. Where such links had not yet been settled (as they had been in Fragment I), tales must exist on sequences of leaves more or less loosely bound with one another, while short passages such as links (or, discrete sections such as the Adam stanza, or some passages in the *Wife of Bath's Prologue*) might exist on single sheets of paper, even scraps. Into this mixture we introduce the possibility of Chaucer revising passages or whole links on yet more sheets of paper, with the sheets carrying the new passages coexisting in the same pile of paper as the sheets carrying the first version of these passages. From this, we have a sense of the problem that faced the first scribes to work on this material – and that continues to trouble (or ought to trouble) Chaucer's editors and readers, six hundred years on.

One might expect, following this hypothesis of a somewhat disordered set of working papers, that these originals might also have contained alterations to single words, deriving from Chaucer's own tinkering with the text. Thus one might find evidence of such revisions scattered almost at random through the first generation copies and their descendants (collectively, the **O** witnesses). Surprisingly to modern readers used to heroic authorial revision operating at the level of the individual word, there appears little or no evidence of such revision. Indeed, the extraordinary unanimity of the **O** witnesses

word by word throughout the text remains the strongest argument for the descent of all extant manuscripts from a single set of originals, rather than from multiple sets of originals "published" at different times. Further, the suggestion that Chaucer may have rewritten (or added, or deleted) whole links or distinct passages, but did not bother with methodically revising the text word by word, casts an interesting light on his revision practices (the two versions of the Prologue to the *Legend of Good Women* are also relevant here).

Conclusions about the value of individual manuscripts

The identification of the project with Norman Blake, who is well known for his forceful advocacy of the Hengwrt manuscript, and the fact that our first "single-text" publication was Estelle Stubbs's Research Edition of *The Hengwrt Chaucer Digital Facsimile*, might encourage the supposition that we had prejudged the issue and decided the merits of Hengwrt in advance of any analysis.[19] Rather the reverse is the case: while Hengwrt retains the "first among equals" status, which it has had at least since Tatlock's advocacy in 1935, our work has qualified its position.[20] First, there is the perennial question: Hengwrt or Ellesmere? Our research suggests that the relationship between the two manuscripts cannot be simply categorized.[21] For long stretches of the text, the manuscripts are extraordinarily close in their substantive readings: for the *General Prologue* and the *Miller's Tale*, around one substantive difference only every eight or so lines, with even these rarely affecting meaning or meter. What is more telling is the number of times a reading shared by Ellesmere and Hengwrt with (sometimes) very few other manuscripts, and against the overwhelming agreement of the other manuscripts, is

patently *lectio difficilior*, and patently the original: a nice instance is *Miller* 605, where just eight manuscripts have "my derelyng" against forty-six with the easier "thy derelyng." Where the manuscripts do differ, the difference defies easy explanation. Our analysis is unable to support either the assertion (first made by J. Koch and often repeated) that Ellesmere is "metrically smoother,"[22] or the assertion that its text shows consistent evidence of editorial sophistication (as Manly and Rickert suggest, 1:150). Solopova demonstrated that for the *Wife of Bath's Prologue* at least the first is incorrect.[23] The random nature of the substantive differences we have found, most of which appear to be the result of simple copying error, argue against the second.[24] This closeness of the two for most of the text suggests that they were copies descended from a single original. However, there are famous differences between the two: Ellesmere includes a whole tale and many lines and passages not present in Hengwrt; the tale orders in Ellesmere and Hengwrt differ. It appears that the scribe of the Hengwrt manuscript was rigorous about not including text marked for deletion, and did not include (generally) text marked for addition (if, indeed, this text had at that time actually been written onto the originals). When the same scribe came to copy the Ellesmere manuscript, he appears to have followed a different policy, of including all the text that was available. Further, the aspect of the hands and many differences of spelling, some marked variations in the text and particularly the quite different "production values" of the two argue for some gap in time between the copying of the two manuscripts. Even though the two appear independent copies of the one set of papers, this set of papers might itself have changed in key respects in the period between the writing of the two manuscripts.

In the Research Edition of *The Hengwrt Chaucer Digital Facsimile*, Estelle Stubbs narrates how the Hengwrt scribe appears to have struggled with the material as it came to him: leaving space for the ending of incomplete tales; having tales arriving without links or other indication of precise order and so having to copy them in a particular order, leaving space for the links which were then copied in when they arrived. On this view, the significance of the Hengwrt Chaucer is that, of all the surviving manuscripts, it appears in many respects closest to this lost set of originals: it reflects, more than any other, the discontinuities and ambiguities we think present in the papers Chaucer left behind him. Accordingly, the Hengwrt Chaucer offers a unique help towards understanding what is the single most important issue in Chaucer textual scholarship: what were these originals? However, the very terms of this assertion undermine any claim to unique authority for Hengwrt: it is a remarkable view of these originals, but there are indications that it is a partial view only. Indeed, the contradictory nature of these originals, as we presume them, suggests that no one manuscript could ever offer more than a partial view. Hengwrt excludes much text – the whole *Canon's Yeoman's Tale*, and many shorter passages – commonly agreed to have been composed by Chaucer for the *Tales*;[25] it has others in an apparently confused ordering (notably, in the sequence Squire-Merchant-Franklin); and there are certainly copying errors. All this suggests that we need to look outside Hengwrt; and look, too, not just at Ellesmere.

Indeed, close analysis by other project researchers of other manuscripts has suggested that this picture – of scribes struggling to make as complete and coherent a text as they could from whatever they began with – is general across the *Tales*: thus Orietta Da Rold's thesis on Cambridge UL Dd.4.24, and Simon Horobin's on BL Additional MS

35286. In both these, glimpses of the older originals can be seen. Particularly remarkable is William Caxton's second edition of the *Tales*, probably published in 1482. According to Caxton's preface, he used a different manuscript in preparation of this second edition to that he used for the first edition published six years earlier. In fact, he used this manuscript to correct his first edition, rather than simply setting the whole text anew from the manuscript. In a recent doctoral dissertation Barbara Bordalejo has studied the 3,000-plus significant changes Caxton made, and concluded from them that this manuscript was indeed one of remarkably high quality, extremely close on many points to other **O** witnesses.[26] At a few places it seems to have contained readings different from those in the Hengwrt and Ellesmere manuscripts, but whose distribution in other **O** witnesses and the nature of the readings suggest that they are more likely to have been present in the originals. A notable instance is the *Reeve's Tale*, line 9, where Hengwrt and Ellesmere (and many others) read

> And by his belt he bar a long panade

Cx2 and a significant number of other **O** manuscripts read "Ay" for "And": both a sharper and more difficult reading, and highly likely to have stood in the originals. Once more, such instances defeat easy summary of our conclusions. In most cases, the consensus of Hengwrt and Ellesmere defines the likely original, but not in every case.

This is what we have found in answer to the first question posed above: what can we deduce from the tradition? A project of this scale must change as it progresses, and this is particularly so of our attempts to answer the second question: how should we

present what we have found to the reader? From the first, we determined that we would publish what we could, when we could. We were doing all the work of data collection and analysis by computer. Just as we were embarking on this work, the first CD-ROM publications were appearing, and the emergence of SGML in the Text Encoding Initiative implementation offered an encoding framework into which all our data could be fitted, to enable such publication. Already, our work on the *Wife of Bath's Prologue* had generated more information than could possibly be contained in any kind of printed form. Naturally, then, we looked toward some form of electronic publication. In 1993, providentially, Cambridge University Press became interested in our work, partly as a means of exploring the possibility of electronic publication for themselves. With their considerable support, we were able to achieve publication of *The Wife of Bath's Prologue on CD-ROM* (1996) and *The General Prologue on CD-ROM* (2000).

Both of these are "single-tale" publications, presenting the whole text of all the witnesses to one section of the *Tales*. However, we were very conscious of the deficiencies of both these. By their nature, one does not expect to "read" such collections of data. They appeared more like repositories of information, from which skilled scholars might quarry what they need. But this presumes high levels of expertise and interest among readers. Would it be possible to present these in a manner that would make them much more accessible to a much wider range of readers? Who are, or should be, or can be, our readers? As we considered this, we were becoming aware that the revolution in publishing brought about by the advent of the computer was being matched by a revolution in the theory of textual scholarship, principally through the work of Jerome McGann, Peter Shillingsburg, and D. F. McKenzie.[27] The effect of their work has been to

undermine the establishment of a "single," "authoritative," or "definitive" text as the key objective of scholarly editing. This is, of course, in harmony with the movement in modernist critical theories against authority in all its forms. But it also happens to be common sense: the notion that we can, in textual situations of any complexity, reconstruct the "original form" of the text, or what Chaucer actually wrote or intended to write, is obviously absurd and always was. However, what is much less clear is what should replace the "single-text plus apparatus" edition. These and other authors stress the integrity of all versions of the text, and stress too that text is much more than the words themselves: the history and material circumstances of its creation, presentation, transmission, and reception are all factors in how texts communicate. At the same time, the rise of the electronic media in the 1990s, with its promise of near-infinite capacity for image, sound, all kinds of multimedia, and dramatic interfaces, seemed to offer tools that could enact this vision of what texts are.

But it is easier to describe all this than to make editions that actually do it. The danger of editions that seek to join many different texts with images, commentaries, and background materials is that they may become accumulations rather than editions: arrays of information, presented in the mass. Faced with such overwhelming quantities of data, where is a reader to start? As one scholar observed, when it was proposed for Coleridge's "Eolian Harp" that we should read at least sixteen versions, "we had better make sure we have plenty of time on our hands."[28] More likely, readers faced with such overwhelming agglomerations of information will find better uses of their time. We were rather aware that the first of our electronic publications (*The Wife of Bath's Prologue on CD-ROM*) presented huge quantities of information – for example, some 400,000 words in the

spelling databases, all categorized by headword, spelling, and grammatical function – but gave hardly any guidance on what all this might mean. If we want to reach beyond a very few readers (or rather, "users") this is not the way to do it. In the four years before our next major publication, *The General Prologue on CD-ROM*, we came to see that if we wanted to fashion our editions for more than the few, we should aim not to help editors edit, but to help readers read.[29] This meant creating editions that seek not just to present, but also to explain; which offer fruitful ways into the text; which draw the reader in to search further, to find his or her own pathways to understanding. We tried, rather modestly, to move towards this in the *General Prologue* CD-ROM, by providing an "Analysis Workshop" and a "Stemmatic Commentary."[30] However, the limitations of the interface we then had to use made it difficult for us to advance further.

Since the *General Prologue*, we have been able to use an electronic publishing program, the Anastasia system, which allows us to exploit the potential of the medium far more than we could with the DynaText software used for our first publications.[31] We first used this to present the Hengwrt manuscript in a manner that might do justice to the eccentric beauty of the manuscript itself, remarkably captured in Leith Haarhof's digital photographs, and to Stubbs's careful and innovative commentary on the manuscript and its making (in her edition of *The Hengwrt Chaucer Digital Facsimile*). Through this, we hoped to go some way towards giving an impression of Hengwrt as a physical object, stains, rat chewings, and all; and give too a sense of the many discontinuities in the inscription of the text in the manuscript, with the light this casts on the first states of the text.[32] A greater challenge yet lies ahead of us: to create a way into the fifty-five manuscripts and early print editions of the *Miller's Tale* (and other "single-tale"

publications to follow) that invites the reader into the many texts, finding understanding and not confusion. One way we will do this is by use of mouse-over "pop-ups" on every word in every manuscript, giving instantly a compact view of the variants at that word. Another way is the "variant map" feature we have developed. This shows the table of manuscript relations we have hypothesized for the tradition, using phylogenetic methods. For each variant, we superimpose all the manuscripts colored differently according to the variants they have. Thus, at a glance, the reader can see how the readings at any point are distributed across the manuscript families: which reading appears characteristic of which group; which seem particularly unstable; which most likely to have stood at the head of the tradition. In this publication we are also removing the concept of a "base text": readers can compare the readings in any number of manuscripts direct with each other, without (if they choose) any mediating edited text. In fact, in later publications we are considering including our own edited text. This is intended not as any kind of "definitive" text, nor as a "reconstruction" of some lost original, but as a useful way into the whole tradition: a hypothetical constructed text which might serve to explain all the texts.

I asked earlier: who should be our readers? My answer is: every reader of Chaucer should find something useful in the editions we seek to make. It is customary for readers to approach Chaucer through consideration of theme, character, genre, narrative method, historical background. The common practice of the classroom, as reflected in "reading" editions, and even editions for advanced students, has presumed that there is no place at this table for exploration of the manuscript culture behind the text, or for explication of the uncertainties in the text itself. There should be. It is unreasonable to expect that every reader should become an editor. But well-made editions, which offer readers the

opportunity to check efficiently the stability of the text at critical points, which offer an agreeable means of discovery of how the text came to be how it is, and which invite rather than baffle, will help us all to be better readers.

De Montfort University

Leicester, England

(peter.robinson@dmu.ac.uk)

I cannot here acknowledge all those who have contributed to this project over the years: the title page of *The Hengwrt Chaucer Digital Facsimile* alone lists nineteen people. But I owe a particular debt to the members of the Project's steering committee over the years: Norman Blake (the chairman), Anne Hudson, Derek Pearsall, Oliver Pickering, Ceridwen Lloyd-Morgan, and Toshiyuki Takamiya. What has been well done, was done with their counsel; the missteps are mine alone.

1. The exact number depends on how one counts manuscripts now physically divided. For example, one part of Manly and Rickert's Ox is Manchester, John Rylands Library English MS 63, the other part Philadelphia, Rosenbach Museum and Library 1084/2, and we follow the *Riverside Chaucer* in treating these as two separate manuscripts, Ox^1 and Ox^2. The details of our count of 84 manuscripts and 4 incunables are given in various project publications, including *The Hengwrt Chaucer Digital Facsimile*, ed. Estelle Stubbs (Leicester, Eng., 2000).

2. Before Tyrwhitt, all editors after Wynkyn de Worde simply reproduced the text of a preceding edition. De Worde himself appears to have followed Caxton's second edition, but he used a manuscript directly for the prose and perhaps some other sections

(T. J. Garbàty, "Wynkyn de Worde's 'Sir Topas' and Other Tales," *Studies in Bibliography* 31 (1978): 57-67. For Caxton, see notes 26 and 32 below. For a survey of editions of the *Tales* before 1980, see Paul G. Ruggiers, *Editing Chaucer: The Great Tradition* (Norman, Okla., 1984).

3. John M. Manly and Edith Rickert, eds., *The Text of the Canterbury Tales Studied on the Basis of All Known Manuscripts*, 8 vols. (Chicago, 1940). For a full account of their work, see R. Vance Ramsey, *The Manly-Rickert Text of the Canterbury Tales* (Lewiston, N.Y., 1994). There is an astringent critique of their work by George Kane in his article "John M. Manly and Edith Rickert," in Ruggiers, *Editing Chaucer*, 207-29.

4. For example: J. Zupitza and F. J. Furnivall, *Specimens of All the Accessible Unprinted Manuscripts of the Canterbury Tales*, Chaucer Society 1, 81, 85-86, 90-94, 97 (London, 1892-1902); J. Koch, *A Detailed Comparison of the Eight Manuscripts of Chaucer's Canterbury Tales As Completely Printed in the Publications of the Chaucer Society*, Chaucer Society 2, 47 (London, 1913 for 1907).

5. This one manuscript is Oxford, Trinity College MS 29, identified by Kate Harris as containing extracts from *Melibee* and the *Parson's Tale* ("John Gower's *Confessio Amantis*: The Virtues of Bad Texts," in Derek Pearsall, ed., *Manuscripts and Readers in Fifteenth-Century England* [Bury St. Edmonds, 1983], 27-40, at 33.

6. The project moved its base from the University of Sheffield in 1998 (upon the retirement of Professor Blake) to De Montfort University, Leicester, Eng.

7. Our collaborators in other universities are: Paul Thomas, Brigham Young University, who is preparing editions based on a transcription of the whole of Fragment

VII; Daniel W. Mosser, Virginia Tech, who is responsible for the witness descriptions; Gabriele Müller-Oberhäuser, University of Münster, preparing an edition of the *Pardoner's Tale*; Toshiyuki Takamiya, Keio University, with whom we are working on an edition of the incunabula; David Hoover and Martha Rust, New York University, preparing an edition of the *Clerk's Tale*. With the aid of these scholars and their associates, another 25% of the text will have been transcribed and partially prepared for publication in May 2004. This puts us in reach of completing the whole project in a further five years.

8. Publications resulting from this collaboration include: A. Barbrook, N. F. Blake, C. J. Howe, and P. M. Robinson, "The Phylogeny of *The Canterbury Tales*," *Nature* 394 (1998): 839; C. J. Howe, A. C. Barbrook, M. Spencer, P. Robinson, B. Bordalejo, and L. R. Mooney, "Manuscript Evolution," *Trends in Genetics* 17 (2001): 147-52, rpt. *Endeavour* 25 (2001): 121-26; M. Spencer, B. Bordalejo, L.-S. Wang, A. C. Barbrook, L. R. Mooney, P. M. Robinson, T. Warnow, and C. J. Howe, "Analyzing the Order of Items in Manuscripts of *The Canterbury Tales*," *Computers and the Humanities* (forthcoming).

9. Published are: P. M. Robinson, ed., *The Wife of Bath's Prologue on CD-ROM* (Cambridge, Eng., 1996); Elizabeth Solopova, ed., *The General Prologue on CD-ROM* (Cambridge, Eng., 2000). Scheduled for publication in the series of "single-tale" publications before May 2004 are the *MilT*, *NPT*, *FranT*, and *MerT*.

10. These are: Claire Thomson, London, British Library Lansdowne MS 851; Simon Horobin, London, British Library Additional MS 35286; Orietta Da Rold, Cambridge, University Library MS Dd.4.24. Pip Willcox is presently working on Oxford,

Christ Church MS 152, and Estelle Stubbs on a study of all the early manuscripts (including also Hengwrt, Ellesmere, Corpus Christi College MS 198, and British Library MS Harley 7334). Alongside these, Barbara Bordalejo's De Montfort doctoral thesis studied the manuscript source of Caxton's second edition (supervised by myself); see note 26 below.

11. The **O** witnesses were so labelled by myself in "A Stemmatic Analysis of the Fifteenth-Century Witnesses to the Wife of Bath's Prologue," in P. M. Robinson and N. F. Blake, eds., *The Canterbury Tales Project: Occasional Papers*, Volume 2 (Oxford, 1997), 69-132.

12. Barbara Bordalejo, "The Phylogeny of the Tale-Order in the *Canterbury Tales*," Ph.D. diss., New York University, 2003.

13. A. I. Doyle and M. B. Parkes, "The Production of Copies of the *Canterbury Tales* and the *Confessio Amantis* in the Early Fifteenth Century," in M. B. Parkes and A. G. Watson, eds., *Medieval Scribes, Manuscripts and Libraries: Essays Presented to N. R. Ker* (London, 1978), 163-210.

14. Another area of research opened up by the project is the spellings of the manuscripts, principally through the massive spelling databases we prepare for the "single-tale" publications. These permit fine-grained analyses of the many different patterns of spelling, and possible categorization by scribe, date, and region. Doctoral dissertations using these are being carried forward currently at Leiden by Louisa Caon and at Leicester by Jacob Thaisen. However, we are currently reviewing the making of the spelling databases: see note 30 below.

15. Even on this point, one can find conflicting signals in the manuscripts: Hengwrt appears to have a change of ink at folio 42, in the Prologue to the Miller's Tale (Link 1 in our numeration), coinciding with a break in the quires of eights up to this point. Stubbs takes this as a sign of discontinuity in the set of exemplars at this point (*Hengwrt Chaucer Digital Facsimile*: see her 'Observations,' on section I).

16. This view of the nature of the originals – colloquially, the "messy desk" theory – appears to have been first proposed by J. S. P. Tatlock, most fully in his article "*The Canterbury Tales* in 1400," *PMLA* 50 (1935): 100-39.

17. Stubbs, ed., *Hengwrt Chaucer Digital Facsimile*. This is the "research edition"; a version of the facsimile intended for general readers was published in 2003: Ceridwen Lloyd-Morgan, ed., *The Hengwrt Chaucer Standard Edition on CD-ROM* (Leicester, Eng., 2003).

18. Thus: *Miller*, 666 lines; *Summoner* with link, 630; *Nun's Priest* with link, 660; *Squire*, 664; *Pardoner*, 640.

19. This appears to be the opinion of Jill Mann, "Chaucer's Meter and the Myth of the Ellesmere Editor," *Studies in the Age of Chaucer* 23 (2001): 71-107, especially 107.

20. Tatlock, "*The Canterbury Tales* in 1400," 100-39.

21. Darin Merrill is currently preparing a doctoral dissertation on the differences between Ellesmere and Hengwrt at Arizona State University (supervisor, Robert Bjork).

22. Koch, *Detailed Comparison*, 410.

23. Elizabeth Solopova, “Chaucer’s Metre and Scribal Editing in the Early Manuscripts of the *Canterbury Tales*,” in Robinson and Blake, eds., *Canterbury Tales Project*, 143-164.

24. This is also the opinion of Mann, “Chaucer’s Meter.” In my “Stemmatic Commentary” for the *General Prologue CD-ROM*, I suggested that of the approximately 90 instances in which Hengwrt and Ellesmere differ substantively in the 858 lines of this text, on approximately 45 occasions it appears that Hengwrt has the likely original reading; on approximately 35 it appears to be Ellesmere; while for some 10 instances I could not judge either way.

25. Compare N. F. Blake’s well-known (though often misinterpreted) position that while text contained in Hengwrt may be accepted a priori as by Chaucer, the case for text outside Hengwrt being by Chaucer must be made on its merits. See his own account in “The Ellesmere Text in the Light of the Hengwrt Manuscript,” in *The Ellesmere Chaucer: Essays in Interpretation*, ed. Martin Stevens and Daniel Woodward (San Marino, Calif., 1995), 205-24, and my discussion in “Can We Trust Hengwrt?,” in *Chaucer in Perspective: Middle English Essays in Honour of Norman Blake*, ed. G. A. Lester (Sheffield, Eng., 1999), 194-218.

26. Barbara Bordalejo, “The Manuscript Source of Caxton’s Second Edition of the *Canterbury Tales* and Its Place in the Textual Tradition of the *Tales*,” Ph.D. diss., De Montfort University, 2003.

27. For example: Jerome McGann, *A Critique of Modern Textual Criticism* (Chicago, 1983); Peter Shillingsburg, *Scholarly Editing in the Computer Age: Theory and Practice* (Athens, Ga., 1986); D. F. McKenzie, *Bibliography and the Sociology of Texts*

(Cambridge, Eng., 1999; 2nd edn. of the Panizzi Lectures [London, 1985]). One should also mention David Parker, *The Living Text of the Gospels* (Cambridge, Eng., 1997).

28. Peter Barry, "Coleridge the Revisionary: Surrogacy and Structure in the Conversation Poems," *Review of English Studies*, n.s. 51 (2000) 600-16, at 603.

29. So formulated in the "General Editors' Introduction" to the *General Prologue CD-ROM*, ed. Solopova. Elsewhere, we have called this approach the "new stemmatics": thus in the "Analysis Workshop" (*General Prologue CD-ROM*).

30. According to this line of thinking, that we should allow our perception of what might be most useful to our readers to determine how and what we present, we are considering dropping the spelling databases from CD-ROM publications from the *FranT* on. Although these are much admired, we are all too conscious of the inconsistencies within them, and they took enormous amounts of time to prepare, check, correct, recheck and correct.

31. On the Anastasia system, see http://www.sd-editions.com/anastasia. A view of the interfaces for many-text editions which can be made with Anastasia may be seen functioning with biblical texts at http://nestlealand.uni-muenster.de.

32. Further publications in this "single-text" series of publications scheduled before May 2004 are *Caxton's Canterbury Tales: The British Library Copies* edited by Barbara Bordalejo, and a digital facsimile of Cambridge UL Dd.4.24 edited by Orietta Da Rold. We are also preparing a "multi-text" publication, amalgamating full transcripts of the whole text of the *Tales* in some twelve witnesses. Publication of this before May 2004 is contingent on the agreement of the University of Sheffield.