From Medieval Manuscripts to Electronic Text: A Transcriber's Tale

Michael Pidd and Estelle Stubbs

On the first day we began transcribing manuscripts into electronic form for The *Canterbury Tales* Project we had never actually seen a medieval manuscript before. Some of us had never even seen a computer. In order to assist us, we were given a set of guidelines for transcription, prepared by Peter Robinson and Elizabeth Solopova (Robinson and Solopova 1993), and intended to outline some of the principles established by the authors during their own initial transcriptions of the manuscripts of The Wife of Bath's Prologue. However, we soon discovered that a manuscript is more than simply gatherings of vellum covered in writing, and the transcription of its text would require more from us than a slavish following of the principles outlined by Robinson and Solopova. We would, as any transcriber does, have to confront and comprehend the problems of de-coding one medium of writing (the manuscript) in order to re-encode it within another (the computer.) How this is achieved and what it might tell us about some of the problems which fifteenth-century scribes may have experienced will be the subject of this paper.

A manuscript page is an infinitely complex visual experience; from the sublime calligraphy and glaring blunders of the scribe's hand to the exquisite tracery or hurried application of the illuminator's paint, from stitched-up holes in the vellum to trailing globules of ink across the text. On the other hand, a transcript of a manuscript in electronic form possesses a clarity and tidiness which hides the complexity of the object of its imitation. Any attempt to re-present the manuscript in electronic form has to be done within the limits established by the computer and, inevitably, this involves stripping away some of its complexity and individuality, such as the precise appearance of a scribal hand. Transcription is, consequently, a 'fundamentally interpretative activity, composed of a series of acts of translation from one system of signs (that of the manuscript) to another (that of the computer)' and thus 'Like all acts of translation, it must be seen as fundamentally incomplete and fundamentally interpretative' (Robinson and Solopova 1993, 19.) To state it simply:'the text in the computer system will not be the same as the text of the primary source' (Robinson and Solopova 1993, 21.)

Consequently, the ideal level of transcription towards which one should aim is 'graphemic,' whereby 'every manuscript spelling is preserved (as: `she`, `sche`) without distinction of separate letter forms' (Robinson and Solopova 1993, 22.) We reject, for reasons of time, computer feasibility and accuracy, 'graphic' representation in which 'every mark in the manuscript, every space, is represented in the transcription' and 'graphetic' representation in which 'every distinct letter-type is distinguished (as: r 'short' is transcribed apart from r 'round' and r 'long descender' etc.)' (Robinson and Solopova 1993, 22.) Four

THE *CANTERBURY TALES* PROJECT
OCCASIONAL PAPERS II, 1997, 55-59

graphetic types of the letter r can be seen below (reproduced from British Library, MS Harley 7334.)

Also rejected is the 'regularized' level of transcription, in which 'all manuscript spellings are regularized to a particular norm, perhaps the spelling of a manuscript considered authoritative' (Robinson and Solopova 1993, 22.) Nevertheless, although rejected from a manuscript's representation on the computer, the graphic and graphetic levels both constitute a part of the transcriber's interpretative act, for it is only by studying a manuscript on both these levels that a transcriber is able to interpret the manuscript's complex semiotic system and then translate it, graphemically, into the computer's.

When confronted with a folio for transcription, the transcriber's first task is invariably to try to distinguish the actual script from its often worn, stained, and blotched vellum background. Interpreting the manuscript graphically is essential from the outset, for the transcriber can never hope to produce a fair representation of the script without first confronting and sifting out those features, such as the materiality of the vellum and the idiosyncracies of the scribe's handwriting, which could distract and so distort one's perception of the script. A common problem during this stage of interpretation is the quality of the manuscript images. When unable to work from the original manuscripts the transcriber often has to use photocopies of the microfilmed editions, and so an old microfilm or an old microfilm machine can often produce poor and misleading images, usually with a much lower level of detail, but frequently with detail which does not actually exist on the manuscript itself. The first time we ever saw a manuscript we were surprised to discover that it was not the filthy and illegible book which our microfilm photocopy had led us to believe, and experience has taught us that the majority of dubious scribal characters can be resolved by consulting the actual manuscript, where it is possible to see a clearer image of their form.

The next level of manuscript interpretation can be considered 'graphetic,' for once the actual script has been filtered out from the morass of physical detail, the transcriber has to identify the individual characters. With the view to representing the script graphemically, of preserving its spelling, it is essential that the graphetic nature of the individual characters be understood, because recognizing that the many different forms of `e` are all significant of `e`, although not transcribed individually, is fundamental to not mistaking one of its forms, such as the 'round' `e`, as actually being an `o`. Some forms of `o` and `e` are almost identical, especially in the secretary hands of manuscripts such as Delamere (now Tokyo, Takamiya MS 32) in which the word `continent` (WB 106 in the Project's lineation system) will also appear to be `centinont`, `continont` and `centinent`. Occasionally it is also difficult to distinguish `o` and `e` from the 'round' `s`. We can see an example of this problem below, taken from British Library, MS Harley 7334; these charatcers are, from left to right, `o` `e` and `s`, but are only distinguishable within the context of their use.

Furthermore, the terrain of a manuscript text always appears to be changing, and even if the entire extant manuscripts of *The Canterbury Tales* had been traversed by the transcriber, the significance of graphetic features would still often appear uncertain and be seen in a different light with every subsequent journey; the transcriber's opinion of past terrain would always be changing, revising itself. For example, an initial transcription of British Library, MS Harley 7334 appears to stumble constantly upon tails attached to the final f, whereas on a second reading the tails can appear to be absent, merely an illusion which had been suggested by the manner in which the scribe draws the bar across his f. As to which of the following examples should be transcribed with or without a tail is, for the most part, purely at the transcriber's discretion.

All of this suggests that transcription is an interpretative act which takes place primarily on the scriptive level, with very little consciousness of the text's narrative meaning. One becomes totally absorbed by the nature of the script and only occasionally, in short flashes, does one become aware of the story which it is intended to communicate. Having made complete transcriptions of nine manuscript witnesses of *The Canterbury Tales*, we still cannot say for certain what all the *Tales* are about. This obliviousness to the narrative itself makes eye skip a common problem, especially when transcribing a work of prose such as The Parson's Tale where the formulaic nature of his sermon can easily deceive a transcriber who may have momentarily lost his or her place. Yet the intensely interpretative nature of transcription endows the visual appearance of the script with its own meaning as we gradually begin to perceive the siginificance of a scribal characteristic within the wider context of its appearance throughout the whole manuscript. For example, when we first transcribed The Parson's Tale in Hengwrt we were aware of the occasional occurrence of an unfamiliar manuscript feature within the body of the text: :

At first we believed it to be a punctuation mark, perhaps an incomplete semi-colon. This seemed a credible idea because we expected that the potentially more complex syntax of the prose tales would have required an equally more sophisticated punctuation than is present in the verse, which is largely unpunctuated. However, upon revising our transcriptions of the Parson's Tale we discovered a blindspot, for we had consistently failed to record a similar manuscript feature which appears down the right-hand margin on the ends of lines, abundantly in Ellesmere, and occasionally in Hengwrt:

Its use at the end of prose lines is probably as a filler, employed to justify the text by filling in any spaces which are left by words which fall short of the right-hand margin, just as the hyphen is used for words which overrun the margin. We therefore began to conclude that the medial occurrence of this feature was perhaps not a punctuation mark at all, but rather it too was a filler, possibly copied accidentally by the scribe from an exemplar whose lines of prose were different in length from his own. In Hengwrt the medial filler is common, whereas its occurrence at the end of a line is rare, perhaps because the scribe was unaware of its purpose and because, in Hengwrt, presentational features such as justified margins do not seem to have been an important concern. However, in Ellesmere, where presentation was of the essence and the scribe more experienced, the filler is common but its medial occurence is rare, suggesting that the scribe knew how it was supposed to be employed. Alternatively, it is possible that the medial filler has a completely different significance from the one which regularly occurs down the right-hand margins of the prose, suggesting that it might in fact indicate 'where the sense of one clause is dependent upon what has gone before, as, for example, in correlative or adversative relationships' (Burnley 1983, 89.)

Another feature of the transcriber's interpretative act is that the scriptive nature of the manuscript attains a familiarity and remembrance: retracing our transcriptions of The Parson's Tale in Hengwrt and Ellesmere, the nature of the hand was familiar, whilst ambiguous flourishes and superscript hooks sparked off vivid recollections of our previous encounter with them. However the sermon itself has taught us nothing. This suggests that transcription is a conscious attention to that pleasure we all gain from the visual nature of the text during an act of reading, whilst also suggesting that some of the presumed manuscript errors might possibly have been because the scribe was, likewise, only interpreting his exemplar on the scriptive level and did not consider whether the words which he thought he saw actually made any sense to the story itself. For example, in the Knight's Tale there is the misreading of the word sterres (KN 1179 in the Project's lineation system.) Many manuscripts have the reading sertres, or some other variant, and Blake (1985, 169-170) attributes this to a misplaced abbreviation for er after the letter s in the original copytext (thus: s tres), suggesting that many scribes automatically expanded the abbreviation without considering whether or not it actually made any sense in this position. Its correct place is after the letter t (thus: st es)

However, the similarity between a twentieth-century transcriber and a fifteenth-century scribe is greater than the mere failure to read a manuscript text on anything other than its visual level. It soon becomes clear that the difficulties of interpretation and the frequent errors of judgement are essentially the same, for possibly the same reasons. Although the tools of the trade have changed—a VDU instead of vellum, a keyboard instead of the quill, and a mouse to guide one's hand, instead of a knife—nevertheless a fifteenth-century scribe would probably still recognize his modern counterpart as being essentially the same person with the same afflictions. Both the scribe and the transcriber probably suffer from a concentration time of approximately 3.5 hours, eye strain due to the close scrutiny of texts, the glare of computer monitors or the flickering of candle light in the evenings; backache, finger cramps and incompetent tools, whether it be splitting quills or stupid computers; as well as the desire to maintain one's average quota of lines, and external pressures in the form of the scriptorium master or Project Director. Added to the unfamiliarity of the exemplar's language, all of these can contribute to inaccuracy and consequently determine the interpretation, certainly for the transcriber, and possibly for the scribe. Even resorting to the opinion of someone with greater experience, or subjecting the manuscripts to a constant revision, does not guarantee that all the inaccuracies will be corrected, or dubious features understood. In repeating the scribal act, complete with the afflictions and inaccuracies, we are not only substantiating but also adding to the irrecoverability of *The Canterbury Tales* as a single, authorial text.

Robinson and Solopova finally have to confess that it is all fundamental to the interpretative nature of transcription that 'we have missed some tails and virgules which are there in the manuscripts, added some which are not in the manuscripts, and misplaced others' (1993, 46.) Although the intention of The *Canterbury Tales* Project is to transcribe and collate all extant manuscripts in order that its scribal dissemination and evolution be studied, in so doing the transcriber is actually repeating the scribal act, adding to and thus altering the extantness of *The Canterbury Tales*. Perhaps in another four centuries, when electronic publishing in CD-ROM form has become an antiquated curiosity, stored under lock and key within the archives of the Bodleian Library, or sold at auctions to eccentric and titled collectors of electronic texts, occasionally allowing academics the joy of seeing and handling them, then perhaps The *Canterbury Tales* Project will be viewed as more of a part of the general textual tradition than as an undertaking to try and comprehend the tradition. Indeed, perhaps the only clear way of evaluating the interpretative nature and accuracy of our own transcription and its lonely pilgrim transcriber will be by considering it to be an act which takes place within the overall textual tradition of *The Canterbury Tales*.

Bibliography

Blake, N. F. 1985. *The Textual Tradition of the Canterbury Tales*. London and Baltimore, Edward Arnold.

Burnley, D. 1983. *A Guide to Chaucer's Language*. London, Macmillan.

Robinson, P. M. W. and E. Solopova. 1993. 'Guidelines for Transcription of the Manuscripts of the Wife of Bath's Prologue.' In *The Canterbury Tales Project Occasional Papers* 1, ed. N. F. Blake and P. M. W. Robinson. Oxford, Office for Humanities Communication: 19-52.

For Figures 1 and 2, see Plate I in the front pages.