TR9 M8: Minutes of the Work Group on Manuscripts held at the Wittgenstein Archives, University of Bergen, on 18-20 September 1992

Lou Burnard

23 Sept 92

The following were present throughout the three days of this meeting, hosted by the University of Bergen's Wittgenstein Archive.

LB: Lou Burnard, Oxford, TEI Editor
DB: Dino Buzzetti, Bologna
CH: Claus Huitfeldt, Bergen, Wittgenstein Archives
MK: Mary Keeler, Washington & Indiana, Peirce Telecommunity Project
CK: Christian Kloesel, Indiana, Peirce Telecommunity Project
OL: Ole Letnes, Bergen, Wittgenstein Archives
AP: Alois Pichler, Bergen, Wittgenstein Archives
AR: Allen Renear, Brown, Peirce Telecommunity Project, TEI Advisory Board
PR: Peter Robinson, Oxford, Chair of TEI TR2
DS: Donald Spaeth, Glasgow, CTICH, Member TEI AI4
MSM: C.M. Sperberg-McQueen, Chicago, TEI Editor

The following documents were distributed before and at the meeting:

TEI TR2 D5: Draft for P2 chapter 47
TR9 T5: Sample pages from the Peirce mss
TR9 M2: Minutes from Louvain (in English and French)
TR9 W4: Diplomatic Transcription of Modern Manuscripts (submitted to Myrdal meeting in Nov 1991)
EXWA W1: Excerpt from MECS: Multi-Element Code system (Draft, forthcoming in Working Papers from the Wittgenstein Archives... no 3, 1992)
TR9 T6: Collection of materials from PR
TR9 T7: Collection of materials from DS
EXDB W1: Masters and books in 14th c. Bologna: an edition as database (two conference abstracts in English and Italian)
TR9 P1: Work group on Manuscripts and Codicology: objectives and deadlines

Chapter 1

Introductory

Welcoming the group, CH noted that the previous chair, Jacqueline Hamesse, had had to withdraw due to pressure of other work. He briefly outlined ancillary social events and summarized the previously distributed agenda before calling on MSM to introduce the proposed work plan for the group.

Responding, MSM explained that the charge to this Workgroup was simply to ensure that the TEI Guidelines address those encoding problems affecting people who work with manuscripts. The group might propose new elements or modifications to existing elements for the purpose of encoding information it felt to be necessary for manuscript encoders. It should aim to produce a report on the problems unique to the encoding of mss, which, even if it proposed no new elements, would demonstrate the applicability of the existing TEI scheme. If the report were ready before the end of the year it would be possible to include it in P2. The existing chapter on text criticism (TR2 D5) would be in P2, though not in its present form, and so comments on it from this group would be particularly effective if made now.

MSM then briefly summarized the existing TEI committee structure and status of the current draft Guidelines (TEI P2). This workgroup is one of twelve reporting to the Text Representation Committee, covering such topics as character sets, textual criticism, hypertext, graphics, manuscripts, poetry, drama, prose. On P2,[1] he noted that “base tagsets” are mutually exclusive, whereas “additional tagsets”are not. Because of their widespread applicability therefore, several of the elements discussed in the TR2 D5 would be moved to the “core” tagset. On the timetable for completion of P2, he reported that, despite unanticipated delays, it was currently hoped to publish most of parts 1 to 4 by the end of 1992, with completion of the whole by late spring or early summer of 1993. At that time, any comments or updates would be integrated to form the final draft, P3, which would be submitted to the TEI Advisory Board for endorsement. The current grant would last at least till then, and would be followed by a transitional period during which further funding would be sought and existing projects (such as the present workgroup) would be continued.

CH proposed and it was agreed that the two objectives for the meeting should therefore be to assess the draft from the Text Criticism Group (TR2 D5), and to produce an outline for a new chapter on the transcription of manuscripts.

Chapter 2

Transcription issues: TR9 T6 and TR2 D5

Following coffee, and a brief review of SGML jargon, the group turned to the issue of manuscript transcription. PR explained that his workgroup, TR2, had found it necessary to address transcription although this was not strictly within its remit. P1 had been criticized for failing to address the issue in the context of critical editions, although there were relevant proposals in the section on core tags. PR then presented a short series of examples (TR9 T6) to demonstrate just how complex and difficult the process of mss transcription is.

In example 1, taken from the Ormulum ms (one of the rare authorial autographs from ME period), the appended commentary demonstrates how much can be read from the original, much of which would be impossible to transcribe systematically. Parkes is able to find evidence for the dating of the ms in such details as the shape of the strokes making up individual letters; this however implies the need to attach annotation rather than necessarily to transcribe at such a level of detail. For PR, transcription is an interpretive act, contingent on the transcribers' primary purposes. The particular problem when transcribing for computer, is that transcripts may be used for unanticipated ends, which may be needlessly difficult to attain if transcription principles are inconsistent.

---

[1] For convenience an existing document (TEI EDJ6) summarizing its proposed structure, contents and publication method is attached (LB).

PR then presented a number of examples of the diversity of methods adopted by transcribers; Bowers' notation was very obscure;[2] the dots in the Hebrew scheme indicate differing degrees of confidence in the reading of the letter below; Godden's simple bracketting scheme lacks explicitness; the Langland edition relies on a partition of information between notes and sigla in the text (for example, note 2 is attached to every case where barred H is possibly a final e); Gabler's edition of Ulysses captures a single work's complex and chaotic textual history in terms of “overlays” within a single document; the sheet of “signes conventionnels” initially defined for the Aquinas edition imply the existence of a collation base.

Surveying these and others it should be possible to identify a manageable set of recurrent problems, notably: discontinuities, additions, deletions, interlineations etc, both within one ms and across several; editorial corrections; uncertainty of reading. PR and his group had proposed a core set of tags which they believed would be equally usable for manuscript, typescript and printed texts (even dictionaries).

Ensuing discussion in the group focussed initially on the use of the “mirror” tags in the draft. It was agreed that the choice as to whether (for example) **<sic>** or **<corr>** should be used in a given context was a matter of editorial convenience and point of view: there was no implication that one or other of the pair must be used throughout a transcript, though there should be consistency as to the circumstances in which one or other alternative was used.

It was agreed that these and other basic transcriptional principles (e.g. what features were included, multi-layered single source vs multi-source single-layers) should be included in the header[3].

It was also agreed to defer discussion of issues such as whether or not a text-critical approach and a transcriptional approach needed a distinct set of elements or whether discussion should be organised on a period basis. It was agreed that the notion of a complete, application-independent, transcription was chimerical and that the most practical approach was to adopt a non-specific general scheme which could be scaled up or specialized for particular purposes.

Chapter 3

Comments on Chapter 47 (TR2 D5)

After lunch the group started to work through TR2 D5 as a way of focussing discussion. Several points emerged, and were noted by MSM for inclusion in the version of the draft currently being worked on for P2. Some specifics are listed below:

In the preface, the PMLA quotation should not imply that scrutiny is only word-based. In the last sentence (Similarly...) some explanation of what is meant by “multiple authorial mss” was necessary.

On section 47.1, discussion of the phrase “mirror” tags resumed. Some felt the concept was so confusing that all but the first paragraph of this section should be dropped. It was agreed that only the phrase itself was questionable. Some guidance as to usage should be given, along the lines proposed earlier, i.e. that people might choose to use only one of the paired sets (possibly modifying the DTD to enforce this), or to mix, or to use only one alternative in a given situation. The choice made should be documented in the **<encodingDesc>** element of the header. There was some feeling that the text here should call

---

[2] See however appendix A to these minutes (LB)

[3] Presumably within the **<encodingDesc>** in the header. (LB)

attention to the implications of using attribute values rather than element content; and also that specific reference should be made to the use of critical apparatus tags for more complex situations. Use of an entity reference to stand for the tagged **<abbr>** element and its content was also suggested.

The *resp* attribute on the **<corr>** and **<sic>** elements should include “authorial” as a possible value, to make clearer that these elements were applicable in non text-critical situations.

Some discussion of what was meant by regularisation was needed at the start of 47.1.3, and a simpler example should be included. The discussion of retroversion was inappropriate here, and there was some feeling that retroversion would in any case be better represented by a different tagging. This led to some discussion of the appropriateness or otherwise of the term “regularisation” for the very general notions for which this element was designed: a *type* attribute to subdivide these was proposed. CK called attention to the existence of a set of standard names and definitions for various kinds of normalizing practices prepared by the MLA Committee on Scholarly Editions.

It was agreed that the present organization of the chapter needed some attention, with various parts of it migrating to other parts of P2. For some, simple “literatim” transcription was not adequately distinguished from editorial description; for others this was not a meaningful distinction.

In section 47.1.4, the description of **<add>**, **<del>** and related elements should make explicit that letters might be added or deleted as well as words or phrases. The *place* attribute for **<add.span>** did not seem to cover the case where parts of the addition were in different locations. The *resp* attribute should include “authorial” as a possibility. It was not clear how over-writing should be encoded, nor how the insertion point of an addition (etc.) should be marked, nor what policy should be adopted with respect to the siglum (e.g. a caret) sometimes used to mark insertion points. MSM suggested use of attributes *place* and *anchored* by analogy with those on the **<note>** tag. Uncertainty about the insertion point should be marked with the **<uncertainty>** element.

MSM reported that it was currently planned to reorganize the chapter as follows:

- 47.1.1 to 47.1.4 and possibly 47.1.7 would move to the Core Tags
- 47.1.5 and 47.1.6 would move to a separate section on the physical description of manuscripts
- 47.2 would remain as the bulk of the chapter on textual criticism

It was hoped that information for the proposed chapter on physical description would be one of the major outputs from this workgroup.

Chapter 4

DB on Possible uses of transcriptions of medieval mss

DB began his presentation by underlining the need to involve the community of working mediaevalists if the scheme was to be acceptable. Hamesse had provided the workgroup with an invaluable contact point with this working community of mediaevalists. The example cited by PR of the encoding scheme used by the Dominicans working on Aquinas was a good example of a model with which this community would be familiar.

This model was appropriate however only for manuscripts, like those of Aquinas or those of Giles of Rome (on whom DB had himself worked), where works circulated in many copies. In this “peciae” system, mss were made by a simple mechanical process, all deriving from a single *archetype*, which we *know* existed because we know the university stationer had an approved exemplar for copying. When transcribing such mss, variation must be handled in a very specific way.

The mss forming the basis of DB's study of 14c teaching at Bologna (described in EXDB W1) were however of a different tradition. In Bologna, the practice of the “repetitio” in which an assistant (the *repetitor*) had to repeat the master's morning lecture in the evening, possibly with additions, meant that different mss exhibited equally important alternative readings. Because repetitors became masters and wrote their own commentaries, it could be very difficult to decide whether a ms was a copy of the master's text, with marginalia by the repetitor, or an entirely new work. For mss in the Parisian tradition, collation was a matter of identifying corruption and divergence from an authoritative text; for mss in the Bolognese tradition we were faced with something more like the problems of modern authorial mss: versions which converge into final text. Mediaevalists usually think in terms of the first tradition, because there are few authorial mss; those working with modern mss tend to think in terms of the second, but the distinction has nothing to do with modernity as such.

DB and his colleagues were using Manfred Thaller's software system Kleio to produce a database reflecting this situation, in which transcribed texts and images can be combined and displayed in parallel to help the scholar understand the (possibly unusual) situation in which the texts were produced.

For DB an encoded transcription should include as many features of the original as possible, because of the difficulty of determining in advance what will be relevant. Transcription is a way of encoding one's own experience or perception of the text.

In discussion, the importance of linking text and image was noted, if only because of the technical difficulties it involved. It was agreed that consideration of how the goals of a transcription affected its nature were equally worthy of consideration.

MSM summarized the general normative principles described in the Guidelines: elements were categorized as mandatory, recommended, optional, mandatory when applicable, and recommended when applicable. Very few elements are mandatory; when elements categorised as recommended or optional are used, the consistency with which they are used should be documented. The consensus of the group was that it was difficult to recommend any particular set of elements as mandatory for all applications. The Wittgenstein Archives for example had decided to ignore line-endings. On the other hand, if people were encoding mss at all, clearly some consideration of their physical features was essential. Encoding some textual features, however relevant, might be too expensive. The importance of potential re-usability of encodings was reiterated: where encoding a feature cost little it should be encouraged. A well-organised encoding would also be extensible. MK suggested that the production of a transcription should be seen as an archival rather than an editorial process.

There was some discussion of various methods for encoding changes of hand or rendering within a ms. Hand changes were often combined with the presence of other features such as deletion, addition, underscore etc. Possible mechanisms included the global *rend* attribute, milestone tags, concurrent structures, and linked feature structures.

Chapter 5

DS on early modern mss

As an historian, DS was not interested in close examination of a single source, but rather the ability to scan large quantities of documentation for small details. Many historians transcribe mss, but usually with extensive normalization and incompletely. Typical cases would be official documents of various kinds, rather than texts with identifiable authorship. In the same way, identification of generic hands (business, court etc.) was of more importance than identification of individual scribes.

For the historian, he suggested, mss would typically be characterised along such axes as the following:

- source e.g. official or private
- authorship e.g. single, multiple, none
- hand e.g. for the medieval period, one of a small set of generic hands; for later periods, much more individualism
- abbreviations: in the medieval period although there are many, they are mostly standard; in the modern period there are fewer and they are more idiosyncratic
- language, e.g. Latin, vernacular
- scribe e.g. either unknown or known individual

DS then presented TR9 T7, a collection of sample pages from legal and ecclesiastical records of many kinds, drawing attention to the continuity of the issues which needed to be addressed across both mediaeval and early modern material. Brief notes follow:

- (1) The same court hand used here is found in English legal documents up to as recently as 1733, when the use of Latin abbreviations was formally outlawed by the British courts. It exhibits a well-developed system of abbreviation, not all of which is uncontentiously captured in the sample transcription
- (2) Two 16th c. indictments in secretary hand, bearing annotations in a different (court) hand which express the outcome of the indictment, and are thus clearly subsequent additions.
- (3) Two pages of a 17th c. ecclesiastical court record. Abbreviations have become so conventionalized that it's doubtful whether the clerk would have known what their expansion was. The second page has a drawing for the place where the seal should be attached; the hand used combines features of court and secretary hands indiscriminately (e.g. in the shape of the letter h)
- (5) This late 17th c exemplar uses a more modern hand, in which only a few abbreviations appear, and are not always used consistently (for example ‘parishe' vs. ‘perish'). The mixture of different letter shapes for S and the use of underlining are also notable. Such features are mostly uninteresting except as a means of dating a document.
- (6/7) Similar documents, using copperplate hand
- (8) A letter written to a Bishop in a neat 18th c hand in which only one abbreviation ('d) appears.
- (9) Another, in which a bar is used to abbreviate the word 'ecclesiastical'.
- (9) Another letter: is it important to encode that the end of the letter has been written in the left margin rotated 90 degrees?
- (10) A 'vera copia' made for Sadleir of a letter written by Woodward.
- (15) Note use of two hands: formulaic Latin text in secretary and the English in a different one, even using a different pen. A note is added amongst the signatures: several signatures use ambiguous glyphs as “marks”.
- (18) This has annotation in a different hand, crossed-out underlining and more conventional glyphs for marks.

This lead to a discussion on the feasibility or otherwise of identifying a standard set of abbreviations, such as mediaeval Latin ‘per'. DS felt this to be a gray area: standardized and systematic lists existed but, MSM pointed out, these had not been prepared for the use of mediaeval scribes but for the use of modern scholars. CH felt it might still be helpful to define an entity set for those abbreviations commonly recognised. Unrecognised abbreviations would not of course be encodable in the same way, nor would marks. Guidance as to whether names should be based on the appearance of an abbreviation or its (presumed) significance might also be helpful.

Chapter 6

CH on problems of modern mss

The objectives of the Wittgenstein Archives were to provide accurate versions of the manuscripts in a variety of formats, and to support both text retrieval and analysis. Features such as insertion, deletion, overwriting, substitution etc. characterised all modern and (as we had just heard) early modern mss; they were not peculiar to Wittgenstein. CH referred to the recent Hölderlin edition in which the appearance and layout of a ms is closely mimicked in print: For that, CH and his colleagues felt a facsimile did a better job. Even so, the very same layout features did need to be encoded because they conveyed structural information of importance.

CH cited as model *mehrfach besetzte Funktionspositionen* or “multiply saturated functional positions” quoting from Herbert Kraft the phrase “In den räumlichen Relationen von chronologisch differenten, aber strukturell äqivalenten Texteinheiten ist der Entstehungsprozess als textliche Struktur fixiert.” (“in spatial relations of chronologically different but structurally equivalent textual units, the creative process of becoming is fixed as a textual structure.”) DB remarked that the spatial layout of a text (necessarily itself ordered in time) is a topological projection of the temporal sequence of its content. PR observed that Shillingsburg defined a text as “a state of becoming”.

The practice of the Wittgenstein Archives in case of uncertainty was always to choose the reading which made best grammatical sense. The basic policy is to transcribe every letter left to right, top to bottom, leaf to leaf, making exceptions only where needed to preserve grammaticality and in the following predefined order

- rearrangement (moving)
- simple substitution (of “alternatives” or “variants”)
- reiterative substitution (e.g. where a series of substitutions mutually presuppose or exclude specific choices of 'variants')
- extension (adding text)
- suspension (i.e. marking “ungrammatical” text)

The process of recording these features goes on at the same time as the transcription itself. It is not therefore a strictly diplomatic transcript[4], but one which facilitated its production. The policy of the Wittgenstein Archive was to mark all deviations from the literatim text in such a way that they might be

[4] This term, described by CK as a term in which authorial instructions are obeyed but not recorded, remained controversial throughout the meeting, both PR and DS pointing out that the retention of authorial or scribal instructions was often of immense importance in older or historical materials. In their fields, a diplomatic transcription recorded deletions etc. rather than silently obeying them.

accepted or rejected in producing reading versions.

Chapter 7

CK and MK on the notion of electronic facsimile

Charles Sanders Peirce (1839-1914), regarded as the last great American polymath, was the father of pragmatism and semiotics, a scientist, mathematician, and egyptologist. He published about 10,000 pages during his lifetime, but left ten times as many unpublished notes, now mostly held by the Houghton Library at Harvard. The subject matter is very diverse, with articles on linguistics including detailed discussions of chemistry. They are also characterised by a close integration of text and images. Many of the papers present variant texts with widely different degrees of detail and expansion. The papers are partially catalogued, but perhaps 20,000 remain loose and incorrectly collated.

CK and others at the University of Indiana have been working on a conventional critical edition, due for completion in 30 volumes in the year 2030. They use computers but only as an aid in the preparation of the published papers. MK, Joseph Ransdale and others are working on an electronic edition of the writings in parallel. Now known as the "Peirce Telecommunity Project", this aims to make the whole of Peirce available online, using images and hypertextual tools to realise the electronic community of coworkers envisaged by Peirce.

The graphic images are particularly important because Peirce's semiotic has many modes of reference. Representation-schemes are his subject matter, and his notes exhibit endless experiments with the use of graphics and graphic formalisms to convey ideas very difficult to express in text. Text is contoured around graphics, graphics are used within text. Freeform graphic symbols become graphemic.

Some but not all of the papers were microfilmed in the sixties, rather inadequately, and in black and white. Colour is also important for an understanding of the material, and archival capture on digital image is now under consideration. The new catalogue records some 36 features including physical features such as the kind of paper, ink etc. Like other mss, the Peirce papers include annotations by others, notably earlier editors and instructions by Peirce to his typist.

Summarizing, CH noted the following issues for further consideration:

- to what extent are all these features codable?
- to what extent does the researcher want or need them?
- if they are transcribed, how, and how do you find them?
- how should mixed text and graphics be encoded?
- what are the appropriate formats for capture, storage, and display?
- will automatic image recognition techniques assist in data capture?

Chapter 8

Identification of topics for further discussion

A list of several specific items for discussion was drawn up in plenary session. After initially attempting to categorize each of these as of ‘local', ‘global' or ‘other', it was decided to divide the topics into two groups: those concerned primarily with matters relating to physical description, and those relating to editorial matters, as listed below. The meeting divided into two corresponding subgroups, which met separately to discuss each set of issues listed below.

The following editorial issues were discussed in a group composed of MSM, PR, CH, AR, and AP.

- abbreviations: is a suggested entity set for these desirable or feasible?
- what recommendations could be made about transcription practices in general, with reference to interpretive vs literatim transcription?
- was an apparatus criticus or some other structure appropriate for the encoding of divergent readings or interpretations
- how should authorial instructions be encoded
- how should conjectural readings be encoded
- relationship of transcription to general issues of historical interpretation

The following topics relating to physical description were discussed in a group composed of LB, DB, MK, CK, OL and DS.

- What features of manuscript hands are of interest and how should they be encoded?
- To what extent and how should physical layout be encoded?
- What other physical characteristics of a manuscript source are encodable and how?
- To what extent are images encodable? How about images intermixed with text? What storage media or formats might be recommended? What about automatic recognition or encoding?
- How should underlining and marginalia be encoded?
- How should illegibility of the source be encoded?

The subgroups met over the remainder of the meeting, reporting back in two final plenary sessions as described in the next two sections.

Chapter 9

Report of the Physical Description Subgroup

As regards physical information, the group proposed the following features for encoding:

- type of carrier (paper, parchment, wax tablet...)
- other properties of the carrier such as watermarks
- instrument used for rendering (pen, quill, pencil, brush...)
- medium used for rendering (ink type and colour, lead, ...)
- dimensions of the carrier (sheet size etc.)
- information about the *assemblage* (e.g. binding, roll, fascicle, folio, sheets, leaves, notebook) possibly including here pagination or foliation

- ancillary information about the carrier, e.g. posession notes for a codex, lists of incipits or explicits

It was noted that these features would need to be specified at both high (applying to whole assemblage) and low (applying to individual parts of a ms) levels, with some mechanism for inheritance of unspecified features.

On layout, a need was identified to define the layout of a page independently of its content and then link the two, perhaps like the alignment system proposed for spoken texts in TEI P2 chapter 34. The neutral term *zone* was proposed to identify parts of physical layout. The following attributes were identified for such zones:

- shape (rectangular, circular, square ...)
- dimensions
- location relative to a fixed co-ordinate system
- location relative to other zones e.g. above, below, left, right, within, overlapping

Additionally, some way of identifying a default layout, e.g. all pages in a codex, was needed, with the ability to specify minor departures from it. Because relative location of zones was felt to be more important than absolute positioning, some way of specifying that a given zone acted as the *root* was needed. In discussion, it was stressed that relative locations should not need to be specified exhaustively. The group felt that this method of specifying layout would be adequate for the general case, but that special techniques would probably be needed for more complex layouts such as those exhibited in the Peirce materials, where figures and text are closely integrated.

On underlining and other forms of line, it was necessary to distinguish underlining added subsequently (which the **<add>** tag should be able to specify) from lines included as part of the text or as comments. Identifying the function of a line was an editorial matter.

To represent information about hands, two elements were proposed. A **<hand>** element should be used to identify any hand distinguished by the encoder. One such element should appear within the header for each hand distinguished within a text. Attributes recorded about it would include one or more of the following:

- identifier (compulsory)
- identifier of the scribe, where known
- type of writing (e.g. secretary, French, Italian...)
- descriptive characteristics of the writing (palsied, regular, irregular...)
- physical characteristics (ink colour, thickness etc)

At places in a ms where the hand changed, a specialised **<handShift>** milestone tag should be used, with an attribute *new* specifying the identifier of the new hand. It was noted that changes of hand often coincided with other tagged elements such as **<add>**: the need to additionally specify a **<handShift>** was felt to be a lesser evil than extending all such elements to allow for specification of this attribute, or folding it into the *rend* attribute.

The group noted that the proposed **<omit>**, **<damage>** and **<unclear>** elements all had very similar attributes. As a prelude to determining whether they might be combined, they focussed on each in turn, but unfortunately ran out of time.

On damage (for example blots, stains, torn off, cut off, rats,water etc) the group first considered whether such information was needed in its own right or simply where it was associated with illegibility. It was felt that such information might be important (e.g. for dating purposes) even where it did not affect the text. The following features were identified as potentially useful:

- location, e.g. to binding, page, zone etc.
- date or sequence of damage
- type e.g. faded, water, coffee, blot, overbound, cutout
- scope, specifying in more detail the pages or zones affected
- effect, i.e. whether or not text is lost or illegible as a result
- repair i.e. conjectural reconstruction of the undamaged text

It was suggested that a mirror tag **<repair>** might alternatively be used to specify a conjectural reconstruction. There was some doubt as to the appropriateness of the term, but no alternative was proposed. The group queried the need for the *resp* attribute proposed for this element in TR2 D4.

Finally the group noted that illegibility was not the same as a lacuna. Some way of distinguishing whether a reading was simply unclear or unquestionably unreadable might be needed.

Chapter 10

Report of the subgroup on editorial matters

AP reported on the consensus reached by the group with reference to recommended interpretative practice. Every scholarly act involved a degree of interpretation. There was a continuum between a literatim and an editorial transcription, and another set of decisions depending on whether one was dealing with individual unique mss or groups of variants. The purpose for which transcription was undertaken determined its position along both scales. The TEI could provide suggestions only about types of transcription practices appropriate to particular objectives.

PR reported that the group had decided against the notion of a set of standard entity names for commonly-occurring abbreviations, largely because of the difficulty of defining a set either generally applicable or of manageable size. Instead guidance should be given as to how such sets might be prepared for particular purposes, highlighting for example the need to choose between names indicative of appearance or of semantic reference. The set used would (of necessity) be documented within a dtd. The expansion of entities for abbreviations could point to various types of image, to tags or even to a group of tags.

On images, the group noted that these are only “encodable” to the extent that they are conventionalized symbols. If it is possible to formalize the list of information one wishes to record about an image (e.g. colors of ink used, style of the ms initial, animal(s) depicted, then that information may usefully be recorded with feature structures. Actual images can also be captured as external files, referred to from the transcription using standard TEI pointer mechanisms, for example by a **<fig>** or **<image>** tag, bearing an attribute of type *ENTITY* which is declared as containing an externally stored graphic image. No specific recommendations on the choice of a graphic storage format could or indeed should be made. Instead, the TEI relied on SGML's ability to coexist with any graphic notation declared by a user.

On uncertainty in general, the group noted that no Decidability attribute was needed on the **<add>** element since the existing **<certainty>** element did the same thing more generally. An attribute should however be added to the **<add>** element to specify the device or siglum used to anchor the addition within the text, for which the name *anchor* was proposed, with sample values such as caret, arrow, circle and arrow, circled A etc. This would not however work for insertion-point markers which contained other tags, for example to indicate that they were in another hand or were deleted etc. Such cases should be treated as editorial instructions, for which a new element **<instruction>** was proposed. This was a distinct type of **<note>** bearing a *resp* attribute to indicate its source. Transcribers should be free to choose whether to encode the insertion point at its physical or its logical point of occurrence, using existing pointer mechanisms to locate it precisely as well as the more general *place* to describe its location.

Chapter 11

Concluding Recommendations

Two chapters were proposed for inclusion in P2. The first, on critical editions, would comprise "licensing" text pointing to other relevant parts in P2 such as the sections on core tags and on physical description, followed by the bulk of what is currently in the second part of chapter TR2 D5. The second, on manuscript transcription, should contain similar licensing text, a typology of common transcriptional practices and sections on physical description of the copy text and on images. It was agreed that both chapters should make clear their applicability to printed as well as manuscript sources, and that the second should explicitly distinguish its concerns from those of critical editions. This general framework was approved.

The topics on which material for inclusion in these chapters should be prepared were itemized and allocated to members of the group for drafting as indicated in the following list

- Physical medium: MK
- Layout: AR/CK
- Lines: PR
- Hands: DS
- Damage: CK/PR
- Legibility : PR
- Substitution & authorial/editorial instructions: CH/DB
- Typology: AR
- Abbrevns: PR/DS
- Images: MK
- Addition anchoring: in core
- Text conjectured for lacuna: PR

Drafts should be prepared using a format similar to that of the existing draft for chapter 47 (TR2 D5), but including real-life examples, with citation of their source. Due to pressure of other work the editors were unable to help with drafting, but would be very willing to answer any specific problems which might arise. Drafts could be circulated for discussion either within the group or with the larger community served by the TEI's technical discussion list (see appendix B) as soon as they were ready, but a complete draft should be available by 1st November. CH undertook to finalise the draft by that date.

CH, as host, thanked all those who had participated, in particular those who had presented their own work. The group expressed their thanks to him and all of his colleagues at the Wittgenstein Archive and the NCCH for their unstinting hospitality.

Appendix A: An example to us all

Amongst the materials presented for discussion by PR was a particularly rebarbative-looking example of Bowers' transcription from a ms of Henry James (see TR9 XPR p2). During discussion PR incautiously opined that this presented incomprehensible complexities beyond the scope of the existing TEI proposals, whereupon MSM was silent for some time before producing the following bravura demonstration that this supposition was false:

First, the original, as transcribed by MSM from TR9 T6:

```
 But [intrl.] One [unreduced in error] must have lived
*some time [insrtd. for del. 'longer']
 with
*a [aft. del. 'such' ab. del. 'this']
 system, to appreciate
[del.
        'its
       *merits
       [ab. del. 'advantages']
        I might now undertake to
       *ingratiate my audience by exhibiting the latter seriatum, but
              **the
               [over 'its']
                advantages
             ***of the system will be
               [intrl.]
       [ab. del. 'describing ['ing' over 'e'] them, but they will appear']
        more
       *intelligible ['e' over 'y']
        after
       *its
       [ab. del. 'the']
        details
       ['of the system' del.]
        have been worked out, and can be dealt with then'
]
```

This might be transcribed as follows, using the tagset of TR2 D5:

```
<add loc='sup'>But</> <sic corr='one'>One</> must
have lived
<app>
    <lem><add loc='sup'>some time</></>
    <rdg><del>longer</></>
</app>
with
<app>
        <lem><del>such</del> a</lem>
        <rdg>this</>
</app>
system, to appreciate
<del>its
```

```
<app>
        <lem><add loc='sup'>merits</>.</>
        <rdg><del>advantages</>.</>
</app>
I might now undertake to<app>
<lem><add loc=sup>ingratiate my audience by exhibiting the
        latter seriatum, but
        <app>
                        <lem><add loc=sup>the</add></lem>
                        <rdg><del>its</del></rdg>
        </app>
        advantages
        <add loc=sup>of the system will be</add>
</add>
</lem>
<rdg><del>describ<app><lem><add loc=sup>ing</></>
                      <rdg><del meth=ow>e</del></rdg></app>
          them, but they will appear</del>
</rdg></app>
more
intelligibl<app><lem>e</><rdg>y</></app>
after
<app>
        <lem><add loc=sup>its</add></lem>
        <rdg><del>the</del></rdg>
</app>
details
<del>of the system</del>
have been worked out, and can be dealt with then
</del>.
```

QED

Appendix B: TEI Electronic Discussion Lists

The TEI maintains two electronic discussion lists, one open and one closed. Both are maintained using the standard LISTSERV software at UICVM, and are accessible in essentially the same way, with messages sent to the list being automatically redistributed to all subscribers as ordinary mail messages.

The open list, TEI-L, is used to announce availability of new drafts of the Guidelines and is intended for general public comment and discussion of the drafts and other matters of interest to the TEI community such as the availability of software. It has several hundred, mostly passive, subscribers. If you are not already subscribed, you should do so by sending a mail message to the following address:

```
LISTSERV@UICVM.UIC.EDU (Internet) or LISTSERV@UICVM (Bitnet)
```

The text of the message you send should contain simply the following line

```
SUBSCRIBE TEI-L Your Name
```

where "Your Name" is your real name e.g. Jane Dough. The software will work out your electronic address from the header of the message you send. It will send you by return a set of introductory messages outlining the purposes of the discussion list and how you can use it, for example, to download copies of TEI drafts and other useful materials. To send a message to everyone on this list, you should address the note to

```
TEI-L@UICVM.UIC.EDU
```

-- bearing in mind that any note so addressed will be automatically redistributed to several hundred people worldwide.

The closed discussion list, TEI-TECH, operates in a very similar way except that only active members of TEI Work Groups and committees are eligible for subscription. At present about thirty people, all of whom are or have been active in the formulation of the contents of TEI P2, are subscribed to the list. Specific technical issues are discussed here, and draft proposals floated in a comparatively informal manner. At the meeting it was agreed to add all current attendees to this list, so that they might benefit as far as possible from informed TEI-aware discussion of their proposals. Unless you let us know otherwise therefore, you will be added to this list within the next few weeks.