18 Transcription of Primary Sources

This chapter defines an optional additional tag set intended for use in the transcription of primary sources, in particular manuscripts, and describes how some elements defined in the core tag set should be used for this work. It is expected that this tag set will also be useful in the preparation of critical editions, but the tag set defined here is distinct from that defined in chapter 19 *Critical Apparatus*, and may be used independently of it.

Scholars may wish to record information concerning individual readings of letters, words or larger units, both within transcriptions and within editions. They may also wish to include other editorial material within transcriptions, such as comments on the status or possible origin of particular readings, corrections, or text supplied to fill lacunae. Further, it is customary in transcriptions to register certain features of the source, such as ornamentation, underlining, deletion, areas of damage and lacunae. This chapter indicates means to record such information:

- first, the problem of recording editorial or other alterations to the text, such as expansion of abbreviations, corrections, conjectures, etc. (section 18.1 *Altered, Corrected, and Erroneous Texts*)
- then, methods of describing important extra-linguistic phenomena in the source: unusual spaces, lines, page and line breaks, change of manuscript hand, etc. (section 18.2 *Non-Linguistic Phenomena in the Source*)
- finally, a method of recording material such as running heads, catch-words, and the like (section 18.3 *Headers, Footers, and Similar Matter*)

These recommendations are not intended to meet every transcriptional circumstance likely to be faced by any scholar. Rather, they should be regarded as a base which can be elaborated if necessary by different scholars in different disciplines, with distinct scholarly domains eventually developing their own document types. In time, the feature structure notation developed in chapter 16 *Feature Structures*, may also permit scholars to tailor the encoding of complex transcriptional information in ways not here anticipated.

It should be noted that this chapter focuses primarily upon problems associated with the transcription of manuscript materials, and that consequently problems of codicology other matters peculiar to early printed materials are not specifically addressed here. Nevertheless, many of the recommendations presented may — mutatis mutandis — also be applied in the encoding of printed matter. We are conscious that a great deal of work remains to done in these areas, and that the encoder will need to take even more individual responsibility than usual in applying the recommendations of this chapter in such contexts, but believe that these recommendations form a good basis for such future work.

Many of the descriptions below use terms like 'scribe', 'author', 'editor', 'annotator', 'corrector', 'transcriber', and 'encoder', to make clear how they apply in cases where these roles are distinct. To the extent that these roles are not distinct (for example, in authorial manuscripts where the author and the scribe are the same person) the interpretation of the markup should be adjusted appropriately. Many of the elements defined here apply (within limits) also in cases of printed materials, so 'compositor', etc., may also be understood as applying where appropriate.

As a rule, all elements which may be used in the course of a transcription of a single witness may also be used in a critical apparatus, i.e. within the elements proposed in chapter 19 *Critical Apparatus*. This can generally be achieved by nesting a particular reading containing tagged elements from a particular witness within the `<rdg>` element in an `<app>` structure.

Just as a critical apparatus may contain transcriptional elements within its record of variant readings in various witnesses, one may record variant readings in an individual witness by use of the apparatus mechanisms `<app>` and `<rdg>`. This is discussed in section 19.3 *Using Apparatus Elements in Transcriptions*.

The tag set defined in this chapter may be selected using the mechanisms described in section 3.3 *Invocation of the TEI DTD*; in a document using this tag set, the document-type-declaration subset should contain the following declaration of the parameter entity TEI.transcr, or the equivalent:

```
<!ENTITY % TEI.transcr 'INCLUDE' >
```

In an XML document using this tag set together with that for textual criticism and the base tag set for verse, the entire document type declaration might resemble the following:

```
<!DOCTYPE TEI.2 PUBLIC "-//TEI P4//DTD Main Document Type//EN"
                         "tei2.dtd" [
   <!ENTITY % TEI.XML 'INCLUDE' >
   <!ENTITY % TEI.prose 'INCLUDE' >
   <!ENTITY % TEI.transcr 'INCLUDE' >
   <!ENTITY % TEI.textcrit 'INCLUDE' >
]>
```

The overall structure of the tag set defined by this chapter is as follows:

```
<!-- 18.: Transcription of Primary Sources-->
<!--Text Encoding Initiative Consortium:
Guidelines for Electronic Text Encoding and Interchange.
Document TEI P4, 2002.
Copyright (c) 2002 TEI Consortium. Permission to copy in any form
is granted, provided this notice is included in all copies.
These materials may not be altered; modifications to these DTDs should
be performed only as specified by the Guidelines, for example in the
chapter entitled 'Modifying the TEI DTD'
These materials are subject to revision by the TEI Consortium. Current versions
are available from the Consortium website at http://www.tei-c.org-->
<!--declarations from 18.1.4: Added and Deleted Spans inserted here -->
<!--declarations from 18.1.6: Cancelled Deletions inserted here -->
<!--declarations from 18.1.7: Supplied Text inserted here -->
<!--declarations from 18.2.1: Hand Shifts inserted here -->
<!--declarations from 18.2.3: Damage and Illegiblity inserted here -->
<!--declarations from 18.2.5: Spaces in the source inserted here -->
<!--declarations from 18.3: Headers and footers inserted here -->
<!-- end of 18.-->
```

This tag set modifies the element class edit by declaring two extra attributes for members of the class:

```
<!-- 18.: Attributes for Transcription of Primary Sources-->
<!--Text Encoding Initiative Consortium:
Guidelines for Electronic Text Encoding and Interchange.
Document TEI P4, 2002.
Copyright (c) 2002 TEI Consortium. Permission to copy in any form
is granted, provided this notice is included in all copies.
These materials may not be altered; modifications to these DTDs should
be performed only as specified by the Guidelines, for example in the
chapter entitled 'Modifying the TEI DTD'
These materials are subject to revision by the TEI Consortium. Current versions
are available from the Consortium website at http://www.tei-c.org-->
<!ENTITY % a.edit '
     resp IDREF %INHERITED;
     cert CDATA #IMPLIED'>
<!-- end of 18.-->
```

18.1 Altered, Corrected, and Erroneous Texts

In the detailed transcription of any source, it may prove necessary to record various types of actual or potential alteration of the text: expansion of abbreviations, correction of the text (by the author, by a scribe, by a later hand, by previous editors or scholars, or by the current editor or encoder), addition, deletion, or substitution of material, and the like. The sections below describe how such phenomena may be encoded using either elements defined in the core tag set (defined in chapter 6 *Elements Available in All TEI Documents*) or specialized elements available only when the additional tag set described in this chapter is available.

18.1.1 Use of Core Tags for Transcriptional Work

In transcribing individual sources of any type, encoders may record their corrections, normalizations, expansions of abbreviations, additions, and omissions using the elements described in section 6.5 *Simple Editorial Changes*. Those particularly relevant to this chapter include:

<abbr> contains an abbreviation of any sort. Attributes include:

expan (expansion) gives an expansion of the abbreviation.

Values any string of characters

resp (responsibility) signifies the editor or transcriber responsible for supplying the expansion of the abbreviation held as the value of the expan attribute.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the expansion of the abbreviation.

type allows the encoder to classify the abbreviation according to some convenient typology.

Sample values include:

suspension the abbreviation provides the first letter(s) of the word or phrase, omitting the remainder.

contraction the abbreviation omits some letter(s) in the middle.

brevigraph the abbreviation comprises a special symbol or mark.

superscription the abbreviation includes writing above the line.

acronym the abbreviation comprises the initial letters of the words of a phrase.

title the abbreviation is for a title of address (Dr, Ms, Mr, ...)

organization the abbreviation is for the name of an organization.

geographic the abbreviation is for a geographic name.

\<expan\> contains the expansion of an abbreviation. Attributes include:

abbr (abbreviation) gives the abbreviation in its unexpanded form.

Values any string of characters

resp (responsibility) signifies the editor or transcriber responsible for supplying the expansion of the abbreviation held as the content of the \<expan\> element.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the expansion of the abbreviation.

type allows the encoder to classify the abbreviation according to some convenient typology.

Values any useful classification name, e.g. 'suspension', 'contraction', 'brevigraph', 'title', 'organization', 'geographic', etc.

\<sic\> contains text reproduced although apparently incorrect or inaccurate. Attributes include:

corr (correction) gives a correction for the apparent error in the copy text.

Values any string of characters

resp (responsibility) signifies the editor or transcriber responsible for suggesting the correction held as the value of the corr attribute.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the correction held as the value of the corr attribute.

\<corr\> contains the correct form of a passage apparently erroneous in the copy text. Attributes include:

sic gives the original form of the apparent error in the copy text.

Values any string of characters

resp (responsibility) signifies the editor or transcriber responsible for suggesting the correction held as the content of the \<corr\> element.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the correction held as the content of the <corr> element.

<add> contains letters, words, or phrases inserted in the text by an author, scribe, annotator, or corrector. Attributes include:

place if the the addition is written into the copy text, indicates where the additional text is written.

Suggested values include:

    `inline` addition is made in a space left in the witness by an earlier scribe

    `supralinear` addition is made above the line

    `infralinear` addition is made below the line

    `left` addition is made in left margin

    `right` addition is made in right margin

    `top` addition is made in top margin

    `bottom` addition is made in bottom margin

    `opposite` addition is made on opposite page

    `verso` addition is made on verso of sheet

    `mixed` addition is made somewhere, one or more of other values

resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the addition.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the addition.

hand signifies the hand of the agent which made the addition.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

<del> contains a letter, word or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, annotator, or corrector. Attributes include:

type classifies the type of deletion using any convenient typology.

Values any string identifying the class of deletion.

status may be used to indicate faulty deletions, e.g. strikeouts which include too much or too little text.

Values any description of flaws in the marking of a deletion, e.g. 'excess left', 'excess right', 'short left', 'short right'.

resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the deletion.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the deletion.

hand signifies the hand of the agent which made the deletion.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

<hi> marks a word or phrase as graphically distinct from the surrounding text, for reasons concerning which no claim is made.

<gap> indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible or inaudible. Attributes include:

desc (description) gives a description of the omitted text.

Values a prose description of the material omitted.

reason gives the reason for omission. Sample values include 'sampling', 'illegible', 'inaudible', 'irrelevant', 'cancelled', 'cancelled and illegible'.

Values any short indication of the reason for the omission.

resp (responsibility) indicates the editor, transcriber or encoder responsible for the decision not to provide any transcription of the text and hence the application of the `<gap>` tag.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

hand in the case of text omitted from the transcription because of deliberate deletion by an identifiable hand, signifies the hand which made the deletion.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

agent In the case of text omitted from the transcription because of damage or other phenomenon resulting from an identifiable cause, signifies the causative agent.

Values any prose description of the agency of damage.

extent indicates approximately how much text has been omitted from the transcription, in letters, minims, inches, or any appropriate unit, either because of editorial policy or because a deletion, damage, or other cause has rendered transcription impossible.

Values any string of characters

When the additional tag set for transcription of primary sources is selected, these elements all gain two specialized attributes for specifying who is responsible for certain aspects of the interpretation and markup, and the certainty attributed to the interpretation:

cert signifies the degree of certainty ascribed to some specific aspect of the markup: the identification of the hand of an addition or deletion, the correctness of the expansion of an abbreviation, the correction of an error, or the regularization of a non-standard form; or the correctness of the transcription of unclear material.

resp signifies the editor or transcriber responsible for the salient information conveyed by a particular tag: the hand of an addition or deletion, the expansion of an abbreviation, the correction of an apparent error, the regularization of a non-standard form, the transcription of unclear material, or the decision not to transcribe some portion of the text.

The specific aspect of the markup described by these attributes differs on different elements; for further discussion, see the relevant sections below, especially section 18.2.2 *Hand, Responsibility, and Certainty Attributes*.

The following sections describe how the core elements just named may be used in the transcription of primary source materials. Examples of more complex application in scholarly transcriptions of these core elements are given, and of their extension by linkage with the `<note>`, `<respons>`, and `<certainty>` elements. Where the core elements do not satisfy the needs of scholarly transcription, additional elements are defined.

18.1.2 Abbreviation and Expansion

The writing of manuscripts by hand lends itself to the use of abbreviation to shorten scribal labour. Commonly occurring letters, groups of letters, words or even whole phrases, may be represented by significant marks. This phenomenon of manuscript abbreviation is so widespread and so various that no taxonomy of it is here attempted. Instead, methods are shown which allow abbreviations to be encoded using the core elements mentioned above.

A manuscript abbreviation may be viewed in two ways. One may transcribe it as a particular sequence of letters or marks upon the page: thus, a "p with a bar through the descender", a "superscript hook", a "macron". One may also interpret the abbreviation in terms of the letter or letters it is seen as standing for: thus, "per", "re", "n". Both of these views are supported by these Guidelines. The entity reference system allows the encoder to declare whatever entities are needed, using entity names like p-underbar, sup-hook, or macron. Furthermore, each entity reference may be linked to an image of the abbreviation

itself, so that the reader might see a rendering of the text's appearance. Alternatively, the encoder may transcribe the letter or letters he or she believes the abbreviation stands for, as the content of an `<expan>` element: thus

```
<expan>per</expan> <expan>re</expan> <expan>n</expan>
```

These two methods of coding abbreviation may also be combined. An encoder may record, for any abbreviation, both the sequence of letters or marks which constitutes it, and its sense, that is, the letter or letters for which it is believed to stand. For example, the abbreviations of 'euery persone' in the following fragment[133] may be transcribed as follows, using the `<expan>` element, with the abbr attribute to hold an entity reference for the brevigraph or other sign indicating the abbreviation in the manuscript:

```
eu<expan abbr="&er;" resp="mp">er</expan>y
<expan abbr="&p-underbar;">per</expan>sone that
loketh after heuen hath a place in this ladder
```

Alternatively, the abbreviations may be encoded using the `<abbr>` element.

```
eu<abbr expan="er" resp="mp">&er;</abbr>y
<abbr expan="per">&p-underbar;</abbr>sone that
loketh after heuen hath a place in this ladder
```

The choice between the `<expan>` and `<abbr>` elements is left to the encoder. As a rule, the `<abbr>` element should be preferred where it is wished to signify that the content of the element is an abbreviation, without necessarily indicating what the abbreviation may stand for. The `<expan>` element should be used where it is wished to signify that the content of the element is an expanded text, without necessarily indicating the abbreviation used in the original. The decision as to which (`<abbr>` or `<expan>`) to use may vary from abbreviation to abbreviation; there is no requirement that the one system be used throughout a transcription. However, processing may be simplified if one only of these is used throughout a transcription. The choice is likely to be a matter of editorial policy, which might be applied consistently throughout. If the highest priority is to transcribe the text literatim, while indicating the presence of abbreviations, the choice will be to use `<abbr>` throughout. If the highest priority is to present a reading transcription, while indicating that some letters or words are expansions of abbreviations, the choice will be to use `<expan>` throughout.

Further information may be attached to instances of these elements by the `<note>` element, on which see section 6.8 *Notes, Annotation, and Indexing*, and by use of the resp and cert attributes. In this instance from the English *Brut*,[134] a note is attached to an editorial expansion of the tail on the final d of 'good' to 'goode':

```
For alle the while that I had
good<expan id="exp01" abbr="&tail;">e</expan>
I was welbeloued
```

Then the note:

```
<note target="exp01">The stroke added to
the final d could signify the plural ending (-es, -is, -ys&gt;)
but the singular <hi rend="it">good</hi> was used with the meaning
<q>property</q>, <q>wealth</q>, at this time (v. examples
quoted in OED, sb. Good, C. 7, b, c, d and 8 spec.)</note>
```

The editor might declare a degree of certainty for this expansion, based on the OED examples, and state the responsibility for the expansion:

```
For alle the while that I had
good<expan abbr="&tail;" resp="mp" cert="90">e</expan>
I was welbeloued
```

Observe that the cert and resp attributes may be used with the `<expan>` element only to indicate respectively confidence in the content of the element (i.e. the expansion), and confidence in the responsibility for suggesting this expansion. In the case of the use of these attributes with the `<abbr>`,

---

[133] From *The Manere of Good Lyuynge*, fol 126v of Bodleian MS Laud Misc. 517, plate 8(ii) in *English Cursive Book Hands 1250–1500* by M. B. Parkes (Clarendon Press: Oxford, 1969).
[134] On fol 65v of Bodleian MS. Rawlinson Poetry 32; in Parkes 12(ii).

the cert and resp attributes are defined as indicating respectively confidence in the expansion held in the expan attribute and the responsibility for suggesting this expansion. The above example could be encoded using the `<abbr>` element as follows:

```
For alle the while that I had
good<abbr expan="e" resp="mp" cert="90">&tail;</abbr>
I was welbeloued
```

If it is desired to express aspects of certainty and responsibility for some other aspect of the use of these elements, then the mechanisms discussed in chapter 17 *Certainty and Responsibility* should be used. See also 18.2.2 *Hand, Responsibility, and Certainty Attributes* for discussion of the issues of certainty and responsibility in the context of transcription.

If more than one expansion for the same abbreviation is to be recorded, multiple notes may be supplied. It may also be appropriate to use the markup for critical apparatus; an example is given in section 19.3 *Using Apparatus Elements in Transcriptions*.

18.1.3 Correction and Conjecture

The `<sic>` and `<corr>` elements, defined in the core tag set, may be used to register authorial or scribal corrections within a witness. For example, in the manuscript of William James's *A Pluralistic Universe*, edited by Fredson Bowers (Cambridge: Harvard University Press, 1977) a sentence first written

One must have lived longer with this system, to appreciate its advantages.

has been modified by James to begin "But one must ...", without the inital capital O having been reduced to lowercase. This non-standard orthography could be recorded and corrected thus:

```
But <sic corr="one">One</sic> must have lived ...
```

The same information could be conveyed by the `<corr>` element:

```
But <corr sic="One">one</corr> must have lived ...
```

In this example from Albertus Magnus,[135] both the manuscript error 'angues' and its correction 'augens' are registered by the `<sic>` element:

```
Nos autem iam ostendimus quod nutrimentum
et <sic corr="augens">angues</sic>.
```

The same information could be conveyed by the `<corr>` element:

```
Nos autem iam ostendimus quod nutrimentum
et <corr sic="angues">augens</corr>.
```

As with the choice between `<expan>` and `<abbr>`, the choice between the synonymous `<sic>` and `<corr>` elements is left to the encoder. As a rule, the `<sic>` element allows the encoding to retain the original text as the content of the element, while simultaneously signifying that the contents of the element require correction, but without necessarily indicating what the correction may be. The `<corr>` element allows the text to be corrected, possibly without recording the details of the faulty source, while still marking explicitly the fact that the contents of the element have been corrected. The choice is likely to be a matter of editorial policy, which might be applied consistently throughout or decided case by case. If the highest priority is to present an uncorrected transcription while noting perceived errors in the original, the choice will typically be to use `<sic>` throughout. If the highest priority is to present a reading transcription, while indicating that perceived errors in the original have been corrected, the choice will be to use `<corr>` throughout.

Further information may be attached to instances of these elements by the `<note>` element and resp and cert attributes. Here, two separate corrections in Dudo of S. Quentin[136] are assigned the same note. First the corrections, held in the attribute value of the `<sic>` elements:

```
quamuis <sic id="sic01" corr="iners">mens</sic> que nutu dei
gesta sunt ... unde esset uiriliter
<sic id="sic02" corr="uegetata">negata</sic>
```

then the note, linked to the id of the `<sic>` element for each of the two corrections:

---

[135] *De Nutrimento et Nutribili, Tractatus 1*, fol 217r col b of Merton College Oxford MS O.2.1 (Parkes pl. 16).
[136] *De moribus et actis primorum Normannie ducum*, in fol 4v of British Library MS Harley 3742, Parkes pl 6(i).

```
<note target="sic01 sic02">Substitution of a more
familiar word which resembles graphically what the
scribe should be copying but which
does not make sense in the context.</note>
```

The cert attribute may also be used with the `<corr>` element to signify the conjectural status of a particular editorial reading, with the resp attribute used to identify the scholar responsible for the conjecture. In this example, editorial confidence in E. Talbot Donaldson's emendation of the Hengwrt manuscript reading 'wight' to 'wright' in line 117 of Chaucer's *The Wife of Bath's Prologue* may be marked as follows:

```
Telle me also, to what conclusioun
Were membres maad, of generacioun
And of so parfit wis a
<corr id="c117" sic="wight" resp="ETD" cert="70">wright</corr>
ywroght?
```

The editor might also conveniently add a note referring to Donaldson's discussion of this passage:

```
<note target="c117">This emendation of the Hengwrt copy text,
based on a Latin source and on the reading of three late
and usually unauthoritative manuscripts, was proposed
by E. Talbot Donaldson in <bibl><title>Speculum</title> 40 (1965)
626&ndash;33.</bibl></note>
```

Alternative corrections within a transcription of a single witness may be held within an `<app>` structure, in the same way that alternative expansions are so grouped in the example given in section 19.3 *Using Apparatus Elements in Transcriptions*. Here, Donaldson's conjectured emendation of the Hengwrt manuscript may be recorded not only alongside the editorial transcription but also alongside another conjecture:

```
And of so parfit wis a
<app>
  <rdg wit="Hg">wight</rdg>
  <rdg wit="Ln Ry2 Ld" resp="ETD"> <corr>wright</corr> </rdg>
  <rdg wit="Gg" resp="PR"> <corr>wyf</corr> </rdg>
</app>
```

Observe that no resp attribute is necessary for the base transcription: by default, responsibility is assigned to the scholar(s) responsible for the transcription, as identified in the TEI header. The conjectures are held within `<corr>` elements, contained within the `<rdg>` elements. The resp attribute identifying responsibility for each correction is attached to the outer `<rdg>`, and inherited by the inner `<corr>` element. Note too that the support for these conjectures in other manuscripts can be noted in the wit attribute in the `<rdg>` element.

The cert and resp attributes may be used with the `<corr>` element only to indicate respectively confidence in the content of the element (i.e. the correction), and confidence in the responsibility for suggesting this correction or conjecture. In the case of the use of these attributes with the `<sic>` element, the cert and resp attributes are defined as indicating respectively confidence in the conjecture held in the corr attribute and the responsibility for suggesting this conjecture. The above example could be encoded using the `<sic>` element as follows:

```
And of so parfit wis a
<sic corr="wright" resp="etd" cert="70">wight</sic>
ywroght?
```

If it is desired to express aspects of certainty and responsibility for some other aspect of the use of these elements, then the mechanisms discussed in chapter 17 *Certainty and Responsibility* should be used. See also 18.2.2 *Hand, Responsibility, and Certainty Attributes* for discussion of the issues of certainty and responsibility in the context of transcription.

18.1.4 Additions and Deletions

Additions and deletions to a text may be described using the following elements:

<add> contains letters, words, or phrases inserted in the text by an author, scribe, annotator, or corrector. Attributes include:

place if the the addition is written into the copy text, indicates where the additional text is written.

Suggested values include:

    `inline`   addition is made in a space left in the witness by an earlier scribe

    `supralinear`   addition is made above the line

    `infralinear`   addition is made below the line

    `left`   addition is made in left margin

    `right`   addition is made in right margin

    `top`   addition is made in top margin

    `bottom`   addition is made in bottom margin

    `opposite`   addition is made on opposite page

    `verso`   addition is made on verso of sheet

    `mixed`   addition is made somewhere, one or more of other values

resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the addition.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the addition.

hand signifies the hand of the agent which made the addition.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

<addSpan> marks the beginning of a longer sequence of text added by an author, scribe, annotator or corrector (see also `<add>`). Attributes include:

place indicates where the addition is made.

Suggested values include:

    `inline`   addition is made in a space left in the witness by an earlier scribe.

    `supralinear`   addition is made above the line.

    `infralinear`   addition is made below the line.

    `marginleft`   addition is made in left margin.

    `marginright`   addition is made in right margin.

    `margintop`   addition is made in top margin.

    `marginbot`   addition is made in bottom margin.

    `overleaf`   addition is made on the other side of the leaf.

resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the addition.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the addition.

hand signifies the hand of the agent which made the addition.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

to indicates the endpoint of the added passage, by supplying the value of the `id` attribute of an `<anchor>` or other empty element placed there.

Values any valid identifier.

<del> contains a letter, word or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, annotator, or corrector. Attributes include:

type classifies the type of deletion using any convenient typology.

    Values any string identifying the class of deletion.

status may be used to indicate faulty deletions, e.g. strikeouts which include too much or too little text.

    Values any description of flaws in the marking of a deletion, e.g. 'excess left', 'excess right', 'short left', 'short right'.

resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the deletion.

    Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the deletion.

hand signifies the hand of the agent which made the deletion.

    Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

<delSpan> marks the beginning of a longer sequence of text deleted, marked as deleted, or otherwise signaled as superfluous or spurious by an author, scribe, annotator, or corrector. Attributes include:

type classifies the deletion, using any convenient typology.

    Sample values include:

        `overstrike` deletion indicated by line crossing out the text.

        `erasure` deletion indicated by erasure of the text.

        `bracketed` deletion indicated by brackets in the text or margin.

        `subpunction` deletion indicated by dots beneath the letters deleted.

status indicates whether the deletion is faulty, e.g. by including too much or too little text.

    Sample values include:

        `excess start` some text at the beginning of the deletion is marked as deleted even though it clearly should not be deleted.

        `excess end` some text at the end of the deletion is marked as deleted even though it clearly should not be deleted.

        `short start` some text at the beginning of the deletion is not marked as deleted even though it clearly should be.

        `short end` some text at the end of the deletion is not marked as deleted even though it clearly should be.

        `unremarkable` the deletion is not faulty.

resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the deletion.

    Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the deletion.

hand signifies the hand of the agent which made the deletion.

    Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

to identifies the endpoint of the deleted passage, by supplying the value of the id attribute of an `<anchor>` or other empty element placed there.

    Values any valid identifier.

Of these, `<add>` and `<del>` are included in the core tag set, while `<addSpan>` and `<delSpan>` are available only when using the additional tag set defined in this chapter.

As described in section 6.5 *Simple Editorial Changes*, the `<add>` element indicating material added may be used to signify manuscript additions or insertions, be they authorial or scribal. In the autograph

manuscript of Max Beerbohm's *The Golden Drugget*,[137] the author's addition of "do ever" may be recorded as follows, with the hand attribute indicating that the addition was Beerbohm's:

```
Some things are best at first sight. Others &#x2014; and
here is one of them &#x2014; <add hand="mb">do ever</add>
improve by recognition
```

Similarly, the `<del>` element indicating material deleted may be used to signify manuscript deletions. In the autograph manuscript of D. H. Lawrence's *Eloi, Eloi, lama sabachthani*[138], the author's deletion of 'my' may be recorded as follows. As well as the hand attribute indicating that the deletion was Lawrence's, the rend attribute indicates that the deletion was by strike-through:

```
For I hate this <del rend="strikethrough" hand="dhl">my</del> body,
which is so dear to me
```

If deletions are classified systematically, the type attribute should normally be used to indicate the classification; when they are classified by the manner in which they were effected, or by their appearance, however, this will lead to a certain arbitrariness in deciding whether to use the type or the rend attribute to hold the information. In general, it is recommended that the rend attribute be used for description of the appearance or method of deletion, and that the type attribute be reserved for higher level or more abstract classifications.

Further characteristics of the addition and deletion, e.g. the date, or ink, may be needed for detailed transcription of manuscripts. Such characteristics may conveniently be recorded as attributes of the `<add>` or `<del>` element. The specific attributes required may be added to the formal declaration of these elements by using the techniques described in chapter 29 *Modifying and Customizing the TEI DTD*.

The `<add>` and `<del>` elements defined in the core tag set available in all TEI documents will suffice for describing typically brief additions and deletions in the text being transcribed. On occasion, it will be necessary to record an addition or deletion which crosses a structural boundary in the text being encoded, for example the addition or deletion from a manuscript of a section containing several distinct structural subdivisions, such as poems or prose items. These are most conveniently encoded using the `<addSpan>` and `<delSpan>` elements, available in the additional tag set defined in this chapter. In this example of the use of `<addSpan>`, the insertion of a gathering containing four neo-Eddic poems into *Landsbókasafn*[139] by Helgi Ólafsson is recorded as follows. A `<hand>` element is first declared, within the header of the document, to associate the identifier HEOL with Helgi. In the body of the text, an `<addSpan>` element is placed to mark the beginning of the span of added text. The hand attribute ascribes the responsibility for the addition to the manuscript to Helgi, and the to attribute declares the identifier for the anchor which marks the end of the added text:

```
<hand id="heol" n="Helgi &Oacute;lafsson"/>
<!-- text of the original material ... -->
<addSpan type="added gathering" hand="heol" to="p025"/>
<!-- text of the four neo-Eddic poems added... -->
<anchor id="p025"/>
<!-- text of the original material continues... -->
```

In this example of the use of the `<delSpan>` element, a full two lines of Thomas Moore's autograph of the second version of *Lalla Rookh*[140] are marked for omission by vertical strike-through. The two lines cross the structural line division marked `<l n='2'>`, so it would not be possible to use a single `<del>` element, since it would have to span the `<l>` marker. The lines also themselves include a further deletion and addition. The `<delSpan>` element indicates the begining of the span marked for deletion, with the to attribute giving the identifier delend01 for an `<anchor>` element which marks the end of the span of text so marked:

```
<l n="1">
    <delSpan rend="vertical strike" to="delend01"/>
```

[137] In Pierpont Morgan MA 3391 (Klinkenborg 123).
[138] In Pierpont Morgan MA 1892, (Klinkenborg 129).
[139] In Reykjavík, Lbs 1562 4to
[140] In Pierpont Morgan MA 310, (Klinkenborg 23).

```
       Tis moonlight <del>upon</del> <add>over</add> Oman's sky</l>
  <l n="2">Her isles of pearl look lovelily<anchor id="delend01"/></l>
```

The text deleted must be at least partially legible, in order for the encoder to be able to transcribe it. If it is not legible at all, the `<gap>` element should be used to signal that the text was not transcribed, because it could not be; the `reason` attribute can give the cause of the omission from the transcription as "deletion, illegible". The `<gap>` element may optionally be enclosed by a `<del>` element, if it is thought useful to record the deletion explicitly using this element. If the deleted text is partially legible, the `<unclear>` element described in section 18.2.3 *Damage, Illegibility, and Supplied Text* should be used to signal the areas of text which cannot be read with confidence; it too may be enclosed within a `<del>` element. See further section 18.1.7 *Text Omitted from or Supplied in the Transcription* and section 18.2.3 *Damage, Illegibility, and Supplied Text*.

The elements `<add>`, `<del>`, and `<gap>` are defined in the core tag set and are available in all TEI documents. The elements `<addSpan>` and `<delSpan>` have the following formal declarations:

```
<!-- 18.1.4: Added and Deleted Spans-->
<!ELEMENT addSpan %om.RO;  EMPTY>
<!ATTLIST addSpan
      %a.global;
      type CDATA #IMPLIED
      place CDATA #IMPLIED
      resp IDREF %INHERITED;
      cert CDATA #IMPLIED
      hand IDREF %INHERITED;
      to IDREF #REQUIRED
      TEIform CDATA 'addSpan'  >
<!ELEMENT delSpan %om.RO;  EMPTY>
<!ATTLIST delSpan
      %a.global;
      type CDATA #IMPLIED
      resp IDREF %INHERITED;
      cert CDATA #IMPLIED
      hand IDREF %INHERITED;
      to IDREF #REQUIRED
      status CDATA "unremarkable"
      TEIform CDATA 'delSpan'  >
<!-- end of 18.1.4-->
```

18.1.5 Substitutions

Substitution of one word or phrase for another is perhaps the most common of all phenomena requiring special treatment in transcription of primary textual sources. It may be simply one word overwriting another, or deletion of one word and its replacement by another written above it by the same hand at the one time; the deletion and replacement may be done by different hands at different times; there may be a long chain of substitutions on the one stretch of text, with uncertainty as to the order of substitution and as to the final reading.

Three different methods may be used to express substitution of one stretch of text by another:

- the `<sic>` and `<corr>` elements, either individually to encode a single substitution or nested to encode a sequence of substitutions;
- the `<del>` and `<add>` elements, used in sequence to show that text was first deleted then other text inserted;
- the `<del>` and `<add>` elements, used within an `<app>` structure (as defined in chapter 19 *Critical Apparatus*) to indicate that the deleted and added text within the individual reading elements making up the `<app>` structure are variants of one another.

The use of all three of these is illustrated in the following encodings of the second line of *Eloi, Eloi, lama sabachthani* from the Lawrence manuscript mentioned above. Lawrence first wrote "How it galls me, what a galling shadow". Subsequently, he deleted 'galls' and wrote 'dogs' above the deletion.

This substitution could be registered using the first method outlined above, as a correction using the `<sic>` or `<corr>` elements. Note the use of the resp attribute on the `<corr>` element to assign the correction to Lawrence. (For further information on the hand and resp attributes, see section 18.2.2 *Hand, Responsibility, and Certainty Attributes*.)

```
How it <corr sic="galls" resp="DHL">dogs</corr>
me, what a galling shadow
```

This substitution could be registered using the second method outlined above, using the `<del>` and `<add>` elements in sequence to reflect the fact that text was first deleted then other text inserted:

```
How it <del type="overstrike" hand="dhl">galls</del>
<add place="supralinear" hand="dhl">dogs</add>
me, what a galling shadow
```

This substitution could be registered using the third method outlined above, using the `<del>` and `<add>` elements within an `<app>` structure to indicate that the deleted and added texts are variants of one another. Note that within the `<app>` structure the hand attribute is moved from the inner `<del>` and `<add>` elements to the outer `<rdg>` element:

```
How it
  <app>
    <rdg hand="dhl"> <del type="overstrike"> galls</del> </rdg>
    <rdg hand="dhl"> <add place="supralinear"> dogs</add> </rdg>
  </app>
me, what a galling shadow
```

Each of these three methods has its particular advantages and disadvantages. The first method (use of `<sic>` or `<corr>`) is compact and indicates clearly that one text is a substitute for another. However, it provides no clear means of stating how the substitution is effected: whether by deletion through strike-through, or underdotting, or erasure, followed by interlinear insertion, or marginal insertion. (The global rend attribute might conceivably be used, but this may not be thought an obvious place to put such information.) In a transcription where this information is not felt to be important, however, this method will suffice to indicate simple cases of direct substitution of one text for another.

The second method (use of a `<del>` and `<add>` sequence) is also compact and provides means for exact declaration of how the deletion and insertion are effected. However, it does not indicate explicitly that one text is a substitute for another. It is left for the reader or the application to infer from the `<del>` and `<add>` sequence that the insertion is to be taken as a substitution for the deletion. In many transcriptions, the inference may be safely drawn for simple cases of direct substitution of one text for another. In other transcriptions, for example of complex authorial manuscripts, this inference may prove fragile; those who desire to express clearly that an adjacent addition and deletion are not independent but constitute a single act of substitution will therefore wish to avoid this method. Others, of course, may prefer it for precisely the same reason, namely that it avoids prejudging the issue of whether adjacent deletions and additions are independent or joined.

The third method (use of the `<del>` and `<add>` elements within an `<app>` structure) provides means both for exact declaration of how the deletion and insertion are effected and for explicit indication that one text is a substitute for another. Further, the exact sequence of readings may also be declared by use of the varSeq attribute on the `<rdg>` element, as follows:

```
How it
  <app>
    <rdg varSeq="1" hand="dhl"> <del>galls</del> </rdg>
    <rdg varSeq="2" hand="dhl"> <add>dogs</add> </rdg>
  </app>
me, what a galling shadow
```

Here, the combination of the hand and varSeq attributes suffices to inform the reader of the authorial substitution of 'dogs' for 'galls'.

Similarly, the varSeq attribute might be used in a transcription of the manuscripts of James Joyce's *Ulysses* to indicate the sequence of Joyce's corrections which is implicit in Hans Walther Gabler's reconstruction of the "overlay" levels of Joyce's transcriptions. This third method is the most powerful

and unambiguous of the three methods and enables the widest range of processing possibilities, at the expense of introducing a heavier burden of markup into the text. Production of such documents should therefore not be undertaken without markup-aware editors. Applications of some sophistication may be needed to make full use of all the information that may be held within an <app> structure. In the absence of such applications, scholars may feel that the present cost of the more informative coding using <app> structures outweighs the future benefits. In making such decisions, it should however be kept in mind that the capabilities of software at the time a project begins will often be wholly irrelevant when the project is completed some years later.

The Lawrence example above shows the three methods used for encoding a single substitution of one reading for another. The same three methods may also be used to encode longer sequences of substitutions. In the example from William James, first written out by James as "One must have lived longer with this system, to appreciate its advantages" the word 'this' is first replaced by 'such a' and this is then replaced by 'a'. [141] This may be encoded using the first method, with the sequence of substitutions shown by the nesting of <corr> elements:

```
One must have lived longer with
<corr sic="this"><corr sic="such a">a</corr></corr> system,
to appreciate its advantages.
```

It may be encoded using the second method, with the two changes being treated as a sequence of additions and deletions:

```
One must have lived longer with
<del>this</del> <del><add>such a</add></del>
<add>a</add> system, to appreciate its advantages.
```

Note the nesting of an <add> element within a <del> to record text first added, then deleted in the source.

It may be encoded using the third method, with each reading in the series contained in a <rdg> element within an <app> structure:

```
One must have lived longer with
  <app>
    <rdg varSeq="1"><del>this</del></rdg>
    <rdg varSeq="2"><del><add>such a</add></del></rdg>
    <rdg varSeq="3"><add>a</add></rdg>
  </app>
system, to appreciate its advantages.
```

The three encodings of this slightly more complex example illustrate the general truth that the more information involving substitutions there is to be encoded, the clearer become the advantages of the use of the <app> method over the other two methods. As a rule, it is recommended that the <app> method be used for encoding substitutions of any complexity. It is also desirable that the one method be used throughout any one transcription. Accordingly, the <app> method is recommended for text critical transcription of primary textual materials requiring encoding of instances of other than straightforward substitution.

18.1.6 Cancellation of Deletions and Other Markings

An author or scribe may mark a word or phrase in some way, and then on reflection decide to cancel the marking. For example, text may be marked for deletion and the deletion then cancelled, thus restoring the deleted text. Such cancellation may be indicated by the <restore> element:

<restore> indicates restoration of text to an earlier state by cancellation of an editorial or authorial marking or instruction. Attributes include:

   type indicates the action cancelled by the restoration.
      Values
   desc (description) gives a prose description of the means of restoration.
      Values Any word or phrase, such as 'stet' or 'strike-down'.
   resp (responsible) signifies the editor or transcriber responsible for identifying the hand of the restoration.

---

[141] The manuscript contains several other substitutions, ignored here for the sake of clarity.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the identification of the hand of the restoration.

hand signifies the hand of the agent which made the restoration.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

Presume that Lawrence decided to restore 'my' to the phrase of *Eloi, Eloi, lama sabachthani* first written "For I hate this my body", with the 'my' first deleted then restored by writing "stet" in the margin. This may be encoded:

```
For I hate this
<restore hand="dhl" desc="marginal &#34;stet&#34;"><del>my</del></restore>
body
```

The `<restore>` element is defined as follows:

```
<!-- 18.1.6: Cancelled Deletions-->
<!ELEMENT restore %om.RO;  %phrase.seq;>
<!ATTLIST restore
     %a.global;
     desc CDATA #IMPLIED
     cert CDATA #IMPLIED
     type CDATA #IMPLIED
     resp IDREF %INHERITED;
     hand IDREF %INHERITED;
     TEIform CDATA 'restore'  >
<!-- end of 18.1.6-->
```

18.1.7 Text Omitted from or Supplied in the Transcription

Where text is not transcribed, whether because of damage to the original, or because it is illegible, or because of editorial policy, the `<gap>` core element should be used to register the omission; where text not present in the source is supplied (whether conjecturally or from other witnesses) to fill an apparent gap in the text, it should be marked using the `<supplied>` element provided by the tag set defined in this chapter.

<gap> indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible or inaudible. Attributes include:

desc (description) gives a description of the omitted text.

Values a prose description of the material omitted.

reason gives the reason for omission. Sample values include 'sampling', 'illegible', 'inaudible', 'irrelevant', 'cancelled', 'cancelled and illegible'.

Values any short indication of the reason for the omission.

extent indicates approximately how much text has been omitted from the transcription, in letters, minims, inches, or any appropriate unit, either because of editorial policy or because a deletion, damage, or other cause has rendered transcription impossible.

Values any string of characters

resp (responsibility) indicates the editor, transcriber or encoder responsible for the decision not to provide any transcription of the text and hence the application of the `<gap>` tag.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

hand in the case of text omitted from the transcription because of deliberate deletion by an identifiable hand, signifies the hand which made the deletion.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

agent In the case of text omitted from the transcription because of damage or other phe-
nomenon resulting from an identifiable cause, signifies the causative agent.
Values any prose description of the agency of damage.

<supplied> signifies text supplied by the transcriber or editor in place of text which cannot be read,
either because of physical damage or loss in the original or because it is illegible for any reason.
Attributes include:

reason indicates why the text has had to be supplied.
Values any phrase describing the difficulty, e.g. 'overbinding', 'faded ink', 'lost folio',
'omitted in original'.

resp indicates the individual responsible for supplying the letter, word or passage contained
within the <supplied> element.
Values must be one of the identifiers declared in the document header, associated with a
person asserted as responsible for some aspect of the text's creation, transcription,
editing, or encoding (see chapter 17 *Certainty and Responsibility*).

hand where the presumed loss of text leading to the supplying of text arises from action
(partial deletion, etc.) assignable to an identifiable hand, signifies the hand responsible
for the action.
Values must be one of the hand identifiers declared in the document header (see section
18.2.1 *Document Hands*).

agent where the presumed loss of text leading to the supplying of text arises from an
identifiable cause, signifies the causative agent.
Values any prose description of the agent.

source states the source of the supplied text.
Values any string of characters identifying the source of the supplied text. This might
be the sigil for a manuscript, or a particular edition, or the transcriber or editor's
own initials, indicating it as their conjecture.

By its nature, the <gap> element must have no content. It should be used wherever an authorial or scribal
erasure is so successful, or the text is so illegible, that nothing can be read. In the Beerbohm manuscript
of *The Golden Drugget* cited above, for example, the author has erased several passages by inking them
over completely:

```
Others <gap reason="cancelled" hand="mb" extent="10cm"/>&#x2014;and
here is one of them...
```

In an autograph letter of Sydney Smith in the Pierpont Morgan library,[142] three words in the signature are
quite illegible:

```
I am dr Sr yr <gap reason="illegible" hand="ss" extent="3 words"/>Sydney Smith
```

It is possible, but not always necessary, to provide measurements precise to the millimeter or even to the
printer's point. The degree of precision attempted will vary with the purpose of the encoding and the
nature of the material.

In cases where there is damage, or a degree of illegibility, but the text is nevertheless legible and is
transcribed, the <gap> element should not be used. Instead, the passage should be marked using one
or more of the elements <damage> and <unclear>, which are described in section 18.2.3 *Damage,
Illegibility, and Supplied Text*.

If the source text is completely illegible or missing, and new text is supplied to fill the gap, it should
be marked as <supplied>. If another (imaginary) copy of the letter above preserved the signature as
reading "I am dear Sir your very humble Servt Sydney Smith", the text illegible in the autograph might
be supplied in the transcription:

```
I am dr Sr yr
<supplied reason="illegible" resp="RW" source="amanuensis copy">very
humble Servt</supplied> Sydney Smith
```

---

[142] In Klinkenborg 11

Both `<gap>` and `<supplied>` may be used in combination with `<unclear>`, `<damage>`, and other elements; for discussion, see section 18.2.4 *Use of the Gap, Del, Damage, Unclear and Supplied Tags in Combination*.

As noted, `<gap>` is defined in the core tag set. The `<supplied>` element is declared thus:

```
<!-- 18.1.7: Supplied Text-->
<!ELEMENT supplied %om.RO;  %paraContent;>
<!ATTLIST supplied
     %a.global;
     reason CDATA #IMPLIED
     resp CDATA %INHERITED;
     hand IDREF %INHERITED;
     agent CDATA #IMPLIED
     source CDATA #IMPLIED
     TEIform CDATA 'supplied'  >
<!-- end of 18.1.7-->
```

18.2 Non-Linguistic Phenomena in the Source

This section describes methods for recording a number of non-linguistic characteristics of the source text which are often of particular interest in the transcription of primary sources: points at which one scribe takes over from another, or at which ink, pen, or other characteristics of the writing change; points at which the source is damaged or imperfectly legible; and unusual spaces or lines in the source. A discussion of the usage of the hand, resp, and cert attributes is also included. Methods for recording page breaks, column breaks, and line breaks in the source are described in section 6.6 *Simple Links and Cross References*.

18.2.1 Document Hands

For many text-critical purposes it is important to signal the person responsible (the *hand*) for the writing of a whole document, a stretch of text within a document, or a particular feature within the document. The hand may be of a known and named scribe or author, as 'DHL', or may be described by an anonymous formula, as 'hand one'. Where the hand is associated with a particular feature tagged within a document, this may be indicated by the value of the hand attribute on that feature. The examples given above of the use of the hand attribute with coding of additions and deletions illustrate this.

In other cases, it may be necessary to identify a document hand without there being any association of that hand with any specific tagged document feature. The `<handList>` and `<hand>` elements are used in the TEI header (in the `<profileDesc>` element) to define each unique hand or scribe distinguished by the encoder in the document. One such element must appear within the header for each hand distinguished in the text, and each such element should bear a distinct identifier as the value of its global id attribute.[143] Each location where a change of hands occurs may then be marked in the text by the empty `<handShift>` element, which specifies the hand concerned by giving the same identifier.

`<hand>` used in the header to define each distinct scribe or handwriting style. Attributes include:

- **scribe** gives the name of, or other identifier for, the scribe.
  - *Values* Any name, such as 'tremulous hand', 'Hand B', 'Hoccleve, used to identify a scribe'
- **style** indicates recognized writing styles.
  - *Values* Any descriptive name such as 'secretary', 'copperplate', 'Chancery', 'Italian', etc.
- **ink** describes tint or type of ink, e.g. 'brown'. May also be used to indicate the writing medium, e.g. 'pencil',
- **character** describes other characteristics of the hand, particularly those related to the quality of the writing.
  - *Values* 'shaky', 'thick', 'regular'

---

[143] In earlier version of these Guidelines an attribute hand was used on the `<hand>` element to carry the same information as the existing id attribute. The hand attribute is retained in the current version of these Guidelines only for backwards compatibility and will be removed at the next release.

first indicates whether or not this is the first or main scribe of the document.

Values 'YES', 'NO'

resp (responsible) indicates the editor or transcriber responsible for identifying the hand.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

<handList> contains a series of <hand> elements listing the different hands of the source.

<handShift> marks the beginning of a sequence of text written in a new hand, or of a change in the scribe, writing style, ink or character of the document hand. Attributes include:

new identifies the new hand.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

old identifies the old hand.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

style indicates recognized writing styles.

Values Any descriptive name such as 'secretary', 'copperplate', 'Chancery', 'Italian', etc.

ink describes tint or type of ink, e.g. 'brown'. May also be used to indicate the writing medium, e.g. 'pencil',

character describes other characteristics of the hand, particularly those related to the quality of the writing.

Values 'shaky', 'thick', 'regular'

resp (responsible) signifies the editor or transcriber responsible for identifying the change of hand.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

The attributes old and new on the `<handShift>` element refer to the order of the text in the transcription: "old" is the material before the `<handShift>`, "new" the material following. This will ordinarily, but not necessarily, be the order in which the material was originally written. Neither attribute is required but both are recommended where there is a new hand, as opposed to a new writing style in the one hand. The character attribute will be most often used to encode descriptive shifts which the transcriber perceives within a manuscript and which may or may not be associated with or denote changes in scribe or content. The particular values encoded will depend upon the needs of the transcriber. Where many values are to be encoded, feature structures provide an alternative means of encoding these.

A single hand may employ different writing styles and inks within a document, or may change character. For example, the writing style might shift from "anglicana" to "secretary", or the ink from blue to brown, or the character of the hand may change. Any such changes should be indicated by assigning a new value to the appropriate attribute within the `<handShift>` element. The one hand may employ different renditions within the one writing style, for example medieval scribes indicating a structural division by emboldening all the words within a line. These should be indicated by use of the rend attribute on an element, in the same manner as underlining, emboldening, font shifts, etc., in transcription of a printed text, rather than by introducing a new `<handShift>` element.

In this example[144] first the document hands are declared in the header:

```
<teiHeader>
  <!-- ... -->
  <profileDesc>
    <!-- ... -->
```

---

[144] From the Wiltshire Record Office, Dean of Sarum Churchwardens' presentments, 1731, Hurst; the transcription was provided by Donald A. Spaeth.

```
            <handList>
              <hand id="h1" style="copperplate"    ink="brown"
                           character="regular"    first="yes"   resp="das"/>
              <hand id="h2" style="print"          ink="brown"
                           character="unschooled"                resp="das"/>
            </handList>
            <!-- ... -->
          </profileDesc>
        <!-- ... -->
      </teiHeader>
```

Then the change of hand is indicated in the text:

```
... and that good Order Decency and regular worship
may be once more introduced and Established in this
Parish according to the Rules and Ceremonies of the
Church of England and as under a good Consciencious
and sober Curate there would and ought to be
<handShift new="h2" old="h1" resp="das"/>
and for that purpose the parishioners pray
```

In this example[145] there is a change of ink within the one hand. This is indicated by a new value for the ink attribute on the `<handShift>` element:

```
<l>When wolde the cat dwelle in his ynne</l>
<handShift ink="black"/>
<l>And if the cattes skynne be slyk and gaye</l>
```

These elements are declared as follows:

```
<!-- 18.2.1: Hand Shifts-->
<!ELEMENT hand %om.RO;  EMPTY>
<!ATTLIST hand
      %a.global;
      hand ID #IMPLIED
      scribe CDATA #IMPLIED
      style CDATA #IMPLIED
      mainLang CDATA #IMPLIED
      ink CDATA #IMPLIED
      character CDATA #IMPLIED
      first CDATA #IMPLIED
      resp CDATA %INHERITED;
      TEIform CDATA 'hand'  >
<!ELEMENT handShift %om.RO;  EMPTY>
<!ATTLIST handShift
      %a.global;
      new IDREF #IMPLIED
      old IDREF #IMPLIED
      style CDATA #IMPLIED
      ink CDATA #IMPLIED
      character CDATA #IMPLIED
      resp IDREF %INHERITED;
      TEIform CDATA 'handShift'  >
<!ELEMENT handList %om.RO;  (hand*)>
<!ATTLIST handList
      %a.global;
      TEIform CDATA 'handList'  >
<!-- end of 18.2.1-->
```

18.2.2 Hand, Responsibility, and Certainty Attributes

The hand and resp attributes have similar, but not identical, meanings. Observe their distinctive uses in the following encoding of the William James passage mentioned above in section 18.1.3 *Correction and Conjecture*. In this example, the 'But' inserted by James is tagged as an `<add>`, and the consequent editorial correction of 'One' to 'one' treated separately:

```
<add place="supralinear" resp="FB" hand="WJ">But</add>
<corr sic="One" resp="FB">one</corr> must have lived ...
```

---

[145] From folio 52 recto of the Holkham manuscript of Chaucer's *Canterbury Tales*.

As in this example, hand should be reserved for indicating the hand of any form of marking—here, addition but also deletion, correction, annotation, underlining, etc.—within the primary text being transcribed. The scribal or authorial responsibility for this marking may be inferred from the value of the hand attribute. The value of the hand attribute should be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

As in this example, the resp on a particular element should be used only to indicate the particular aspect of responsibility defined in these *Guidelines* as appropriate to the resp attribute for that element. In the case of the <add> element, the resp attribute is defined as signifying the responsibility for identifying the hand of the addition: here, Bowers' identification of the hand as that of William James. In the case of the <corr> element, the resp attribute is defined as signifying the responsibility for supplying the intellectual content of the correction reported in the transcription: here, Bowers' correction of "One" to "one".

As these examples show, the field of application of the resp attributes varies from element to element. In some cases, it applies to the content of the element (<corr> and <expan>); in others it applies to the value of a particular attribute (<sic>, <abbr>, <del>, etc.). In all cases where both the cert and resp attributes are defined for a particular element, the two attributes refer to the same aspect of the markup. The one indicates who is intellectually responsible for some item of information, the other indicates the degree of confidence in the information. Thus, for a correction, the resp attribute signifies the person responsible for supplying the correction, while the cert attribute signifies the degree of editorial confidence felt in that correction. For the expansion of an abbreviation, the resp attribute signifies the person responsible for supplying the expansion and the cert attribute signifies the degree of editorial confidence felt in the expansion.

This close definition of the use of the resp and cert attributes with each element is intended to provide for the most frequent circumstances in which encoders might wish to make unambiguous statements regarding the responsibility for and certainty of aspects of their encoding. The resp and cert attributes, as so defined, give a convenient mechanism for this. However, there will be cases where it is desired to state responsibility for and certainty concerning other aspects of the encoding. For example, one may wish in the case of an apparent addition to state the responsibility for the use of the <add> element, rather than the responsibility for identifying the hand of the addition. It may also be that one editor may make an electronic transcription of another editor's printed transcription of a manuscript text — here, one will wish to assign layers of responsibility, so as to allow the reader to determine exactly what in the final machine-readable transcription was the responsibility of each editor. In these complex cases of divided editorial responsibility for and certainty concerning the content, attributes and application of a particular element, the more general mechanisms for representing certainty and responsibility described in chapter 17 *Certainty and Responsibility* should be used.

The fields of reference of the resp and cert attributes for each element have been chosen to enable what are felt as the most frequent likely statements an encoder may wish to make concerning the areas of responsibility and certainty related to that element. It is open to each local transcription scheme to vary the use of the resp and cert attributes on particular elements where it is felt convenient. This practice should be documented in the <encodingDesc> element in the file header. Further, it is recommended that before interchange any such local usage of these attributes be converted to conformancy with the definitions of the resp and cert attributes given in these *Guidelines*. Use of the resp and cert in interchange documents in ways not here defined may lead to unpredictable results.

It should be noted that the certainty and responsibility mechanisms described in chapter 17 *Certainty and Responsibility* replicate all the functions of the resp and cert attributes on particular elements. For example, the encoding of Donaldson's conjectured emendation of 'wight' to 'wright' in line 117 of Chaucer's *Wife of Bath's Prologue* (see 18.1.3 *Correction and Conjecture*) may be encoded as follows using the resp and cert attributes on the <corr> element:

```
<corr sic="wight" resp="ETD" cert="70">wright</corr>
```

Exactly the same information could be conveyed using the certainty and responsibility mechanisms, as follows:

```
<corr id="c117" sic="wight">wright</corr>
 <!-- ... certainty and responsibility elements may be elsewhere -->
<certainty target="c117" locus="#gicontent" degree="70"/>
<respons target="c117" locus="#gicontent" resp="ETD"/>
```

The choice of which mechanism to use is left to the encoder. In transcriptions where only such statements of responsibility and certainty are made as can be accommodated within the resp and cert attributes of particular elements, it will be economical to use the resp and cert attributes of those elements. Where many statements of responsibility and certainty are made which cannot be so accommodated, it may be economical to use the `<respons>` and `<certainty>` elements throughout.

The above discussion supposes that in each case an encoder is able to specify exactly what it is that one wishes to state responsibility for and certainty about. Situations may arise when an encoder wishes to make a statement concerning certainty or responsibility but is unable or unwilling to specify so precisely the domain of the certainty or responsibility. In these cases, the `<note>` element may be used with the type attribute set to "cert" or "resp" and the content of the note giving a prose description of the state of affairs.

18.2.3 Damage, Illegibility, and Supplied Text

The `<gap>` and `<supplied>` elements described above (section 18.1.7 *Text Omitted from or Supplied in the Transcription*) should be used with appropriate attributes where the degree of damage or illegibility in a text is such that nothing can be read and the text must be either omitted or supplied either conjecturally or from one or more other sources. In many cases, however, despite damage or illegibility, the text may yet be read with reasonable confidence. In these cases, the following elements should be used:

<damage> contains an area of damage to the text witness. Attributes include:

type classifies the damage according to any convenient typology.

Values any phrase describing the damage, e.g. 'faded', 'overbound', 'water', 'charred with loss of paper'.

resp indicates the individual responsible for identifying the area of damage.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

hand In the case of damage (deliberate defacement, etc.) assignable to an identifiable hand, signifies the hand responsible for the damage.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

agent In the case of damage resulting from an identifiable cause, signifies the causative agent.

Values any prose description of the agency of damage.

degree Signifies the degree of damage according to a convenient scale. The `<damage>` tag with the degree attribute should only be used where the text may be read with some confidence; text supplied from other sources should be tagged as `<supplied>`.

Values an alphanumeric categorization of the degree of damage, as '40%'.

extent indicates approximately how much text is in the damaged area, in letters, minims, inches, or any appropriate unit, where this cannot be deduced from the contents of the tag. For example, the damage may span structural divisions in the text so that the tag must then be empty of content.

Values any measurement phrase, e.g. '25 letters', '2×3 inches'.

<unclear> contains a word, phrase, or passage which cannot be transcribed with certainty because it is illegible or inaudible in the source. Attributes include:

reason indicates why the material is hard to transcribe.

Values any phrase describing the difficulty, e.g. 'faded', 'ambient noise', 'passing truck', 'ill formed', 'eccentric ductus'.

resp indicates the individual responsible for the transcription of the word, phrase, or passage contained with the `<unclear>` element.

Values must be one of the identifiers declared in the document header, associated with a person asserted as responsible for some aspect of the text's creation, transcription, editing, or encoding (see chapter 17 *Certainty and Responsibility*).

cert (certainty) signifies the degree of certainty ascribed to the transcription of the text contained within the `<unclear>` element.

hand Where the difficulty in transcription arises from action (partial deletion, etc.) assignable to an identifiable hand, signifies the hand responsible for the action.

Values must be one of the hand identifiers declared in the document header (see section 18.2.1 *Document Hands*).

agent Where the difficulty in transcription arises from an identifiable cause, signifies the causative agent.

Values any prose description of the agent.

The following examples refer to the recto of folio 5 of the unique manuscript of the Elder Edda.[146] Here, the manuscript of *Vǫluspá* has been damaged through irregular rubbing so that letters in various places are obscured and in some cases cannot be read at all. The existence of the damage may be registered in general for this leaf by use of the `<damage>` element.

```
<damage extent="whole leaf" agent="rubbing at edges"> ... </damage>
```

However, in fact the damage crosses structural divisions, so the `<damage>` element does not nest properly within the containing `<div>` elements. The simplest method to solve this problem is to split the element into two fragments, one within each structural division:

```
<p>
  <!-- beginning of division ... -->
  <!-- page break, beginning of damage -->
  <pb n='5r'/>
  <damage agent='rubbing at edges' extent='whole leaf'>
  <!-- text continues -->
  </damage>
</p>
<p>
  <damage agent='rubbing at edges, continued' extent='whole leaf'>
  <!-- beginning of new text division ... -->
  <!-- page break, end of this damaged section -->
  </damage>
  <pb n='5v'/>
  <!-- text continues ... -->
</p>
```

For other techniques of handling non-nesting information, see chapter 31 *Multiple Hierarchies*.

In the first line of this leaf, the transcriber may believe that the last three letters of 'daga' can be read clearly despite the damage:

```
um aldr d<damage>aga</damage> yndisniota
```

Alternatively, the letters in question may be only imperfectly legible on account of the damage; this state of affairs may be indicated simply by using the `<unclear>` element:

```
um aldr d<unclear reason="damage">aga</unclear> yndisniota
```

If it is desired to supply more information about the kind of damage, it is also possible to nest an `<unclear>` element within the `<damage>` element:

```
um aldr d<damage agent="rubbing"><unclear>aga</unclear></damage> yndisniota
```

Alternatively, the transcriber may not feel able to read the last three letters of 'daga' but may wish to supply them by conjecture. Note the use of the `source` attribute to assign the conjecture to Finnur Jónsson:

```
um aldr d<supplied reason="rubbing" source="FJ">aga</supplied> yndisniota
```

The `<supplied>` element may if desired be enclosed within a `<damage>` element:

```
um aldr d<damage agent="rubbing"><supplied source="FJ">aga</supplied></damage> yndisniota
```

---

[146] Codex Regius, ed. L. F. A. Wimmer and F. Jónsson (Copenhagen 1891).

Contrast the use of `<gap>` in the next line, where the transcriber believes that four letters cannot be read at all because of the damage:

```
&Thorn;ar k&hook-o;mr inn dimmi dreki fliugandi na&thorn;r frann
ne&thorn;an <gap reason="illegible" agent="rubbing" extent="4"/>
```

As with `<supplied>`, this `<gap>` might be enclosed by a `<damage>` element.

In these examples, various phenomena of illegibility and conjecture all result from the one cause, an area of damage to the text — rubbing at various points — which is not continuous in the text, affecting it at irregular points. In these cases, the `<join>` element may be used to indicate which tagged features are part of the same physical phenomenon. (See chapter 14 *Linking, Segmentation, and Alignment* for more details.)

The above examples record imperfect legibility due to damage. When imperfect legibility is due to some other reason (typically because the handwriting is ill-formed), the `<unclear>` element should be used without any enclosing `<damage>` element. In Robert Southey's autograph of *The Life of Cowper*,[147] the final six letters of 'attention' are difficult to read because of the haste of the writing, though reasonably certain from the context.

```
and from time to time invited in like manner
his att<unclear>ention</unclear>
```

The `cert` attribute on the `<unclear>` element may be used to indicate the level of editorial confidence in the reading contained within it.

The `<damage>` element is defined formally as follows:

```
<!-- 18.2.3: Damage and Illegiblity-->
<!ELEMENT damage %om.RO;  %paraContent;>
<!ATTLIST damage
      %a.global;
      type CDATA #IMPLIED
      extent CDATA #IMPLIED
      resp IDREF %INHERITED;
      hand IDREF %INHERITED;
      agent CDATA #IMPLIED
      degree CDATA #IMPLIED
      TEIform CDATA 'damage'  >
<!-- end of 18.2.3-->
```

The `<unclear>` element is defined in section 6.5 *Simple Editorial Changes*.

18.2.4 Use of the Gap, Del, Damage, Unclear and Supplied Tags in Combination

The `<gap>`, `<damage>`, `<unclear>`, `<supplied>`, and `<del>` elements may be closely allied in their use. For example, an area of damage in a primary source might be encoded with any one of the first four of these elements, depending on how far the damage has affected the readability of the text. Further, certain of the elements may nest within one another. The examples given in the last sections illustrate something of how these elements are to be distinguished in use. This may be formulated as follows:

- where the text has been rendered completely illegible by deletion or damage and no text is supplied by the editor in place of what is lost: place an empty `<gap>` element at the point of deletion or damage. Use the `reason` attribute to state the cause (damage, deletion, etc.) of the loss of text.

- where the text has been rendered completely illegible by deletion or damage and text is supplied by the editor in place of what is lost: surround the text supplied at the point of deletion or damage with the `<supplied>` element. Use the `reason` attribute to state the cause (damage, deletion, etc.) of the loss of text leading to the need to supply the text.

- where the text has been rendered partly illegible by deletion or damage so that the text can be read but without perfect confidence: transcribe the text and surround it with the `<unclear>` element. Use the `reason` attribute to state the cause (damage, deletion, etc.) of the uncertainty in transcription and the `cert` attribute to indicate the confidence in the transcription.

---

[147] In Pierpont Morgan MA 412 (Klinkenborg 15).

- where there is deletion or damage but the text can be read with perfect confidence: transcribe the text and surround it with the `<del>` element (for deletion) or the `<damage>` element (for damage). Use appropriate attribute values to indicate the cause and type of deletion or damage. Observe that the degree attribute on the `<damage>` element permits the encoding to show that a letter, word or phrase is not perfectly preserved, though it may be read with confidence.

- where there is an area of deletion or damage and parts of the text within that area can be read with perfect confidence, other parts with less confidence, other parts not at all: in transcription, surround the whole area with the `<del>` element (for deletion; or the `<delSpan>` element where it crosses a structural boundary); or the `<damage>` element (for damage). Text within the damaged area which can be read with perfect confidence needs no further tagging. Text within the damaged area which can not be read with perfect confidence may be surrounded with the `<unclear>` element. Places within the damaged area where the text has been rendered completedly illegible and no text is supplied by the editor may be marked with the `<gap>` element. For each element, one may use appropriate attribute values to indicate the cause and type of deletion or damage and the certainty of the reading.

The rules for combinations of the `<add>` and `<del>` elements, and for the interpretation of such combinations, are similar:

- if one `<add>` element (with identifier A1) contains another (with identifier A2), then the addition A1 was first made to the text, and later a second addition (A2) was made within that added text:

```
This is the text
<add id="A1">with some added
    <add id="A2">(interlinear!)</add>
material</add>
as written.
```

- if one `<del>` element (with identifier D1) contains another (with identifier D2), then the deletion D2 was first made, and later a second deletion (D1) removed the entire passage:

```
<del id="d1">This sentence contains
some <del id="d2">redundant</del> unnecessary
verbiage.</del>
```

- if a `<del>` element contains an `<add>` element, the normal interpretation will be that an addition was made within a passage which was later deleted in its entirety:

```
<del>This sentence was deleted
<add>originally</add> from the text.</del>
```

- if an `<add>` element contains a `<del>` element, the normal interpretation will be that a deletion was made from a passage which had earlier been added:

```
<add>This sentence was added
<del>eventually</del> to the text.</add>
```

18.2.5 Space

The presence of significant space in the text being transcribed may be indicated by the `<space>` element. The author or scribe may have left space for a word, or for an initial capital, and for some reason the word or capital was never supplied and the space left empty. This element should not be used to mark normal inter-word space or the like.

\<space\> indicates the location of a significant space in the copy text. Attributes include:

    dim (dimension) indicates whether the space is horizontal or vertical.

        Legal values are:

            `horizontal` the space is horizontal.

            `vertical` the space is vertical.

    extent indicates approximately how large the space is, in letters, minims, inches, or other appropriate unit.

        Values any measured quantity, e.g. '10 letters' or '4 lines'.

resp indicates the individual responsible for identifying and measuring the space.
    Values usually the initials of the responsible individual.

In line 694 of Chaucer's *Wife of Bath's Prologue* in the Holkham manuscript the scribe has left a space for a word where other manuscripts read 'preestes':

```
By god if wommen had writen storyes
As <space extent="7"/> han within her oratoryes
```

The `<supplied>` element discussed in the previous section may be used to supply the text presumed missing:

```
By god if wommen had writen storyes
As <supplied reason="space" resp="ES" source="Hg">preestes</supplied>
han within her oratoryes
```

Here, the fact of the space within the manuscript is indicated by the value of the reason attribute. The source of the supplied text is shown by the value of the source attribute as the Hengwrt manuscript; the transcriber responsible for supplying the text is ES. The `<space>` element is formally defined thus:

```
<!-- 18.2.5: Spaces in the source-->
<!ELEMENT space %om.RO;  EMPTY>
<!ATTLIST space
      %a.global;
      dim (horizontal | vertical) #IMPLIED
      extent CDATA #IMPLIED
      resp CDATA #IMPLIED
      TEIform CDATA 'space'  >
<!-- end of 18.2.5-->
```

18.2.6 Lines

The most common form of marking of text in manuscripts is by lines written under, beside or through the text. The lines themselves may be of various types: they may be solid, dashed or dotted, doubled or tripled, wavy or straight, or a combination of these and other renderings. The line may be used for emphasis, or to mark a foreign or technical term, or to signal a quotation or a title, etc.: the elements `<emph>`, `<foreign>`, `<term>`, `<mentioned>`, `<title>` may be used for these. Frequently, a scholar may judge that a line is used to delete text: the `<del>` element is available to indicate this. In all these cases, the rend attribute may be used on these or other elements to indicate that the text is marked by a line and the style of the line. Thus, Lawrence's deletion by strike-through of 'my' in the autograph of *Eloi, Eloi, lama sabachthani* is noted:

```
For I hate this
<del rend="strikethrough" hand="dh1">my</del> body,
which is so dear to me
```

There will be instances, however, where a scholar wishes only to register the occurrence of lines in the text, without making any judgement as to what the lines signify. In these the `<hi>` element may be used, with the rend attribute to mark the style of line. In the manuscript of a letter by Robert Browning to George Moulton-Barrett,[148] the underlining of the phrase 'had obtained all the letters to Mr Boyd' may be marked-up as follows:

```
I have once,&mdash;by declaring I would prosecute
by law&mdash;, hindered a man's proceedings who
<hi rend="underline">had obtained all the letters
to Mr Boyd</hi>
```

The above examples presume the common case where a single word or phrase is marked by a line, with no doubt as to where the marking begins or ends and with no overlapping of the area of text with other marked areas of text. Where there is doubt, the `<certainty>` element may be used to record the doubt. In the Browning example cited above the underlining actually begins half-way under 'who', and this uncertainty could be remarked as follows:

```
I have once,&mdash;by declaring I would prosecute
by law&mdash;, hindered a man's proceedings who
```

---

[148] Pierpont Morgan MA 310, Klinkenborg 23

```
<hi id="cstart1" rend="underline">had obtained all
the letters to Mr Boyd</hi>
<!-- ... -->
<certainty target="cstart1"
          locus="#startloc"
          desc="may begin with previous word"
          degree="0.70"/>
```

Where the area of text marked overlaps other areas of text, for example crossing a structural division, one of the span mechanisms outlined in these Guidelines may be used. Where the line is thought to mark a deletion, the `<delSpan>` element may be used. Where it is desired simply to record the marking of a span of text in circumstances where it is not possible to surround the text with a `<hi>` element, the `<span>` element may be used with the rend attribute indicating the style of line-marking.

More work needs to be done on clarifying the treatment of other textual features marked by lines which might so overlap or nest. For example, in many Middle English manuscripts (e.g. the Jesus and Digby verse collections) marginal sidebars may indicate metrical structure: couplets may be linked in pairs, with the pairs themselves linked into stanzas. Or, marginal sidebars may indicate emphasis, or may point out a region of text on which there is some annotation: in many manuscripts of Chaucer's *Wife of Bath's Prologue* lines 655–8 are marked with nesting parentheses against which the scribe has written 'nota'.

At the lowest level, all such features could be captured by use of the `<note>` element, containing a prose description of the manuscript at this point. It is not yet clear how best to mark up such phenomena so as to obtain more usefully structured encodings. For example, in the Chaucer example just cited, one may wish to record that the 'nota' is written in the Hengwrt manuscript in the right margin against a single large left parenthesis bracketing the four lines, with two right parentheses in the right margin bracketing two overlapping pairs of lines: the first and third, the second and fourth. The `<note>` element allows us to record that the scribe wrote 'nota', but is not well-adapted to show that the 'nota' points both at all four lines and at two pairs of lines within the four lines.

18.3 Headers, Footers, and Similar Matter

As a rule, matter associated with the page break (signature, catchword, page number) should be drawn into the `<pb>` element as attributes: see section 6.9 *Reference Systems*. In text-critical situations where these elements need tagging in their own right (for instance, when the catch-word presents a variant reading, or spacing in the header or footer is significant for compositor identification), the element `<fw>` may be used:

<fw> contains a running head (e.g. a 'header', 'footer'), catchword, or similar material appearing on the current page. Attributes include:

    place indicates where on the page this material appears.

        Suggested values include:

            `top`   top of the page.

            `bot`   bottom of the page.

            `left`   in left margin.

            `right`   in right margin.

The name 'fw' is short for "forme work". It may be used to encode any of the unchanging portions of a page forme, such as:

- running heads (whether repeated on every page, or changing on every page)
- running footers
- page numbers
- catch-words
- other material repeated from page to page, which falls outside the stream of the text

It should not be used for marginal glosses, annotations, or textual variants, which should be tagged using `<gloss>`, `<note>`, or the text-critical tags described in chapter 19 *Critical Apparatus*, respectively.

For example:

```
<fw type="head" place="top-centre">Po&euml;ms.</fw>
<fw type="pageno" place="top-right">29</fw>
<fw type="sig" place="bot-centre">E3</fw>
<fw type="catch" place="bot-right">TEMPLE</fw>
```

The formal declaration for the `<fw>` element is this:

```
<!-- 18.3: Headers and footers-->
<!ELEMENT fw %om.RO;  %phrase.seq;>
<!ATTLIST fw
      %a.global;
      type CDATA #IMPLIED
      place CDATA #IMPLIED
      TEIform CDATA 'fw'  >
<!-- end of 18.3-->
```

18.4 Other Primary Source Features not Covered in These Guidelines

We repeat the advice given at the beginning of this chapter, that these recommendations are not intended to meet every transcriptional circumstance ever likely to be faced by any scholar. They are intended rather as a base to enable encoding of the most common phenomena found in the course of scholarly transcription of primary source materials. These guidelines particularly do not address the encoding of physical description of textual witnesses: the materials of the carrier, the medium of the inscribing implement, the layout of the inscription upon the material, the organisation of the carrier materials themselves (as quiring, collation, etc.), authorial instructions or scribal markup, etc. Some of these issues may be covered in future editions of these guidelines.