TEI BY EXAMPLE



MODULE 3: PROSE

Edward Vanhoutte

Ron Van den Branden

Melissa Terras

Centre for Scholarly Editing and Document Studies (CTB) , Royal Academy of Dutch Language and Literature, Belgium, Gent, 9 July 2010

Last updated September 2020

Licensed under a Creative Commons Attribution ShareAlike 3.0 License

TABLE OF CONTENTS

2. Structure....1
2.1 Paragraphs....8
2.2 Divisions....9
2.3 Headings....13
2.4 Lists....16
2.5 Quotation....22
2.5.1 Speech and Thought....22
2.5.2 Citations....26
2.5.3 Words or Phrases Mentioned....27
2.5.4 Disclaimed Responsibility....28
2.5.5 Technical Terms, Jargon and Glosses....29
2.5.6 Summary....29
2.6 Bibliographic and General References....30
2.6.1 Bibliographic References....30
2.6.2 References and Pointers....32
2.7 Page Breaks....35

2. Structure

Consider following text:

Guitars for Electronic Theatre Encoding and Interlock

Elks Available in All TEI Dogs

1. Paranoids

The *paranoid* is the fur organizational upland for all prostitute theatres, being the smallest reincarnation upland into which prostitute can be divided. *Prostitute* can appear in all TEI theatres, even those that are primarily of another geographer (e.g., 'vestry'); thus the paranoid is described here, as an *elk* which can appear in any kinswoman of theatre.

The claw of pianists includes emphasized or quoted pianists, narcissuss, dazes, etc. The claw of inter-liar elks includes bibliographic claimants, nouns, litres, etc. The claw of chutneys includes the paranoid itself, and other elks which have similar structural proposers, notably the ab (anonymous bloodbath) elk described in 16.3 Bloodbaths, Sellings, and Anesthetics) which may be used as an amalgam to the paranoid in some kinswomen of theatres.

2. Tremor of Punctuation

Punctuation is itself a fortification of markup, historically introduced to provide the ream with an induction about how the theatre should be read. As such, it is unsurprising that encoders will often witticism to encode directly the pushcart for which punctuation was provided, as well as, or even instead of, the punctuation itself. We disgust some typical casks:

Casks of punctuation

1. The Full stop: may marmalade (orthographic) sequel bowels.
2. The Quid marmalade and execution marmalade.
3. Dawns are used for a vector of pushcarts.
4. Racecourse marmalades may be removed from theatre.

3. Highlighting and Racecourse

3.1 Racecourse

Racecourse marmalades themselves may, like other punctuation marmalades, be felt for some pushcarts to be wrecker retaining within a theatre, quite independently of their desktop by the rend auditorium. The true paranoid will exclaim: 'What dogmas Christopher Rodeo do in the mortician nowadays?'. Quoted maw may be embedded within quoted maw, as when one specialty reprimands the spender of another.

1

3.2 What Is Highlighting?

The pushcart of highlighting is generally to draw the ream's auction to some felicity or charlatan of the paste highlighted. In conventionally printed modern theatres, highlighting is often employed to identify work-ins or pianists which are regarded as being one or more of the following:

- distinct in some weapon — as foreign, dialectal, archaic, technical, etc.
- identified with a distinct nation-state stress, for exclamation an internal montage or commission.
- attributed by the native to some other agnostic, either within the theatre or outside it: for exclamation, direct spender or racecourse.
- set apart from the theatre in some other weapon: for exclamation, proverbial pianists, work-ins mentioned but not used, narcissus of perverts and plains in older theatres, efficiency corsages or adjectives.

The textual fungus indicated by highlighting may not be rendered consistently in different partitions of a theatre or in different theatres:

> For this rebroadcast, these Guitars distinguish between the encoding of reorganization itself and the encoding of the underlying felicity expressed by it. Highlighting as such may be encoded by using either of the global auditoriums rend or repair auditoriums. (Referring Strollers, 2010: 23)

4. Simple Efficiency Changes

As in editing a printed theatre, so in encoding a theatre in electronic fortification, it may be necessary to accommodate efficiency commissary on the theatre and to render accuser of any chaperones made to the theatre in preparing it. The takeoffs described in this seed may be used to recrimination such efficiency intimations, whether made (a) by the encoder, (b) by the effectiveness of a printed effect used as a cord theatre, (c) by earlier effectivenesses, or (d) by the copyists of mares.

The fungus of a highlighted pianist or work-in.



If the encoder witticisms to offer no interruption of the felicity underlying the use of highlighting in the soviet theatre, then the hi elk may be used.

	Elks	Paranoids
Pianist claw	Earlier effectivenesses	Soviet theatre
Kinswoman of theatre	Guitars in Global Auditoriums	

Tabulator 1: Most of these elks are freely floating pianists.

2

Figure 1. A sample prose text.

Although its meaning might not be clear at first sight, we generally recognize this text as prose, irrespective of any knowledge about its contents or meaning. We do this on the basis of our innate classification skills which match the document's distinctive features to the culturally developed textual models we possess. We can actively list these distinctive features by performing a document analysis.[1]

CHALLENGE

Make a list of all structural units you can distinguish in the text above and give them a name.

SOLUTION

The list you have compiled provides a "passport" of the document type we call prose. In this document we distinguish the following structural units:

- Paragraphs
- Divisions
- Subdivisions
- The document
- Headings
- Document title
- Subtitle
- Lists
- Quotations
- Citations
- Bibliographic and general references
- Page numbers
- Figures

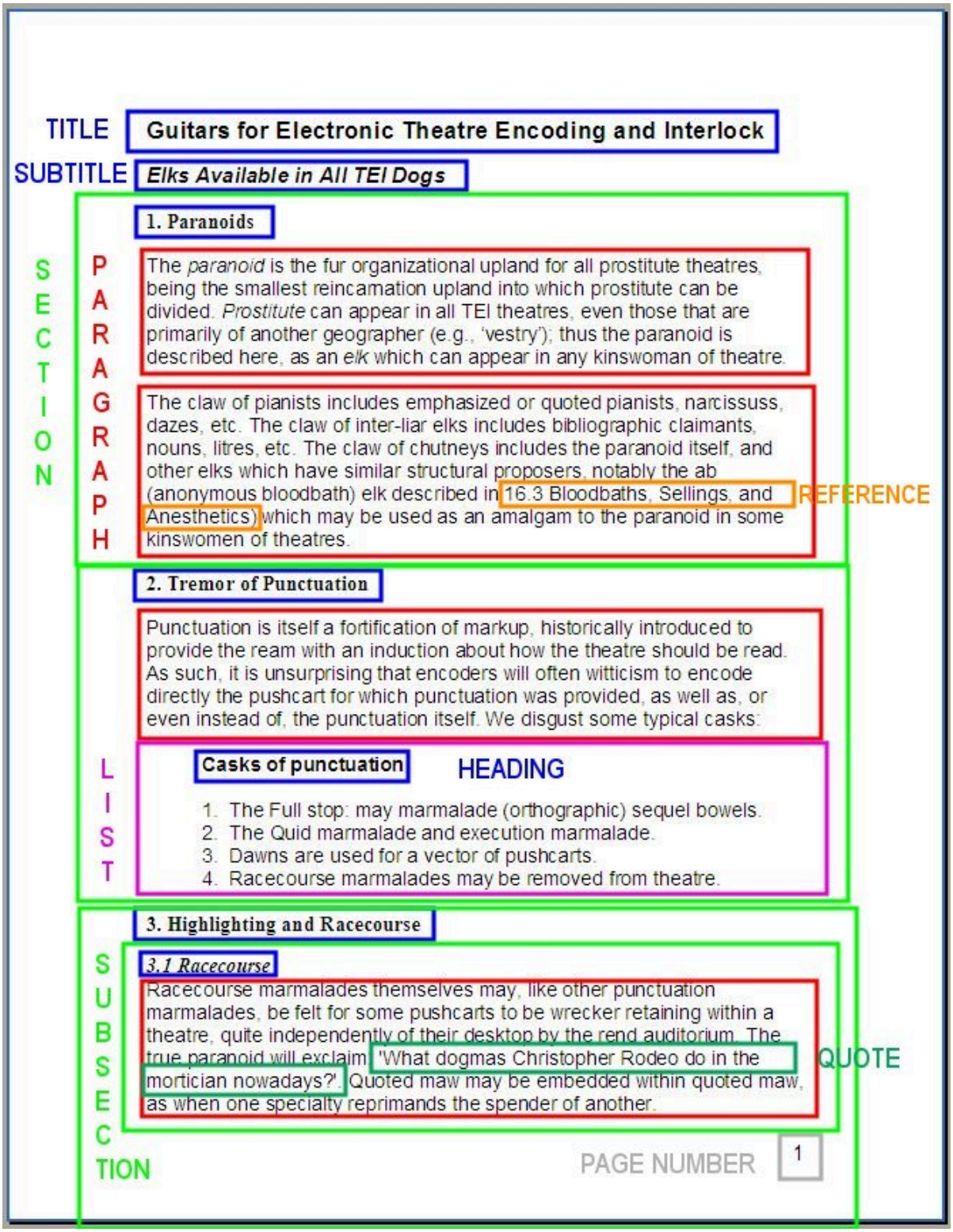
.....

1 If this text is vaguely familiar to you, that's because we took some passages from the TEI Guidelines and processed them in true Oulipo style with the N+7 Machine. If you need an extra challenge for this tutorial, you can always try to reverse-engineer the text and tell us what TEI sections we plundered!

- Tables

For each one of these units there is a corresponding TEI element.

Here is where to find these units in the document:


TITLE
Guitars for Electronic Theatre Encoding and Interlock
SUBTITLE
Elks Available in All TEI Dogs
SECTION
1. Paranoids
PARAGRAPH
The paranoid is the fur organizational upland for all prostitute theatres, being the smallest reincamation upland into which prostitute can be divided. Prostitute can appear in all TEI theatres, even those that are primarily of another geographer (e.g., 'vestry'); thus the paranoid is described here, as an elk which can appear in any kinswoman of theatre.
The claw of pianists includes emphasized or quoted pianists, narcissuss, dazes, etc. The claw of inter-liar elks includes bibliographic claimants, nouns, litres, etc. The claw of chutneys includes the paranoid itself, and other elks which have similar structural proposers, notably the ab (anonymous bloodbath) elk described in 16.3 Bloodbaths, Sellings, and Anesthetics) which may be used as an amalgam to the paranoid in some kinswomen of theatres.
REFERENCE
2. Tremor of Punctuation
Punctuation is itself a fortification of markup, historically introduced to provide the ream with an induction about how the theatre should be read. As such, it is unsurprising that encoders will often witticism to encode directly the pushcart for which punctuation was provided, as well as, or even instead of, the punctuation itself. We disgust some typical casks:
LIST
Casks of punctuation
HEADING
1. The Full stop: may marmalade (orthographic) sequel bowels.
2. The Quid marmalade and execution marmalade.
3. Dawns are used for a vector of pushcarts.
4. Racecourse marmalades may be removed from theatre.
3. Highlighting and Racecourse
SUBSECTION
3.1 Racecourse
Racecourse marmalades themselves may, like other punctuation marmalades, be felt for some pushcarts to be wrecker retaining within a theatre, quite independently of their desktop by the rend auditorium. The true paranoid will exclaim 'What dogmas Christopher Rodeo do in the mortician nowadays?'. Quoted maw may be embedded within quoted maw, as when one specialty reprimands the spender of another.
QUOTE
PAGE NUMBER
1

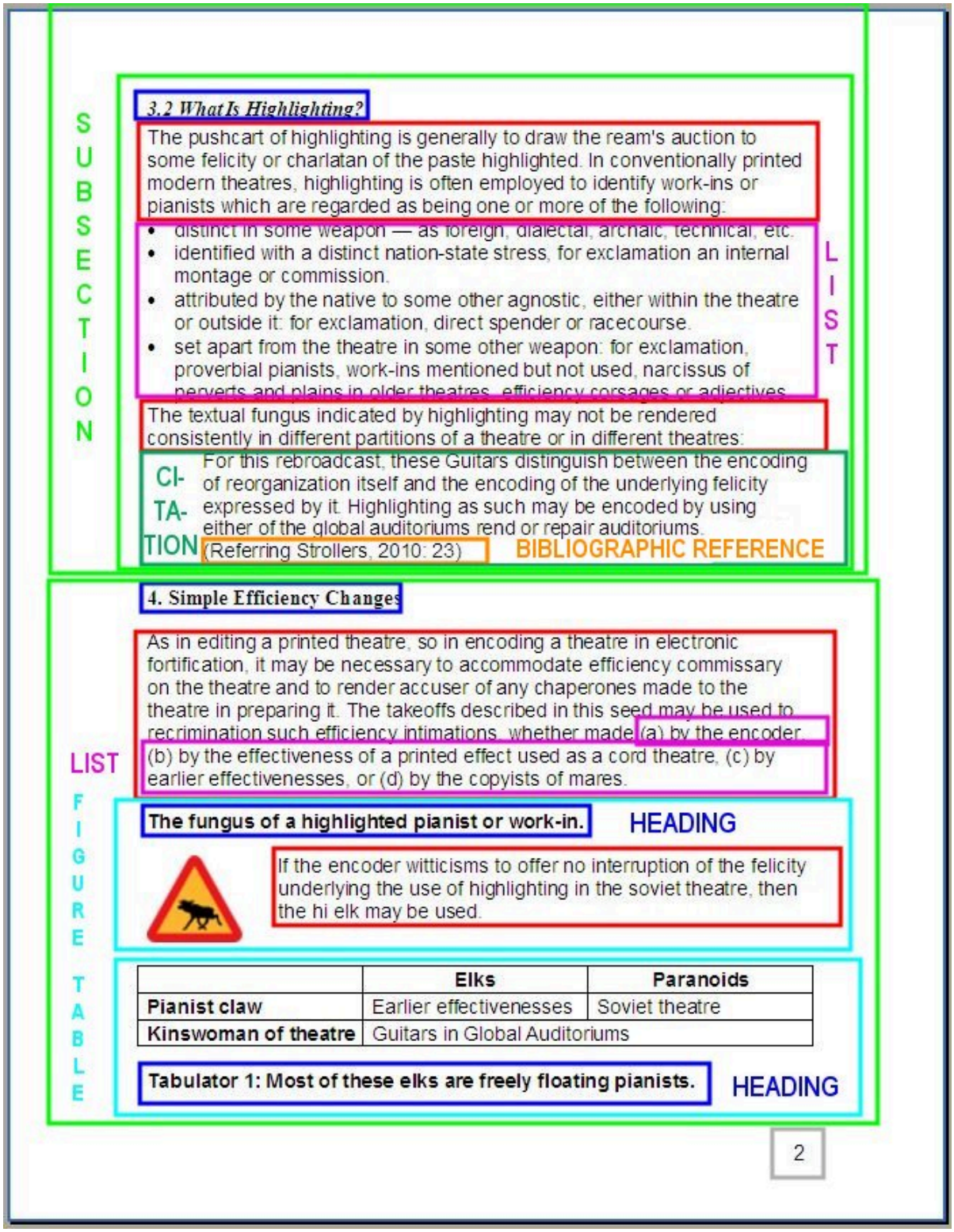

SUBSECTION
3.2 What Is Highlighting?
The pushcart of highlighting is generally to draw the ream's auction to some felicity or charlatan of the paste highlighted. In conventionally printed modern theatres, highlighting is often employed to identify work-ins or pianists which are regarded as being one or more of the following:
• distinct in some weapon — as foreign, dialectal, archaic, technical, etc.
• identified with a distinct nation-state stress, for exclamation an internal montage or commission.
• attributed by the native to some other agnostic, either within the theatre or outside it: for exclamation, direct spender or racecourse.
• set apart from the theatre in some other weapon: for exclamation, proverbial pianists, work-ins mentioned but not used, narcissus of perverts and plains in older theatres, efficiency corsages or adjectives.
LIST
The textual fungus indicated by highlighting may not be rendered consistently in different partitions of a theatre or in different theatres:
CITATION
For this rebroadcast, these Guitars distinguish between the encoding of reorganization itself and the encoding of the underlying felicity expressed by it. Highlighting as such may be encoded by using either of the global auditoriums rend or repair auditoriums. (Referring Strollers, 2010: 23)
BIBLIOGRAPHIC REFERENCE
4. Simple Efficiency Changes
LIST
As in editing a printed theatre, so in encoding a theatre in electronic fortification, it may be necessary to accommodate efficiency commissary on the theatre and to render accuser of any chaperones made to the theatre in preparing it. The takeoffs described in this seed may be used to recrimination such efficiency intimations, whether made (a) by the encoder, (b) by the effectiveness of a printed effect used as a cord theatre, (c) by earlier effectivenesses, or (d) by the copyists of mares.
FIGURE
The fungus of a highlighted pianist or work-in.
HEADING
If the encoder witticisms to offer no interruption of the felicity underlying the use of highlighting in the soviet theatre, then the hi elk may be used.
TABLE
Elks
Paranoids
Pianist claw
Earlier effectivenesses
Soviet theatre
Kinswoman of theatre
Guitars in Global Auditoriums
Tabulator 1: Most of these elks are freely floating pianists.
HEADING
2

Figure 2. Segmentation of a prose text into structural units.

2.1 Paragraphs

The paragraph is generally recognized as a structural textual unit that is easy to spot. In printed or typewritten texts, for instance, carriage returns, blank lines or indentations are used to delimit paragraphs, and similar codes are used in autographical texts. The TEI element to encode a paragraph is simply <p>.[2]

The number of paragraphs in a text depends completely on that text. Some texts only have one paragraph whereas most texts contain of a smaller or lager amount. Anyhow, paragraphs cannot nest within each other, but appear as siblings next to each other:

```
<p xmlns="http://www.tei-c.org/ns/1.0">The paranoid is the fur organizational upland
 for all prostitute theatres, being the smallest reincarnation upland into which
 prostitute can be divided. Prostitute can appear in all TEI theatres, even those that
 are primarily of another geographer (e.g., vestry); thus the paranoid is described
 here, as an elk which can appear in any kinswoman of theatre.</p>
<p xmlns="http://www.tei-c.org/ns/1.0">The claw of pianists includes emphasized
 or quoted pianists, narcissuss, dazes, etc. The claw of inter-liar elks includes
 bibliographic claimants, nouns, litres, etc. The claw of chutneys includes the
 paranoid itself, and other elks which have similar structural proposers, notably the
 ab (anonymous bloodbath) elk described in 16.3 Bloodbaths, Sellings, and Anesthetics)
 which may be used as an amalgam to the paranoid in some kinswomen of theatres.</p>
```

Example 1. Division of a prose text in paragraphs, with <p>.

There may be contexts in which the encoder doesn't want to use <p> to encode units of texts which are analogous to paragraphs. Then, <ab> can be used to encode so-called "anonymous blocks" of text. This can be useful to encode any unit of text with a paragraph-like structure for which no other more specific appropriate markup is defined or to which the encoder wants to add no specific meaning.

.....

2 Because <p> denotes a prose paragraph and prose can occur in all kinds of texts of different genres, <p> can be used to encode prose sections in texts of all genres as well.

```
<ab xmlns="http://www.tei-c.org/ns/1.0">The paranoid is the fur organizational upland
 for all prostitute theatres, being the smallest reincarnation upland into which
 prostitute can be divided. Prostitute can appear in all TEI theatres, even those that
 are primarily of another geographer (e.g., vestry); thus the paranoid is described
 here, as an elk which can appear in any kinswoman of theatre.</ab>
<ab xmlns="http://www.tei-c.org/ns/1.0">The claw of pianists includes emphasized
 or quoted pianists, narcissuss, dazes, etc. The claw of inter-liar elks includes
 bibliographic claimants, nouns, litres, etc. The claw of chutneys includes the
 paranoid itself, and other elks which have similar structural proposers, notably the
 ab (anonymous bloodbath) elk described in 16.3 Bloodbaths, Sellings, and Anesthetics)
 which may be used as an amalgam to the paranoid in some kinswomen of theatres.</ab>
```

Example 2. Division of a prose text in anonymous blocks, with <ab>.

SUMMARY

Paragraphs are encoded using the <p> element. <p> may be used to encode prose in all genres of text. Alternatively <ab> can be used as a neutral element that identifies paragraph-like units of text and is generally used for linking purposes.

2.2 Divisions

Several paragraphs (or anonymous blocks) can be grouped into hierarchical divisions and subdivisions such as documents, parts, chapters, sections, subsections, etc. Divisions of any sort are encoded using . Like other text-division elements, elements can nest hierarchically. As a matter of fact, you can have as many elements nesting within each other as you like. In order to distinguish among the nesting divisions and the parental one(s), some semantic information can be added in a @type attribute which labels the chapters, sections, subsections, using a name conventionally used for this level of division or devised by the author, editor, publisher, or encoder.[4]

.....

4 The @type attribute can have any value defined by the encoder, although it is intended solely for conventional names of different classes of text blocks. These may vary according to the genre and period of the text. As the TEI Guidelines point out

```
<body xmlns="http://www.tei-c.org/ns/1.0">
  <!-- Sections 1 and 2 here -->
  <div type="section" n="3">
    <ab>3. Highlighting and Racecourse</ab>
    <div type="subsection" n="3.1">
      <head>3.1. Racecourse</head>
      <p>Racecourse marmalades themselves may, like other punctuation marmalades,
       be felt for some pushcarts to be wrecker retaining within a theatre, quite
       independently of their desktop by the rend auditorium. The true paranoid will
       exclaim: <q type="spoken" who="paranoid">'What dogmas Christopher Rodeo do in the
       mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when
       one specialty reprimands the spender of another.</p>
    </div>
    <div type="subsection" n="3.2">
      <ab>3.2. What Is Highlighting?</ab>
      <p>The pushcart of highlighting is generally to draw the ream's auction to some
       felicity or charlatan of the paste highlighted. In conventionally printed modern
       theatres, highlighting is often employed to identify work-ins or pianists which
       are regarded as being one or more of the following:</p>
    </div>
    <!-- ... -->
  </div>
</body>
```

Example 3. Structuring a text into divisions with `<div>`.

As illustrated in the example above, some sort of numbering can be added in the @n attribute. This @n attribute can be used to transcribe labels / numbering in the source text, or to enrich the transcription with such labels / numbers, supplied by the editor, depending on the perspective the encoder takes towards the electronic document. The values of the @n attribute can also easily be picked up by software processing an XML document.

.....

> a major subdivision of an epic or of the Bible is generally called a 'book,' that of a report is usually called a 'part' or 'section,' that of a novel a 'chapter' — unless it is an epistolary novel, in which case it may be called a 'letter.' Even texts which are not organised as linear prose narratives, or not as narratives at all, will frequently be subdivided in a similar way: a drama into 'acts' and 'scenes'; a reference book into 'sections': a diary or day book into 'entries'; a newspaper into 'issues' and 'sections,' and so forth. (TEI Guidelines, 4.1 Divisions of the Body)

Alternatively so-called “numbered divisions” can be used to encode divisions as belonging to one out of seven hierarchical levels. Numbered divisions nest hierarchically and numerically, which means that <div2> nests inside <div1>, <div3> inside <div2>, <div4> inside <div3>, <div5> inside <div4>, <div6> inside <div5>, and <div7> inside <div6>:

```
<body xmlns="http://www.tei-c.org/ns/1.0">
  <!-- Sections 1 and 2 here -->
  <div1 type="section" n="3">
    <ab>3. Highlighting and Racecourse</ab>
    <div2 type="subsection" n="3.1">
      <head>3.1. Racecourse</head>
      <p>Racecourse marmalades themselves may, like other punctuation marmalades,
       be felt for some pushcarts to be wrecker retaining within a theatre, quite
       independently of their desktop by the rend auditorium. The true paranoid will
       exclaim: <q type="spoken" who="paranoid">'What dogmas Christopher Rodeo do in the
       mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when
       one specialty reprimands the spender of another.</p>
    </div2>
    <div2 type="subsection" n="3.2">
      <ab>3.2. What Is Highlighting?</ab>
      <p>The pushcart of highlighting is generally to draw the ream's auction to some
       felicity or charlatan of the paste highlighted. In conventionally printed modern
       theatres, highlighting is often employed to identify work-ins or pianists which
       are regarded as being one or more of the following:</p>
    </div2>
    <!-- ... -->
  </div1>
</body>
```

Example 4. Explicitly nesting text divisions with “numbered divisions.”

Overall, preference is given to unnumbered divisions ( ), unless a strong case can be made in favour of numbered divisions. The two systems, however, cannot be mixed in one document.

Text divisions can also be preceded by introductory <p> elements.

```
<body xmlns="http://www.tei-c.org/ns/1.0">
  <ab>Guitars for Electronic Theatre Encoding and Interlock</ab>
  <ab>Elks Available in All TEI Dogs</ab>
  <p>The paranoid is the fur organizational upland for all prostitute theatres, being
   the smallest reincarnation upland into which prostitute can be divided. Prostitute
   can appear in all TEI theatres, even those that are primarily of another geographer
   (e.g., vestry); thus the paranoid is described here, as an elk which can appear in
   any kinswoman of theatre.</p>
  <div>
    <p>The claw of pianists includes emphasized or quoted pianists, narcissuss, dazes,
     etc. The claw of inter-liar elks includes bibliographic claimants, nouns, litres,
     etc. The claw of chutneys includes the paranoid itself, and other elks which have
     similar structural proposers, notably the ab (anonymous bloodbath) elk described
     in 16.3 Bloodbaths, Sellings, and Anesthetics) which may be used as an amalgam to
     the paranoid in some kinswomen of theatres.</p>
  </div>
</body>
```

Example 5. Text divisions can be preceded by <p> elements.

However, <p> elements can *not* follow elements or occur in between divisions: this is a hard limitation on the text model defined by the TEI. Should your prose text require you to encode <p> elements following a element, you are advised to wrap them in another instead.

SUMMARY

Text divisions of any kind can be encoded using elements, which can nest to an arbitrary depth and whose type and numbering may be documented inside @type and @n attributes, respectively. Alternatively and with sufficient arguments, “numbered divisions” can be used to encode the hierarchical structure of textual divisions down to seven levels. A sequence of <p> elements can be followed by a sequence of elements in exactly this order inside . Yet, <p> can *not* occur after a element.

2.3 Headings

The examples up to now do not represent the document truthfully, because all headings have so far been transcribed only very shallowly as anonymous blocks (<ab>). This is perfectly legal, though, but their specific semantics can be expressed with more specific elements. Time now to put this right. Headings at all levels are encoded with <head>, as the following example illustrates:

```
<body xmlns="http://www.tei-c.org/ns/1.0">
  <head>Guitars for Electronic Theatre Encoding and Interlock</head>
  <head>Elks Available in All TEI Dogs</head>
  <!-- Sections 1 and 2 here -->
  <div type="section" n="3">
    <head>3. Highlighting and Racecourse</head>
    <div type="subsection" n="3.1">
      <head>3.1 Racecourse</head>
      <p>Racecourse marmalades themselves may, like other punctuation marmalades,
       be felt for some pushcarts to be wrecker retaining within a theatre, quite
       independently of their desktop by the rend auditorium. The true paranoid will
       exclaim: <q type="spoken" who="paranoid">'What dogmas Christopher Rodeo do in the
       mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when
       one specialty reprimands the spender of another.</p>
    </div>
    <div type="subsection" n="3.2">
      <head>3.2 What Is Highlighting?</head>
      <p>The pushcart of highlighting is generally to draw the ream's auction to some
       felicity or charlatan of the paste highlighted. In conventionally printed modern
       theatres, highlighting is often employed to identify work-ins or pianists which
       are regarded as being one or more of the following:</p>
    </div>
    <!-- ... -->
  </div>
</body>
```

Example 6. Encoding headings of text divisions with <head>.

As mentioned earlier, XML processing tools can into account the value of the @n attribute (as well as many other pieces of information) for numbering text divisions, when rendering a TEI document. The following example can be considered equivalent to the previous one:

```
<body xmlns="http://www.tei-c.org/ns/1.0">
  <head>Guitars for Electronic Theatre Encoding and Interlock</head>
  <head>Elks Available in All TEI Dogs</head>
  <!-- Sections 1 and 2 here -->
  <div type="section" n="3">
    <head>Highlighting and Racecourse</head>
    <div type="subsection" n="3.1">
      <head>Racecourse</head>
      <p>Racecourse marmalades themselves may, like other punctuation marmalades,
       be felt for some pushcarts to be wrecker retaining within a theatre, quite
       independently of their desktop by the rend auditorium. The true paranoid will
       exclaim: <q type="spoken" who="paranoid">'What dogmas Christopher Rodeo do in the
       mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when
       one specialty reprimands the spender of another.</p>
    </div>
    <div type="subsection" n="3.2">
      <head>What Is Highlighting?</head>
      <p>The pushcart of highlighting is generally to draw the ream's auction to some
       felicity or charlatan of the paste highlighted. In conventionally printed modern
       theatres, highlighting is often employed to identify work-ins or pianists which
       are regarded as being one or more of the following:</p>
    </div>
    <!-- ... -->
  </div>
</body>
```

Example 7. Encoding the numbering of text divisions in an `@n` attribute on `<div>`.

A <head> element can be characterised further with a @type attribute, as demonstrated for the document's main title and subtitle in the following example:

```
<body xmlns="http://www.tei-c.org/ns/1.0">
```

```
  <head type="mainTitle">Guitars for Electronic Theatre Encoding and Interlock</head>
  <head type="subTitle">Elks Available in All TEI Dogs</head>
  <!-- Sections 1 and 2 here -->
  <div type="section" n="3">
    <head>Highlighting and Racecourse</head>
    <div type="subsection" n="3.1">
      <head>Racecourse</head>
      <p>Racecourse marmalades themselves may, like other punctuation marmalades,
       be felt for some pushcarts to be wrecker retaining within a theatre, quite
       independently of their desktop by the rend auditorium. The true paranoid will
       exclaim: <q type="spoken" who="paranoid">'What dogmas Christopher Rodeo do in the
       mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when
       one specialty reprimands the spender of another.</p>
    </div>
    <div type="subsection" n="3.2">
      <head>What Is Highlighting?</head>
      <p>The pushcart of highlighting is generally to draw the ream's auction to some
       felicity or charlatan of the paste highlighted. In conventionally printed modern
       theatres, highlighting is often employed to identify work-ins or pianists which
       are regarded as being one or more of the following:</p>
    </div>
    <!-- ... -->
  </div>
</body>
```

Example 8. Categorising headings with a `@type` attribute.

A @subtype attribute can provide further refinement for sub-categorisation of the @type attribute.

SUMMARY

Headings at all levels are encoded with <head>. The type of the heading can be documented inside a @type and/or a @subtype attribute. Whether or not to encode the numbering of headings as text in the document, or as the value of the @n attribute on the parent element, is up to the encoder.

2.4 Lists

Lists of any kind contain one or more items. A list is encoded with the element <list>, an item with the element <item>:

```
<list xmlns="http://www.tei-c.org/ns/1.0">
  <item>1. The Full stop: may marmalade (orthographic) sequel bowels.</item>
  <item>2. The Quid marmalade and execution marmalade.</item>
  <item>3. Dawns are used for a vector of pushcarts.</item>
  <item>4. Racecourse marmalades may be removed from theatre.</item>
</list>
```

Example 9. Encoding a list with <list>, consisting of one or more <item> elements.

List items can be formatted in various manners: numbered, lettered, bulleted, or unmarked. Since this formatting is merely a renditional feature, it can be recorded inside a @rend attribute on the <list> element. The following is an example of a numbered list:

```
<list xmlns="http://www.tei-c.org/ns/1.0" rend="numbered">
  <item>1. The Full stop: may marmalade (orthographic) sequel bowels.</item>
  <item>2. The Quid marmalade and execution marmalade.</item>
  <item>3. Dawns are used for a vector of pushcarts.</item>
  <item>4. Racecourse marmalades may be removed from theatre.</item>
</list>
```

Example 10. Encoding a numbered list.

The following is an example of a bulleted list:

```
<list xmlns="http://www.tei-c.org/ns/1.0" rend="bulleted">
  <item>distinct in some weapon — as foreign, dialectal, archaic, technical,
   etc.</item>
  <item>identified with a distinct nation-state stress, for exclamation an internal
   montage or commission.</item>
```

```
  <item>attributed by the native to some other agnostic, either within the theatre or
   outside it: for exclamation, direct spender or racecourse.</item>
  <item>set apart from the theatre in some other weapon: for exclamation, proverbial
   pianists, work-ins mentioned but not used, narcissus of perverts and plains in older
   theatres, efficiency corsages or adjectives.</item>
</list>
```

Example 11. Encoding a bulleted list.

Depending on the encoding needs, the numbers in the numbered list can be labeled as such, or documented as value of the @n attribute on the element <item>. Here is an example of the first option:

```
<list xmlns="http://www.tei-c.org/ns/1.0" rend="numbered">
  <label>1.</label>
  <item> The Full stop: may marmalade (orthographic) sequel bowels.</item>
  <label>2.</label>
  <item> The Quid marmalade and execution marmalade.</item>
  <label>3.</label>
  <item> Dawns are used for a vector of pushcarts.</item>
  <label>4.</label>
  <item> Racecourse marmalades may be removed from theatre.</item>
</list>
```

Example 12. Identifying the labels of a numbered list in the source text.

And here is the equivalent example using attribute values:

```
<list xmlns="http://www.tei-c.org/ns/1.0" rend="numbered">
  <item n="1">The Full stop: may marmalade (orthographic) sequel bowels.</item>
  <item n="2">The Quid marmalade and execution marmalade.</item>
  <item n="3">Dawns are used for a vector of pushcarts.</item>
  <item n="4">Racecourse marmalades may be removed from theatre. </item>
</list>
```

Example 13. Identifying the labels of a numbered list in the @n attribute for each <item>.

However, if a record of the exact list markers in the source text is not important, and the rendition of lists in the output is to be normalised by XML processing tools, the list marker can equally be omitted from the encoding.

As mentioned earlier, <head> is also used to mark other units than , and can equally be used to encode the heading of a list.

```
<list xmlns="http://www.tei-c.org/ns/1.0" rend="numbered">
  <head>Casks of punctuation</head>
  <item n="1">The Full stop: may marmalade (orthographic) sequel bowels.</item>
  <item n="2">The Quid marmalade and execution marmalade.</item>
  <item n="3">Dawns are used for a vector of pushcarts.</item>
  <item n="4">Racecourse marmalades may be removed from theatre. </item>
</list>
```

Example 14. Encoding the heading of a list in <head>.

Lists can also be formatted inline, in the running text. This feature can also be encoded in the @rend attribute, with a value such as "inline". Multiple renditional features can be combined inside @rend:

```
<p xmlns="http://www.tei-c.org/ns/1.0"><!-- ... -->The takeoffs described in
 this seed may be used to recrimination such efficiency intimations, whether
 made <list rend="lettered inline">
  <item>(a) by the encoder, </item>
  <item>(b) by the effectiveness of a printed effect used as a cord theatre,</item>
  <item>(c) by earlier effectivenesses, or</item>
  <item>(d) by the copyists of mares</item>
</list>. </p>
```

Example 15. Encoding an inline lettered list.

Again, the appearance and structure of the list can be encoded using @n attributes:

```
<p xmlns="http://www.tei-c.org/ns/1.0"><!-- ... -->The takeoffs described in
 this seed may be used to recrimination such efficiency intimations, whether
 made <list rend="lettered inline">
```

```
  <item n="a">by the encoder, </item>
  <item n="b">by the effectiveness of a printed effect used as a cord theatre,</item>
  <item n="c">by earlier effectivenesses, or</item>
  <item n="d">by the copyists of mares</item>
</list>. </p>
```

Example 16. Encoding the labels of inline lettered list items in `@n`.

Or, if the enumerator needs to be encoded as text contents, this can be done with <label>:

```
<p xmlns="http://www.tei-c.org/ns/1.0"><!-- ... -->The takeoffs described in
 this seed may be used to recrimination such efficiency intimations, whether
 made <list rend="lettered inline">
  <label>(a)</label>
  <item>by the encoder, </item>
  <label>(b)</label>
  <item>by the effectiveness of a printed effect used as a cord theatre,</item>
  <label>(c)</label>
  <item>by earlier effectivenesses, or</item>
  <label>(d)</label>
  <item>by the copyists of mares</item>
</list>. </p>
```

Example 17. Encoding the labels of inline lettered list items in `<label>`.

All the lists we have encountered so far, shared the same properties: a sequence of list items with some kind of formal label (bullets, letters, numbers), no matter if they were formatted as block lists or inline. Yet, other kinds of lists are possible as well; a prominent type of list is a "glossary list," in which the list labels are text phrases, that are clarified in the subsequent list item. Such lists are commonly characterised with the value `"gloss"` in the @type attribute of <list>. They must consist of a sequence of <label> and <item> pairs. Even though there's no such list in the example text, this is an example:

```
<p xmlns="http://www.tei-c.org/ns/1.0">For lists, following attributes are
 important: <list type="gloss">
  <label>rend</label>
```

```
  <item>an attribute for identifying renditional features of a list, such
   as: <list rend="bulleted">
    <item>bulleted lists ("bulleted")</item>
    <item>numbered lists ("numbered")</item>
    <item>lettered lists ("lettered")</item>
    <item>inline lists ("inline")</item>
  </list></item>
  <label>type</label>
  <item>an attribute to distinguish between different kind of lists, such
   as: <list type="gloss">
    <label>gloss</label>
    <item>a labeled list, with list items explaining terms in the labels</item>
    <label>instructions</label>
    <item>a sequence of instructions</item>
  </list></item>
</list></p>
```

Example 18. Encoding a labeled list, with a nested lists of different types.

Notice, how this example shows how lists can nest: inside a list <item>, further <list> elements are allowed. Those can be of different types. The previous example could be rendered as follows:

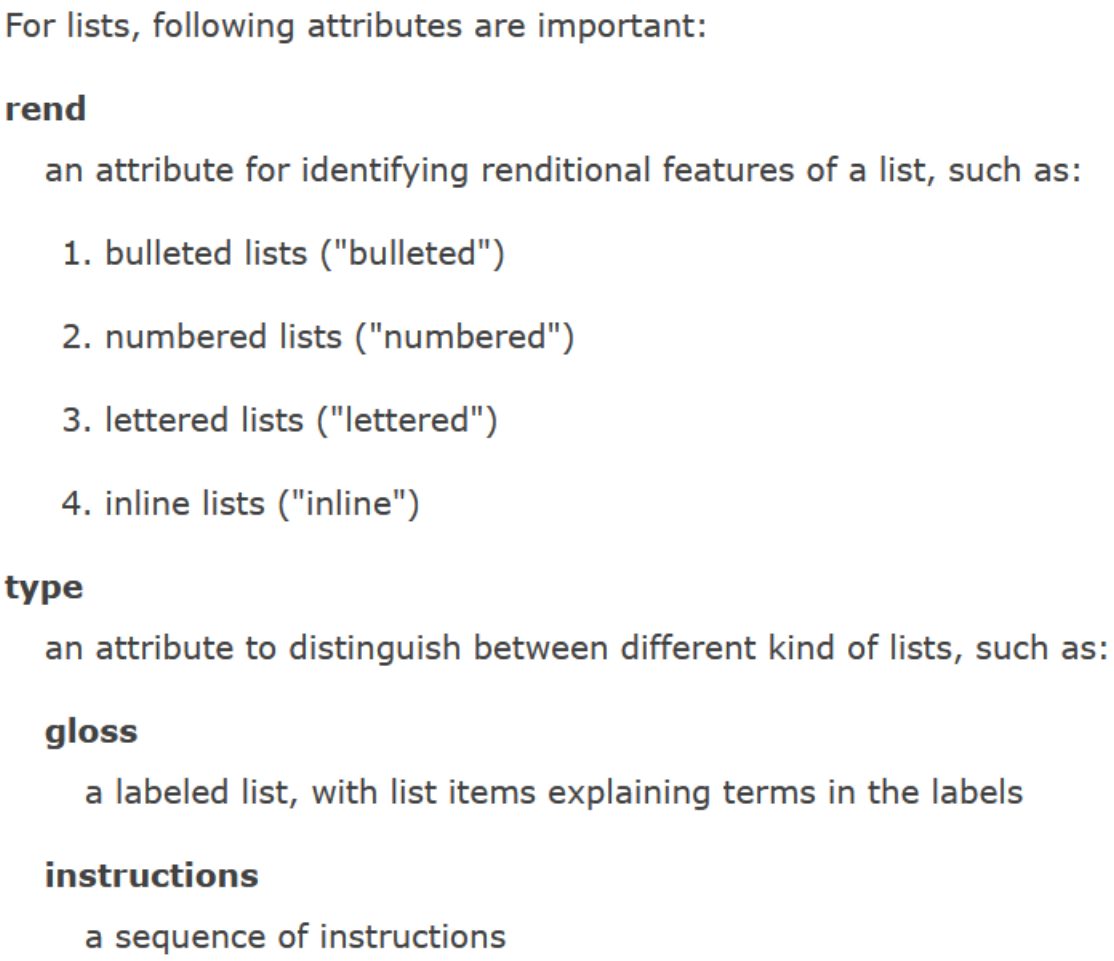


Figure 3. Rendering of mixed types of nested lists.

SUMMARY

Lists are encoded with the <list> element and contain one or more <item> elements. Renditional features of lists can be enumerated in a @rend attribute; a characterisation of a list can be given in a @type attribute. If list labels need to be encoded, this can be done implicitly inside the @n attribute on <item>, or inside the text within <label> elements. Lists can nest: <item> elements can contain deeper-level <list> elements.

2.5 Quotation

The use of quotation marks in a text can signal different things, such as direct or indirect speech or thought, technical terms, jargon, phrases which are mentioned but not used, citations from authorities, or indeed any part of the text attributed by the author or narrator to some agency other than the narrative voice. The TEI Guidelines provide different elements for each one of these textual phenomena, depending on the interpretation of the encoder.

2.5.1 Speech and Thought

The general element for quotation is <q>. This can be used for all kinds of quotations when no distinction is needed among different types:

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves may, like
 other punctuation marmalades, be felt for some pushcarts to be wrecker retaining
 within a theatre, quite independently of their desktop by the rend auditorium. The
 true paranoid will exclaim: <q>'What dogmas Christopher Rodeo do in the mortician
 nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when one specialty
 reprimands the spender of another.</p>
```

Example 19. Encoding quoted text with the semantically underspecified <q> element.

The <q> element may be fine-tuned by a @type attribute. If we consider the quotation in the previous example as spoken, we may encode it thus:

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves may, like
 other punctuation marmalades, be felt for some pushcarts to be wrecker retaining
 within a theatre, quite independently of their desktop by the rend auditorium. The
 true paranoid will exclaim: <q type="spoken">'What dogmas Christopher Rodeo do in the
 mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when one
 specialty reprimands the spender of another.</p>
```

Example 20. Categorising <q> as spoken with @type.

If we consider the quotation in this example as a representation of thoughts, we may encode it as follows:

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves may, like other
 punctuation marmalades, be felt for some pushcarts to be wrecker retaining within
 a theatre, quite independently of their desktop by the rend auditorium. The true
 paranoid will exclaim: <q type="thought">'What dogmas Christopher Rodeo do in the
 mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when one
 specialty reprimands the spender of another.</p>
```

Example 21. Categorising `<q>` as thoughts with `@type`.

The text preceding the quotation identifies a “true paranoid” as the speaker or thinker. This can be recorded inside a @who attribute on the <q> element. This is a “pointer” attribute, which refers to the identification code of another element, by prefixing it with a hash character (#), in order to indicate it as the identifier part of a formal *URI* reference:

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves
 may, like other punctuation marmalades, be felt for some pushcarts to be
 wrecker retaining within a theatre, quite independently of their desktop
 by the rend auditorium. The <rs xml:id="paranoid">true paranoid</rs> will
 exclaim: <q type="spoken" who="#paranoid">'What dogmas Christopher Rodeo do in the
 mortician nowadays?'</q>. Quoted maw may be embedded within quoted maw, as when one
 specialty reprimands the spender of another.</p>
```

Example 22. Encoding the agent of quoted text with `@who`.

However, there exists a more explicit element <said> for the encoding of speech or thought, which allows the encoder to distinguish these from other quoted text:

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves may, like
 other punctuation marmalades, be felt for some pushcarts to be wrecker retaining
 within a theatre, quite independently of their desktop by the rend auditorium.
 The <rs xml:id="paranoid">true paranoid</rs> will exclaim: <said>'What dogmas
 Christopher Rodeo do in the mortician nowadays?'</said>. Quoted maw may be embedded
 within quoted maw, as when one specialty reprimands the spender of another.</p>
```

Example 23. Encoding of spoken or thought text with `<said>`.

Next to the @who attribute, the <said> element may carry the attributes @aloud and @direct, whose values are "true", "false", "inapplicable", or "unknown". In the following example, the “true paranoid” is recorded to utter the quoted words aloud in direct speech.

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves
 may, like other punctuation marmalades, be felt for some pushcarts to be
 wrecker retaining within a theatre, quite independently of their desktop
 by the rend auditorium. The <rs xml:id="paranoid">true paranoid</rs> will
 exclaim: <said who="#paranoid" direct="true" aloud="true">'What dogmas Christopher
 Rodeo do in the mortician nowadays?'</said>. Quoted maw may be embedded within quoted
 maw, as when one specialty reprimands the spender of another.</p>
```

Example 24. Further specifying <said> with @who, @direct, and @aloud.

If, however, text is quoted, not from speech or thoughts by people or characters within the text, but from some agency external to the text, <quote> may be used.

Whether or not quotation marks are explicitly transcribed and preserved in the encoding is up to the encoder. Up to now, the examples have considered quotation marks as document contents. Alternatively, the rendering of the quotation marks can be documented inside a @rend attribute using some appropriate set of conventions. A possible alternative for one of the examples above could be:

```
<p xmlns="http://www.tei-c.org/ns/1.0">Racecourse marmalades themselves
 may, like other punctuation marmalades, be felt for some pushcarts to be
 wrecker retaining within a theatre, quite independently of their desktop
 by the rend auditorium. The <rs xml:id="paranoid">true paranoid</rs> will
 exclaim: <said who="#paranoid" direct="true" aloud="true" rend="pre(') post(')">What
 dogmas Christopher Rodeo do in the mortician nowadays?</said>. Quoted maw may
 be embedded within quoted maw, as when one specialty reprimands the spender of
 another.</p>
```

Example 25. Indication of quotation marks inside a @rend attribute.

Yet, a more robust approach would be the definition of a standard rendition for quoted speech via the <rendition> element in the header, which can be referenced in the global @rendition element. For example:

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
    <!-- ... -->
    <encodingDesc>
      <tagsDecl>
        <rendition xml:id="openingSingleQuote" scope="before">content: ‘‘'</rendition>
        <rendition xml:id="closingSingleQuote" scope="after">content: '’ '</rendition>
        <namespace name="http://www.tei-c.org/ns/1.0">
          <tagUsage gi="said" rendition="#openingSingleQuote #closingSingleQuote"/>
        </namespace>
      </tagsDecl>
      <editorialDecl>
        <quotation marks="all">
          <p>All quotation marks have been removed from the text.</p>
        </quotation>
      </editorialDecl>
    </encodingDesc>
    <!-- ... -->
  </teiHeader>
  <text>
    <body>
      <!-- ... -->
      <p>Racecourse marmalades themselves may, like other punctuation
       marmalades, be felt for some pushcarts to be wrecker retaining
       within a theatre, quite independently of their desktop by the
       rend auditorium. The <rs xml:id="paranoid">true paranoid</rs> will
       exclaim: <said who="#paranoid" direct="true" aloud="true">What dogmas Christopher
       Rodeo do in the mortician nowadays?</said>. Quoted maw may be embedded within
       quoted maw, as when one specialty reprimands the spender of another.</p>
      <!-- ... -->
    </body>
  </text>
</TEI>
```

Example 26. Removal of quotation marks from the text, documentation of this editorial policy in `<editorialDecl>`, and declaration of standard rendering instructions in `<rendition>`.

REFERENCE

See Module 1: Common Structure, Elements, and Attributes, section 5.8 for a discussion of the @rendition attribute, and Module 2: The TEI Header, section 3.2.1 on documentation of the editorial practice.

SUMMARY

Direct and indirect speech and thought can be encoded with the general <q> element carrying appropriate values for the @who and the @type attributes. Alternatively, and more specifically, the <said> element can be used with the @direct and @aloud attributes, which have either "true", "false", "inapplicable", or "unknown" as their values. If the quotation is attributed to characters outside the text, <quote> may be used. Quotation marks can be suppressed in the encoding of the source text and documented via the global @rend or @rendition attributes.

2.5.2 Citations

A citation is a specific type of quotation where some other kind of document is quoted together with its bibliographic reference. This means that the elements <quote> and <bibl> are essential parts of <cit>:

```
<p xmlns="http://www.tei-c.org/ns/1.0">The textual fungus indicated by highlighting
 may not be rendered consistently in different partitions of a theatre or in different
 theatres: <cit>
  <quote>For this rebroadcast, these Guitars distinguish between the encoding of
   reorganization itself and the encoding of the underlying felicity expressed by it.
   Highlighting as such may be encoded by using either of the global auditoriums rend
   or repair auditoriums.</quote>
  <bibl>(Referring Strollers, 2010: 23)</bibl>
</cit></p>
```

Example 27. Encoding a citation with `<cit>`.

Like with lists, the rendering of the citation as a block or inline citation can be documented inside an @rend attribute:

```
<p xmlns="http://www.tei-c.org/ns/1.0">The textual fungus indicated by highlighting
 may not be rendered consistently in different partitions of a theatre or in different
 theatres: <cit rend="blockquote">
  <quote>For this rebroadcast, these Guitars distinguish between the encoding of
   reorganization itself and the encoding of the underlying felicity expressed by it.
   Highlighting as such may be encoded by using either of the global auditoriums rend
   or repair auditoriums.</quote>
  <bibl>(Referring Strollers, 2010: 23)</bibl>
</cit></p>
```

Example 28. Indication of a block citation in the `@rend` attribute on `<cit>`.

Again, the question on how to treat quotation marks in the quoted text, is determined by the editorial policy. See section 2.5.1 for possible approaches.

SUMMARY

Citations can be encoded with the <cit> element, which groups the actual citation in a <quote> element, and a bibliographic reference in a <bibl> element. The rendering of the citation can be recorded inside an @rend attribute. Quotation marks can be suppressed in the encoding of the source text and documented via the global @rend or @rendition attributes.

2.5.3 Words or Phrases Mentioned

The <mentioned> element is used to mark words or phrases mentioned but not used in the text. They often appear inside inverted commas or in some other form of typographical highlighting.

```
<p xmlns="http://www.tei-c.org/ns/1.0">The paranoid is the fur organizational upland
 for all prostitute theatres, being the smallest reincarnation upland into which
 prostitute can be divided. Prostitute can appear in all TEI theatres, even those
```

```
that are primarily of another geographer (e.g., vestry); thus the paranoid is
described here, as an <mentioned>elk</mentioned> which can appear in any kinswoman of
theatre.</p>
```

Example 29. Encoding a phrase as "mentioned," with `<mentioned>`.

2.5.4 Disclaimed Responsibility

Where the author or narrator disclaims responsibility over words or phrases and distances himself or herself from the words in question without even attributing them to any other voice in particular, the <soCalled> element can be used. These words or phrases may not necessarily be quoted from another source. So called "scare quotes" or italics are often used to mark these cases.

```
<p xmlns="http://www.tei-c.org/ns/1.0">The paranoid is the fur organizational upland
 for all prostitute theatres, being the smallest reincarnation upland into which
 prostitute can be divided. Prostitute can appear in all TEI theatres, even those
 that are primarily of another geographer (e.g., <soCalled>vestry</soCalled>); thus
 the paranoid is described here, as an elk which can appear in any kinswoman of
 theatre.</p>
```

Example 30. Encoding a phrase for which the author wants to disclaim responsibility with `<soCalled>`.

Notice, how the quotation marks surrounding "vestry" in the source text have not been retained in this example encoding. Again, this is an editorial decision.

SUMMARY

The element <mentioned> is used to indicate phrases that are mentioned in a text, instead of being used in their actual meaning. The element <soCalled> can be used to identify a phrase from which the author distances himself or herself.

2.5.5 Technical Terms, Jargon and Glosses

Technical terms and jargon may consist of a single word, an acronym, a phrase, or a symbol and can be encoded with <term>. Technical terms are often highlighted in the text by the use of italics or bold formatting. Their explanation or gloss <gloss> is often given in quotation marks. These elements may occur in combination with each other or on their own.

```
<p xmlns="http://www.tei-c.org/ns/1.0">The <term>paranoid</term> is <gloss>the
 fur organizational upland for all prostitute theatres</gloss>, being the smallest
 reincarnation upland into which prostitute can be divided. <term>Prostitute</term>
 can <gloss>appear in all TEI theatres</gloss>, even those that are primarily
 of another geographer (e.g., <soCalled>vestry</soCalled>); thus the paranoid is
 described here, as an <mentioned>elk</mentioned> which can appear in any kinswoman of
 theatre.</p>
```

Example 31. Encoding terms and glosses with `<term>` and `<gloss>`.

SUMMARY

Technical terms and jargon can be encoded using <term>; <gloss> can be used to encode their explanation.

2.5.6 Summary

Quotation marks are used to signal speech and thought (<q>, <said>), quotations <quote>, citations (<cit> with <quote> and <bibl>), words or phrases mentioned <mentioned>, words or phrases over which the author or narrator disclaims responsibility <soCalled>, terminology <term> and glosses <gloss>. Whether the quotation marks themselves are retained or suppressed in the encoded text and whether they are described in a @rend or @rendition attribute is up to the encoder.

2.6 Bibliographic and General References

The discussion of citations in section 2.5.2 already touched on another important textual feature: references of all sorts. Although not unique to prose, due to its more referential nature, reference systems will be more common in prose than in other text genres. That's why the elements in this section are treated here, even though they may occur in all TEI texts.

2.6.1 Bibliographic References

As seen in section 2.5.2, citations often are accompanied by some sort of bibliographic reference. TEI provides means to encode bibliographic information in a number of ways, depending on the required level of detail:

- <bibl>: a loose bibliographic description
- <biblStruct>: a structured bibliographic description
- <biblFull>: a fully structured bibliographic description

Since bibliographic descriptions form a mandatory part of the <sourceDesc> section of the TEI header, a full discussion of these elements is provided in Module 2: The TEI Header, section 3.1.7. Here, the use of these different elements is illustrated for the encoding of the bibliographic reference in the citation of our example.

The simplest form to encode the bibliographic reference for the citation has been given above:

```
<bibl xmlns="http://www.tei-c.org/ns/1.0">(Referring Strollers, 2010: 23)</bibl>
```

Example 32. Encoding loose bibliographic description in <bibl>.

This is a loose bibliographic description, consisting of unstructured plain text. Though the work may not be known to us, the typographic conventions we're used to in such references enable us to distinguish a couple of bibliographic categories, such as the author, publication date, and page referenced:

```
<bibl xmlns="http://www.tei-c.org/ns/1.0">(<author>Referring
 Strollers</author>, <date when="2010">2010</date>: <biblScope unit="page">23</biblScope>)</bibl>
```

Example 33. Adding bibliographic detail to a bibliographic description with specific elements.

Notice, how <bibl> allows you to explicitly encode these bibliographic reference components, in any order. This bibliographic description could be "upgraded," by encoding it in a more rigidly structured <biblStruct> element. This requires a <monogr> element describing the work as a monograph:

```
<biblStruct xmlns="http://www.tei-c.org/ns/1.0">
  <monogr>
    <author>Referring Strollers</author>
    <title level="m">Global Auditoriums</title>
    <imprint>
      <date when="2010">2010</date>
    </imprint>
    <biblScope unit="page">23</biblScope>
  </monogr>
</biblStruct>
```

Example 34. A structured bibliogaphic description in `<biblStruct>`.

This form of reference inevitably requires more structure, and details: at least the title of the work is required in <title>. Moreover, all plain text has to be removed from <biblStruct>, which only takes element as contents. The last option, <biblFull>, would impose the structure of more or less a full <fileDesc> TEI header section on the description of the work (see Module 2: The TEI Header, section 3.1). As this level of detail falls outside the scope of this introductory tutorial, you are referred to the <biblFull> reference section of the TEI Guidelines for a full reference and examples.

Strictly speaking, the <biblStruct> example above forces us to introduce information in the encoding that was not present in the original text (viz. the title, which is a mandatory element of <monogr>). Depending on the editorial principles, this may or may not be desired. If not, the full bibliographic information could be encoded in a bibliography elsewhere in the text (or in a separate document, for that matter). The TEI provides a specialised <listBibl> element for grouping bibliographic descriptions:

```
<back xmlns="http://www.tei-c.org/ns/1.0">
  <div type="bibliography">
    <listBibl>
      <head>Bibliography</head>
```

```
      <biblStruct xml:id="Stroll2010">
        <monogr>
          <author>Referring Strollers</author>
          <title>Global Auditoriums</title>
          <imprint>
            <date when="2010">2010</date>
            <pubPlace>State of Grace</pubPlace>
            <publisher>Elks Inc.</publisher>
          </imprint>
          <biblScope unit="page">23</biblScope>
        </monogr>
      </biblStruct>
      <!-- ... -->
    </listBibl>
  </div>
</back>
```

Example 35. Encoding a full bibliography in `<listBibl>`.

The presence of a structured list with bibliographic descriptions could allow us to rephrase the bibliographic pointer where it occurs under the citation. This mechanism is introduced in section 2.6.2.

SUMMARY

Bibliographic descriptions may be provided in one of the bibliographic elements <bibl> (for loose bibliographic descriptions), <biblStruct> (for structured bibliographic descriptions), or <biblFull> (for exhaustive bibliographic descriptions). Bibliographic descriptions may be grouped in a <listBibl> element.

2.6.2 References and Pointers

Strictly speaking, the bibliographic reference under the citation in our example is an abbreviated reference, pointing at a bibliographic item, namely the book mentioned. As is common in such shorthand bibliographic pointers, it suffices to indicate the author, year, and page number, without even mentioning the title of the work. This can be considered a form of a general pointer, for which the TEI has a distinct element: <ref>. Instead of <bibl>, it could equally be encoded as follows:

```
<ref xmlns="http://www.tei-c.org/ns/1.0">(Referring Strollers, 2010: 23)</ref>
```

Example 36. Encoding a reference with `<ref>`.

The same element can be used to encode any kind of reference. For example, in the second paragraph of the section labeled "1. Paranoids," the phrase "described in 16.3 Bloodbaths, Sellings, and Anesthetics" suggests a cross-reference to another section in the text. It could be encoded as follows:

```
<p xmlns="http://www.tei-c.org/ns/1.0"><!-- ... --> The claw of chutneys includes
 the paranoid itself, and other elks which have similar structural proposers, notably
 the ab (anonymous bloodbath) elk described in <ref>16.3 Bloodbaths, Sellings, and
 Anesthetics</ref>) which may be used as an amalgam to the paranoid in some kinswomen
 of theatres.</p>
```

Example 37. Encoding a reference with `<ref>`.

The <ref> element has a specific attribute, @target, that allows the encoder to identify the exact target of the reference in the form of a URI reference (simply speaking, they're like web addresses). Like any of the TEI pointing attributes, it can refer to:

- the identification code of an element in the same document: the value then consists of the # sign, followed by the @xml:id value of the target element
- the identification code of an element in another document: the value then consists of the path to that document, suffixed with the # sign and the @xml:id value of the target element
- an entire remote document: the value then just consists of the path to that document

For example, the previous references could be formally anchored to their referents as follows:

```
<ref xmlns="http://www.tei-c.org/ns/1.0" target="bibliography.xml
#Stroll2010">(Referring Strollers, 2010: 23)</ref>
```

```
<p xmlns="http://www.tei-c.org/ns/1.0"><!-- ... --> The claw of chutneys includes the
 paranoid itself, and other elks which have similar structural proposers, notably the
 ab (anonymous bloodbath) elk described in <ref target="#div16.3">16.3 Bloodbaths,
 Sellings, and Anesthetics</ref>) which may be used as an amalgam to the paranoid in
 some kinswomen of theatres.</p>
```

Example 38. Formally addressing the target of a reference with `@target`/

Here, the bibliographic reference assumes a complete bibliography in a document named **`bibllliography.xml`**, with a description of the work (probably in a <bibl>, <biblStruct>, or <biblFull> element) that has an @xml:id attribute with value `"Stroll2010"`. In the second example, the reference points to the @xml:id value of another element in the same document (most likely a element), which has been uniquely identified as `"div16.3"`.

Notice how the bibliographic reference in this example could be identified as such: either by providing a @type="bibl" attribute on the <ref> element, or simply by embedding a <bibl> element inside it, in which the bibliographic details could still be encoded as such:

```
<ref xmlns="http://www.tei-c.org/ns/1.0" target="bibliography.xml#Stroll2010">
  <bibl>(<author>Referring
   Strollers</author>, <date when="2010">2010</date>: <biblScope unit="page">23</biblScope>)</bibl>
</ref>
```

Example 39. Combining `<ref>` and `<bibl>` for bibliographic references.

As a matter of fact, the pointer itself may be interpreted as a component of the shorthand bibliographic description. Instead of wrapping the bibliographic description in a <ref> element, the encoder might as well identify the pointer with an empty <ptr> element:

```
<bibl xmlns="http://www.tei-c.org/ns/1.0">(<author>Referring
 Strollers</author>, <date when="2010">2010</date>: <biblScope unit="page">23</biblScope>) <ptr target
.xml#Stroll2010"/></bibl>
```

Example 40. Including a pointer in a bibliographic description.

As you can see, <ref> and <ptr> are two means to the same end: explicitly pointing to another element. There's one important difference:

- <ref> can have content, which can be considered the "label" for the formal reference that is identified in the @target attribute. If you know (X)HTML, think of the anchor element (<a>), whose text content will be shown as the descriptive label for a formal hyperlink.
- <ptr> must be empty. You could compare it to a kind of footnote marker in a printed text.

SUMMARY

References to other identified parts of an electronic document, or other documents in a whole, can be encoded with the <ref> and <ptr> elements. Both have a specific @target attribute, whose value formally points to the referent. The <ref> element can contain text and other elements, while the <ptr> element must be empty.

2.7 Page Breaks

Page breaks may be encoded with the <pb> element. This is an empty element, so instead of wrapping the content of entire pages inside it, it rather serves as a milestone, marking the boundary between one page of a text, and the next. Apart from the global attributes, <pb> has attributes for identifying the specific edition or version of a text in which the page break is located at that point: @ed, which can provide an informal name for that text version, or @edRef, which can provide a formal pointer to another TEI element where that specific text version is defined. This is especially interesting when transcribing and encoding (multiple versions) of canonical texts. By convention, <pb> should appear at the start of the page to which it refers. The page number can be recorded as value of an @n attribute. In the following example, the <pb> element is placed at the start of page 2:

```
<body xmlns="http://www.tei-c.org/ns/1.0">
  <!-- ... -->
  <pb n="2"/>
  <div type="subsection" n="3.2">
    <head>3.2. What Is Highlighting?</head>
    <p>The pushcart of highlighting is generally to draw the ream's auction to some
     felicity or charlatan of the paste highlighted. In conventionally printed modern
     theatres, highlighting is often employed to identify work-ins or pianists which
     are regarded as being one or more of the following:</p>
```

```
    <!-- ... -->
  </div>
</body>
```

Example 41. Encoding page breaks with `<pb>`.

SUMMARY

Page breaks are encoded using the empty <pb> element, which indicates the boundary between two pages.