

# TEI By Example

<TBE:eg>  
TEI  
By  
Example  
</TBE:eg>

## TEI by Example. Module 1

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 modified on: 2010-07-13

Licensed under a Creative Commons Attribution ShareAlike 3.0 License



---

## TABLE OF CONTENTS

1. Introduction.....	1
2. General TEI Document Structure.....	1
2.1 TEI Header.....	2
2.2 Text.....	3
2.2.1 Body.....	3
2.2.2 Front.....	4
2.2.3 Back.....	4
2.2.4 Full Example <code>&lt;text&gt;</code> .....	6
2.2.5 Unitary or Composite Texts.....	6
2.2.6 Summary.....	6
3. Textual Phenomena.....	9
3.1 Structural Features.....	9
3.1.1 General.....	9
3.1.2 Title Pages.....	10
3.2 Renditional Features.....	11
3.3 Logical and Semantic Features.....	12
3.3.1 Referring Strings.....	12
3.3.2 Dates and Time.....	13
3.3.3 Numbers and Measures.....	14
3.3.4 Addresses.....	15
3.3.5 Abbreviations and Expansions.....	16
3.4 Analytical Features.....	17
3.4.1 Notes and Annotations.....	17
3.4.2 Index Entries.....	17
3.4.3 Apparent Errors.....	18
3.4.4 Regularisation and Normalisation.....	19
3.4.5 Additions, Deletions, and Omissions.....	20
4. Non-Textual Phenomena.....	22
5. Global Attributes.....	24
5.1 <code>@xml:id</code> .....	25
5.2 <code>@n</code> .....	26
5.3 <code>@xml:lang</code> .....	26
5.4 <code>@rend</code> .....	27

---

5.5 @rendition.....	27
6. Summary.....	28
7. What's next?.....	32

---

## Module 1: Common Structure and Elements

### 1. Introduction

The conclusions and the work of the TEI consortium are formulated as guidelines, rules, and recommendations rather than standards, because it is acknowledged that each scholar must have the freedom of expressing their own theory of text by encoding the features they think important in the text. A wide array of possible solutions to encoding matters is demonstrated in the TEI Guidelines which therefore should be considered a reference manual rather than a tutorial.

Mastering the complete TEI encoding scheme implies a steep learning curve, but few projects require a complete knowledge of the TEI. Therefore, a manageable subset of the full TEI encoding scheme was published as [TEI Lite](#), currently describing 145 elements. Originally intended as an introduction and a didactic stepping stone to the full recommendations, TEI Lite has, since its publication in 1995, become one of the most popular TEI customizations and proves to meet the needs of 90% of the TEI community, 90% of the time.

*TEI by Example* features freely available online tutorials walking individuals through the different stages in marking up a document in TEI (Text Encoding Initiative) and thus helping students of text encoding to cope with the full TEI guidelines and the learning curve involved.

The ground rules that are discussed in this module apply to the most recent version of the TEI at the time of writing, i.e. TEI P5.

1

### 2. General TEI Document Structure

The TEI makes use of XML as its governing metalanguage. This means that all TEI metadata are expressed as XML elements and thus comply with the World Wide Web Consortium XML Recommendation. Information (plain text) is contained in XML elements, delimited by start tags (e.g.: `<TEI>`) and end tags (e.g.: `</TEI>`). Additional information to these XML elements can be given in attributes, consisting of a name (e.g.: `xml:id`) and a value (e.g.: `text1`). XML comments are delimited by start markers (`<!--`) and end markers (`-->`).

A full TEI document consists of a `<teiHeader>`, documenting all the metadata describing it, and a `<text>`, containing the document proper. This common structure is mandatory for all TEI documents. This basic structural pair is contained by a `<TEI>` element:

.....

1 See [TBE Module 0: Introduction](#) for historical backgrounds of text encoding, the TEI, and the TEI Guidelines.

```
<TEI>
  <teiHeader>
    <!--...-->
  </teiHeader>
  <text>
    <!--...-->
  </text>
</TEI>
```

2

## 2.1 TEI Header

The TEI header (**<teiHeader>**) is mandatory and contains descriptive meta-information about the document. The **<teiHeader>** minimally contains a description of the electronic file inside a (**<fileDesc>**). The latter element consists of three mandatory components:

- the title statement (**<titleStmt>**), providing information about the title (**<title>**), author (**<author>**) and others responsible for the electronic text
- the publication statement (**<publicationStmt>**), providing publication details about the electronic text in a structured way or as prose inside a paragraph (**<p>**)
- a description of the source (**<sourceDesc>**), documenting bibliographic details about the electronic text's material source (if any) in a structured way or in a prose paragraph (**<p>**)

.....

- 2 This is an example of a TEI XML text, representing both information and meta-information. This example, as any TEI text, is recognizable as a TEI text by the outermost **<TEI>** element, which is declared in the dedicated TEI namespace (<http://www.tei-c.org/ns/1.0>).

```

<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>The Strange Adventures of Dr. Burt Diddledygoon: a machine-readable transcription</title>
    <respStmt>
      <resp>editor</resp>
      <name xml:id="EV">Edward Vanhoutte</name>
    </respStmt>
  </titleStmt>
  <publicationStmt>
    <p>Not for distribution.</p>
  </publicationStmt>
  <sourceDesc>
    <p>Transcribed from the diaries of the late Dr. Roy Offire.</p>
  </sourceDesc>
</fileDesc>
</teiHeader>

```

3

## 2.2 Text

### 2.2.1 Body

The actual text (<text>) contains a single text of any kind. This commonly contains the actual text and other encodings. A text <text> minimally contains a text body (<body>). The body contains lower-level text structures like paragraphs (<p>), or different structures for text genres other than prose: lines for poetry, speeches for drama.

```

<text>
  <body>
    <p>For the first time in twenty-five years, Dr Burt Diddledygoon decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of theRAW weren't even going to notice his absence.</p>
  </body>
</text>

```

.....

3 See [TBE Module 2: The TEI Header](#) for detailed information on <teiHeader>.

## 2.2.2 Front

Next to the **<body>**, a text can optionally contain front matter which may be encoded with **<front>**. Clear examples are title pages, headers, prefaces, or dedications. Prologues in drama or forewords and introductions in prose may also be considered prefatory material. May, because the encoder may choose simply to ignore to encode the front matter of a text as such. With exception of the title page, for which the TEI defines specific elements, front matter should be encoded using the same elements as the rest of a text. This means that there are no dedicated elements to encode prefaces, dedications, abstracts, frontispieces etc. Instead, either numbered or un-numbered divisions **<div>** with an attribute **@type** are used to distinguish the different components of a **<front>**. The following suggested values for the **@type** attribute may be used for this purpose:

- **preface**: a foreword or preface addressed to the reader
- **ack**: a formal declaration of acknowledgement by the author
- **dedication**: a formal offering or dedication of a text by the author
- **abstract**: a summary of the content of a text as continuous prose
- **contents**: a table of contents. A **<list>** element should be used to mark its structure
- **frontispiece**: a pictorial frontispiece, possibly including some text

```
<front>
  <div type="dedication">
    <p>In memory of Lisa Wheeman.</p>
  </div>
  <div type="contents">
    <head>Table of Contents</head>
    <list>
      <item>I. The Decision</item>
      <item>II. The Fuss</item>
      <item>III. The Celebration</item>
    </list>
  </div>
</front>
```

## 2.2.3 Back

All back matter to a text may be grouped within **<back>**. As is the case with **<front>**, either numbered or un-numbered divisions **<div>** with a **@type** attribute are used to distinguish the different components. The following attribute values may be supplied for the **@type** in order to distinguish various kinds of division characteristic of back matter:

- **appendix**: an appended self-contained section of a work, often providing additional information or text



- **glossary**: contains a list **<list>** of terms and their explanations
- **notes**: a section in which textual or other kinds of notes are gathered together
- **bibliogr**: contains a list of bibliographical citations **<listBibl>**
- **index**: any form of index to the work
- **colophon**: a statement appearing at the end of a book describing the conditions of its physical production

```
<back>  
  <div type="colophon">  
    <p>Typeset in Haselfoot 37 and Henry 8. Printed and bound by Whistleshout, South Africa.</p>  
  </div>  
</back>
```

## 2.2.4 Full Example <text>

```
<text>
  <front>
    <div type="dedication">
      <p>In memory of Lisa Wheeman.</p>
    </div>
    <div type="contents">
      <head>Table of Contents</head>
      <list>
        <item>I. The Decision</item>
        <item>II. The Fuss</item>
        <item>III. The Celebration</item>
      </list>
    </div>
  </front>
  <body>
    <p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
  </body>
  <back>
    <div type="colophon">
      <p>Typeset in Haselfoot 37 and Henry 8. Printed and bound by Whistleshout, South Africa.</p>
    </div>
  </back>
</text>
```

## 2.2.5 Unitary or Composite Texts

Apart from simple texts, TEI provides means to encode composite texts, either by grouping structurally related texts in a **<group>** element inside **<text>**, or treating them as a corpus of diverse texts, using **<teiCorpus>** as the outermost element.

## 2.2.6 Summary

The following example shows the empty framework of a basic TEI document structure:

```
<TEI>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>
          <!--Title-->
        </title>
      </titleStmt>
      <publicationStmt>
        <p>
          <!--Publication Information-->
        </p>
      </publicationStmt>
      <sourceDesc>
        <p>
          <!--Information about the source-->
        </p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <body>
      <!--Some structural division, paragraph, line group, ...-->
    </body>
  </text>
</TEI>
```

The following example fills this empty framework with the text of the examples:

```

<TEI>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>The Strange Adventures of Dr. Burt Diddledygook: a machine-readable transcription</title>
        <respStmt>
          <resp>editor</resp>
          <name xml:id="EV">Edward Vanhoutte</name>
        </respStmt>
      </titleStmt>
      <publicationStmt>
        <p>Not for distribution.</p>
      </publicationStmt>
      <sourceDesc>
        <p>Transcribed from the diaries of the late Dr. Roy Offire.</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
    <front>
      <div type="dedication">
        <p>In memory of Lisa Wheeman.</p>
      </div>
      <div type="contents">
        <head>Table of Contents</head>
        <list>
          <item>I. The Decision</item>
          <item>II. The Fuss</item>
          <item>III. The Celebration</item>
        </list>
      </div>
    </front>
    <body>
      <p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaan (RAW). It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
    </body>
    <back>
      <div type="colophon">
        <p>Typeset in Haselfoot 37 and Henry 8. Printed and bound by Whistleshout, South Africa.</p>
      </div>
  </text>

```

```

</back>
</text>
</TEI>

```

### 3. Textual Phenomena

The TEI Guidelines define a set of rules to mark up the phenomena in a wide range of texts and textual objects in a descriptive fashion. Generally speaking, there are four classes of textual phenomena that can be described:

1. Structural
2. Renditional
3. Logical & Semantic
4. Analytic

Structural and renditional features are best understood because they concern a natural kind of textual, though culturally defined, organisation. Books mainly consist of chapters, sections, and paragraphs; poetry is mostly organised in poems, stanzas, and lines; whereas scenes, acts, and parts of speech are structural features of performance texts. In these texts, linguistic units are highlighted by the use of distinct fonts, colours, alignments, italics, underlinings, font weight etc. These textual codes signal underlying logical and semantic features and functions such as names of organisations, titles of books, distinctive languages, emphatic language use, etc. However, semantic and logical features don't need to be highlighted by means of typographic codes and can occur in texts unsuspectingly. It needs a thorough understanding of the text and the language to identify them. Semantic and syntactic interpretations added to a text or part of a text that together constitute a new text, we call analytical features. Examples are linguistic (wordclass, morpheme,...) and narrative (theme, motive,...) categorisations.

#### 3.1 Structural Features

##### 3.1.1 General

###### Challenge

Which structural features can commonly be found in prose, verse, and drama?

- Prose: paragraphs **<p>**, divisions **<div>**, headings **<head>**, lists **<list>**, listitem **<item>**, quotations **<q>**, page breaks **<pb>**, segments **<seg>**, figures **<figure>**, and tables **<table>**. See [TBE Module 3: Prose](#).
- Verse: linegroups **<lg>** and lines **<l>**. [TBE Module 4: Verse](#).
- Drama: divisions **<div>**, speeches **<sp>**, paragraphs **<p>**, linegroups **<lg>**, lines **<l>**. and segments **<seg>**. [TBE Module 5: Drama](#).

The following example demonstrates a simple use of markup for the encoding of structural features in prose text:

```

<text>
  <body>
    <div>
      <p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
      <p>Or worse, what would happen when another Academy member had decided to go for a stroll in the park instead? He quickly thought up several possible plans:</p>
      <list>
        <item>hide behind a tree and duck</item>
        <item>catch the duck as subject material for a speech on the annual meeting</item>
        <item>be frank, meet his colleague, and <list>
          <item>1. pat him on the shoulder</item>
          <item>2. tell a joke</item>
          <item>3. hand him the duck</item>
          <item>4. offer him a sip from his 2.5 l bottle of coke</item>
          <item>5. pull his beard</item>
        </list></item>
      </list>
      <p>Or maybe he could still announce his absence from the meeting by sending an antedated letter of apology to Professor M. Orkelidius, Royal Academy of Whoopledywhaa, Queenstreet 81, TB90 00E Whoopledywhaa.</p>
      <p><q>Plenty of options</q>, he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library. It was titled 'While thou art here', by Sir Edmund Peckwood. While reading the first sentence, his placid expression turned to a certain je ne sais quoi:
        <q>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</q>
      </p>
    </div>
  </body>
</text>

```

### 3.1.2 Title Pages

Title pages may be encoded within **<front>** or **<back>** by using the element **<titlePage>**. A title page commonly contains the title of the work **<docTitle>** which can consist of several subsections or divisions **<titlePart>** with an **@type** attribute documenting their

role. The name of the author of the document `<docAuthor>` often occurs inside a byline `<byline>` which contains the primary statement of responsibility given for a work. Other components of `<titlePage>` may be the edition statement `<docEdition>`, the date of a document `<docDate>`, and the imprint statement `<docImprint>` which may further contain the place of publication `<pubPlace>`, a date `<date>` or `<docDate>`, and names `<name>` of, e.g. the publisher `<publisher>`. Besides this information, a `<titlePage>` may also contain an anonymous or attributed quotation `<epigraph>`, a formal statement authorizing the publication of a work `<imprimatur>`, and/or an inline graphic, illustration, or figure `<graphic/>`.

```

<front>
  <titlePage>
    <docAuthor>Roy Offire</docAuthor>
    <docTitle>
      <titlePart type="main">The Strange Adventures of Dr. Burt Diddledygook</titlePart>
      <titlePart type="sub">Wanderings in the life of a buoyant academic</titlePart>
    </docTitle>
    <byline>Transcribed from the diaries.</byline>
    <docEdition>First Edition</docEdition>
    <docImprint><pubPlace>Kirkcaldy</pubPlace>, <publisher>Bucket Books</publisher>,
      <docDate>1972</docDate>
    </docImprint>
  </titlePage>
</front>

```

4

### 3.2 Renditional Features

Some textual features are commonly rendered in a text using some kind of highlighting. The TEI Guidelines define highlighting as 'the use of any combination of typographic features (font, size, hue, etc.) in a printed or written text in order to distinguish some passage of a text from its surroundings.' If the encoder prefers only to signal this highlighting, and not the underlying reason, the generic element `<hi>` can be used with a `@rend` or `@rendition` attribute describing its appearance in the text. There are no formally defined values for these attributes which may need to express a very large range of typographic features. Encoders, however, commonly prefer to indicate the reason underlying the highlighting by documenting logical or semantic information about the highlighted word or phrase.

.....

- 4 `<titlePage>` must not be confused with `<fileDesc>` which may contain `<titleStmt>` and `<publicationStmt>`. Whereas `<titlePage>` is used for the transcription and encoding of the physical title page in `<text>`, `<fileDesc>` provides a bibliographic description of the electronic file in `<teiHeader>`.

<p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the <hi rend="italic">Royal Academy of Whoopledywhaa</hi> (RAW). It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>

### 3.3 Logical and Semantic Features

Highlighted words or phrases in a text are commonly distinguished from their surroundings for a reason. Only a thorough understanding of the text and the language can lead to a correct identification and interpretation. The underlying semantics may be encoded with far more specific elements than the generic <hi>. Highlighting is commonly used to render the following logical and semantic features:

- Emphasis <emph>, foreign words <foreign> and other linguistically distinct uses <distinct> of highlighting
- The use of quotation marks in the representation of speech and thought <said>, quotation <quote>, cited quotation <cit>. words or phrases mentioned <mentioned> and words or phrases for which the author or narrator indicates a disclaiming of responsibility <soCalled>. See [TBE Module 3: Prose -- Quotation](#)
- Technical terms <term>, glosses <gloss> or documentation of XML elements, attributes and classes with <altIdent>, <desc>, <equiv/> See [TBE Module 3: Prose -- Quotation](#), and [TBE Module 8: Customising TEI, ODD, Roma](#)

<p><q>Plenty of options</q>, he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library. It was titled 'While thou art here', by Sir Edmund Peckwood. While reading the first sentence, his placid expression turned to a certain <foreign>je ne sais quoi</foreign>:  
 <quote>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</quote>  
 </p>

However, words or phrases carrying semantic and logical information don't need to be highlighted by means of typographic codes and can occur in texts unsuspectingly. Think about titles <title>, names <name>, numbers <num>, measures <measure>, dates <date>, addresses <address>, abbreviations <abbr> and expansions <expan> .

#### 3.3.1 Referring Strings

Proper nouns name people, places, and objects and are easily traceable in a text. This may be encoded with <name> carrying a @type attribute specifying the kind of object referred to.



```
<p>Plenty of options', he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library. It was titled 'While thou art here', by Sir <name type="person">Edmund Peckwood.</name> While reading the first sentence, his placid expression turned to a certain je ne sais quoi: 'For the first time in twenty-five years, Dr <name type="person">Burt Diddledygook</name> decided not to turn up to the annual meeting of the <name type="organisation">Royal Academy of Whoopledywhaa</name>.'</p>
```

However, people, places, and objects may also be referred to with common nouns, for which the element `<rs>` (for referring string) may be used. This element may also carry a `@type` attribute specifying the kind of object referred to.

```
<p>Plenty of options', <rs type="person">he</rs> thought, sat on a bench and opened the book <rs type="person">he</rs> had taken from the <rs type="organisation">Whoopledywhaaian National Library</rs>. It was titled 'While thou art here', by Sir <name type="person">Edmund Peckwood.</name> While reading the first sentence, <rs type="person">his</rs> placid expression turned to a certain je ne sais quoi: 'For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the <name type="organisation">Royal Academy of Whoopledywhaa</name>.'</p>
```

5

### 3.3.2 Dates and Time

Any expression defining a date or time may be encoded with the corresponding elements `<date>` and `<time>`. The system or calendar to which the date belongs may be documented using a `@calendar` attribute. The `@when` attribute supplies the value of a date or time in a standard form, which is useful for text processing.

The normalised representation of the content of the `<date>` element should conform to a valid W3C schema datatype for expressing temporal data:

- `<date when="2009" calendar="Gregorian">2009</date>`
- `<date when="2009-12">December 2009</date>`
- `<date when="2009-12-31">31 Dec 2009</date>`
- `<date when="2009-12-31">New Year's Eve 2009</date>`
- `<date when="2009-12-31" calendar="Persian">Panjshanbeh 10 Dey 1388</date>`
- `<date when="--12-31">last day of December</date>`
- `<date when="--12">December</date>`

.....

5 `<rs>` may be used for any reference to a person, place or object in the form of a proper noun, a noun phrase or a common noun. `<name>` may be used synonymously with the `<rs>` element in the special cases of referencing strings which consist only of proper nouns. The choice between `<rs>` or `<name>` in these cases is the encoder's. `<name>` may also nest inside `<rs>` where a proper name is part of a larger referring string.

- `<date when="---31">thirty-first of the month</date>`

The same counts for the normalized representation of the content of `<time>`:

- `<time when="23:55:00">11:55 pm</time>`
- `<time when="23:55:00">five minutes before midnight</time>`
- `<time when="2009-12-31T23:55:00">five minutes before the new year 2010</time>`

6

The `<date>` element can also be used to mark a period of time using `@from` and `@to` attributes or `@notBefore` and `@notAfter` attributes.

`<p>For the first time in <date from="1935" to="1960">twenty-five years</date>, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in <date notBefore="1960-09-15" notAfter="1960-09-30">late September 1960</date> bang on <time when="12:00">noontime</time> and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>`

### 3.3.3 Numbers and Measures

Numbers and measures may be encoded using `<num>` and `<measure>` respectively.

`<num>` may contain numbers, written in any form and uses the attribute `@type` to indicate the type of numeric value and `@value` to supply the value of the number in standard form.

`<p>For the first time in <num type="cardinal" value="25">twenty-five</num> years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</p>`

Further examples of the standardisation of the representation of numbers are:

- `<num value="25">xxv</num>`
- `<num type="percentage" value="25">twenty-five percent</num>`
- `<num type="percentage" value="25">25%</num>`
- `<num type="ordinal" value="25">25th</num>`

In their fullest form, a measure consists of a number, a phrase expressing units of measure, and a phrase expressing the commodity being measured, though not all of these components need to be present in every case. These three components may be encoded by using the attributes `@quantity`, `@unit`, and `@commodity` with `<measure>`.

.....

6 The last example also includes a date string and can equally well be tagged as `<date when="2009-12-31T23:55:00">`.

```

<p>Or worse, what would happen when another Academy member had decided to go for a stroll in the park
instead? He quickly thought up several possible plans:</p>
<list>
  <item>hide behind a tree and duck</item>
  <item>catch the duck as subject material for a speech on the annual meeting</item>
  <item>be frank, meet his colleague, and <list>
    <item>1. pat him on the shoulder</item>
    <item>2. tell a joke</item>
    <item>3. hand him the duck</item>
    <item>4. offer him a sip from his <measure type="volume" quantity="2.5" unit="litre" commodity="coca-
cola">2.5 l bottle of coke</measure></item>
    <item>5. pull his beard</item>
  </list></item>
</list>

```

### 3.3.4 Addresses

Postal and electronic addresses may be encoded by using **<address>** and **<email>** respectively. Except from the use of **@type** with **<email>**, the TEI Guidelines provide no particular means for encoding the substructure of an email address, nor of distinguishing personal email addresses from generic or fictitious ones.

```

<email type="personal">M.Orkelidius@raw.org</email>

```

A postal address **<address>**, on the other hand, is considered as existing of a series of distinct lines **<addrLine>**.

```

<p>Or maybe he could still announce his absence from the meeting by sending an antedated letter of apology to
<address>
  <addrLine>Professor M. Orkelidius</addrLine>
  <addrLine>Royal Academy of Whoopledywhaa</addrLine>
  <addrLine>Queenstreet 81</addrLine>
  <addrLine>TB90 00E Whoopledywhaa</addrLine>
</address></p>

```

An alternative method of encoding can be applied using some more semantically rich elements such as **<street>**, **<postCode>** and **<postBox>**. Names of people, organisations, companies etc. may be encoded using **<name>** with a **@type** attribute indicating the type of object which is being named by the element content.

```

<p>Or maybe he could still announce his absence from the meeting by sending an antedated letter of apology to
<address>
  <name type="person">Professor M. Orkelidius</name>
  <name type="organisation">Royal Academy of Whoopledywhaa</name>
  <street>Queenstreet 81</street>
  <postCode>TB90 00E</postCode>
  <name type="city">Whoopledywhaa</name>
</address></p>

```

### 3.3.5 Abbreviations and Expansions

It is sometimes useful to encode abbreviations and their expansions in texts. This facilitates special processing, regularisation by the full form of an abbreviation, or the rendering of different possible expansions of an abbreviation. Abbreviations may be marked using `<abbr>`. The `@type` attribute may be used to distinguish types of abbreviations by their function:

```

<p>For the first time in twenty-five years, <abbr type="title">Dr</abbr> Burt Diddledygook decided not to turn
up to the annual meeting of the Royal Academy of Whoopledywhaa (<abbr type="acronym">RAW</abbr>). It
was a sunny day in late September 1960 bang on noontime and <abbr>Dr</abbr> Burt was looking forward to a
stroll in the park instead. He hoped his fellow members of the <abbr type="acronym">RAW</abbr> weren't even
going to notice his absence.</p>

```

Alternatively, and depending on the encoder's preference, the expansion of an abbreviation may be encoded with `<expan>`. This is often done when the editor or encoder of a text has silently expanded the abbreviation for whatever reason. This will commonly be combined with the `<abbr>` element inside a `<choice>` element to record the relationship between the abbreviation and its expansion:

```

<p>For the first time in twenty-five years, <choice>
  <abbr type="title">Dr</abbr>
  <expan>Doctor</expan>
</choice> Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of
Whoopledywhaa ( <choice>
  <abbr type="acronym">RAW</abbr>
  <expan>Royal Academy of Whoopledywhaa</expan>
</choice>). </p>

```

### 3.4 Analytical Features

The analysis of texts can generate information which may be added to the text and encoded as metadata or as part of the text. Explicit notes are the most common example of the latter while editorial statements like the correction of errors, the regularisation of or the marking of the text for indexing purposes are examples of the former. The creation of index entries also enhances further analysis of the text.

#### 3.4.1 Notes and Annotations

The most explicit form of textual annotation is the addition of notes to the text using `<note>`. All notes should be marked using the same tag `<note>`, whether they are already present in the text or supplied by the editor, whether they appear as block notes in the main text area, at the foot of the page, at the end of the chapter or volume, in the margin, or in some other place. The `@type` attribute distinguishes the different types of annotations in use in a text. In a `@resp` attribute, the responsible subject for a note can be documented. Where possible, a note can be inserted in the text at the point at which its identifier or mark first appears. The location of the note may be documented using a `@place` attribute.

```
<p>Plenty of options', he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library<note n="1" place="foot" type="authorial">The National Library of Whoopledywhaa was found in 1886 with the acquisition of the library of the late King Anthony.</note>. It was titled 'While thou art here', by Sir Edmund Peckwood<note type="editorial" resp="EV">The manuscript reads 'Petwood'.</note>. While reading the first sentence, his placid expression turned to a certain je ne sais quoi: 'For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.'</p>
```

7

#### 3.4.2 Index Entries

Pre-existing indexes may be encoded as a `<list>` inside `<div>` in `<front>` or `<back>`, for example. On the other hand, new indexes can be generated from machine readable text when the location to be indexed is marked with `<index>` with a headword encoded as `<term>` .

.....

- 7 See section [3.6. Simple Links and Cross-References](#) in the TEI Guidelines for a full discussion of notes which are encoded not at the point of attachment but at the point of appearance, e.g. at the end of a chapter or a volume. See chapter [16. Linking, Segmentation, and Alignment](#) for mechanisms to encode multiple views of larger or heterogeneous spans of text. See section [17.3. Spans and Interpretation](#) for a discussion of advanced interpretive annotations.

```

<p>'Plenty of options', he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian
National Library<index>
  <term>Library, National</term>
</index>. It was titled 'While thou art here', by Sir Edmund Peckwood. While reading the first sentence, his
placid expression turned to a certain je ne sais quoi: 'For the first time in twenty-five years, Dr Burt Diddledygook
decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaaw<index>
  <term>Academy, Royal</term>
</index>.'</p>

```

The effect of this will be to generate an index entry for the terms 'Library' and 'Academy', referencing the location of the original **<index>** element.

8

### 3.4.3 Apparent Errors

Apparent errors in the text may be indicated using **<sic>** or corrected inside **<corr>**.

```

<p>It was titled 'While thou art here', by Sir Edmund <sic>Petwood</sic></p>

```

```

<p>It was titled 'While thou art here', by Sir Edmund <corr>Peckwood</corr></p>

```

Alternatively, the encoder may both record the original source text and provide a correction by using both **<sic>** and **<corr>** in either order wrapped in a **<choice>**.

```

<p>It was titled 'While thou art here', by Sir Edmund <choice>
  <corr>Peckwood</corr>
  <sic>Petwood</sic>
</choice></p>

```

The encoder may encode the degree of certainty associated with the intervention or interpretation using a **@cert** attribute and indicate the agency responsible for the intervention or interpretation, for instance an editor or transcriber, using **@resp**. The value of **@resp** is a pointer to an element in the document header that is associated with a person responsible for the intervention.

.....

8 See section [3.8.3 Index Entries](#) in the TEI Guidelines for a full discussion of the TEI encoding strategies applied to indexes.

```
<p>It was titled 'While thou art here', by Sir Edmund <choice>
  <corr cert="high" resp="#EV">Peckwood</corr>
  <sic>Petwood</sic>
</choice></p>
```

The attribute value '#EV' points to a **<name>** element in the **<teiHeader>**, for example in the **<respStmt>**.

```
<respStmt>
  <resp>editor</resp>
  <name xml:id="EV">Edward Vanhoutte</name>
</respStmt>
```

9

### 3.4.4 Regularisation and Normalisation

Standard or regularised forms for variant forms or non-standard spelling may be provided for a number of reasons. This is called regularisation or normalisation. The original, non-normalized form may be flagged using **<orig>**.

```
<p>It was titled 'While <orig>thou</orig> <orig>art</orig> here', by Sir Edmund Peckwood</p>
```

If the encoder wants to indicate that certain words have been normalised, which means modernisation of spelling in this example, **<reg>** may be used.

```
<p>It was titled 'While <reg>you</reg> <reg>are</reg> here', by Sir Edmund Peckwood</p>
```

Alternatively the encoder may decide to record both the original form **<orig>** and the regularised form **<reg>** wrapped inside a **<choice>**. In the case of the modernisation of spelling, an electronic text could thus serve as the basis of an old- or new-spelling edition.

.....

9 See [Section 6.4 Editorial Interventions](#) of [TBE Module 6: Primary Sources](#) for a fuller treatment of editorial interventions.

```
<p>It was titled 'While <choice>
  <orig>thou</orig>
  <reg>you</reg>
</choice> <choice>
  <orig>art</orig>
  <reg>are</reg>
</choice> here', by Sir Edmund Peckwood</p>
```

The **@resp** attribute may be used to specify the agency responsible for the regularisation or normalisation.

```
<p>It was titled 'While <choice>
  <orig>thou</orig>
  <reg resp="#EV">you</reg>
</choice> <choice>
  <orig>art</orig>
  <reg resp="#EV">are</reg>
</choice> here', by Sir Edmund Peckwood</p>
```

### 3.4.5 Additions, Deletions, and Omissions

Another editorial intervention in the text may be the documentation and creation of additions, deletions and omissions. When transcribing a source document, **<gap>** may be used to indicate a point where material has been omitted both because the material is illegible, invisible or inaudible in the source and because the editor or transcriber has decided to omit material for editorial reasons or as part of sampling practice. The reason for omission may be given in a **@reason** attribute. Sample values include *sampling*, *illegible*, *inaudible*, *irrelevant*, *cancelled*. Additional attributes like **@extent** and **@unit** may document the amount of characters, words, lines or any other unit omitted.

10

```
<p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). <gap reason="irrelevant" unit="words" extent="32"/>It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
```

.....

- 10 If the omission is an editorial policy decision, e.g. the systematic exclusion of marginal commentaries from an encoding, the full details of the policy should be documented in **<editorialDecl>** inside the **<encodingDesc>** of the TEI Header. See section [3.2. The Encoding Description](#) in [TBE Module 2: The TEI Header](#).



The `<gap>` element may appear as an empty element, but may also contain a description of the material omitted using `<desc>`.

```
<p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). <gap reason="irrelevant" unit="words" extent="32">
  <desc>Commentary on the founding charter of the RAW</desc>
</gap> It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
```

Where words or phrases of moderate lengths have been added or deleted in the copy text., this may be recorded using `<add>` and `<del>`. As with all TEI elements, information on the actual rendition of the additions and deletions can be provided in the global `@rend` attribute. Additionally, the place of the addition may also be recorded using `@place`. See section [3.1.1. Simple additions and deletions](#) for a detailed discussion of these elements and their attributes.

11

```
<p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in <add place="supralinear">late</add> September 1960 bang on noontime and Dr Burt was looking forward to a <del rend="overstrike">walk</del> <add place="infralinear">stroll</add> in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
```

12

Additions and deletions with a causal relationship may be grouped by the `<subst>` element.

```
<p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in <add place="supralinear">late</add> September 1960 bang on noontime and Dr Burt was looking forward to a <subst>
  <del rend="overstrike">walk</del>
  <add place="infralinear">stroll</add>
</subst> in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
```

.....

- 11 When an editor wants to mark his or her own additions as editorial interventions in the text `<corr>` or `<supplied>` should be used, not `<add>`. See Section [4. Editorial interventions](#) in [TBE Module 6: Primary Sources](#).
- 12 For longer passages `<addSpan/>` and `<delSpan/>` may be used. See section [3.1.2. Complex additions and Deletions](#) in [TBE Module 6: Primary Sources](#).

Where deletions in the copy text cannot be read with confidence, **<unclear>** should be used with the **@reason** attribute indicating that the difficulty of transcription is due to deletion. See Section [4.1. Unclear, supplied, omitted text](#) in [TBE Module 6: Primary Sources](#).

#### 4. Non-Textual Phenomena

Textual documents often include non-textual phenomena such as images and graphics (illustrations, diagrams, drawings, artwork...) . These non-textual phenomena serve different purposes: some are an integral part of the text, e.g. in comic books and graphic novels, others just function as illustrations to the text; some are essential for a good understanding of the text, others add very little to that text. The decision how to encode these non-textual materials is once more up to the encoder and the encoding policy in force.

From a structural point of view, images and graphics may be anchored to a particular point in the text. This inline location can be indicated by using the empty element **<graphic/>**. Typically, a **@url** attribute will reference the location of the visual information outside the XML document. This can be a local path or a reference to an online image or graphical file.

```
<p>Plenty of options', he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library. <graphic url="http://www.whoopledywhaa.info/library/facade.png"/> It was titled 'While thou art here', by Sir Edmund Peckwood. <graphic url="wtatcover.jpg"/> While reading the first sentence, his placid expression turned to a certain je ne sais quoi: 'For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</p>
```

Alternatively, encoded binary data representing an inline graphic or image may be embedded directly within the document. In this case the **<binaryObject>** element may be used containing some suitable binary format.

An image or a graphic will often be accompanied by associated text such as a caption, a label or a heading which may be encoded using **<head>**. More extensive comments or discussions on the figure or graphic may be given inside one or more **<p>** or **<ab>** elements. Both the graphic or figure **<graphic/>** and its associated text(s) (**<head>**, **<p>** or **<ab>**) are grouped in a wrapping **<figure>** element:

```

<figure>
  <graphic url="http://www.whoopledywhaa.info/library/facade.png"/>
  <head>The National Library of Whoopledywhaa.</head>
</figure>
<figure>
  <graphic url="wtatcover.jpg"/>
  <head>Figure 2:</head>
  <p>The cover of the first print edition of "While thou art here" by Sir Edmund Peckwood from the rare books collection of the National Library of Whoopledywhaa.</p>
</figure>

```

### Summary

The figure **<figure>** element is used to contain images, captions, and textual descriptions of the pictures. The images themselves are specified using the graphic **<graphic/>** element, whose **@url** attribute provides the location of an image and whose optional **@n** and **@xml:id** attribute provide numbering and identifying opportunities.

Figures consisting of several figures or sub-figures can be encoded with nesting **<figure>** elements:

```

<figure n="2">
  <figure n="2a">
    <graphic url="wtatcoverfront.jpg"/>
    <head>Front</head>
  </figure>
  <figure n="2b">
    <graphic url="wtatcoverback.jpg"/>
    <head>Back</head>
  </figure>
  <head>Figure 2:</head>
  <p>Front and back cover of the first print edition of "While thou art here" by Sir Edmund Peckwood from the rare books collection of the National Library of Whoopledywhaa.</p>
</figure>

```

Note, how in the previous example the different nesting figures are numbered in an **@n** attribute. This is one of the global attributes available to all TEI elements. For a discussion of this and other global attributes, see [1.5. Global Attributes \(p. 24\)](#).

For the purpose of reading devices that cannot represent images, e.g. reading software for the visually impaired, a description of the figure or graphic may be supplied in a **<figDesc>** element:

```
<figure>
  <graphic url="http://www.whoopledywhaa.info/library/facade.png"/>
  <head>The National Library of Whoopledywhaa.</head>
  <figDesc>The figure shows the front of the National Library of Whoopledywhaa with the two typical towers in the so called Whooply-Gothic style. The towers are 145 metres high and the facade of the building is 48 metres wide. The 16 windows in the front are made of recycled stained glass windows of the nearby Saint-Morkel's church which now serves as a swimming pool. </figDesc>
</figure>
```

13 14

## 5. Global Attributes

Just as any XML element, TEI elements can carry one or more attributes which provide additional information and function as their qualifiers and quantifiers. There are two kinds of attributes in the TEI world. Specific attributes are defined for specific elements only. They are documented in Appendix C. Elements of the TEI Guidelines. Global attributes, however, are optional attributes which are defined for every TEI element.

There are six global attributes:

- @xml:id** : provides a unique identifier for an element.
- @n** : provides a number or other label for an element. The number or label does not need to be unique within the document.
- @xml:lang** indicates the language of an element using a 'tag' generated according to [BCP 47](#).
- @rend** indicates how the element was rendered or presented in the source text.
- @rendition** points to a description of the rendering or presentation used for this element in the source text.
- @xml:base** provides a base URI reference with which applications can resolve relative URI references into absolute URI references.

.....

- 13 It's, again, up to the encoder to decide whether graphics consisting of large amounts of text, should be encoded as graphics containing the text or as text in which the graphic appears.
- 14 For more information on the treatment of non-textual phenomena in TEI, see [TBE Module 3 -- 4.2 Figures](#), and [14 Tables, Formulæ, and Graphics](#) in the TEI Guidelines.

## 5.1 @xml:id

The **@xml:id** attribute provides a unique identifier for the element bearing the attribute. The identifier must be unique in the whole XML document. If there is another element in the XML document bearing the same unique identifier as a value of this attribute, a validating XML parser will signal a syntax error. Conforming to the World Wide Web Consortium's XML Recommendations, the attribute value must be a legal name, which means that it must start with a letter or the underscore character and contain no characters other than letters, digits, hyphens, underscores, full stops, and certain combining and extension characters. The use of the colon in a unique identifier is forbidden as it has the specific purpose of indicating namespace prefixes in XML.

### Challenge

Which one of the following examples demonstrates a correct use of **@xml:id** and why?

1.

```
<p xml:id="p:1">For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</p>
<p xml:id="p:2">It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead.</p>
<p xml:id="p:2">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>
```

2.

```
<p xml:id="1">For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</p>
<p xml:id="2">It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead.</p>
<p xml:id="3">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>
```

3.

```
<p xml:id="p1">For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</p>
<p xml:id="p2">It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead.</p>
<p xml:id="p3">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>
```

Example 3 demonstrates a correct use of the **@xml:id** attribute: the attribute value is unique, it starts with a letter and contains no illegal characters. The attribute values in example 1 are not unique and the use of a colon is forbidden. The attribute values in example 2 start with a number, which is illegal.

## 5.2 @n

The **@n** attribute also provides an identifier for an element but its value doesn't need to be a legal XML name. This means that they don't have to be unique inside the XML document and they may start with and contain any character. Typically **@n** is used to number or label elements. All **@n** values in the following examples are legal:

```
<p n="1">For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.</p>
<p n="p2">It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead.</p>
<p n="paragraph 3">He hoped his fellow members of the Royal Academy weren't even going to notice his absence.</p>
```

15 16

## 5.3 @xml:lang

The language of the content of a given element may be documented as the value of an **@xml:lang** attribute. If it is not specified, the value is inherited from that of the immediately enclosing element. Therefore, it is simplest to specify the base language of a text on the **<TEI>** element and only specify the language of an element when necessary for a specific purpose and different from the base language.

.....

- 15 Although by no means mandatory, it often makes sense to enrich the structural units of a document (e.g. lines in a poem) with some sort of identification (in **@xml:id**) or reference mechanism (in **@n**). Of course, when dealing with complex and or long documents, this labelling could become a rather demanding task in itself. Fortunately, this job can be done automatically by an XML processor, which can identify the sequential position of one element within another in an XML document without any additional tagging. Instead of manually providing mechanical references for a long poem or collection of poems, you could as well instruct an XML processor to either enrich the TEI encoding and add **@xml:id** or **@n** attributes with appropriate values, or to automatically deduct such reference systems from your markup and present them while rendering the document (e.g. in an HTML version of a poem).
- 16 Both **@xml:id** and **@n** are used to express a reference system. This system can follow the text's canonical system, it may be derived from the structure of the electronic text, or it may be designed by the encoder. See Section [3.10.2 Creating New Reference Systems](#) in the TEI Guidelines for a discussion of the latter.

```
<p xml:lang="en">'Plenty of options', he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library. It was titled 'While thou art here', by Sir Edmund Peckwood. While reading the first sentence, his placid expression turned to a certain <foreign xml:lang="fr">je ne sais quoi</foreign>: 'For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa.'</p>
```

17

#### 5.4 @rend

The **@rend** attribute is used to document information about the physical appearance of the text in the source. In the following example, it is used to indicate that the title, the French phrase and the name of the Royal Academy are printed in italics:

```
<p xml:lang="en">'Plenty of options', he thought, sat on a bench and opened the book he had taken from the Whoopledywhaaian National Library. It was titled <title type="m" rend="italics">While thou art here</title>, by Sir Edmund Peckwood. While reading the first sentence, his placid expression turned to a certain <hi rend="italics">je ne sais quoi</hi>: 'For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the <name type="organisation" rend="italics">Royal Academy of Whoopledywhaa</name>.'</p>
```

#### 5.5 @rendition

Whereas the **@rend** attribute documents the appearance of text locally, i.e. attached to an element, the **@rendition** attribute points to a description of the rendering or appearance in the **teiHeader** **<teiHeader>**, more specifically inside a **<tagsDecl>** inside **<encodingDesc>**. This is done in free text or using a formal language inside **<rendition>**. This way, only one description of the rendering must be given to which the **@rendition** attributes refer. The advantage of this system becomes clear when both **@rendition** and **@rend** are used for occurrences of a given element. While the former refers to an overall description of the appearance of that element in the source, the latter documents the local deviation from that generally imposed rendition.

In the following example, we see a description of the overall rendering of **<hi>** in the **<tagsDecl>** inside **<teiHeader>**, and a documentation of the deviation of that overall rendering in the third occurrence of **<hi>** in the text. The **@gi** attribute of **<tagUsage>** names the elements for which the rendition described in **<rendition>** is documented. The formal namespace in which the tags described in **<tagUsage>** are defined, must be specified in the **@name** attribute of a surrounding **<namespace>** element. The value of the **@render** attribute of **<tagUsage>**

.....

17 The values for the **@xml:lang** attribute must be constructed in a uniform way as explained in Section [vi.1. Language identification](#) of the TEI Guidelines.

refers to `<rendition>` by way of the latter's `@xml:id` attribute. This way, all `<hi>` elements inside `<text>` have 'italic' as their default style. The deviation to that rule is articulated in the third occurrence of `<hi>` by making use of the `@rend` attribute.

```
<teiHeader> <!-- ... --> <encodingDesc>
  <tagsDecl>
    <rendition xml:id="IT">Italic</rendition>
    <namespace name="http://www.tei-c.org/ns/1.0">
      <tagUsage gi="hi" render="#IT"/>
    </namespace>
  </tagsDecl>
</encodingDesc> <!-- ... --> </teiHeader>
<text>
  <body>
    <p>Plenty of options', he thought, sat on a bench and opened the book he had taken from the
    Whoopledywhaaian National Library. It was titled <hi>While thou art here</hi>, by Sir Edmund Peckwood.
    While reading the first sentence, his placid expression turned to a certain <hi>je ne sais quoi</hi>: 'For the
    first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the <hi
    rend="roman">Royal Academy of Whoopledywhaa</hi>.'</p>
  </body>
</text>
```

## 6. Summary

After this overview of the most common structures and elements of a TEI document, it is time to put them all together:



```

<TEI>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>The Strange Adventures of Dr. Burt Diddledycook: a machine-readable transcription</title>
        <respStmt>
          <resp>editor</resp>
          <name xml:id="EV">Edward Vanhoutte</name>
        </respStmt>
      </titleStmt>
      <publicationStmt>
        <p>Not for distribution.</p>
      </publicationStmt>
      <sourceDesc>
        <p>Transcribed from the diaries of the late Dr. Roy Offire.</p>
      </sourceDesc>
    </fileDesc>
    <encodingDesc>
      <tagsDecl>
        <rendition xml:id="IT">Italic</rendition>
        <namespace name="http://www.tei-c.org/ns/1.0">
          <tagUsage gi="hi" render="#IT"/>
        </namespace>
      </tagsDecl>
    </encodingDesc>
  </teiHeader>
  <text>
    <front>
      <div type="dedication">
        <p>In memory of Lisa Wheeman.</p>
      </div>
    <titlePage>
      <docAuthor>Roy Offire</docAuthor>
      <docTitle>
        <titlePart type="main">The Strange Adventures of Dr. Burt Diddledycook</titlePart>
        <titlePart type="sub">Wanderings in the life of a buoyant academic</titlePart>
      </docTitle>
      <byline>Transcribed from the diaries.</byline>
      <docEdition>First Edition</docEdition>
      <docImprint><pubPlace>Kirkcaldy</pubPlace>, <publisher>Bucket Books</publisher>,
        <docDate>1972</docDate>
      </docImprint>

```

```

</titlePage>
<div type="contents">
  <head>Table of Contents</head>
  <list>
    <item>I. The Decision</item>
    <item>II. The Fuss</item>
    <item>III. The Celebration</item>
  </list>
</div>
</front>
<body>
  <p n="1">For the first time in <date from="1935" to="1960">twenty-five years</date>, <choice>
    <abbr type="title">Dr</abbr>
    <expan>Doctor</expan>
  </choice> Burt Diddledygoon decided not to turn up to the annual meeting of the Royal Academy of
  Whoopledywhaa<index>
    <term>Academy, Royal</term>
  </index> (<abbr type="acronym">RAW</abbr>). <gap reason="irrelevant" unit="words" extent="32">
    <desc>Commentary on the founding charter of the RAW</desc>
  </gap> It was a sunny day in <date notBefore="1960-09-15" notAfter="1960-09-30"><add
  place="supralinear">late</add> September 1960</date> bang on <time
  when="12:00:00">noontime</time> and <abbr type="title">Dr</abbr> Burt was looking forward to a
  <subst>
    <del rend="overstrike">walk</del>
    <add place="infralinear">stroll</add>
  </subst> in the park instead. He hoped his fellow members of the <abbr type="acronym">RAW</abbr>
  weren't even going to notice his absence.</p>
  <p n="2">Or worse, what would happen when another Academy member had decided to go for a stroll in
  the park instead? He quickly thought up several possible plans:</p>
  <list>
    <item>hide behind a tree and duck</item>
    <item>catch the duck as subject material for a speech on the annual meeting</item>
    <item>be frank, meet his colleague, and <list>
      <item>1. pat him on the shoulder</item>
      <item>2. tell a joke</item>
      <item>3. hand him the duck</item>
      <item>4. offer him a sip from his <measure type="volume" quantity="2.5" unit="litre"
      commodity="coca-cola">2.5 l bottle of coke</measure></item>
      <item>5. pull his beard</item>
    </list></item>
  </list>

```

```

<p n="3">Or maybe he could still announce his absence from the meeting by sending an antedated letter of
apology to <address>
  <name type="person">Professor M. Orkelidius</name>
  <name type="organisation">Royal Academy of Whoopledywhaa</name>
  <street>Queenstreet 81</street>
  <postCode>TB90 00E</postCode>
  <name type="city">Whoopledywhaa</name>
</address></p>
<p n="4" xml:lang="en">'Plenty of options', he thought, sat on a bench and opened the book he had taken
from the Whoopledywhaaian National Library<index>
  <term>Library, National</term>
</index><note n="1" place="foot" type="authorial">The National Library of Whoopledywhaa was found
in 1886 with the acquisition of the library of the late King Anthony.</note>. It was titled <title type="m"
rend="italics">While <choice>
  <orig>thou</orig>
  <reg resp="#EV">you</reg>
</choice> <choice>
  <orig>art</orig>
  <reg resp="#EV">are</reg>
</choice> here</title>, by Sir Edmund
  <choice>
    <corr>Peckwood</corr>
    <sic>Petwood</sic>
  </choice>
<note type="editorial" resp="EV">The manuscript reads 'Petwood'.</note>. <figure n="2">
  <figure n="2a">
    <graphic url="wtatcoverfront.jpg"/>
    <head>Front</head>
  </figure>
  <figure n="2b">
    <graphic url="wtatcoverback.jpg"/>
    <head>Back</head>
  </figure>
<head>Figure 2:</head>
<p>Front and back cover of the first print edition of "While thou art here" by Sir Edmund Peckwood
from the rare books collection of the National Library of Whoopledywhaa.</p>
</figure>While reading the first sentence, his placid expression turned to a certain <hi xml:lang="fr"
rend="italics">je ne sais quoi</hi>: 'For the first time in twenty-five years,
  <choice>
    <abbr type="title">Dr</abbr>
    <expan>Doctor</expan>
  </choice>

```

```
Burt Diddledygook decided not to turn up to the annual meeting of the
  <name type="organisation" rend="italics">Royal Academy of Whoopledywhaa</name>
<index>
  <term>Academy, Royal</term>
</index>.</p>
</body>
<back>
  <div type="colophon">
    <p>Typeset in Haselfoot 37 and Henry 8. Printed and bound by Whistleshout, South Africa.</p>
  </div>
</back>
</text>
</TEI>
```

## 7. What's next?

You have reached the end of this tutorial module providing an introduction to the TEI and text encoding for the humanities. You can now either

- proceed with other TEI by Example modules
- have a look at the examples section for the general structure and elements module
- take an interactive test. This comes in the form of a set of multiple choice questions, each providing a number of possible answers. Throughout the quiz, your score is recorded and feedback is offered about right and wrong choices. Can you score 100%. Test it here.