The TEI, TEI Structure and Core Elements

James Cummings

April 2009

An Introduction to the TEI

The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts chiefly in the humanities, social sciences and linguistics.

Why the TEI?

The TEI provides

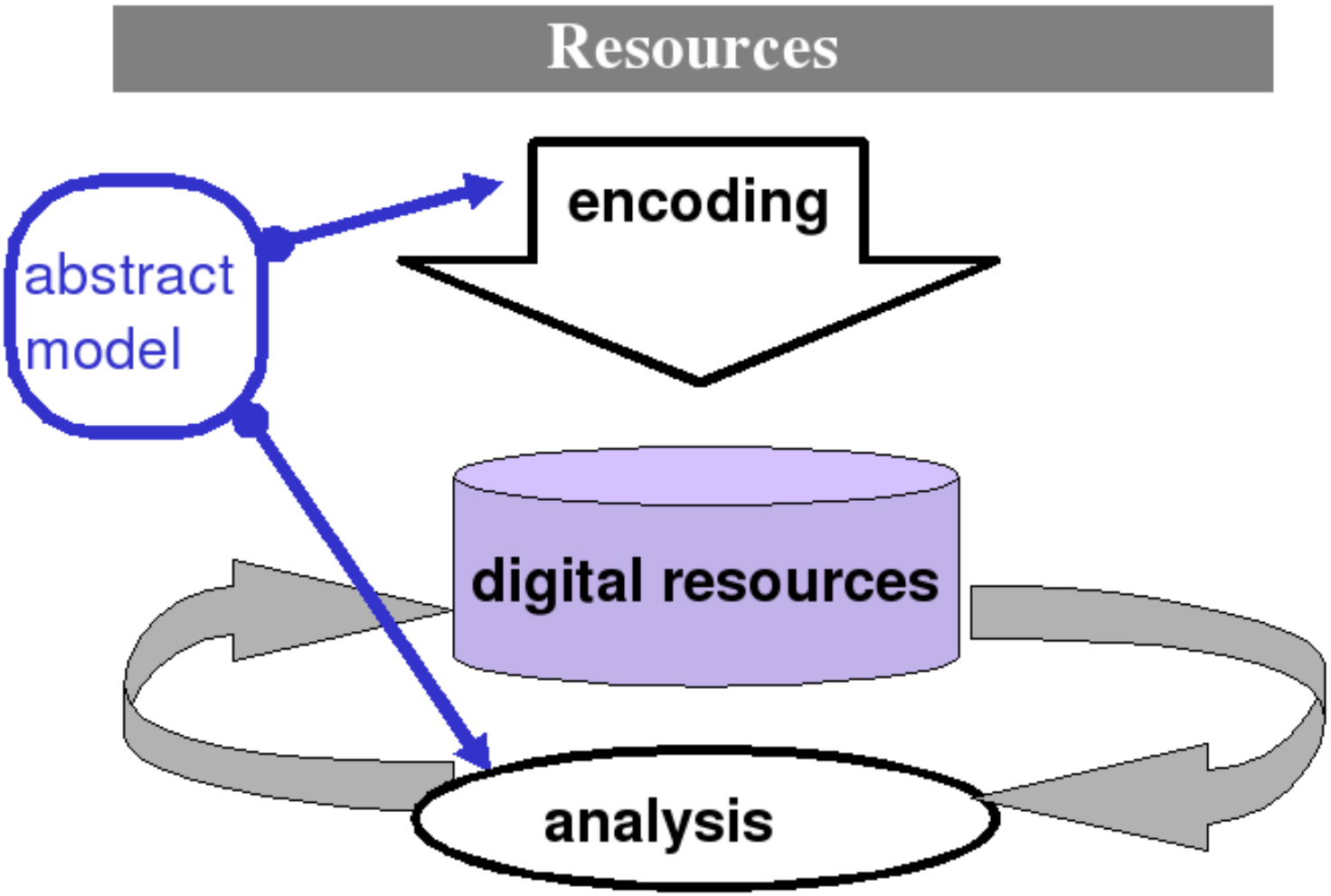
- a language-independent framework for defining markup languages
- a very simple consensus-based way of organizing and structuring textual (and other) resources...
- ... which can be enriched and personalized in highly idiosyncratic or specialised ways
- a very rich library of existing specialised components
- an integrated suite of standard stylesheets for delivering schemas and documentation in various languages and formats
- a large and active open source style user community

Relevance

Why would you want those things?

- because we need to interchange resources
  - between people
  - (increasingly) between machines
- because we need to integrate resources
  - of different media types
  - from different technical contexts
- because we need to preserve resources
  - cryogenics is not the answer!
  - we need to preserve metadata as well as data

The virtuous circle of encoding



The scope of intelligent markup

Even within the original scope of the TEI we have

- basic structural and functional components
- diplomatic transcription, images, annotation
- links, correspondence, alignment
- data-like objects such as dates, times, places, persons, events (named entity recognition)
- meta-textual annotations (correction, deletion, etc)
- linguistic analysis at all levels
- contextual metadata of all kinds
- ... and so on and so forth

Is it possible to delimit encyclopaedically all possible kinds of markup?

Reasons for attempting to define a common framework

- re-usability and repurposing of resources
- modular software development
- lower training costs
- 'frequently answered questions' — common technical solutions for different application areas

The TEI was designed to support multiple views of the same resource

Conformance issues

A document is TEI Conformant if and only if it:

- is a well-formed XML document
- can be validated against a TEI Schema, that is, a schema derived from the TEI Guidelines
- conforms to the TEI Abstract Model
- uses the TEI Namespace (and other namespaces where relevant) correctly
- is documented by means of a TEI Conformant ODD file which refers to the TEI Guidelines

or if it can be transformed automatically using some TEI-defined procedures into such a document (it is then considered TEI-conformable).

Standardization should not mean 'Do what I do', but rather 'Explain what you do in terms I can understand'

Default Text Structure

All TEI documents are structured in a particular manner. This section attempts to describe the different variations on this as briefly as possible.

Structure of a TEI Document

There are two basic structures of a TEI Document:

- `<TEI>` (TEI document) contains a single TEI-conformant document, comprising a TEI header and a text, either in isolation or as part of a teiCorpus element.
- `<teiCorpus>` contains the whole of a TEI encoded corpus, comprising a single corpus header and one or more TEI elements, each containing a single text header and a text.

TEI basic structures (1)

```
<teiCorpus>
 <teiHeader>
<!-- required -->
 </teiHeader>
 <TEI>
<!-- required -->
 </TEI>
</teiCorpus>
```

TEI basic structures (2)

```
<TEI>
 <teiHeader>
<!-- required -->
 </teiHeader>
 <facsimile>
<!-- optional, new in TEI P5 -->
 </facsimile>
 <text>
<!-- required if no facsimile -->
 </text>
</TEI>
```

<text>

What is a text?

- A text may be unitary or composite
  - unitary: forming an organic whole
  - composite: consisting of several components which are in some important sense independent of each other
- a unitary text contains
  - optional front matter
  - <body> (required)
  - optional back matter

<text>

What is a text?

- A text may be unitary or composite
  - unitary: forming an organic whole
  - composite: consisting of several components which are in some important sense independent of each other
- a unitary text contains
  - optional front matter
  - <body> (required)
  - optional back matter

Composite texts

A composite text contains

- optional front matter
- <gr oup> (required)
- optional back matter

A corpus is a collection of text and header pairs. It has its own header.
<gr oup> tags may self-nest.

TEI text structure (1)

```
<text>
 <front>
<!-- optional -->
 </front>
 <body>
<!-- required -->
 </body>
 <back>
<!-- optional -->
 </back>
</text>
```

TEI text structure (2)

```
<text>
 <front>
<!-- ... -->
 </front>
 <group>
  <text>
   <body>
    <p>...</p>
   </body>
  </text>
 </group>
 <back>
<!-- ... -->
 </back>
</text>
```

Another Grouped Text Example

```
<TEI>
 <teiHeader>
<!-- header information for the whole collection -->
 </teiHeader>
 <text>
<!-- optional front matter -->
  <group>
   <text>
<!-- optional front matter -->
    <body>
<!-- First Body -->
    </body>
   </text>
   <text>
<!-- optional front matter -->
    <body>
<!-- Second Body-->
    </body>
   </text>
  </group>
 </text>
</TEI>
```

The Imaginary Punch Project

- Punch is a famous English humorous journal, published regularly between 1841 and 1992: see http://www.punch.co.uk/historyofpunch.html.
- The IPP plans to make available fully marked up texts of the journal, in conjunction with page images...
  - for social historians
  - for librarians
  - for linguists
- How will the TEI help? And which parts of the TEI will we use?
- Although we won't have time to accomplish this in this workshop, we've provided the Punch text and images in your materials directory.

Punch example page 1

JULY 1, 1914.] PUNCH, OR THE LONDON CHARIVARI. 1

PROGRESS.

["Giving evidence recently before a Select Committee of the House of Commons, Miss C. E. Collet, of the Home Office, said the commercial laundry was killing the small hand laundry."—*Evening News.*]

THE little crafts! How soon they die!
In cottage doors no shuttle clicks;
The hand-loom has been ousted by
A large concern with lots more sticks.

The throb of pistons beats around;
Great chimneys rise on Thames's banks;
The same phenomena are found
In Sheffield (Yorks) and Oldham (Lancs).

No longer now the housewife makes
Her rare preserves, for what 's the good?
The factory round the corner fakes
Raspberry jam with chips of wood.

'Tis so with what we eat and wear,
Our bread, the boots wherein we splosh;
'Tis so with what I deemed most fair,
Most virginal of all—the Wash.

'Tis this that chiefly, when I chant,
Fulfils my breast with sighs of ruth,
To think that engines can supplant
The Amazons I loved in youth.

That not with tender care, as erst
By spinster females fancy-free,
These button-holes of mine get burst
Before the shift comes back to me;

That mere machines, and not a maid
With fingers fatuously plied,
The collars and the cuffs have frayed
That still excoriate my hide;

That steam reduces to such states
What once was marred by human skill;
That socks are sundered from their mates
By means of an electric mill;

That not by Cupid's coy advance
(Some crone conniving at the fraud),
But simply by mechanic chance,
I get this handkerchief marked "Maud."

This is, indeed, a striking change;
I sometimes wonder if the world
Gets better as the skies grow strange
With coils of smoke about them curled.

If the old days were not the best
Ere printed formulas conveyed
Sorrow about that silken vest
For all eternity mislaid;

Ere yet the unwieldy motor-van
Came clattering round the kerbstone's brink,
Its driver dreaming some new plan
To make my mauve pyjamas shrink.

EVOE.

Punch example page 2

have it in the guide-book)—one young man, "sobbing, buries his head in his hands, upsetting thereby a dish of fruit." As for *Potiphar*, it failed to stir the sombre depths of his abysmal boredom, but his wife, whose ennui had hitherto been of the most profound, began to sit up and take notice, and at the end of the dance she sent for *Joseph* and supplemented his rather exiguous costume with a gross necklace of jewels, letting her hand linger awhile on his bare neck. Already, it will be seen, she was intrigued with the "unknown divine." *Joseph*, on the contrary, received her attentions without *empressement*.

In the next scene—after a rather woolly and unintelligible interlude—we see *Joseph* retiring to his couch in an alcove behind the place where the banqueting-table had been. You will judge how urgent was the lady's keenness to probe the mysteries of his divine nature when I tell you that she could not wait till the morning to pursue her enquiries, but must needs visit him in his chamber at dead of night, and wearing the one garment of the hour. At first, still half dreaming, he mistakes her for an angel (he had already seen one in his sleep), but subsequently, growing suspicious, he repels her with a dignified disdain. For I must tell you that, whatever the guide-book may allege about the loftiness of her designs, the music gave her away. It reverted, in fact, to the motive of those passages which had already accompanied and illustrated the nuptial dance, the dance (as Herr TIESSEN calls it) of "burning Love-longing."

At this juncture, *Potiphar* and his minions break upon the scene. His wife, after denouncing *Joseph*, is distracted between passion of hatred and passion of love, and there is some play (reminding one of *L'Après-midi d'un Faune*) with the purple cloak which *Joseph* had discarded. Presently she eludes her dilemma by fainting.

Meanwhile it has been the work of a moment to order up a brazier, a pair of pincers, a poker, a headsman and an axe. The instruments of torture waste no time in getting red-hot; and we anticipate the worst. *Joseph*, however, who has ignored these preparations and maintained an attitude of superbly indifferent aloofness, suddenly becomes luminous under great pressure of limelight; and most of the cast, including a ballet of female dervishes, are abashed to the ground.

Now appears, on the open-work entresol at the back of the stage, an archangel. The guide-book is in error where it says that he glides downwards on a shaft of light radiating from a star. As a matter of fact he walks down the main staircase to the ground floor. Approaching *Joseph* he takes him by the hand and "leads him heavenwards" by the same flight of steps; and we are to understand that, in the opinion of Herr STRAUSS, the boy's subsequent career, as recorded in the Hebraic Scriptures, may be treated as negligible.

I should like, in excuse of my own flippancy, to assume the same detachment, and to regard this ballet-theme as having practically no relation whatever to Biblical history, but being just one of many themes out of Oriental lore, mostly secular, that lend themselves to the drama of disappointed passion. My only serious protest is against the hypocrisy which pretends, with regard to *Potiphar's Wife*, to see a spiritual significance in what is mere vulgar animalism.

I ought, by the way, to have said that, in a spasm of chagrin, she chokes herself with the pearl necklace which lent the only touch of superfluity to her night attire, and was carried out—but not up the main staircase. Thus ends this sordid tragedy that so well illustrates that quality in Herr STRAUSS to which my guide refers when he speaks of his realization of a "poignant longing for divine cheerfulness." O. S.

"EXCUSE ME, SIR, BUT WOULD YOU LIKE TO BUY A NICE LITTLE DAWG?"

"NO, THANKS VERY MUCH. HE LOOKS AS THOUGH HE WOULD BITE."

"'E WON'T BITE YER *IF YOU BUY 'IM*, GUV'NER."

ENIGMA.

My love to me is cold,
And no more seeks my gaze;
I wonder why!
The smile of welcome that I loved of old
No longer lights her eye.

One little week ago
I asked no surer guide than Cupid's chart;
I said, "Your eyes reveal the depths below,
And I can read your heart."

She let her shy gaze fall,
And smiling asked, "Is then my face a screed,
My brow an open love-letter, where all
The world my thoughts may read?"

Said I, "The world, I'll vow,
Is blind! Myself alone may see the signs,
And know the message written on your brow:
I read between the lines."

My dear to me is cold:
Gone somewhere is the love-light from her eye;
And, when our ways meet, stately she doth hold
Her course. I wonder why.

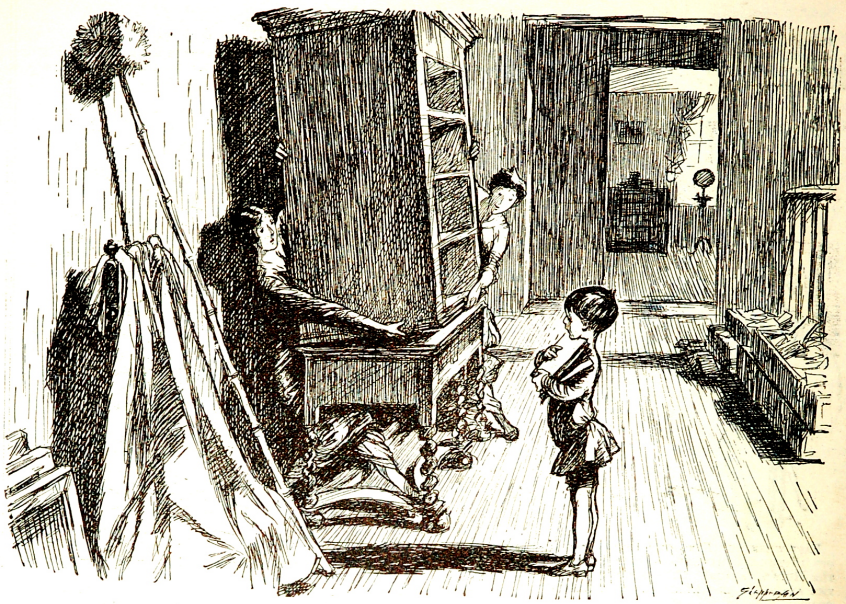
"Curiously, the Australian Minister of Defence in the last Parliament bore the same name as the Prime Minister in that which has just been dissolved."
Westminster Gazette.

A similar curious coincidence happened in England, the War Minister in the last Parliament bearing the same name as the present Lord Chancellor.

"MEN FOR THE ANTARCTIC.
105 CANADIAN DOGS TO GO WITH SIR E. SHACKLETON."
Daily Express.

A gay lot, these Canadians.

Punch example page 3



Alan (to his mother, who is busy with a heavy house-cleaning). "PLEASE, MOTHER, READ ME A STORY."

THE WALKERS.

THERE were eight pretty walkers who went up a hill;
They were Jessamine, Joseph and Japhet and Jill,
And Allie and Sally and Tumbledown Bill,
And Farnaby Fullerton Rigby.

They were all in good training and all of them keen,
And their chief wore a coat and a waistcoat of green;
He was always a proud man and kept himself clean,
Did Farnaby Fullerton Rigby.

They intended to lunch when they got to the top
On a sandwich apiece and a biscuit and chop.
The provisions were carefully bought in a shop
By Farnaby Fullerton Rigby.

They were jesters of merit—the sort who can poke
Funny tales in your ribs till you splutter and choke;
But the best of the lot at a jibe or a joke
Was Farnaby Fullerton Rigby.

It was ten of the clock when the walking began,
And they started with Tumbledown Bill in the van;
And the rear was brought up by that excellent man,
By Farnaby Fullerton Rigby.

They went off at a pace I am bound to deplore,
For they did twenty yards in a minute or more
And a yard or two over, a capital score
For Farnaby Fullerton Rigby.

They had all that pedestrians fairly can ask:
Smooth roads, sunny weather and beer in a cask,
And a friend who could teach them to stick to their task,
Viz.: Farnaby Fullerton Rigby.

Yet I somehow suppose that they hadn't the knack,
For in spite of it all they have never come back,
And I own that the future looks dismally black
For Farnaby Fullerton Rigby.

Now the walkers who seem to be stuck on the hill,
They are Jessamine, Joseph and Japhet and Jill,
And Allie and Sally and Tumbledown Bill,
And Farnaby Fullerton Rigby. R. C. L.

King Peter of Servia.

(From *The Daily Mirror.*)

"The proclamation, however, as given in a later message, reads thus:—To My Beloved People: As I shall be prevented by illness from exercising my royal power for some time, I order, by Article 69 of the Constitution, that so long as my cure lasts the Crown Prince Alexander shall govern in my name. On this occasion I recommend my dear fatherland to the care of the Almighty.
(Signed) PETER."

"On this occasion" is perhaps a little invidious.

Two consecutive books in *The Western Daily Press* list of publications received:—

"RING STRATEGY AND TACTICS.
CHARLES DICKENS IN CHANCERY."

The boxing boom continues.

Looking at Punch, what do we need to mark up?

- issue information and page number for reference purposes
- "chunks" or divisions of text, which may contain a picture, a poem, some prose, some drama, or a combination
- within the chunks, we can identify formal units such as
  - a picture, a caption
  - stanzas, lines
  - paragraphs
  - speeches and stage-directions
- and more...

Macrostructure

All the issues of *Punch* for one year make up a volume. We could regard the volume as a single <text>, and each issue as a within it. Or we could use the <group> element:

```
<text xml:id="v147">
 <front>
<!-- introductory materials for volume 147 here -->
 </front>
 <group>
  <text xml:id="I1914-07-01">
   <body>
<!-- first issue (1 July) -->
   </body>
  </text>
  <text xml:id="I1914-07-15">
   <body>
<!-- second issue (15 July) -->
   </body>
  </text>
<!-- etc...  -->
 </group>
 <back>
<!-- volume index, appendix etc.  -->
 </back>
</text>
```

TEI tags for the high level structure

We will treat each issue as a single `<text>` element, and each identifiable chunk within it as a `<div>` element of a particular type (e.g. cartoon, verse, prose)
For example, page 1 has two divisions,

```
<pb n="1"/>
<div type="cartoon">
 <p>....</p>
</div>
<div type="poem">
 <head>Progress</head>
 <lg>
  <l>....</l>
 </lg>
</div>
```

More high level structure

page 2 also has two, of different types:

```
<pb n="2"/>
<div type="prose">
 <head>The enchanted castle</head>
 <p>....</p>
</div>
<div type="snippet">
 <head>Correspondence</head>
 <p>....</p>
</div>
```

Why divisions rather than pages?

Because a division can start on one page (page 5 for example) and finish on another (page 6)
We use an empty element<pb> to mark the boundary between pages, rather than enclosing each page in a .

```
<pb n="5"/>
<div type="cartoon">
 <p>...</p>
</div>
<div type="review">
 <head>Egypt in Venice</head>
 <p>...</p>
 <pb n="6"/>
 <p>...</p>
</div>
<div type="cartoon">
 <p>...</p>
</div>
```

Divisions can contain divisions...

```
<div type="snippets">
 <div type="snippet">
  <p>Curiously....Chancellor</p>
 </div>
 <div type="snippet">
  <p>Men for the Antarctic...  Canadians</p>
 </div>
</div>
```

- TEI also provides division elements with names that indicate their degree of nesting (<div1>, <div2> etc.) which some people prefer
- Divisions must always tessellate: once "down" a level, you cannot pop "up" again within the same division.

More about divisions

`<div>`

- generic, hierarchic subdivisions, each incomplete
- the *@type* attribute is used to label a particular level e.g. as 'part' or 'chapter'
- the *@n* attribute gives a particular division a name or number
- the *@xml:id* attribute gives a particular division a unique identifier

More about divisions

`<div>`

- generic, hierarchic subdivisions, each incomplete
- the *@type* attribute is used to label a particular level e.g. as 'part' or 'chapter'
- the *@n* attribute gives a particular division a name or number
- the *@xml:id* attribute gives a particular division a unique identifier

More about divisions

`<div>`

- generic, hierarchic subdivisions, each incomplete
- the *@type* attribute is used to label a particular level e.g. as 'part' or 'chapter'
- the *@n* attribute gives a particular division a name or number
- the *@xml:id* attribute gives a particular division a unique identifier

More about divisions

`<div>`

- generic, hierarchic subdivisions, each incomplete
- the *@type* attribute is used to label a particular level e.g. as 'part' or 'chapter'
- the *@n* attribute gives a particular division a name or number
- the *@xml:id* attribute gives a particular division a unique identifier

Divisions may have heads and trailers

```
<div>
 <head>Chapter 1</head>
 <p>
<!-- content of the div -->
 </p>
 <trailer>...</trailer>
</div>
```

Numbered and unnumbered divisions

The level can be made explicit by using 'numbered' divs (div1, div2).
Opinions vary:
`<div1>` vs. `<div n="1">`

- numbered: the number indicates the depth of this particular division within the hierarchy, the largest such division being ‘div1’, any subdivision within it being ‘div2’, etc.
- unnumbered: nest recursively to indicate their hierarchic depth. (And computers can count very well!)

The two styles must not be combined within a single `<front>`, `<body>`, or `<back>` element.
N.B. Divisions always tessellate

Groups vs Floating Texts

The `<group>` element should be used to represent a collection of independent texts which is to be regarded as a single unit for processing or other purposes.
`<floatingText>` contains a single text of any kind, whether unitary or composite, which interrupts the text containing it at any point and after which the surrounding text resumes.

Floating Text (1)

s must tesselate over the entire text

```
<div1>
 <div2>
<!-- content -->
 </div2>
 <div2>
<!-- content -->
 </div2>
</div1>
```

is valid, while

```
<div1>
<!-- content -->
 <div2>
<!-- content -->
 </div2>
<!-- content -->
</div1>
```

is not valid.

Floating Text (2)

In the second case, div2 is a 'floating' text and its content must be encoded using the `<floatingText>` element.
The `<floatingText>` element is a member of the model.divPart class, and can thus appear within any division level element in the same way as a paragraph.

Floating Text Example

```
<p>She was thus ruminating, when a Gentleman enter'd the
Room, the Door being a jar... calling for a Candle, she
beg'd a thousand Pardons, engaged him to sit down, and
let her know, what had so long conceal'd him from her
Correspondence. </p>
<pb n="5"/>
<floatingText>
 <body>
  <head>The Story of <hi>Captain Manly</hi>
  </head>
  <p>
<!-- Captain Manly's store here -->
  </p>
 </body>
</floatingText>
<pb n="37"/>
<p>The Gentleman having finish'd his Story ...
<!-- more -->
</p>
```

Virtual divisions

Where the whole of a division can be automatically generated, for example because it is derived from another part of this or another document, an encoder may prefer not to represent it explicitly but instead simply mark its location by means of a processing instruction, or by using the special purpose <divGen> element:

```
<front>
 <divGen type="toc"/>
 <div>
  <head>Preface</head>
  <p>...</p>
 </div>
</front>
```

(intended primarily for use in document production or manipulation, rather than in transcription of pre-existing material)

Back to Punch

Page 3 contains a figure and a dialogue...

```
<div type="cartoon">
 <figure>
  <head>When the ships come home</head>
  <figDesc>A man in Turkish dress lounges on a sofa,
     smoking a cigarette and consulting a book labelled
     "Naval ledger".  Another man, in traditional Greek
     costume, stands beside him, also reading a
     notebook.</figDesc>
  <graphic url="materials/Punch/XML/Graphics/003.png"/>
 </figure>
 <sp>
  <speaker>Greece.</speaker>
  <p> Isn't it time we started fighting again?</p>
 </sp>
 <sp>
  <speaker>Turkey.</speaker>
  <p> Yes, I daresay.  How soon could you begin?</p>
 </sp>
 <sp>
  <speaker>Greece.</speaker>
  <p> Oh, in a few weeks.</p>
 </sp>
 <sp>
```

Punch example page 3

PUNCH, OR THE LONDON CHARIVARI.—JULY 1, 1914.

NAVAL LEDGER

WHEN THE SHIPS COME HOME.

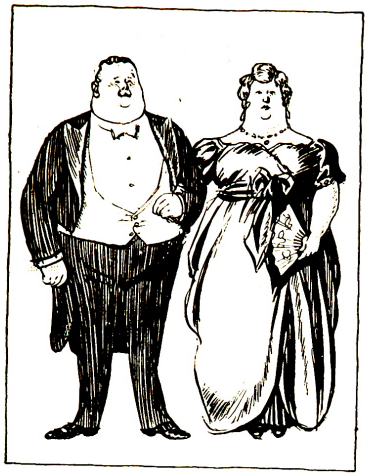
GREECE. "ISN'T IT TIME WE STARTED FIGHTING AGAIN?"
TURKEY. "YES, I DARESAY. HOW SOON COULD YOU BEGIN?"
GREECE. "OH, IN A FEW WEEKS."
TURKEY. "NO GOOD FOR ME. SHAN'T BE READY TILL THE AUTUMN."

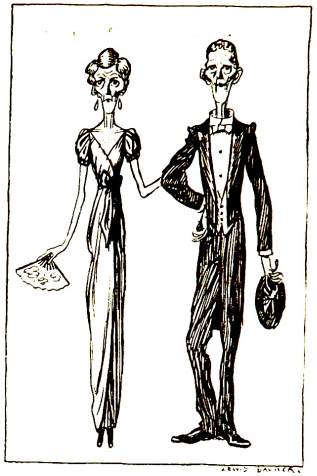
For example...

The militants' tariff (on Page 15) contains headings, paragraphs, and a table...

```
<div type="prose">
 <head>THE MILITANTS' TARIFF.</head>
 <head rend="right">Etna Lodge, W.</head>
 <p>Mrs. Bangham Smasher, having entered into
partnership with the
    Misses Burnham Blazer, as General Agents of
Destruction, begs to
    inform the public that the firm will be prepared to
execute
    commissions of all kinds, at the shortest notice, on
the very moderate
    terms given below : -- </p>
 <table>
  <row role="label">
   <cell/>
   <cell>£</cell>
   <cell>s.</cell>
   <cell>d.</cell>
  </row>
  <row>
   <cell>For breaking windows, per window ...</cell>
   <cell>0</cell>
   <cell>7</cell>
```

Punch example page 15





GREAT ECONOMY EFFECTED BY CO-OPERATION IN ADVERTISEMENT.

THE MILITANTS' TARIFF.

Etna Lodge, W.

Mrs. Bangham Smasher, having entered into partnership with the Misses Burnham Blazer, as General Agents of Destruction, begs to inform the public that the firm will be prepared to execute commissions of all kinds, at the shortest notice, on the very moderate terms given below:—

	£	s.	d.
For breaking windows, per window	0	7	6
For howling, kicking, or biting during service in church, per howl, kick, or bite	0	10	6
For sitting on doorsteps of obnoxious persons, per hour, if fine	0	15	0
For sitting on doorsteps of obnoxious persons, per hour, if wet	1	1	0
For damaging golf greens, per green	1	11	6
For throwing shoes at magistrates in court, according to size and weight of shoe, from	2	2	0
For beating officials connected with gaols	3	3	0
For slashing and hacking valuable pictures, from	7	7	0
For bombs not intended to explode	8	8	0
For burning down a house, according to value and social position of owner, from	10	0	0
For insulting exalted Personages, per insult	10	10	0
For burning down a modern red-brick church	15	15	0
For burning down a specially valuable and interesting ancient one (eleventh and twelfth centuries extra)	21	0	0
For bombs warranted to destroy an ordinary church	30	0	0
For bombs suited to wreck really superior buildings, such as Westminster Abbey and St. Paul's	50	0	0
For disturbing public meetings and the general harassing and annoyance of all peaceable and decent people	No charge.		

Bangham Smasher, Burnham Blazer & Co. beg to assure their patrons that all the choppers, hammers, bombs, stones, etc., employed in their business are of the very best quality, and only refined paraffin and wax matches will be used in burning down any building.

Being in a position to offer such exceptional advantages they trust to receive a large measure of support in their elevating and enlightening work.

If none of the above is found suitable to the needs of intending clients, a further list of assorted outrages will be supplied on application.

LOVE'S LOGIC.

My happiness is in another's keeping,
My heart delivered to a maiden's care,
And she can cast it down or set it leaping
(The latter process is extremely rare);
Ah, would that love indeed had made me blind,
That I might put her image out of mind!

Yet if I looked at her with eyes unseeing
Her voice and laughter would not pass unheard;
I should not be a reasonable being,
I still should tremble at her lightest word:
How could I then gain freedom from the spell
Unless I turned completely deaf as well?

So, blind and deaf, I might perhaps recover
A partial peace of mind, but all in vain,
For memories pursue the luckless lover,
And only death can ease him of his pain.
Thus, having proved that I were better dead,
I think I 'll go and talk to her instead.

For example...

Egypt in Venice (on Page 5) begins with two headings, one in French....

```
<div type="prose" xml:lang="en" xml:id="I1914-07-
01_05_02">
 <head>Egypt in Venice.</head>
 <head xml:lang="fr" rend="it">"La Légende de
Joseph."</head>
 <p>Those who know the kind of attractions that the
Russian ballet
    offers in so many of its themes ....</p>
</div>
```

Each stanza of the poem on page 10 has a last line which is significantly indented:

```
<lg>
 <l>There were eight pretty walkers who went up a
hill;</l>
 <l>They were Jessamine, Joseph and Japhet and Jill,</l>
 <l>And Allie and Sally and Tumbledown Bill,</l>
 <l rend="indent">And Farnaby Fullerton Rigby.</l>
</lg>
```

Punch example page 5

"WE'RE GIVING OUR PASTOR A NEW DRAWING-ROOM CARPET ON THE OCCASION OF HIS JUBILEE. SHOW ME SOMETHING THAT LOOKS NICE BUT ISN'T TOO EXPENSIVE." "HERE IS THE VERY THING, MADAME—REAL KIDDERMINSTER."

EGYPT IN VENICE.

"*La Légende de Joseph.*"

THOSE who know the kind of attractions that the Russian ballet offers in so many of its themes could have easily guessed, without previous enlightenment, what episode in the life of JOSEPH had been selected for illustration last week at Drury Lane. But they could never have guessed that Herr TIESSEN, author of a shilling guide to the intentions of the composer, would attach a transcendental significance to the conduct of *Potiphar's Wife*. "Through the unknown divine," he informs us, "which is still new and mysterious to her, an imperious desire awakens in her to fathom, to possess this world"—the world, that is to say, which *Joseph's* imagination creates in the course of an exhibition dance. If this is so, I can only say that her behaviour is strangely misleading.

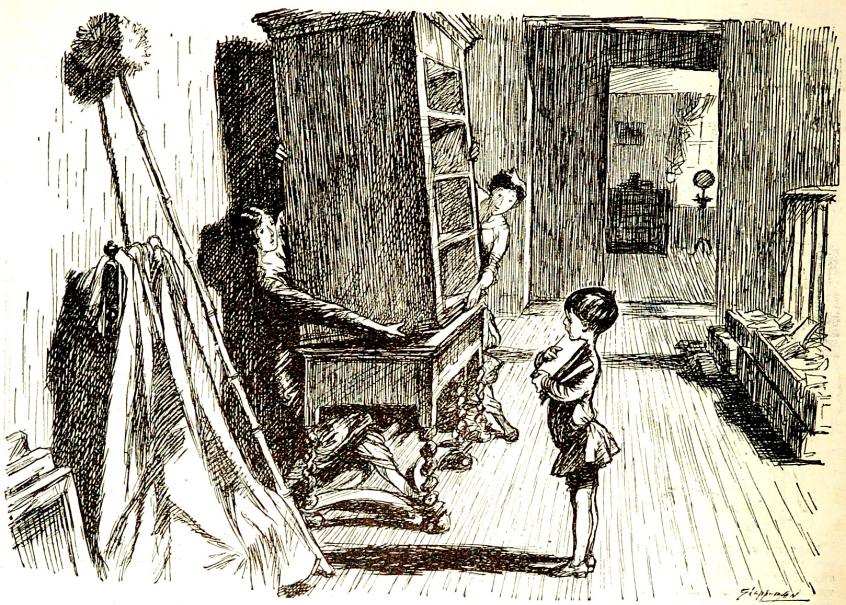
The scene opens at a party given by *Potiphar* in Venice. Venice, of course, was not *Potiphar's* home address; and I marvel a little at the change of *venue* when I think how much more harmony could have been got out of an Egyptian setting. But then I remind myself that the Russian ballet is nothing if not *bizarre*. The long banqueting-table recalls the canvases of VERONESE, but with discordant notes of the Orient and elsewhere. *Potiphar* himself, seated on a däis, has the air of an Assyrian bull. By his side *Mme. Potiphar* wears breeches ending above the knee, with white stockings and high clogs.

For the entertainment of the guests there was a dance of nuptial unveiling and a bout between half-a-dozen Turkish boxers. But it was a decadent and *blazé* company, and something more piquant was needed for their titillation. This was supplied in the shape of an original dance by the fifteen-year-old *Joseph*, whom my guide describes as "graceful, wild and pungent." He was introduced in a recumbent posture, and asleep, on a covered stretcher, and at first I had the clever idea that he was the customary corpse that appeared at Egyptian feasts to remind the company of their liability to die. But when he woke up and began to dance I saw at once that I was wrong.

I now know all about the interpretation of *Joseph's* dance; but I defy anyone to say at sight and without a showman's assistance what precisely he was after. In the Third Figure (according to my guide-book) "there is in his leaps a feeling of heaviness, as if he were bound to earth, and he stumbles once or twice as one who has missed his goal;" but how was I to guess that this signified that his "searching after God" was still ineffectual? or that when in the Fourth Figure he "leaps with light feet" this meant that "Joseph has found God"? I don't blame the boy for not knowing the rule that forbids one art to trespass on the domain of another; but there is no excuse for Herr STRAUSS, who must have been well aware that, for the conveyance of any but the most obvious emotions, mute dancing can never be a satisfactory substitute for articulate poetry.

However, *Potiphar's* guests seemed better instructed than I was, for they threw off their apathy and took quite an intelligent interest in *Joseph's pas seul*. Indeed, one young man (the episode escaped me at the dress rehearsal, but I

Punch example page 10



Alan (to his mother, who is busy with a heavy house-cleaning). "PLEASE, MOTHER, READ ME A STORY."

THE WALKERS.

THERE were eight pretty walkers who went up a hill;
They were Jessamine, Joseph and Japhet and Jill,
And Allie and Sally and Tumbledown Bill,
And Farnaby Fullerton Rigby.

They were all in good training and all of them keen,
And their chief wore a coat and a waistcoat of green;
He was always a proud man and kept himself clean,
Did Farnaby Fullerton Rigby.

They intended to lunch when they got to the top
On a sandwich apiece and a biscuit and chop.
The provisions were carefully bought in a shop
By Farnaby Fullerton Rigby.

They were jesters of merit—the sort who can poke
Funny tales in your ribs till you splutter and choke;
But the best of the lot at a jibe or a joke
Was Farnaby Fullerton Rigby.

It was ten of the clock when the walking began,
And they started with Tumbledown Bill in the van;
And the rear was brought up by that excellent man,
By Farnaby Fullerton Rigby.

They went off at a pace I am bound to deplore,
For they did twenty yards in a minute or more
And a yard or two over, a capital score
For Farnaby Fullerton Rigby.

They had all that pedestrians fairly can ask:
Smooth roads, sunny weather and beer in a cask,
And a friend who could teach them to stick to their task,
Viz.: Farnaby Fullerton Rigby.

Yet I somehow suppose that they hadn't the knack,
For in spite of it all they have never come back,
And I own that the future looks dismally black
For Farnaby Fullerton Rigby.

Now the walkers who seem to be stuck on the hill,
They are Jessamine, Joseph and Japhet and Jill,
And Allie and Sally and Tumbledown Bill,
And Farnaby Fullerton Rigby. R. C. L.

King Peter of Servia.

(From *The Daily Mirror.*)

"The proclamation, however, as given in a later message, reads thus:—To My Beloved People: As I shall be prevented by illness from exercising my royal power for some time, I order, by Article 69 of the Constitution, that so long as my cure lasts the Crown Prince Alexander shall govern in my name. On this occasion I recommend my dear fatherland to the care of the Almighty.
(Signed) PETER."

"On this occasion" is perhaps a little invidious.

Two consecutive books in *The Western Daily Press* list of publications received:—

"RING STRATEGY AND TACTICS.
CHARLES DICKENS IN CHANCERY."

The boxing boom continues.

Elements Available in All TEI Documents

The so-called 'Core' module groups together elements which may appear in any kind of text and the tags used to mark them in all TEI documents. This includes:

- paragraphs
- highlighting, emphasis and quotation
- simple editorial changes
- basic names numbers, dates, addresses
- simple links and cross-references
- lists, notes, annotation, indexing
- graphics
- reference systems, bibliographic citations
- simple verse and drama

Paragraphs

<p> (paragraph) marks paragraphs in prose

- Fundamental unit for prose texts
- <p> can contain all the phrase-level elements in the core
- <p> can appear directly inside <body> or inside (divisions)

```
<p>It was a cottage, the cottage of a
dream.  And by a cottage I mean, not
four plain rooms and a kitchen, but one
surprising room opening into another;
rooms all on different levels and of
different shapes, with delightful places
to bump your head on; open fireplaces;
a large square hall, oak-beamed, where
your guests can hang about after breakfast,
while deciding whether to play
golf or sit in the garden.  Yet all so
cunningly disposed that from outside
it looks only a cottage or, at most, two
cottages persuaded into one.</p>
```

Highlighting

By *highlighting* we mean 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. For words and phrases which are:

- distinct in some way (e.g. foreign, archaic, technical)
- emphatic or stressed when spoken
- not really part of the text (e.g. cross references, titles, headings)
- a distinct narrative stream (e.g. an internal monologue, commentary)
- attributed to some other agency inside or outside the text (e.g. direct speech, quotation)
- set apart in another way (e.g. proverbial phrases, words mentioned but not used)

Highlighting Examples

- <hi> (general purpose highlighting)

```
<p>[ The rest of this communication is
omitted owing to considerations of
space. —<hi rend="sc">Ed</hi>. ]</p>
```

- <distinct> (linguistically distinct)

```
But then I remind myself
that the Russian ballet is nothing if not
<distinct>bizarre</distinct>.
```

- Other similar elements include: <emph>, <mentioned>, <soCalled>, <term> and <gloss>

Quotation

Quotation marks can be used to set off text for many reasons, so the TEI has the following elements:

- <q> (separated from the surrounding text with quotation marks)
- <said> (speech or thought)
- <quote> (passage attributed to an external source)
- <cit> (groups a quotation and citation)

```
<p>
 <said who="#Celia">I know a lovely tin of potted
   grouse,</said> said Celia, and she went off
to cut some sandwiches.  By twelve
o'clock we were getting out of the
train.
</p>
```

Simple Editorial Changes: <choice> and Friends

- <choice> (groups alternative editorial encodings)
- Errors:
  - <sic> (apparent error)
  - <corr> (corrected error)
- Regularization:
  - <orig> (original form)
  - <reg> (regularized form)
- Abbreviation:
  - <abbr> (abbreviated form)
  - <expan> (expanded form)

Choice Example

```
I profess not to know how women's
<choice>
 <orig>heartes</orig>
 <reg>hearts</reg>
</choice> are wooed and won.  To me they have
always been <choice>
 <sic>maters</sic>
 <corr>matters</corr>
</choice> of riddle and <choice>
 <abbr>admirat'n</abbr>
 <expan>admiration</expan>
</choice>.
```

Additions, Deletions, and Omissions

- <add> (addition to the text, e.g. marginal gloss)
- <del> (phrase marked as deleted in the text)
- <gap> (indicates point where material is omitted)
- <unclear> (contains text unable to be transcribed clearly)

Example of <add>, <del>, <gap>, and <unclear>

```
<add place="left">The Cause</add> The immediate
cause, however, of the prevalence of supernatural

<del>tales</del>
<add place="supra">stories</add>
in these parts, was doubtless owing to the
<unclear reason="blood splatter">vicinity</unclear>
of Sleepy Hollow.
<gap reason="illegible">
 <desc>The rest of this paragraph is covered
   in dried blood.</desc>
</gap>
```

Basic Names

- <name> (a name in the text, contains a proper noun or noun phrase)
- <rs> (a general-purpose name or referencing string)

The *@type* attribute is useful for categorizing these, and they both also have *@key*, *@ref*, and *@nymRef* attributes.

Basic Names Example

```
<p>The scene opens at a party given by <name
    nymRef="http://www.meanings-of-
name.com/potiphar.html">Potiphar</name>
in
<name ref="http://en.wikipedia.org/wiki/Venice" type="place">V
</p>
<p>It is when the natural end of the story is reached,
and <name xml:id="SIMON">Simon</name> has come into his
own and has just been
wedded to his proper affinity, that the structure seems
to me to fall
with a crash.  I might perhaps, though not without
reluctance, have
pardoned an impertinent railway accident which leaves
<rs corresp="#SIMON">the young man</rs> apparently
crippled for life.</p>
```

Addresses

- <email> (an electronic mail address)
- <address> (a postal address)
- <addrLine> (a non-specific address line)
- <street> (a full street address)
- <postCode> (a postal (or zip) code)
- <postBox> (a postal box number)
- <name> can also be used
- and the 'namesdates' module extends this with more geographic names

Basic Address Example

```
<email>gbs@heaven.com</email>


<address>
 <name>George Bernard Shaw</name>
 <addrLine>Shaw's Corner</addrLine>
 <settlement>Ayot St Lawrence</settlement>
 <district>Hertfordshire</district>
 <postCode>HE 1 XXX</postCode>
 <country>England.</country>
</address>
```

Basic Numbers and Measures

- <num> (marks a number of any sort)
- <measure> (marks a quantity or commodity)
- <measureGrp> (groups specifications relating to a single object)
- While <num> has simple *@type* and *@value* attributes, <measure> has *@type*, *@quantity*, *@unit* and *@commodity* attributes

Number and Measure examples

```
<l>They went off at a pace I am bound to deplore,</l>
<l>For they did <num value="20">twenty</num> yards in a
minute or more</l>
<l>And a yard or <num value="2">two</num> over, a
capital score</l>
<l>For Farnaby Fullerton Rigby.</l>


<p>If neither of these values is available, a value of
<num>20,35</num>
for ash content can be assumed initially and checked,
after the
sampling has been carried out, using one of the methods
described in
ISO 13909-7.</p>


It is on these days that we travel to our Castle of
Stopes; as the
crow flies, <measure quantity="24140" unit="m">fifteen
miles</measure>
away. Indeed, that is the way we get to it, for it is a
castle in the
air.
```

Dates

- <date> (contains a date in any format and includes a *@when* attribute for a regularised form and a *@calendar* attribute to specify what calendar system)
- <time> (contains a time in any format and includes a *@when* attribute for a regularised form)

```
<p>At <time when="09:30:00">9.30 o'clock</time>,
as the fog lifted somewhat, the rescuing steamer
Lyonnesse had sighted the Gothland, fast on the rocks,
with a bad
list to starboard, and apparently partly filled with
pater.</p>
<p>House of Commons, <date when="1914-06-22">Monday,
June 22, 1914</date>.</p>
```

Simple Linking

- <ptr> (defines a pointer to another location)
- <ref> (defines a reference to another location, with optional linking text)
- Both elements have:
  - *@target* attribute taking a URI reference
  - *@cRef* attribute for canonical referencing schemes
- If the linking text is able to be generated, <ptr> and <ref> might be used in the same place.

Simple Linking Example

```
See <ref target="#Section12">section 12 on page 34</ref>.

See <ptr target="#Section12"/>.
```

Lists

- `<list>` (a sequence of items forming a list)
- `<item>` (one component of a list)
- `<label>` (label associated with an item)
- `<headLabel>` (heading for column of labels)
- `<headItem>` (heading for column of items)

Simple List Example

The previous slide contained only:

```
<div>
 <head>Lists</head>
 <p>
  <list>
   <item>
    <gi>list</gi> (a sequence of items forming a
list)</item>
   <item>
    <gi>item</gi> (one component of a list)</item>
   <item>
    <gi>label</gi> (label associated with an
item)</item>
   <item>
    <gi>headLabel</gi> (heading for column of
labels)</item>
   <item>
    <gi>headItem</gi> (heading for column of
items)</item>
  </list>
 </p>
</div>
```

Notes

- <note> (contains a note or annotation)
- Notes can be those existing in the text, or provided by the editor of the electronic text
- A *@place* attribute can be used to indicate the physical location of the note
- Although notes should usually be encoded where its identifier/mark first appears, notes can also be kept separately and point back to their location with a *@target* attribute

Note Example

```
<p>It is not only misfortune that makes strange
bedfellows.  <note place="foot">By-the-by, it is denied
that Sir <name>Joseph Beecham</name> was in any way
responsible for the Government's <title>Pills for
Earthquakes</title>, by which it was hoped to avert the
Irish crisis.</note>
</p>
```

Indexing

- If converting an existing index, use nested lists. For auto-generated indexes:
- `<index>` (marks an index entry) with optional *@indexName* attribute
- The `<term>` element is used to mark a term inside an `<index>` element
- The `<index>` element can self-nest for hierarchical index entries

Indexing Example

```
<p>… activated sludge treatment<index>
   <term>activated sludge</term>
   <index>
    <term>treatment</term>
   </index>
  </index> process for the biological treatment of
wastewater in which a mixture of wastewaterand
<hi>activated sludge</hi> is agitated and aerated.  The
<hi>activated sludge</hi> is subsequently separated from
the <hi>treated wastewater</hi> by
<term>sedimentation</term>
  <index>
   <term>sedimentation</term>
  </index>, and is removed or returned to the process as
required.</p>
```

Graphics

- <graphic> (indicates the location of an inline graphic, illustration, or figure)
- <binaryObject> (encoded binary data embedding a graphic or other object)
- The *figure* module provides <figure> and <figDesc> for more complex graphics

```
<figure>
 <graphic url="images/014.png"/>
 <head>Garden City Washing-day.</head>
 <p>Our sensitive artist insists on a harmonious
colour-scheme.</p>
 <figDesc>A bearded man sits in a deckchair and wags his
finger at a woman hanging up washing</figDesc>
</figure>
```