Using TEI: workshop for Taiwan, 27-28 March 2009

Sebastian Rahtz

February 25, 2009

1 Summary

This two-day course combines coverage of the latest version of the Text Encoding Initiative (TEI) Recommendations for the encoding of digital text with hands-on practical exercises in their application.

Participants should be be broadly familiar with the idea of marking up text, with (for example) some experience of producing HTML web pages, or of traditional scholarly editing.

At the end of the course we hope to have given you:

1. a good grounding in the theoretical issues underlying the use of text markup, XML in particular;
2. an understanding of the purpose and principles of the Text Encoding Initiative;
3. a review of the full range of modules constituting the TEI's current Recommendations;
4. an understanding of how the TEI scheme can be customized for particular applications, and internationalizated for different languages.

We will also aim to provide you with practical experience in:

- using online tools to build, verify, and document a TEI-conformant schema
- using XML editing software to
  - create new encoded texts
  - standardize existing digital texts
- using a variety of web-based and desktop tools to display and analyse TEI documents

The structure for the workshop is as follows:

Se:	Topic	Exercise
1:	Why and how do we encode texts? Basic choices and decisions. Making a document analysis.	Deciding what to encode in a document. Working from a few sample pages, choose which components to record.
2:	Introduction to the TEI, its class system, default text structure, and core elements and header. Exploring the TEI landscape with Roma.	Work with TEI Roma application to create a schema for the document.; Editing real XML: encode a page of the text using the schema created above.
3:	The TEI header and bibliographies	Editing the text to improve the metadata.
4:	Facsimiles: linking the text to images	Using the image markup tool to link text and pictures.
5:	Names, people and places in the text	Editing the text to mark up names and places.
6:	Accessing TEI texts. Untangling the XML markup and extracting what you need; introducing standards such as XPath, XSLT, and XQuery	Using XPath and TEI XSL stylesheet from within oXygen to access the data we have recorded

2 Resources Catalogue

2.1 TEI Resources

Current events - TEIWiki http://www.tei-c.org/wiki/index.php/Current_events

Roma: generating validators for the TEI http://www.tei-c.org/Roma/

TEI @ Sourceforge http://tei.sourceforge.net/

TEI Guidelines TOC http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index-toc.html

TEI P5: Guidelines for Electronic Text Encoding and Interchange http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html

TEI Wiki: Tools Category http://www.tei-c.org/wiki/index.php/Category:Tools

TEI: Join the TEI http://www.tei-c.org/Membership/join.xml

TEI: Learn the TEI http://www.tei-c.org/Support/Learn/

TEI: P5 Guidelines http://www.tei-c.org/Guidelines/P5/

TEI: Projects Using the TEI http://www.tei-c.org/Activities/Projects/

TEI: Special Interest Groups (SIGs) Mailing Lists http://www.tei-c.org/Activities/SIG/

TEI: Stylesheets http://www.tei-c.org/Tools/Stylesheets/

TEI: Text Encoding Initiative http://www.tei-c.org/index.xml

TEI: Tutorials http://www.tei-c.org/Support/Learn/tutorials.xml

TEIWiki - Main Page http://www.tei-c.org/wiki/index.php/Main_Page

2.2 W3Schools Tutorials

DTD Tutorial http://www.w3schools.com/dtd/default.asp

W3C Tutorial http://www.w3schools.com/w3c/default.asp

W3C XML Pointer, XML Base and XML Linking http://www.w3.org/XML/Linking

XLink and XPointer Tutorial http://www.w3schools.com/xlink/default.asp

XML Schema Tutorial http://www.w3schools.com/schema/default.asp

XML Tutorial http://www.w3schools.com/xml/default.asp

XPath Tutorial http://www.w3schools.com/xpath/default.asp

XQuery Tutorial http://www.w3schools.com/xquery/default.asp

XSLT Tutorial http://www.w3schools.com/xsl/default.asp

2.3 Related W3C Specifications

Cascading Style Sheets http://www.w3.org/Style/CSS/

Extensible Markup Language (XML) http://www.w3.org/XML/

The Extensible Stylesheet Language Family (XSL) http://www.w3.org/Style/XSL/

W3C HTML http://www.w3.org/html/

W3C Internationalization (I18n) Activity http://www.w3.org/International/

W3C Math Home http://www.w3.org/Math/

W3C XHTML2 Working Group Home Page http://www.w3.org/MarkUp/

W3C XML Query (XQuery) http://www.w3.org/XML/Query/

W3C XML Schema http://www.w3.org/XML/Schema

World Wide Web Consortium - Web Standards http://www.w3.org/

2.4 Related Mailing Lists

LISTSERV 15.0 - TEI-L List at LISTSERV.BROWN.EDU http://listserv.brown.edu/archives/cgi-bin/wa?A0=tei-l

Mailing Lists for eXist http://sourceforge.net/mail/?group_id=17691

TEI: Special Interest Groups (SIGs) Mailing Lists http://www.tei-c.org/Activities/SIG/

XSL-List http://www.mulberrytech.com/xsl/xsl-list/

2.5 Other Tools and Resources

Apache Cocoon http://cocoon.apache.org/

Apache FOP http://xmlgraphics.apache.org/fop/

Apache Xindice http://xml.apache.org/xindice/

Free XML Tools http://www.garshol.priv.no/download/xmltools/

Open Source Native XML Database http://exist.sourceforge.net/

RELAX NG http://www.thaiopensource.com/relaxng/

The SAXON XSLT and XQuery Processor http://saxon.sourceforge.net/

The UVic Image Markup Tool Project http://tapor.uvic.ca/~mholmes/image_markup/

Versioning Machine http://v-machine.org/

oXygen XML Editor http://www.oxygenxml.com/

2.6 Books

- http://www.tei-c.org/Activities/SIG/Education/tei_bibliography.xml, a TEI bibliography
- DHQ: Digital Humanities Quarterly <http://www.digitalhumanities.org/dhq/>
- Jeni Tennison, Beginning XSLT (2004)
- Jeni Tennison, Beginning XSLT 2.0: From Novice to Professional, (2005)
- Jeni Tennison, XSLT and XPath on the Edge (Unlimited Edition), (2001)
- Lou Burnard, Katherine O'Brien O'Keefe, and John Unsworth (eds.) Electronic Textual Editing. New York: Modern Languages Association, (2006)
- Micahel Kay, XSLT 2.0 and XPath 2.0 Programmer's Reference, 4th ed., (2008)

2.7 Projects Using TEI

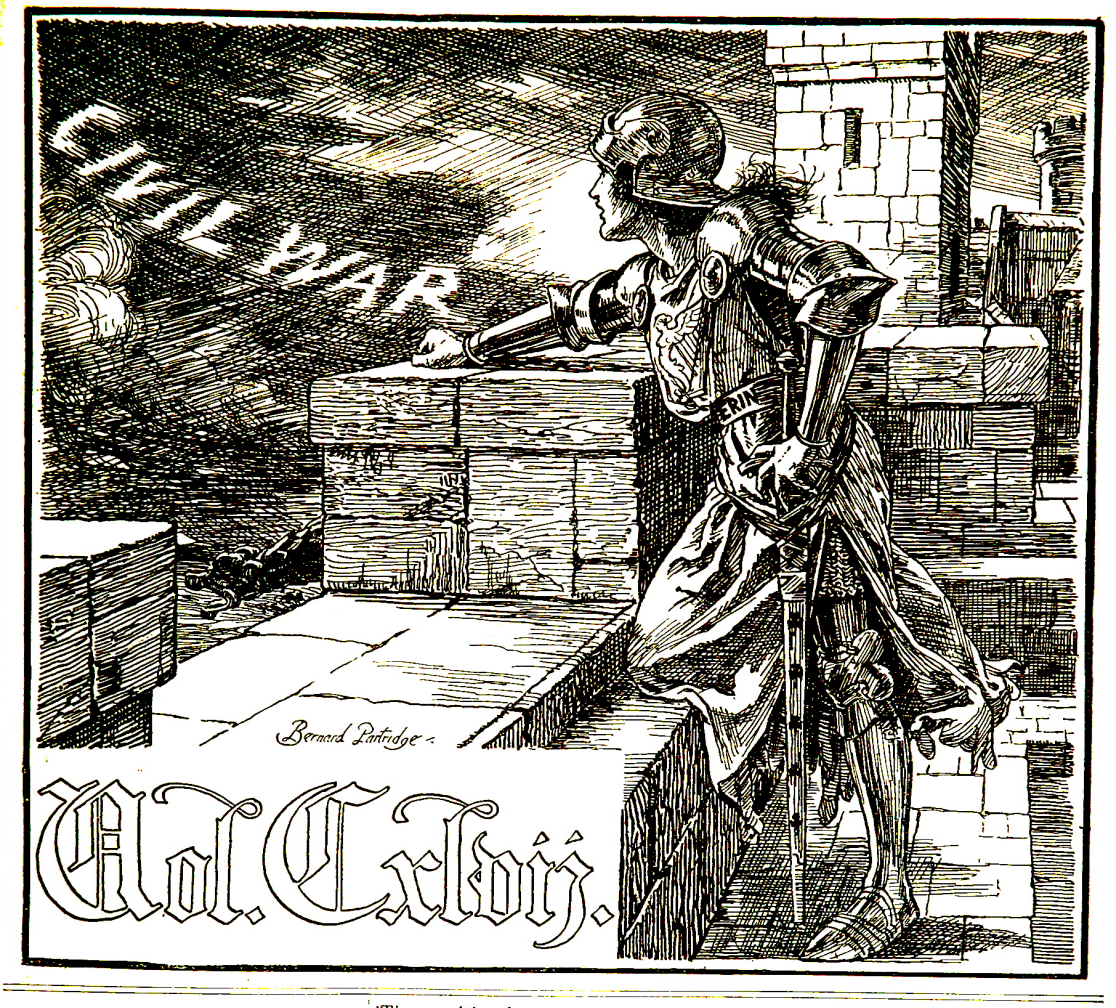
- A list of projects using the TEI (`http://www.tei-c.org/Activities/Projects/index.xml`)
- TEI: Anglo Saxon Poetry (`http://www.tei-c.org/Activities/Projects/an01.xml`)
- TEI: Bocaccio's Decameron (`http://www.tei-c.org/Activities/Projects/bo02.xml`)
- TEI: Book of Martyrs (`http://www.tei-c.org/Activities/Projects/jo01.xml`)
- TEI: British National Corpus (`http://www.tei-c.org/Activities/Projects/br01.xml`)
- TEI: Chinese Buddhist ETA (`http://www.tei-c.org/Activities/Projects/cb01.xml`)
- TEI: Chinese Saṃyukta Āgama (`http://www.tei-c.org/Activities/Projects/di06.xml`)
- TEI: Chymistry of Isaac Newton (`http://www.tei-c.org/Activities/Projects/ch04.xml`)
- TEI: Cursus Project (`http://www.tei-c.org/Activities/Projects/cu01.xml`)
- TEI: Diary of Robert Graves (`http://www.tei-c.org/Activities/Projects/ex01.xml`)
- TEI: DocSouth (`http://www.tei-c.org/Activities/Projects/do01.xml`)
- TEI: Early Americas Digital Archive (`http://www.tei-c.org/Activities/Projects/ea02.xml`)
- TEI: Emblem Project (`http://www.tei-c.org/Activities/Projects/ep02.xml`)
- TEI: English Poetry Database (`http://www.tei-c.org/Activities/Projects/ch01.xml`)

- TEI: EpiDoc
  (http://www.tei-c.org/Activities/Projects/ep01.xml)
- TEI: Newton Manuscript Project
  (http://www.tei-c.org/Activities/Projects/ne02.xml)
- TEI: Perseus Project
  (http://www.tei-c.org/Activities/Projects/pe01.xml)
- TEI: Piers Plowman Archive
  (http://www.tei-c.org/Activities/Projects/pi01.xml)
- TEI: Victorian Women Writers
  (http://www.tei-c.org/Activities/Projects/vi01.xml)
- TEI: Voltaire Foundation
  (http://www.tei-c.org/Activities/Projects/vo01.xml)
- TEI: Women Writers Project
  (http://www.tei-c.org/Activities/Projects/br04.xml)
- TEI: Wright American Fiction
  (http://www.tei-c.org/Activities/Projects/wr01.xml)
- TEI: Writings of J. F. Cooper
  (http://www.tei-c.org/Activities/Projects/th05.xml)

3 Punch

3.1 July 1st 1914

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.

THE ENCHANTED CASTLE.

THERE are warm days in London when even a window-box fails to charm, and one longs for the more open spaces of the country. Besides, one wants to see how the other flowers are getting on. It is on these days that we travel to our Castle of Stopes; as the crow flies, fifteen miles away. Indeed, that is the way we get to it, for it is a castle in the air. And when we are come to it Celia is always in a pink sun-bonnet gathering roses lovingly, and I, not very far off, am speaking strongly to somebody or other about something I want done. By-and-by I shall go into the library and work . . . with an occasional glance through the open window at Celia.

To think that a month ago we were quite happy with a few pink geraniums!

Sunday, a month ago, was hot. "Let's take train somewhere," said Celia, "and have lunch under a hedge."

"I know a lovely place for hedges," I said.

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

The first thing we came to was a golf course, and Celia had to drag me past it. Then we came to a wood, and I had to drag her through it. Another mile along a lane, and then we both stopped together.

"Oh!" we said.

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.

And, of course, we only saw it from outside. The little drive, determined to get there as soon as possible, pushed its way straight through an old barn, and arrived at the door simultaneously with the flagged lavender walk for the humble who came on foot. The rhododendrons were ablaze beneath the south windows; a little orchard was running wild on the west; there was a hint at the back of a clean-cut lawn. Also, you remember, there was a golf course, less than two miles away.

"Oh," said Celia with a deep sigh, "but we must live here."

An Irish terrier ran out to inspect us. I bent down and patted it. "With a dog," I added.

"Isn't it all lovely? I wonder who it belongs to, and if——"

"If he'd like to give it to us."

"Perhaps he would if he saw us and admired us very much," said Celia hopefully.

"I don't think Mr. Barlow is that sort of man," I said. "An excellent fellow, but not one to take these sudden fancies."

"Mr. Barlow? How do you know his name?"

"I have these surprising intuitions," I said modestly. "The way the chimneys stand up——"

"I know," cried Celia. "The dog's collar."

"Right, Watson. And the name of the house is Stopes."

She repeated it to herself with a frown.

"What a disappointing name," she said. "Just Stopes."

"Stopes," I said. "Stopes, Stopes. If you keep on saying it, a certain old-world charm seems to gather round it. Stopes."

"Stopes," said Celia. "It *is* rather jolly."

We said it ten more times each, and it seemed the only possible name for it. Stopes—of course.

"Well?" I asked.

"We must write to Mr. Barlow," said Celia decisively. "'Dear Mr. Barlow, er——Dear Mr. Barlow,——we——' Yes, it will be rather difficult. What do we want to say exactly?"

"'Dear Mr. Barlow,—May we have your house?'"

"Yes," smiled Celia, "but I'm afraid we can hardly ask for it. But we might rent it when—when he doesn't want it any more."

"'Dear Mr. Barlow,'" I amended, "'have you any idea when you're going to die?' No, that wouldn't do either. And there's another thing—we don't know his initials, or even if he's a 'Mr.' Perhaps he's a knight or a—a duke. Think how offended Duke Barlow would be if we put '——Barlow, Esq.' on the envelope."

"We could telegraph, 'Barlow. After you with Stopes.'"

"Perhaps there's a young Barlow, a Barlowette or two with expectations. It may have been in the family for years."

"Then we—— Oh, let's have lunch." She sat down and began to undo the sandwiches. "Dear o' Stopes," she said with her mouth full.

We lunched outside Stopes. Surely if Earl Barlow had seen us he would have asked us in. But no doubt his dining-room looked the other way; towards the east and north, as I pointed out to Celia, thus being pleasantly cool at lunch-time.

"Ha, Barlow," I said dramatically, "a time will come when *we* shall be lunching in there, and *you*——bah!" And I tossed a potted-grouse sandwich to his dog.

However, that didn't get us any nearer.

"Will you *promise*," said Celia, "that we shall have lunch in there one day?"

"I promise," I said readily. That gave me about sixty years to do something in.

"I'm like—who was it who saw something of another man's and wouldn't be happy till he got it?"

"The baby in the soap advertisement."

"No, no, some king in history."

"I believe you are thinking of AHAB, but you aren't a bit like him, really. Besides, we're not coveting Stopes. All we want to know is, does Barlow ever let it in the summer?"

"That's it," said Celia eagerly.

"And, if so," I went on, "will he lend us the money to pay the rent with?"

"Er—yes," said Celia. "That's it."

* * * * *

So for a month we have lived in our Castle of Stopes. I see Celia there in her pink sun-bonnet, gathering the flowers lovingly, bringing an armful of them into the hall, disturbing me sometimes in the library with "*Aren't* they beauties? No, I only just looked in—good luck to you." And she sees me ordering a man about importantly, or waving my hand to her as I ride through the old barn on my road to the golf-course.

But this morning she had an idea.

"Suppose," she said timidly, "you *wrote* about Stopes, and Mr. Barlow happened to see it, and knew how much we wanted it, and——"

"Well?"

"Then," said Celia firmly, "if he were a gentleman he would give it to us."

Very well. Now we shall see if Mr. Barlow is a gentleman. A. A. M.

Correspondence.

"Equal Rights" writes:—

"Dear Sir,—Why are descriptive names confined to boxers, such as Bombardier Wells and Gunboat Smith? Why not Rifleman Redmond, Airman Churchill, Solicitor George, Golfer Asquith, Bushman Wilding, Trundler Hitch, Dude Alexander, Bandsman Beecham, Hunger-Striker Pankhurst? Or, to take Editors——"

[The rest of this communication is omitted owing to considerations of space.—ED.]

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

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."

"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 KIDDERMINISTER."

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

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.

A SCANDALMONGRIAN ROMANCE.

(*By Francis Scribble.*)

[*The following article, specially written for us by the Author of "Ten Frail Beauties of the Restoration," "Tales Told by a Royal Washerwoman," etc., is another important contribution to the literature of the Royal Dirty-Linen Bag.*]

A DAY or two ago a short notice in the papers told of the death of Mrs. Maria Tubbs at Cannes; but few, if any, of those who read that brief announcement will have recognised in it the close of one of the most amazing careers of the nineteenth century. Yet little surprise need be expressed at this general ignorance, for who would think to find under that somewhat commonplace name the ravishingly beautiful Maria Cotherstone, who, forty years ago, was swept by Fate into the track of the late King of Scandalmongria, and well-nigh caused that singularly unstable bark to founder? It is with the kindly object of rescuing her romance from oblivion that this brief chronicle is written.

In 1873 the Scandalmongrian Minister in London was requested to find an English lady to take charge of the two children of his Royal master, and, after searching enquiries, he was successful, and Miss Maria Cotherstone turned her back on England never more to return. She was just twenty-two, fresh and blooming, possessed of the gayest of spirits, delightful manners and the highest accomplishments. Quietly she assumed control of the Royal schoolroom, and by her charm no less than by her firmness she quickly won the respect and love of her charges. Well had it been for her memory if her influence had never spread beyond the walls of her schoolroom; this article had then been unwritten. But alas for human nature! One day His Majesty's eyes fell upon the person of his children's governess, and then began one of the most sordid intrigues it has ever been my pleasure to recall. [A large statement, as readers of our author's *Gleanings from a Royal Dustbin* will readily acknowledge. However, the succeeding three-quarter of a column of details, here omitted, prove that there is at least some foundation for the remark.]

. . . And so their romance ended, and His Majesty returned to the bosom of his family and became once more the righteous upholder of the sanctity of the marriage tie. At first his easy-going Court smiled somewhat at the claim; but, when one or two highly-placed officials presumed to follow in the footsteps of their Sovereign, and were in consequence banished irrevocably from his presence, Scandalmongrian Society realised with a pained surprise that what is venial in a monarch may, in a subject, be a damnable offence.

And what of Maria, the charming, fascinating, much injured Maria? For several years she is lost, and then we hear of her marriage at Rome to "John Tubbs, Esq., of London," and once again she vanishes, only to turn up many years later at Cannes. She is a widow now, and a model of all the virtues. Who so staid and respectable as Madam? Who so charitable to the poor? Few, it is to be feared, will have recognised in that handsome old lady, so regular in her attendance at the services of the English Church, the beauteous Maria Cotherstone whose name was once on the lips of everybody from one end of Europe to the other. It nearly happened, indeed, that she went down to her grave with all her scandalous, feverish past forgotten, leaving behind her only the fragrant memory of her later life. But I have saved her. It is a queer story, quite interesting enough to recall.

THE HIGHER EDUCATION OF WOMEN.

Mistress. "THAT 'S A NICELY-MADE DRESS YOU HAVE ON, JANE. IT 'S LIKE THE NEW PARLOURMAID'S, ISN'T IT?"

Jane (*a close student of the fashion catalogues*). "OH NO, MA'AM, *THIS* IS *QUITE* A DIFFERENT CREATION."

CHARIVARIA.

IT is not only misfortune that makes strange bedfellows. Both Earl BEAUCHAMP and Sir JOSEPH BEECHAM appear in the recent Honours List.

***

By-the-by, it is denied that Sir JOSEPH BEECHAM was in any way responsible for the Government's "Pills for Earthquakes," by which it was hoped to avert the Irish crisis.

***

A New York cable announces that the Duke of MANCHESTER is interesting himself in a cinematograph proposition of a philanthropic nature, and that the company will be known as the "Church and School Social Service Corporation for the Advancement of Moral and Religious Education and Social Uplift Work through the medium of the Higher Art of the Moving Picture." It will of course be possible for the man in a hurry to call it, *tout court*, the "C.&S.S.S.C.F.T.A.O.M.&R.E.&S.U.W.T.T.M.O.T.H.A.O.T.M.P."

***

The penny off the income tax came just in time. It enabled several Liberal plutocrats to buy a rose on Alexandra Day.

***

The balance sheet of the German Company which has been running a Zeppelin airship passenger service has just been issued, and shows a loss of £10,000 on the year's working. This is not surprising. The difficulty which all aircraft experience is to keep their balance.

***

At the launch of the liner *Bismarck* last week, the bottle of wine which was thrown by the Countess HANNAH VON BISMARCK missed the vessel, whereupon the KAISER hauled back the bottle, and with his proverbial good luck hit the target.

***

Five shots were fired last week at Baron HENRI DE ROTHSCHILD. At first it was thought that this was done to stop the author of *Crœsus* from writing more plays, but, when it transpired that the assailant was a man who objected to the "Rothschild Cheap Milk Supply," public sympathy veered round in favour of the Baron.

***

Messrs. SELFRIDGE AND CO. were last week defrauded by a well-dressed man, who obtained two dressing-bags with silver fittings by means of a trick without paying for them. This is really abominable. It is bad enough when merely commercial firms are victimised: to best a philanthropic institution in this way is peculiarly base.

"MEXICAN REBEL SPLIT."
Morning Post.

Now perhaps the other civilised Powers will intervene. We have heard of many inhumanities marking the war in Mexico, but this treatment of a rebel is surely the limit.

***

It is not often, we imagine, that the British Navy is used to enforce a change of diet. H.M.S. *Torch* has just been ordered on a punitive expedition to Malekula Island, where certain of the natives have been eating some of their compatriots.

***

An American woman, according to *The Express*, has a serious complaint about the London policeman. She declares that she walked all the way from Queen's Hall to Piccadilly Circus with three buttons of her blouse undone at the back, and "not a single policeman" offered to do it up for her. No doubt the Force was reluctant to interfere with what might turn out to be the latest fashion. A Boy Scout who offered, the other day, to sew up a split skirt got his ears soundly boxed.

***

Meanwhile the glad tidings reach us that women's skirts and bodices are to fasten in front instead of at the back. Husbands all over the world who have on occasions been pressed into their wives' service as maids, only to learn that they were clumsy boobies, would like to have the name of the arbiter of fashion who is responsible for this innovation, as there is some thought of erecting a statue to him.

***

Some distinguished German professors have been discussing the question of the best place in which to keep a baby in summer. It is characteristic, however, of these unpractical persons that not one of them suggests the obvious ice-safe.

***

"One of the first things the rich should learn," says Dean INGE, "is that money is not put to the best use when it is merely spent on enjoyment." It is hoped that this pronouncement may lead wealthy people to patronise our concert-halls more than they do.

***

"£1,600," a newspaper tells us, "were found hidden in the cork leg of HARRY C. WISE while he was undergoing treatment in a hospital at Denver." And now, we suspect, HARRY'S friends will always be pulling his leg.

***

"Have you seen *Pelleas and Mélisande?*"

"No. Is it as funny as *Potash and Perlmutter?*"

THE COLLECTORS.

MY dinner partner was a self-made man and not ashamed of it.

"Do you take an interest in china, ma'am?" he asked me.

I felt that if I said "Yes" I should have to buy some. So I said "No," but he didn't wait to hear what I said.

"I think I may say," he continued, "that I have the finest collection of old Dresden china in London."

He went into the figures, explaining the cost price and the difficulty of storage.

"Oh," said I, "if you find it a nuisance, I've a parlour-maid I could recommend to you; just the girl to help you to get rid of it."

At this point I think he had some idea of having the finest collection of parlourmaids in Middlesex, but he made it small dogs instead. Was I interested in these? No, but I supposed I'd have to be if he insisted.

"I don't think I should be far wrong," he began, but I hustled him through to the end of his sentence.

"Finest collection in—?" I asked.

"England," he said.

He went over their points, and in an expansive moment I marvelled. This was imprudent, as it caused him to search his mind for some further spectacular triumph wherewith to amaze and delight.

"That," he said, looking up the table, "is my wife."

"Marvellous," said I.

He took this in the best part. "You refer to her diamonds?" he said.

"Did I?" said I.

"The finest collection in Great Britain," he declared, and spread himself over the subject.

Later, in a mood of concession, he inquired as to my specialities. I had none, at least none that I could think of. Determined to extract something noteworthy, he questioned me on every possibility. Was I not married? That was so, I agreed, but then so many women are.

"You have sons, ma'am?" he persisted, with that implacable optimism to which, among other things, he no doubt owed his success in the world.

I thought of Baby. "Ah yes, of course," I said. "The finest collection in Europe."

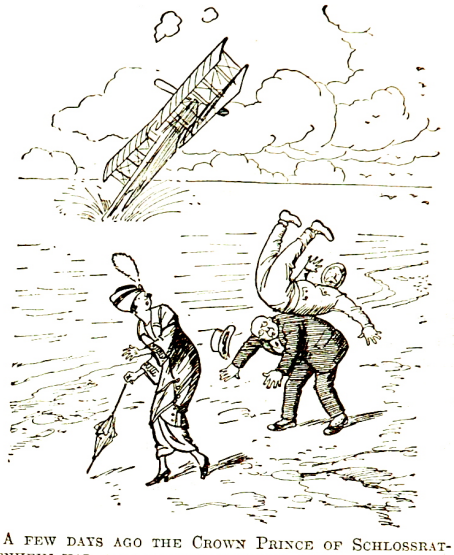
"'In Norway,' she says, 'we do not eat one-third the quantity that the English eat; our meals are simpler and shorter. I believe that this is the cause of the enormous amount of indigestion that is suffered by the English.'"
Daily News and Leader.

So our doctor, who attributed our indigestion to lobster mayonnaise, was wrong again.

KINDNESS TO SUBJECTS.

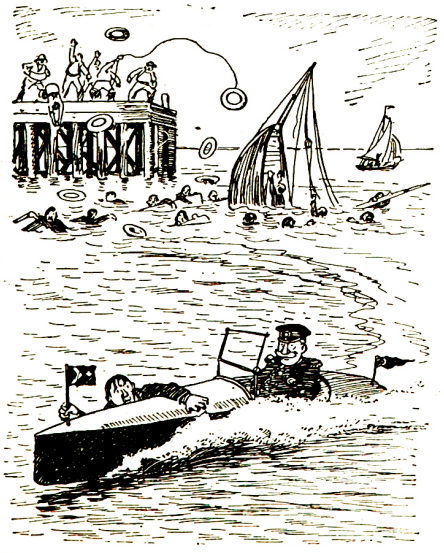
[One of our illustrated papers recently published a picture of the King of SPAIN in a motor-car which had broken down. The car was being pushed along by some helpful people, and the comment on the picture was, "It is these thoughtful little acts that make royalty so popular nowadays." Lest it should be thought that the other potentates of Europe take less trouble to make themselves beloved by their subjects, we hasten to give a few instances which have come to our notice.]

LAST WEEK THE KING OF CADONIA HAD HIS HAT BLOWN OFF IN THE BLUMENGARTEN (THE BEAUTIFUL PARK NEAR THE ROYAL PALACE). THIS KINDLY ACT SHOULD DEEPEN THE AFFECTION IN WHICH THE MONARCH IS HELD BY HIS PEOPLE.



A FEW DAYS AGO THE CROWN PRINCE OF SCHLOSSRATTENHEIM HAD AN ACCIDENT WITH HIS AEROPLANE, WHICH OVERTURNED NEAR SCHUTZMEER. FORTUNATELY HIS ROYAL HIGHNESS FELL ON A RETIRED WUERST-HAENDLER WHO WAS WALKING ON THE BEACH.

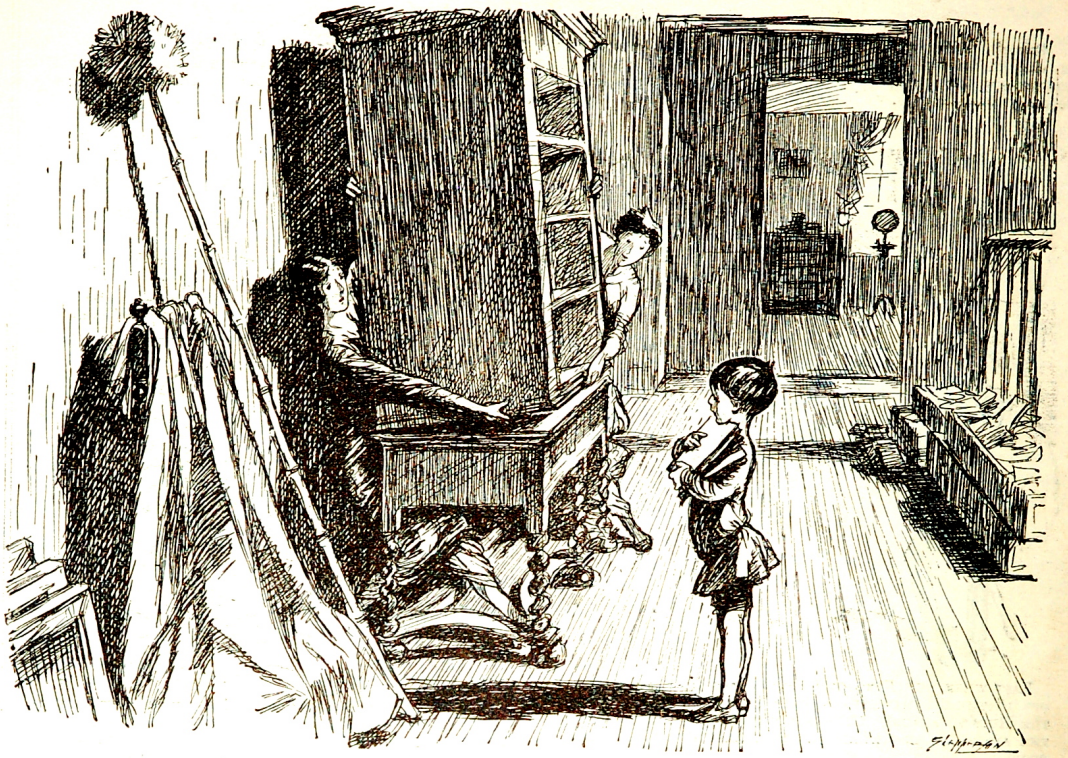
THE CROWN PRINCE'S DEVOTION TO HIS BELOVED SUBJECTS IS WELL KNOWN, AND THIS TACTFUL DEED WAS ONLY ANOTHER INSTANCE OF IT.



YESTERDAY PRINCE JOHN OF PUMPENHOSEN INADVERTENTLY COLLIDED WITH A PLEASURE-YACHT AT THE MOUTH OF THE HARBOUR OF KREBS WHILE TRYING A NEW MOTOR BOAT. ALL THE PASSENGERS WERE SAVED AND THE PRINCE SHOWED NO SIGNS OF FEAR.

THIS SHOULD ENHANCE HIS GREAT POPULARITY, IF SUCH A THING WERE POSSIBLE.

KING STEPHAN III. OF SERVILIA, WHILE PLAYING ON THE LINKS AT NIBLIKSK LAST WEEK, INITIATED ONE OF HIS EQUERRIES INTO THE HUMOUR OF THE GAME. BY THIS THOUGHTFUL ACT HIS MAJESTY ADDS TO THE DESERVED LOVE AND REVERENCE IN WHICH HE IS HELD BY THE SERVILIANS OF ALL CLASSES.



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.

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

THE EMERGENCY EXIT.

SCENE—*A Tight Place.*

CHILD HERBERT (*to "Wicked Baron"*). "MY LORD, I HAVE EVER REGARDED YOU AS A PESTILENT VILLAIN—NAY WORSE, AN HEREDITARY IMBECILE. I THEREFORE RELY ON YOUR BENEFICENT WISDOM TO FIND ME A WAY OUT OF THIS SINISTER WOOD."

ESSENCE OF PARLIAMENT.

(EXTRACTED FROM THE DIARY OF TOBY, M.P.)

House of Commons, Monday, June 22.—Great muster of forces on both sides. Not wholly explained by second reading of Budget Bill standing as first Order. A section of Ministerialists, purists in finance, took exception to proposed procedure. HOLT, spokesman at mouth of new Cave, put down amendment challenging CHANCELLOR OF EXCHEQUER's proposals. Here was chance for watchful Opposition. If some thirty Ministerialists would go with them into Lobby it would not quite suffice to turn out Ministry; but it would be better than a Snap Division, with its personal inconvenience of preliminary hiding in bath-rooms and underground cellars.

CASSEL, adding to Parliamentary reputation studiously attained, raised subject on point of order. Underlying suggestion was that Budget Bill should be withdrawn and reintroduced under amended form of procedure. SPEAKER, whilst admitting irregularity, stopped short of approving extreme course. Pointed out that the matter might be put right by moving fresh resolutions.

This disappointing. Worse to follow. The INFANT SAMUEL, making fresh appearance in new part of understudy of CHANCELLOR OF EXCHEQUER, conceded point of procedure made by Radical Cave. Promised objection should be fully met. HOLT, amid ironical cheers from Opposition, said in these circumstances would not move amendment. Incident reminded WALTER LONG of story of the Colonel and the opossum up a tree.

"Don't shoot!" said the opossum; "I'll come down."

CHANCELLOR OF THE EXCHEQUER had come down. No need for Colonel HOLT to discharge his gun.

Thus threatened crisis blew over. Members, cheered by promise of reduction by one half of proposed increase in Income Tax, got away early to attend various functions in honour of KING'S birthday.

Business done.—Second reading of Budget Bill moved.

House of Lords, Tuesday.—London season in full fling. May be said to reach dizziest height in this birthday week. Social engagements numerous and clashing. To-day House of Lords magnet of attraction of surpassing force. The thing for *grandes dames* to do is to go down to the House and be present at opening of fresh tourney round Home Rule Bill. Accordingly, the peeresses, alive to their responsibility as leaders of high thinking and simple living, flock down to Westminster, filling side-galleries with grace, beauty, and some finely feathered hats.

Seats on floor also crowded. Patriotic peers arriving late, finding no room on the benches where the Union Jack is kept flying, cross over. Temporarily seat themselves among the comparatively scanty flock of discredited Ministerialists. Bishops muster in exceptional number. Their rochets form wedge of spotless white thrust in centre of black-coated laity seated below Gangway on right of Woolsack. Space before Throne thronged with Privy Councillors availing themselves of the privilege their rank confers to come thus closely into contact with what is still an hereditary chamber.

In centre of first row CARSON uplifts his tall figure and surveys a scene he has done much to make possible.

Perhaps in matter of dramatic interest the play did not quite come up to its superb setting. Principal parts taken by CREWE and LANSDOWNE. Neither accustomed to move House to spasms of enthusiasm. LEADER OF HOUSE, introducing what is officially known as Government of Ireland Amending Bill, made it clear in such sentences as were fully audible that scheme does not go a step beyond overture towards settlement proffered by PREMIER last March.

LANSDOWNE expressed profound disappointment at this lack of enterprise. "Rather a shabby and undignified proceeding on the part of a strong Government," he said, "to come down with proposal they know to be wholly inadequate, and to hint that we ought to assist them in converting it into a practical and workable measure."

Actual condition of things could not with equal brevity be more clearly stated. Bill presented to Lords as sort of lay figure, which they may, in accordance with taste and conviction, suitably clothe. No assurance forthcoming that style and fit will be approved when submitted to House of Commons, final arbiters.

Meanwhile Bill read a first time, and ordered to be printed.

Wicket-keeper (*Mr.* CASSEL). "How 's that?"
Umpire (*Mr.* SPEAKER). "Out!"
Batsman (*Mr.* LLOYD GEORGE). "Rotten antiquated rule!"

["I did not expect . . . that hon. members would go rummaging in the dustbins of ancient precedent to find obstacles to place in the way of these proposals."—*Mr.* LLOYD GEORGE *on his Budget.*]

"Bill presented to Lords as a sort of lay-figure, which they may, in accordance with taste and conviction, suitably clothe."

Business done.—The Commons still harping on the Budget. TIM HEALY enlivened proceedings by vigorous personal attack on "the most reckless and incapable CHANCELLOR OF THE EXCHEQUER that ever sat on the Treasury Bench." LLOYD GEORGE'S retort courteous looked forward to with interest.

House of Commons, Wednesday.—Wh. n, shortly after half-past five, CHANCELLOR OF EXCHEQUER rose to take part in debate on new development of Budget Bill, House nearly empty. Interests at stake enormous. Situation enlivened for Opposition by quandary of Government. But afternoon is hot, and from the silver Thames cool air blows over Terrace. Accordingly thither Members repair, leaving House to solitude and CHIOZZA MONEY.

Benches rapidly filled when news went round that CHANCELLOR was on his legs. Soon there was crowded audience. Sound of cheering and counter-cheering, applausive and derisive, frequently broke forth. CHANCELLOR in fine fighting form. Malcontents in his own camp are reconciled. Hereditary foe in front. Went for him accordingly. WALTER LONG seated immediately opposite conveniently served as suitable target for whirling lance. Effectively quoted from speeches made by him at other times, insisting upon relief of the rates so heavily burdened as to make it impossible to carry out social reforms of imperative necessity.

"After these lavish professions of anxiety to help local authorities, I did not," said the CHANCELLOR, "expect the right hon. gentleman and his friends would go rummaging in the dustbins of ancient precedent, to find obstacles to place in the way of proposals of reform."

Carried away by his own eloquence, the CHANCELLOR, whilst sarcastically complimentary to WALTER LONG, went so far as to call him "The Father of Form IV." The putative parent blushed. There were cries of "Order!" and "Withdraw!" SPEAKER did not interpose, and CHANCELLOR hurried on to another point of his argument.

Quite a long time since our old friend Form IV., at one time a familiar impulse to party vituperation, was mentioned in debate. This unexpected disclosure of its paternity made quite a stir.

SON AUSTEN followed CHANCELLOR in brisk speech that led to one or two interludes of angry interruption across the Table. When he made an end of speaking, debate relapsed into former condition of languor. Talk dully kept up till half-past eleven.

Business done.—Further debate on Budget.

Thursday.—CHANCELLOR OF EXCHEQUER admittedly allured by what he describes as "attractive features" of proposal to raise fresh revenue. It is simply the levying of a special tax on all persons using titles.

Idea not absolutely new. Principle established in case of citizens displaying crest or coat-of-arms. What is novel is suggested method of taxation. Differing from the dog-tax, levied at a common rate, it is proposed that our old nobility shall, in this fresh recognition of their lofty estate, be dealt with on a sliding scale. A duke will have his pre-eminence recognised by an exceptionally high rate of taxation. Marquises, earls and a' that will be mulct on a descending scale, till the lowly knight is reached. He will be compensated for comparative obscurity in the glittering throng by being let off for a nominal sum.

CHANCELLOR fears it is too late to adopt proposal this year, a way of putting it which seems to suggest that we may hear more of it in next year's Budget.

Business done.—HAYES FISHER'S Amendment to Budget Bill negatived by 303 votes to 265. Reduction of Ministerial majority to 38 hailed with boisterous burst of cheers and counter-cheers.

GARDEN CITY WASHING-DAY.

OUR SENSITIVE ARTIST INSISTS ON A HARMONIOUS COLOUR-SCHEME.

The LORD MAYOR (on hearing a certain PEEL): "Turn again (in your grave), WHITTINGTON."

New song for old Cantabs.:—

"O. B., what can the maté be?"

RUS IN URBE.

No, this is not the Russian ballet. It is the English Folk Dance Society, and their performances at the Royal Horticultural Hall at Westminster the other day showed that the Russian ballet is not to have things all its own way. I am not going to moralise upon the salacious quality of some of the themes of our exotic visitors, but certainly it would be difficult to find a stronger contrast to their ruling passion than is presented by the purity and simplicity of these country dances.

"Sellinger's Round," danced to an air that lulled *Titania* to sleep all through the winter at the Savoy, was the most popular, with its ring of a dozen dancers, hands joined, running together into the centre of their circle, as if to honour some imaginary deity—possibly Mr. CECIL SHARP, director of the Society, who has collected and revived the airs to which they dance.

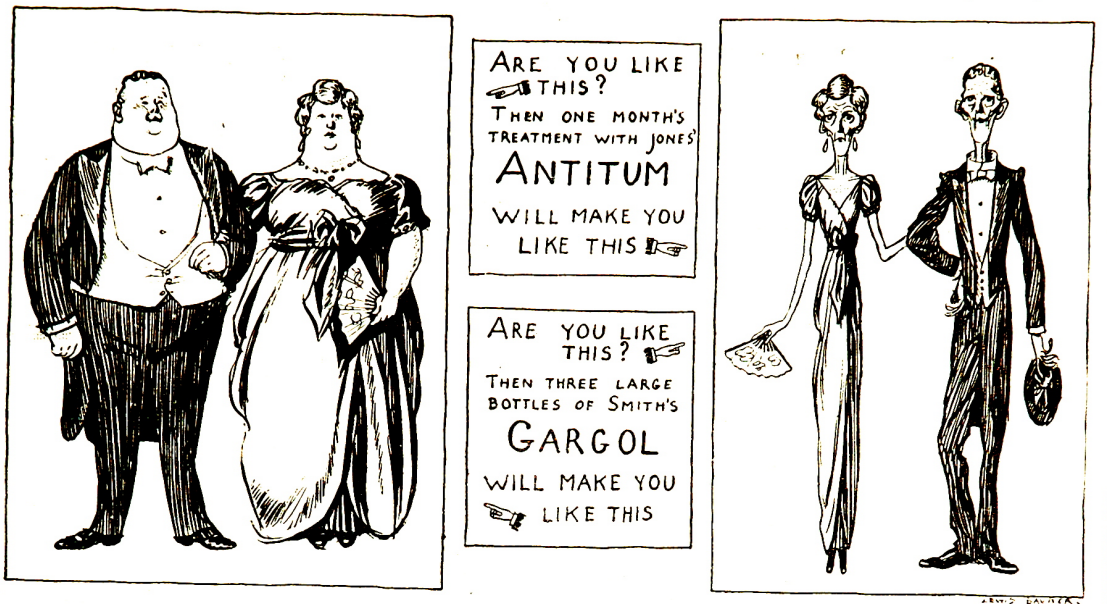
Then there were the Morris-dances, "Shepherd's Hey" (with nothing about a "nonny-nonny" in it), and "Haste to the Wedding." There might perhaps be a greater propriety in the latter if it were confined to men; but at least it raised no apprehension that anybody was going to "repent at leisure." In the "Flamborough Sword" dance, the men (with no Amazon assistance) raced through the figure and out again, eight of them, armed with bloodless wooden swords—a finely ordered riot.

"Lady's Pleasure," a Morris-jig for two men, lays hold of you at the first bar, and again with a fresh grip and a tighter as the music slows up for the dancers to do their "capers"—all to the music of Mr. CECIL SHARP at the piano and Miss AVRIL at the fiddle.

The object of The English Folk Dance Society is to teach rather than to perform in public. Hence the rarity of their displays, and the better reason why we should seize, when they come, our chances of assisting at these delightful exhibitions of an art whose revival has done so much to restore to the countryside the unpretentious joys that gave its name to Merrie England.

"It was the time when Henry III. was batting with Simon de Montfort and his Barons."—*Straits Times.*

But not at Lord's, which has only just celebrated its centenary.



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.

BALM FOR THE BRAINLESS.

["If one man has more brains than another, which enable him to outstrip his fellows, is not that good fortune? What had he got to do with it? If your brain is a bad one, it is not your responsibility. If your brain is a good one it is not your merit. Some men have greater physical, mental, moral strength than others that enables them to win in the race. That is their good fortune and they ought to be grateful for it; and the one way they can best show their gratitude is by helping those who are less fortunate than themselves. Men endowed with any, or most, or all of these fortunate conditions ought not to be stingy in helping others who have not been so fortunate as themselves."—*Mr. LLOYD GEORGE at Denmark Hill, June* 20.]

As a result of Mr. LLOYD GEORGE's vivid and convincing pronouncement on the responsibilities of the fortunate, we have been deluged with appeals from all sorts and conditions of unlucky correspondents. We select the following from among the most deserving cases in the hope that our opulent readers may avail themselves of the chances thus offered of redressing the partiality of fortune.

THE CRY OF THE CRACKSMAN.

The Sanctuary, Crookhaven.

SIR,—Endowed by nature with an imperfect moral sense and a complete inability to discriminate between *meum* and *tuum*, I was irresistibly impelled at an early age to adopt the precarious profession of housebreaker. I have just served a sentence of three years, and was on the point of resuming my career when I read Mr. LLOYD GEORGE's epoch-making speech at Denmark Hill, in which he clearly defines the duty of the State to redress the inequalities of moral as well as material endowment by which so large a proportion of the community is penalised. I am the master of a fine literary style and admirably suited to discharge any secretarial duties, but it is only right that I should clearly explain at the outset that it is no use offering me any post unless it is so well salaried that I should never feel it was worth while to explore or appropriate the contents of my employer's safe.

Respectfully yours,
RAPHAEL BUNNY.

THE LUCK OF THE LAW.

Railway Carriage Bungalow, Shoreham, Sussex.

SIR,—It is precisely thirty years since I was called to the Bar, and several of my contemporaries have already been elevated to the Bench, while Sir JOHN SIMON, who is considerably my junior, is in the receipt of a salary probably double that drawn by an ordinary Judge. My earnings for the last ten years have exempted me from income-tax, but this is but a poor consolation when I consider that were it not for the caprice of fortune I should probably be returning £400 or £500 a year to the Exchequer in super-tax. But not only have I been badly treated in regard to mental equipment; I have been further handicapped by hereditary conscientious objection to pay any bills. An annuity of £500 a year, or only one-tenth of the salary of a Judge, is the minimum that my self-respect will allow me to accept in payment of the State's long-standing debt to

Yours faithfully, WILLIAM WEIR.

THE CRUELTY OF COMPETITION.

SIR,—I confidently appeal for your support in the application for a grant which I am forwarding to the PRIME MINISTER. My son, aged 14, has failed to win an entrance scholarship at Winchester and Charterhouse, not from any fault of his own, but simply owing to the unfair competition of other candidates more liberally endowed with brains. At a modest estimate I calculate that the extra drain on my resources for the next eight years in consequence of this undeserved hardship will amount to at least £600, which I can ill afford owing to unfortunate speculations in Patagonian ruby mines—another example of that bad luck which, in the noble words of the CHANCELLOR OF THE EXCHEQUER, it is the privilege of the prosperous to remedy.

I am, Sir, yours expectantly,
(Rev.) J. STONOR BROOKE.

VIS INERTIÆ.

Lotus Lodge, Limpsfield.

SIR,—A victim since birth to congenital lassitude, which has rendered all labour, whether manual or mental, distasteful, nay, intolerable to me, I find myself at the age of 41 so out of touch with the spirit of strenuous effort which has invaded every corner of our national life that I am anxious to confer on the State or, failing that, some meritorious millionaire the privilege of providing for my modest needs. A snug sinecure with a commodious residence and a good car—cheap American motors are of course barred—represent the indispensable minimum.

I am, Sir, yours faithfully,
EVERLEIGH SLACK.

Some day, says the President of the Aero Club, we shall be able to go into a shop and buy a pair of wings. But we can do that already; the only difficulty is to fly with them.

"Gentleman, middle aged, would be glad of a few correspondents (40 to 60)."
T. P.'s Weekly.

Too many.

THE SILENT CHARMER.

[Speaking of flowers a contemporary recently remarked:—"These careless-looking creatures filling the air with delight, robbing tired brains of tiredness, are a delicate texture of coloured effort that has prevailed out of a thousand chances, aided in all that effort by man. Without man they would be but weeds—a profusion of Nature's quantity."]

My dearest Thomas, I would not
Deny the fact that you are clever;
You've taught Dame Nature what is what
At horticultural endeavour
(She has not got that useful thing,
The shilling book of gardening).

She has her merits, but, of course,
Her wild attempts won't stand comparing
With such a floral *tour de force*
As that geranium you are wearing;
Yon chosen emblem of your skill
Must surely make her wilder still.

But give me Nature; when we meet
She does not prattle of her posies,
Dull facts of what begonias eat,
The dietetic fads of roses,
And how she strove with spade and spud,
Or nipped the green fly on the bud.

'Tis she that really soothes the brain,
Spreading her weeds in bright profusion,
And never troubling to explain
How much they owe to her collusion,
While, Thomas, *your* achievements seem
To be your one and only theme.

Mr. J. C. PARKE, writing in *The Strand Magazine* on the best way to beat WILDING, says:—

"Personally, after close observation and from playing against him, I would suggest a determined attack on the champion's forehead from the base-line."

That ought to learn him.

"His Majesty has been pleased to confer the dignity of an Earldom of the United Kingdom upon Field-Marshal the Viscount Kitchener of Khartoum, P.G.C., B.O.M.G.C., S.I.G.C.M., G.G.C.I.E."
Newcastle Daily Journal.

The old orders change, yielding place to new.

From a magazine cover:—

"This magazine has been the turning point in many a man's career. Spend twopence and half-an-hour on it. . . . Price Threepence."

We would rather pay the threepence.

"In our report of the wedding of Mr. Lee Kwee Law to Miss Chan Siew Cheen we inadvertently left out the following, who also sent presents:——"—*Straits Echo.*

And then they inadvertently left them out again.

THE CURE FOR CRICKET.

THERE is no longer any doubt that golf is threatening the supremacy of our national game. Judged by the only true standard—the amount of space allotted to it in the daily press—it is manifest that the encroachments of this insidious pastime have now reached a point where the cricket reformer must bestir himself before it is too late. We are convinced that so far we have been taking much too narrow a view. The time has come to look for light and leading outside the confines of our own Book of Rules. There are other games besides cricket. Let us call them to our councils.

In the first place a valuable hint may surely be found in the development of Rugby football. It is common knowledge what immense results have followed the introduction, some twenty years ago, of the Four Three-quarter System. No spectator (and we cannot exist without the spectator) would ever dream now of returning to the old formation. Very well. The same principle can be easily adapted to our requirements in the form of the Three Batsmen System. The pitch would become an equilateral triangle, and we should suggest that the bowler have the option of bowling (from his own corner) at either of the two outlying batsmen (at theirs). Lots of interesting developments would follow, as, for instance, the institution of a sort of silly-point-short-mid-on in the centre of the triangle. (Should he be allowed to wear gloves?)

Golf has also a lesson to teach us. We are all familiar with the huge strides that have been made by the introduction of the rubber-cored ball. We don't want to plagiarize, although a rubber-cored cricket ball is a nice idea. Why not aim at the opposite extreme and try a ball "reinforced" with concrete? The tingling of the batsman's fingers which might result could be neutralised by the use of a rubber-faced bat. This reform would, we believe, have one happy consequence. People wouldn't be so keen to play with their legs.

As to lawn tennis—another dangerous rival—we hear a good deal in these days about "foot-faults." That seems to show the trend of modern thought. If we are to be in the swim we shall have to reconsider our no-ball rule. Why not make it a no-ball every time unless the bowler has both feet in the air at the moment when the ball leaves his hand? One might put up a little hurdle — nothing obtrusive — only a matter of a few inches high.

We believe that something might even be done by borrowing from hockey the principle of the semi-circle, outside of which a goal may not be shot. The whole pitch might be enclosed in a circular crease—which would look uncommonly well in Press photographs. (We cannot exist without the Press.) No fielder inside the magic circle would be allowed to stop the ball with his feet.

Finally there is the case of billiards, not a game that is very closely allied to cricket, but one from which much may be learned. How has billiards brightened itself? By adopting the great principle of "barring" certain strokes. Here we have got on to something really valuable. We propose to go one better, and draw up a schedule of the different conditions of barring under which matches may be played. It will only remain for secretaries, when fixtures are made, to arrange the terms by negotiation. In time to come, should we be able to carry our point, we shall all be familiar with such announcements as the following:—

Notts. *v.* Surrey. (Cut-barred.)
Gentlemen *v.* Players. (L.b.w.-barred.)
England *v.* Australia. (Googly-and-yorker-barred.)

We do not pretend to have exhausted the subject, but we have made a start. We must look about us. Something may be learned, we firmly believe, even from skittles and ping-pong. Our national game cannot afford to exclude special features. It should have the best of everything.

"ARE YOU MRS. PILKINGTON-HAYCOCK?"
"NO." "WELL, I AM, AND THIS IS HER PEW."

Professional Candour.

"The sermon over, a collection was taken, and hardly a person present did not contribute. Mgr. Benson's sermon went to the hardest heart there. Even the journalists contributed."
The Universe.

THE HERE, THERE AND LONDON LETTER.

With apologies to "The Westminster Gazette."

THE HOME OF THE SOUTH SAXONS.

Sussex, the county for which Mr. C. B. FRY (who hurt his leg in the Lord's centenary match) used to play before he moved to Hampshire, is an attractive division of the country to the south of London with a long sea border. Mr. KIPLING has praised it in some memorable verses, and among frequent visitors to its principal town, Brighton, is the CHANCELLOR OF THE EXCHEQUER. The word Sussex is a contraction of South Saxon. All will wish the old Oxonian a speedy recovery from his strain.

A MONETARY PROVERB.

The origin of the old saying, "Penny wise, pound foolish," which has come into vogue again in connection with the revised income tax—for who can deny that the saving of the penny is wise?—is lost in obscurity; but there is no doubt that it is very ancient. Many nations have the same proverb in different terms as applied to their own currency. In France the coins to which the saying best applies would be the sou and the louis; in America, the cent and the dollar; and so forth.

CORDIALITY BEFORE PARTY.

The circumstance of Mr. LULU HARCOURT'S unveiling a memorial to Mr. JOSEPH CHAMBERLAIN and Mr. AUSTEN CHAMBERLAIN at the Albert Dock Hospital is not without precedent. On more than one occasion party differences have been similarly forgotten. Thus several golf players contributed to *The Daily Telegraph* shilling fund in honour of the great W. G. GRACE some few years ago. Such sinking of private shibboleths is a very excellent thing and goes far to show how thoroughly sound and healthy English public life really is *au fond*.

THE NAMES OF COLLEGES.

Exeter College, Oxford, which has just celebrated its six hundredth anniversary, is not the only college which bears the same name as that of a city. Pembroke is another. Keble is, of course, named after the hymn-writer and divine; and Balliol, where C. S. C. played the wag so divertingly, after Balliol. *A propos* of Oxford, it is a question whether that extremely amusing book, *Verdant Green*, is still much read by freshers.

THE AUTHOR OF *THE LITTLE MINISTER*.

Sir JAMES BARRIE, who is said to have written a revue for production this autumn at a West-End Theatre, must not be confounded with the French sculptor, BARYE, in spite of the similarity of name. BARYE is famous chiefly for his bronzes of lions, and fortunately, in making his studies of these dangerous animals, he escaped the fate which so often befalls the trainer of wild beasts whose animals suddenly turn upon him.

ONCE UPON A TIME.

THE ALIEN.

ONCE upon a time a poet was sitting at his desk in his cottage near the woods, trying to write.

It was a hot summer day and great fat white clouds were sailing across the sky. He knew that he ought to be out, but still he sat on, pen in hand, trying to write.

Suddenly, among all the other sounds of busy urgent life that were filling the warm sweet air, he heard the new and unaccustomed song of a bird. At least not new and not unaccustomed, but new and unaccustomed there, in this sylvan retreat. The notes poured out, now shrill, now mellow, now bubbling like musical water, but always rich with the joy of life, the fulness of happiness. Where had he heard it before? What bird could it be?

Suddenly the poet's housekeeper hurried in. "Oh, Sir," she exclaimed, "isn't it a pity? Someone's canary has got free, and it's singing out here something beautiful."

"Of course," said the poet—"a canary;" and he hastened out to see it. But before he could get there the bird had flown to a clump of elms a little way off, from which proceeded sweeter and more tumultuously exultant song than they had ever known.

The poet walked to the elms with his field-glasses, and after a while he discerned among the million leaves the little yellow bird, with its throat trembling with rapture.

But the poet and his housekeeper were not the only creatures who had heard the strange melody.

"I say," said one sparrow to another, "did you hear that?"

"What?" inquired the other sparrow, who was busy collecting food for a very greedy family.

"Why, listen," said the first sparrow.

"Bless my soul," said the second. "I never heard that before."

"That's a strange bird," said the first sparrow; "I've seen it. It's all yellow."

"All yellow?" said the other. "What awful cheek!"

"Yes, isn't it?" replied the first sparrow. "Can you understand what it says?"

"Not a note," said the second. "Another of those foreigners, I suppose. We shan't have a tree to call our own soon."

"That's so," said the first. "There's no end to them. Nightingales are bad enough, grumbling all night, and swallows, although there's not so many of them this year as usual; but when it comes to yellow birds—well."

"Hullo," said a passing tit, "what's the trouble now?"

"Listen," said the sparrows.

The tit was all attention for a minute while the gay triumphant song went on.

"Well," he said, "that's a rum go. That's new, that is. Novel, I call it. What is it?"

"It's a yellow foreigner," said the sparrows.

"What's to be done with it?" the tit asked.

"There's only one thing for self-respecting British birds to do," said the first sparrow. "Stop it. Teach it a lesson."

"Absolutely," said the tit. "I'll go and find some others."

"Yes, so will we," said the sparrows; and off they all flew, full of righteous purpose.

Meanwhile the canary sang on and on, and the poet at the foot of the tree listened with delight.

Suddenly, however, he was conscious of a new sound—a noisy chirping and harsh squeaking which seemed to fill the air, and a great cloud of small angry birds assailed the tree. For a while the uproar was immense, and the song ceased; and then, out of the heart of the tumult, pursued almost to the ground where the poet stood, fell the body of a little yellow bird, pecked to death by a thousand avenging furies.

Seeing the poet they made off in a pack, still shrilling and squawking, but conscious of the highest rectitude.

The poet picked up the poor mutilated body. It was still warm and it twitched a little, but never could its life and music return.

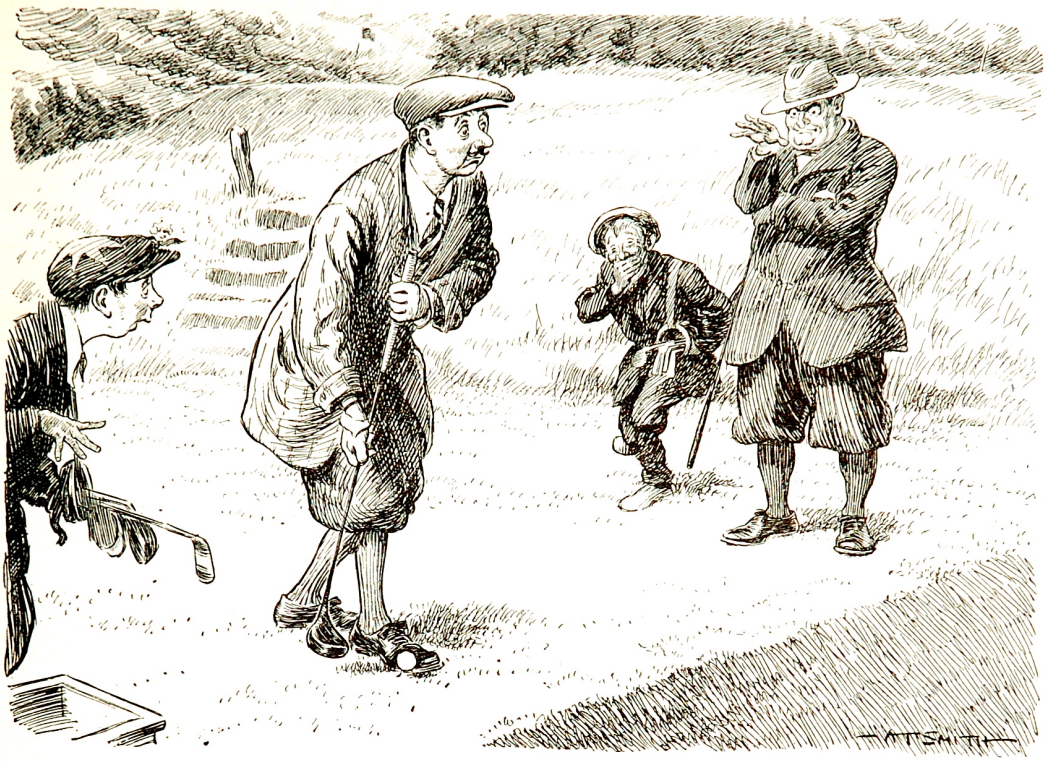
While he stood thoughtfully there an old woman, holding an open cage and followed by half-a-dozen children, hobbled along the path.

"My canary got away," she said. "Have you seen it? It flew in this direction."

"I'm afraid I have seen it," said the poet, and he opened his hand.

"My little pet!" said the old woman. "It sang so beautifully, and it used to feed from my fingers. My little pet."

The poet returned to his work. "'In tooth and claw,'" he muttered to himself, "'In tooth and claw.'"



HOW TO UTILISE THE ART OF "SUGGESTION."

THE DOCTOR, SIX DOWN AT THE TURN, "SUGGESTS" TO HIS OPPONENT THAT THEY ARE PLAYING CROQUET, AND WINS BY TWO AND ONE.

OUR BOOKING-OFFICE.

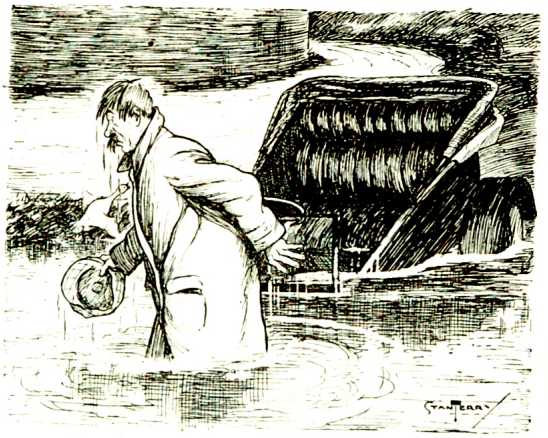
(By Mr. Punch's Staff of Learned Clerks.)

Tents of a Night (SMITH, ELDER) is a quite ordinary story, about entirely commonplace persons, which has however an original twist in it. I never met a story that conveyed so vividly the nastiness of a summer holiday that isn't nice. The holiday was in Brittany, just the common round, Cherbourg, Coutances, Mont St. Michel, and the rest of it; and the holiday-makers were *Mr.* and *Mrs. Hepburn*, their niece *Anne*, and a rather pleasant flapper named *Barbara* whom they had taken in charge. *Anne* is the heroine and central character of the holiday; and certainly whatever discomforts it contained she seems to have done her successful best to add to. "This is a beastly place!" was her written comment upon St. Michel; and it was typical of her attitude throughout. Of course the real trouble with *Anne* was something deeper than drains or crowded hotels or the smell of too many omelettes: she was in love. Apparently she was more or less in love with two men, *Dragotin Voinovich* (whose name was a constant worry to *Anne's* aunt, and I am bound to say that I share her feelings about it) and *Jimmy Fordyce*, a pleasant young Englishman who pulls the girls out of quicksands and makes himself generally agreeable. In the end, however—but on second thoughts the end, emotionally speaking, of *Anne* is just what I shall not tell you, as it is precisely the thing that redeems the book from being commonplace. This you will enjoy; and also those remarkably real descriptions of various plage-hotels in August, the noise, the crowds, the long hot meals, the sunshine and constant wind, the sand on the staircase, and the general atmosphere of wet bathing-gowns—all these are a luxurious delight to read about in a comfortable English room. Miss MARY FINDLATER evidently knows them.

Dippers who have given a new meaning to the classical motto, *Respice finem*, are so common amongst novel readers that PATRICIA WENTWORTH will only have herself to thank if many who are unfamiliar with her work fail to do justice to a book nine-tenths of which is thoroughly interesting and excellently well-written. As a boy, the hero of *Simon Heriot* (MELROSE) is misunderstood, and although *Mr. Martin*, his step-father, is a somewhat stagey specimen of the heavy and vulgar papa, the child's emotions (as, for instance, when he pretends that the storm of his parent's wrath is the ordeal of the Inquisition or some far-away battle of paladins in which he is contending) are finely conceived, and many of the later passages in *Simon's* life—his unhappy love affair with *Maud Courtney*, his relations with his grandmother and with *William Forster*, the schoolmaster—are quite engrossing and give occasion for memorable sketches of character. It is when the natural end of the story is reached, and *Simon* 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 the young man apparently

crippled for life, but the last chapters, in which he finds spiritual comfort and (after the doctors have given up hope) complete anatomical readjustment through the ministrations of faith healing, alienated me entirely. From the outset the obvious scheme of the novel is to bring the hero back happily to the home and, if you will, the rustic church of his ancestors; and, though the science of Christian healing may do all that its adherents claim for it, it has about as much to do with the case of *Simon Heriot* as the dancing dervishes or the rites of Voodoo.

DEMETRA VAKA has melted my literary heart. By way of homage to her I eat the dust and recant all the hard and bitter things I said and thought in my youth concerning Ancient Greece; especially I apologise, on behalf of myself and my pedagogues, for ever regarding its language as a dead one. *A Child of the Orient* (LANE) has taught me better, though the last object the author appears to have in view is to educate. This "Greek girl brought up in a Turkish household" writes to amuse, entertain and charm, and her success is abundant. Whether it is attributable to the romantic particulars of the Turkish household or to the ingenuous personality of the Greek girl, I hesitate to say, since both are so captivating; but this I know, that, considered as descriptive sketches or personal episodes, each of the twenty-two chapters is a separate delight. For the ready writer material is not wanting in the Near East; a fine theme is provided in the national ambition of the Greek, who cannot forget his glorious past and be content with his less conspicuous present. As for the love interest, who should supply this better than the Turk? In these days of cosmopolitanism there are bound to be romantic complications in the lives of a polygamous people situate in a monogamous continent. By way of postscript the authoress travels abroad and deals with alien matters; her impression, I gather, is that if her ancestors of classical times could see our world of to-day and express an opinion upon it the best of their praise would be reserved for the fact of the British Empire, and the worst of their abuse be spent upon what is known as American humour. I am so constituted that I cannot but be prejudiced in favour of a writer gifted with so profound a judgment.



BARGAIN: Two-seater, with most of the accessories; only done fifty miles; water-cooled engine; owner giving up driving.

The creatrix of *Pam* must look to her laurels. Slovenliness is the aptest word to apply to the workmanship of *Maria* (HUTCHINSON), the latest heroine of the Baroness VON HUTTEN. *Maria* has the air of having been contracted for, while that fastidious overseer who lurks at the elbow of every honest craftsman, condemning this or that phrase, readjusting the other faulty piece of construction, has frankly abandoned the contractor. *Maria* was the daughter of an artist cadger (name of *Drello*), friend of the great and seller of their autograph letters, whereby he was astute enough to make a comfortable living. *Maria* had a dull brother named *Laertes*, who accidentally met a highness, who fell very abruptly in love with *Maria* and made her strictly dishonourable proposals. *Maria* drew herself up, compelled him to apologise and go away, until the nineteenth chapter, when she made similar proposals to the highness, now a duly and unhappily married *King of Sarmania*. But she is saved by the chivalrous love-lorn dwarf, *Tomsk*, who, with the irascible singing-master *Sulzer*, is responsible for the chief elements of vitality in this rather suburban romance. And I found myself never believing in *Maria's* wondrous beauty and quite sharing *Sulzer's* poor opinion of her singing. But this of course was mere prejudice.

In *Grizel Married* (MILLS AND BOON) Mrs. GEORGE DE HORNE VAIZEY exhibits the highest-handed method of treating Romance that ever I met. For consider the situation to be resolved. *Dane Peignton* was engaged to *Teresa*, but in love with *Lady Cassandra Raynor*, whose husband, I regret to add, was still alive. *Dane* and *Cassandra* had never told their love, and concealment might have continued to prey on their damask cheeks, if Mrs. VAIZEY had not (very naturally) wished to give us a big emotional scene of avowal. It is the way in which this is done that compels my homage. Off go the characters on a picnic, obviously big with fate. *Teresa* goes, and *Dane* and *Cassandra*, the fourth being *Grizel*, whom you may recall pleasantly from an earlier book; but, though she fills the title *rôle* in this one, she has little to do with its development. Of course I saw that something tragic was going to happen to somebody on that picnic—cliffs or tides or mad bulls or something. But I don't suppose that in twenty guesses you could get at the actual instrument of destiny. *Cassandra* chokes over a fish-bone! That 's what I meant about Mrs. VAIZEY's courage. And the reward of it is that, after your first moment of incredulity, the fish-bone isn't in the least bit absurd. Poor *Cassandra* comes quite near to expiring of it; and *Dane*, having thumped and battered her into safety, sobs out his wild and whirling passion, while *Grizel* and poor *Teresa* have just to sit about and listen. It really is rather a striking and original climax; incidentally it is far the best scene in an otherwise not very brilliant tale. But, having attended that picnic, I shall be astonished if you don't want to go on to the end and see how it all straightens out.

"At 9.30 o'clock, 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."
Daily Chronicle.

"Our Special Correspondent's" father seems to be a big man.

"While the class watches, the teacher pronounces all the words. Then the whole class pronounces them while the teacher points, skipping around."—*Hawaii Educational Review.*

A pretty scene, if the teacher is a man of graceful movements.

4 Talks

Why and how do we encode texts? Basic choices and decisions. The TEI.

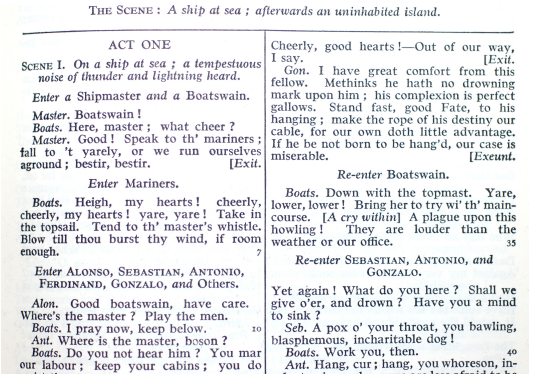
Sebastan Rahtz and other TEI@Oxford authors

February 2009

What is markup?

In order to talk about texts, markup and encoding of texts, we need to understand what we mean by these basic concepts. When we talk about text encoding, what do we mean by a text? What is in a text and what assumptions do we make in reading them?

What's in a text?



THE SCENE: *A ship at sea; afterwards an uninhabited island.*

ACT ONE

SCENE I. *On a ship at sea; a tempestuous noise of thunder and lightning heard.*

Enter a Shipmaster *and a* Boatswain.

Master. Boatswain!
Boats. Here, master; what cheer?
Master. Good! Speak to th' mariners; fall to 't yarely, or we run ourselves aground; bestir, bestir. [*Exit.*

Enter Mariners.

Boats. Heigh, my hearts! cheerly, cheerly, my hearts! yare, yare! Take in the topsail. Tend to th' master's whistle. Blow till thou burst thy wind, if room enough.

Enter ALONSO, SEBASTIAN, ANTONIO, FERDINAND, GONZALO, *and* Others.

Alon. Good boatswain, have care. Where's the master? Play the men.
Boats. I pray now, keep below.
Ant. Where is the master, boson?
Boats. Do you not hear him? You mar our labour; keep your cabins; you do Cheerly, good hearts!—Out of our way, I say. [*Exit.*
Gon. I have great comfort from this fellow. Methinks he hath no drowning mark upon him; his complexion is perfect gallows. Stand fast, good Fate, to his hanging; make the rope of his destiny our cable, for our own doth little advantage. If he be not born to be hang'd, our case is miserable. [*Exeunt.*

Re-enter Boatswain.

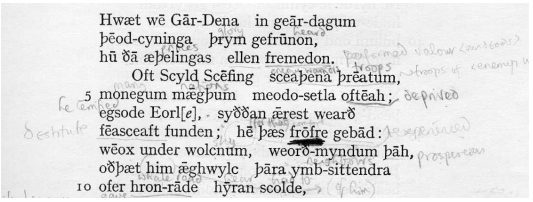
Boats. Down with the topmast. Yare, lower, lower! Bring her to try wi' th' main-course. [*A cry within*] A plague upon this howling! They are louder than the weather or our office.

Re-enter SEBASTIAN, ANTONIO, *and* GONZALO.

Yet again! What do you here? Shall we give o'er, and drown? Have you a mind to sink?
Seb. A pox o' your throat, you bawling, blasphemous, incharitable dog!
Boats. Work you, then.
Ant. Hang, cur; hang, you whoreson, in-

What's in a text (2)?

What's in a text (3)?



Hwæt wē Gār-Dena in geār-dagum
þēod-cyninga þrym gefrūnon,
hū ðā æþelingas ellen fremedon.
Oft Scyld Scēfing sceaþena þrēatum,
monegum mǣgþum meodo-setla oftēah;
egsode Eorl[*e*], syððan ærest wearð
fēasceaft funden; hē þæs frōfre gebād:
wēox under wolcnum, weorð-myndum þāh,
oðþæt him æghwylc þāra ymb-sittendra
ofer hron-rāde hȳran scolde,

The ontology of text

Where is the text?

- in the shape of letters and their layout?
- in the original from which this copy derives?
- in the stories we read into it? or in its author's intentions?

A "text" is an abstraction, created by or for a community of readers. Markup encodes and makes concrete such abstractions.

Encoding of texts

- Texts are more than sequences of encoded glyphs
  - They have structure and content
  - They also have multiple readings
- Encoding, or markup, is a way of making these things explicit

Only that which is explicit can be reliably processed

Styles of markup

- In the beginning there was *procedural* markup
  `RED INK ON; print balance; RED INK OFF`
- which being generalised became *descriptive* markup <balance type='overdrawn'>some numbers</balance>
- also known as encoding or annotation

descriptive markup allows for easier re-use of data

What's the point of markup?

- To make explicit (to a machine) what is implicit (to a person)
- To add value by supplying multiple annotations
- To facilitate re-use of the same material
  - in different formats
  - in different contexts
  - by different users

It's (usually) more useful to markup what we think things *are* than what they *look like*

Separation of form and content

- Presentational markup cares more about fonts and layout than meaning
- Descriptive markup says what things are, and leaves the rendition of them for a separate step
- Separating the form of something from its content makes its re-use more flexible
- It also allows easy changes of presentation across a large number of documents

Markup as a scholarly activity

- The application of markup to a document can be an intellectual activity
- In deciding what markup to apply, and how this represents the original, one is undertaking the task of an editor
- There is (almost) no such thing as neutral markup -- all of it involves interpretation
- Markup can assist in answering research questions, and the deciding what markup is needed to enable such questions to be answered can be a research activity in itself
- Good textual encoding is never as easy or quick as people would believe
- Detailed document analysis is needed before encoding for the resulting markup to be useful

What does markup capture?

Compare

```
<hi rend="dropcap">H</hi>&amp;WYN;ÆT WE GARDE <lb/>na in
gear-dagum þeod-cyninga <lb/>þrym gefrunon,
hu ða æþelingas <lb/>ellen fremedon. oft scyld scefing
sceaþe
<add>na</add>
<lb/>þreatum, moneg<expan>um</expan> mægþum meodo-setl
<add>a</add>
<lb/>of<damage>
 <desc>blot</desc>
</damage>teah ...
```

and

```
<lg>
 <l>Hwæt! we Gar-dena in gear-dagum</l>
 <l>þeod-cyninga þrym gefrunon,</l>
 <l>hu ða æþelingas ellen fremedon,</l>
</lg>
<lg>
 <l>Oft Scyld Scefing sceaþena þreatum,</l>
 <l>monegum mægþum meodo-setla ofteah;</l>
 <l>egsode Eorle, syððan ærest wearþ</l>
 <l>feasceaft funden...</l>
</lg>
```

A useful mental exercise

Imagine you are going to markup several thousand pages of complex material....

- Which features are you going to markup?
- Why are you choosing to markup this feature?
- How reliably and consistently can you do this?

Now, imagine your budget has been halved. Repeat the exercise!

XML, TEI, etc

Extensible Markup Language (XML) is a simple, very flexible text format derived from SGML (ISO 8879). Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere.

Why the TEI?

The TEI provides

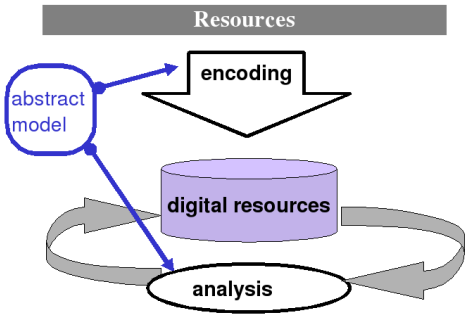
- a language-independent framework for defining markup languages
- a very simple consensus-based way of organizing and structuring textual (and other) resources...
- ... which can be enriched and personalized in highly idiosyncratic or specialised ways
- a very rich library of existing specialised components
- 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

The wrong way of thinking about the TEI

- A traditional (if large) research project with soft funding, driven by academic curiosity
- a codification of best practice, with no formal maintenance method
- uncertain licencing and development practices
- unmanageably complex except by the priesthood — or simultaneously as too simple for real scholarly work
- lack of specific tools to *do* something with a TEI text
- failure to market the advantages of rich markup

The good way of thinking about the TEI

- Proper open source licence, with visible development on Sourceforge
- Architecture rethought to facilitate expansion and integration with other systems
- Self documenting, each release fully validated, delivered using standard mechanisms
- Publicly available processing tools managed together with the Guidelines
- Active developer community, wiki, etc. Test files, exemplars, regular updates...

One Specification Language

- A set of TEI documents is described by an ODD, which is itself a TEI document that combines:
  - references to existing declarations
  - formal declarations for elements and attributes
  - documentation and usage notes
- Underlying this:
  - a conceptual model which abstracts from specific elements to generic classes
  - a modular architecture for combining sets of definitions
- specifications are chainable; modifications are written in ODD with ODD as input and output
- Roma is one interface to this: there will be others

For example

A TEI ODD file is a valid TEI document, containing as much discursive prose as you want, and a `<schemaSpec>` element to define the schema it documents

```
<text>
 <body>
  <div>
   <head>Our Project Manual</head>
   <p>In this project we use the basic TEI structures
with a few minor modifications to
       exclude elements we don't plan to use.</p>
   <schemaSpec ident="TEI-minimal" start="TEI">
    <moduleRef key="tei"/>
    <moduleRef key="header"/>
    <moduleRef key="core"/>
    <moduleRef key="textstructure"/>
<!-- We don't need these drama elements: -->
    <elementSpec ident="sp" mode="delete" mod-
ule="core"/>
    <elementSpec ident="speaker" mode="delete" mod-
ule="core"/>
   </schemaSpec>
  </div>
 </body>
</text>
```

Support for many schema languages

- TEI schemas can be generated for
  - XML DTD language
  - ISO RELAX NG language
  - W3C Schema Language
- Content models are defined using RELAX NG syntax
- Datatypes are defined in terms of W3C datatypes
- Some facilities (e.g. alternation, namespaces) cannot be expressed in DTD
- Additional constraints can be expressed in Schematron

Two reasons why standards fail

- The theory is not yet ripe
- The "not invented here" attitude: the community of users is too diverse

Coping with partially-baked ideas

In a TEI ODD, you can ...

- constrain the domain of a value list
- enforce schematron rules about e.g. codependency
- provide new elements in your own namespace
- remove (non-mandatory) child elements

New elements

A schema is a grammar. How can you add new terminals to an existing syntax?

- All content models are expressed indirectly, by reference to element classes rather than elements
- Hence adding a new element is simply a matter of saying which class/es it belongs to

The TEI schema is also enriched with semantics. How can you explain what a new element means?

- Class membership also conveys some semantics
- ODD includes detailed documentation

Do not re-invent the wheel

- TEI P5 has extensive I18N features for translation of ...
  - schema objects
  - schema documentation
- TEI is hospitable to other namespaces:
  - You can use SVG for graphics, MathML for math, Word Table markup if you like
- ODD also includes an `<equiv>` element for mapping to external ontologies

For example

Embedding SVG within TEI:

```
<figure>
 <svg xmlns="http://www.w3.org/2000/svg"
  width="6cm" height="5cm" viewBox="6 3 6 5">
  <ellipse xmlns="http://www.w3.org/2000/svg"
   style="fill:
#ffffff" cx="9.75" cy="6.35" rx="2.75" ry="2.35"/>
 </svg>
</figure>
```

A user-defined attribute:

```
<div
   xmlns:my="http://www.example.org/ns/nonTEI">
 <p n="12" my:topic="rabbits">Flopsy, Mopsy, Cottontail,
and Peter...</p>
</div>
```

NVDL processors validate against multiple namespace schemas

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 TEI-conformable)

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

Using the basic TEI structural elements

Lou Burnard and other TEI@Oxford authors

February 2009

TEI Infrastructure

- The TEI encoding scheme consists of a number of modules
- These declare XML elements and their attributes
- An element's declaration assigns it to one (or more) model classes
- Another part declares its possible content and attributes with reference to these classes
- This indirection allows strength and flexibility
- It makes it easy to add/exclude new elements by referencing existing classes

What is a module?

- A convenient way of grouping together a number of element declarations
- These are usually on a related topic or specific application
- Most chapters focus on elements drawn from a single module, which that chapter then defines
- A TEI Schema is created by selecting modules and add/removing elements from them as needed

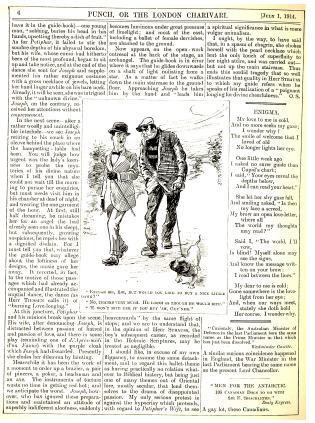
Modules

Module name	Chapter
analysis	Simple Analytic Mechanisms
certainty	Certainty and Responsibility
core	Elements Available in All TEI Documents
corpus	Language Corpora
dictionaries	Dictionaries
drama	Performance Texts
figures	Tables, Formulae, and Graphics
gaiji	Representation of Non-standard Characters and Glyphs
header	The TEI Header
iso-fs	Feature Structures
linking	Linking, Segmentation, and Alignment
msdescription	Manuscript Description
namesdates	Names, Dates, People, and Places
nets	Graphs, Networks, and Trees
spoken	Transcriptions of Speech
tagdocs	Documentation Elements
tei	The TEI Infrastructure
textcrit	Critical Apparatus
textstructure	Default Text Structure
transcr	Representation of Primary Sources
verse	Verse

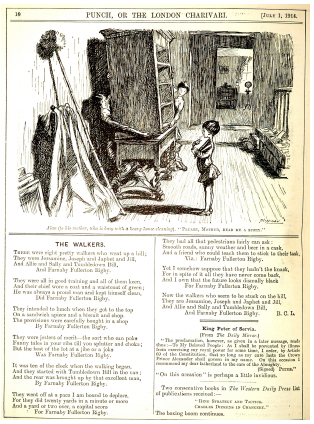
The Imaginary Punch Project

- 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?

Punch example page 1



Punch example page 1



PUNCH, OR THE LONDON CHARIVARI.

THE WALKERS.

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...

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>
```

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 `<div type="page">`.

```
<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>
<div type="poem">
  <head>Enigma</head>
  <lg>
    <l>...</l>
  </lg>
</div>
<div type="snippets">
  <p>...</p>
</div>
```

The sequence in which divisions appear is rather arbitrary.

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.

Floating text

As mentioned above, `<div>`s must tesselate over the entire text

```
<div1>
  <p> ... </p>
  <div2>
    <p> ... </p>
  </div2>
  <div2>
    <p> ... </p>
  </div2>
</div1>
```

is valid BUT

```
<div1>
  <p> ... </p>
  <div2>
    <p> ... </p>
  </div2>
  <p> ... </p>
</div1>
```

is *not valid.*

A special `<floatingText>` element is available for "interruptions"

What are divisions made of?

(apart from other smaller divisions)

- <head> (heading)
- <p> (paragraph)
- <sp> (speech, contains any of the foregoing, also <stage> and <speaker>)
- <list> (contains <head>, <label>, <item>)
- <table>, (contains <row> containing <cell>) ...
- <l> (verse line) optionally grouped into <lg> (line group) stanzas
- <figure> (contains <graphic>, <figDesc>, <head>...)

For example....

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="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>
    <speaker>Turkey.</speaker>
    <p> No good for me. Shan't be ready till the autumn.</p>
  </sp>
</div>
```

Punch example page

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>
      <cell>6</cell>
    </row>
    <row>
      <cell>For howling, kicking, or biting during service in church, per howl,
kick, or
          bite ...</cell>
      <cell>0</cell>
      <cell>10</cell>
      <cell>6</cell>
    </row>
    <row>
      <cell>For sitting on doorsteps of obnoxious persons, per hour, if fine
```

Global attributes

Some features (potentially) apply to everything:

- identity
- language
- rendition

TEI provides global attributes for these:

- *@xml:id* provides a unique identifier for any element;
- *@n* provides a name or number for any element
- *@xml:lang* specifies the language of any element, using an ISO standard code
- *@rend* and *@rendition* provide ways of specifying the visual appearance (rendition) of any element

Punch example page

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>
```

Macrostructure 1

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 `<div>` 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>
```

Macrostructure 2

As well as the texts, we have detailed metadata about each volume, and images of its pages. These are the three parts of a canonical TEI document:

```
<TEI>
  <teiHeader>
<!-- required; provides metadata -->
  </teiHeader>
  <facsimile>
<!-- the text, represented in image form -->
  </facsimile>
  <text>
<!-- the text, transcribed and marked up -->
  </text>
</TEI>
```

Macrostructure 3

If many such documents are grouped together to form a corpus (rather than a collection), it may be useful to factor out the metadata they have in common:

```
<teiCorpus>
  <teiHeader>
<!-- shared metadata -->
  </teiHeader>
  <TEI>
    <teiHeader>
<!-- specific metadata -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
  <TEI>
    <teiHeader>
<!-- specific metadata -->
    </teiHeader>
    <text>
<!-- ... -->
    </text>
  </TEI>
</teiCorpus>
```

What kinds of metadata?

For *IPP* and for any other comparable project, we will need a place for such information as

- identification of the resource itself ("what is this thing?")
- statements of responsibility ("who did what when?")
- indication of source ("what was this derived from?")
- publication statement ("how is this item distributed and by whom?")
- declaration of encoding practice ("what do the codes we added mean?")

The TEI Header supports all these, and more.

The TEI Header

The TEI header was designed with two goals in mind

- needs of bibliographers and librarians trying to document 'electronic books'
- needs of text analysts trying to document 'coding practices' within digital resources

On the one hand, the Librarian's header

- uses standard bibliographic concepts
- respects established mappings to other such records (e.g. MARC)
- has a preference for structured data over loose prose

On the other, Everyman's header

- Supports a (potentially) huge range of very miscellaneous information, organized in fairly ad hoc ways
- Unpredictable combinations of narrowly encoded documentation systems and loose prose descriptions

TEI Header Structure

The TEI header has four main components:

- <fileDesc> (file description) contains a full bibliographic description of an electronic file.
- <encodingDesc> (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived.
- <revisionDesc> (revision description) summarizes the revision history for a file.
- <profileDesc> (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting. (just about everything not covered in the other header elements

Only <fileDesc> is required; the others are optional.

Simple TEI Header for IPP

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Punch, or the London Charivari, Vol. 147, July 1, 1914</title>
    </titleStmt>
    <publicationStmt>
      <idno type="gutenberg">24357</idno>
      <availability>
        <p>This text is freely available for re-use under US and UK law,
              consult your local legal restrictions if elsewhere.</p>
      </availability>
    </publicationStmt>
    <sourceDesc>
      <p>This text is a TEI version of a Project Gutenberg text originally
           located at <ptr
         target="http://www.gutenberg.org/dirs/2/4/3/5/24357/"/>. As per their
          license agreement we have removed all references to the PG
          trademark.</p>
    </sourceDesc>
  </fileDesc>
  <revisionDesc>
    <change when="2008-07-26T23:49:55.968+01:00"/>
  </revisionDesc>
</teiHeader>
```

Below the paragraph...

Within the elements already introduced, TEI offers plenty of scope for mark-up of smaller components. For example:

- boundaries, such as page, column, or line breaks
- highlighting, emphasis and quotation
- editorial changes such as correction, normalization etc.
- names, numbers, dates, addresses...
- links and cross-references
- notes, annotation, indexing
- graphics
- bibliographic citations
- words and other analyses

Highlighting

By highlighting we mean any combination of typographic features (font, size, hue, etc.) which distinguishes the highlighted text from its surroundings. This may be for many reasons...

- to mark foreign, archaic, technical usages
- for emphasis when spoken
- to show something is not part of the text.. (e.g. cross references, titles, headings)
- or is attributed to some other agency inside or outside the text (e.g. direct speech, quotation)

TEI provides both a generic <hi> tag and a large number of specific ones...

A few highlighting examples

- <hi> (highlighted: reason unknown or unimportant)

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

- <emph> (emphasized)

```
<said>'E won't bite yer <emph>if you buy 'im</emph> guv'ner.</said>
```

- <title> and <foreign>:

```
<p>
  <foreign xml:lang="fr">À propos</foreign> of Oxford, it is a question
whether that
extremely amusing book <title>Verdant Green</title> is still much read by
freshers.
</p>
```

- <distinct> (linguistically marked)

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

Quotation

Quotation marks can similarly be used to set off text for many reasons:

- <q> (used if the reason is unknown or unimportant)
- <said> (speech or thought)
- <quote> (attributed to an external source)
- <mentioned> and <soCalled> (nuances of narrative status)

```
<p>
  <said who="#Celia">I know a lovely tin of potted grouse,</said> said Celia,
and she
went off to cut some sandwiches.
</p>
```

```
<head>How to utilise the art of <soCalled>suggestion</soCalled>
</head>
<head>The Doctor, six down at the turn, <soCalled>suggests</soCalled> to his
opponent that
they are playing croquet, and wins by two and one.</head>
```

Note that these elements can nest within one another:

```
<p>The poet returned to his work. <said>
   <quote>In tooth and claw,</quote>
  </said> he muttered to himself, <said>
   <quote>In tooth and claw.</quote>
  </said>
```

Editorial intervention

As a simple example, consider: 'Excuse me sir, but would you like to buy a nice little dawg?' on page 6.
We can:

- use <orig> to show that "dawg" is what it says, even though this is a nonstandard spelling
- use <reg> to show that "dog" is an editorially-supplied regularisation of what it says
- or provide both within a <choice> element to say either is a valid encoding:

```
...a nice little <choice>
 <orig>dawg</orig>
 <reg>dog</reg>
</choice>?
```

Names of persons, places, things...

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

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

Examples of names

Using *@type* to distinguish personal from geographic names:

```
<p>The scene opens at a party given by <name  type="person">Potiphar</name> in
<name  type="place">Venice</name>.  </p>
```

Using *@key* and *@ref* to de-reference names:

```
<p>
  <label>Business done.</label>—The Commons still harping on the
Budget.  <name
     type="person"
     ref="http://en.wikipedia.org/wiki/Timothy_Michael_Healy"> Tim
    Healy</name> enlivened proceedings by vigorous personal attack on
<q>the most reckless and incapable <rs  key="LLG">Chancellor of the
       Exchequer</rs> that ever sat on the Treasury Bench.</q>
  <name  key="LLG">Lloyd George's</name> retort courteous looked forward to
with interest.
</p>
```

Dates

- <date> contains a date and time in any format
- For processing it is convenient to add a normalized version, using the *@when* attribute
- Uncertain dates and times, and ranges, can be indicated by other attributes: *@notBefore*, *@notAfter*, *@from@to*

```
<p>House of Commons, <date  when="1914-06-22"> Monday, June 22,
1914</date>.</p>
<p>
  <date  notAfter="1914-06-01"  notBefore="1914-03-01"> Sunday, a month
ago,</date> was
hot.
</p>
```

Cross references

A cross reference is a link from one point in a text (the source) to another (the target).
TEI provides generic elements <ptr> and <ref> for this purpose.
If the linking text can be automatically generated use <ptr>; otherwise use <ref>.
The source is the location of the <ptr> or <ref>; the target is specified by the *@target* attribute, in the form of a URI reference.

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

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

Bibliographic Citations

TEI provides special elements for bibliographic citations or references:

- <bibl> (loosely structured)
- <biblStruct> (standard bibliographic structure)
- <listBibl> (encloses a bibliography)

These are typically used in preparing bibliographies, or in footnotes. But even in Punch, there are examples.

Simple <bibl> Example

In Punch, bibliographic citations are usually associated with a a quotation from another paper:
The<cit> element groups the two:

```
<cit>
  <quote>It was the time when Henry III. was batting with Simon de Montfort and
his
     Barons.</quote>
  <bibl>
    <title>Straits Times.</title>
  </bibl>
</cit>
```

Embedded notes

Notes, whether appearing in the original source, or added by an editor, can be marked using the <note> element.
We might use this to add biographical details to the Punch transcriptions:

```
<p>By-the-by, it is denied that Sir <name  rend="sc">Joseph Beecham</name>
  <note>Sir Joseph Beecham, 1st Baronet (8 June 1848 - 23 October
1916)...</note>. was
in any way responsible for the Government's "Pills for Earthquakes," by which
it was
hoped to avert the Irish crisis.</p>
```

<note> has attributes *@place* and *@resp*

Linked notes

Since we have several references to the same person, it might be better to put the notes elsewhere and point to them from the names:

```
<div  type="notes">
  <note  xml:id="BEECHJO">Sir Joseph Beecham, 1st Baronet (8 June 1848 -
     23 October 1916) the eldest son of Thomas Beecham (1820-1907) played a
     large part in the growth and expansion of his father's medicinal pill
     business which he joined in 1866....</note>
<!-- other notes -->
</div>
<div  type="snippets">
  <p>... Both Earl <name  rend="sc">Beauchamp</name> and <name>Sir
<ref  target="#BEECHJO">Joseph Beecham</ref>
    </name> appear in the recent
     Honours List.</p>
  <p>By-the-by, it is denied that Sir <name  rend="sc"  ref="#BEECHJO">Joseph
Beecham</name> was in any way responsible...</p>
</div>
```

Could also use specialised <person> element, in this case.
"Elsewhere" can be *anywhere* on the Internet...

The TEI header and bibliographies

TEI@Oxford authors

February 2009

The TEI header

The TEI header is the bibliographic record for the electronic file, storing information about the file itself, how it is was made, categorisations and analytical information about the text, and revision control information. It is one of the most important aspects of any TEI file.

TEI Header Structure

The TEI header has four main components:

- `<fileDesc>` (file description) contains a full bibliographic description of an electronic file.
- `<encodingDesc>` (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived.
- `<profileDesc>` (text-profile description) provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting. (just about everything not covered in the other header elements
- `<revisionDesc>` (revision description) summarises the revision history for a file.

Only `<fileDesc>` is required; the others are optional.

Example Header: Minimal required header

```
<teiHeader>
 <fileDesc>
  <titleStmt>
   <title>A title?</title>
  </titleStmt>
  <publicationStmt>
   <p>Who published?</p>
  </publicationStmt>
  <sourceDesc>
   <p>Where from?</p>
  </sourceDesc>
 </fileDesc>
</teiHeader>
```

Types of content in the TEI header

- free prose
  - prose description: series of paragraphs
  - phrase: character data, interspersed with phrase-level elements, but not paragraphs
- grouping elements: specialised elements recording some structured information
- declarations: Elements whose names end with the suffix Decl (e.g. subjectDecl, refsDecl) enclose information about specific encoding practices applied in the electronic text.
- descriptions: Elements whose names end with the suffix Desc (e.g. `<settingDesc>`, `<projectDesc>`) contain a prose description, possibly, but not necessarily, organised under some specific headings by suggested sub-elements.

File Description

- has some mandatory parts:
  - `<titleStmt>`: provides a title for the resource and any associated statements of responsibility
  - `<sourceDesc>`: documents the sources from which the encoded text derives (if any)
  - `<publicationStmt>`: documents how the encoded text is published or distributed
- and some optional ones:
  - `<editionStmt>`: yes, electronic texts have editions too
  - `<seriesStmt>`: and they also fit into "series".
  - `<extent>`: how many floppy disks, gigabits, files?
  - `<notesStmt>`: notes of various types

NB A "file" may actually correspond with several operating system files.

The File Description

- <titleStmt>: contains a mandatory <title>which identifies the electronic file (not its source!)
- optionally followed by additional titles, and by 'statements of responsibility', as appropriate, using <author>, <editor>, <sponsor>, <funder>, <principal> or the generic <respStmt>
- <publicationStmt>: may contain
  - plain text (e.g. to say the text is unpublished)
  - one or more <publisher>, <distributor>, <authority>, each followed by <pubPlace>, <address>, <availability>, <idno>

The Source Description

Many electronic texts were not 'born digital': their source/s need specification in traditional bibliographic style

- <bibl>, <biblStruct>
- (for texts which were born digital): <biblFull> may contain a nested <fileDesc>
- <listBibl> a list of the foregoing
- prose description
- more specialised elements are available for spoken texts (<recordingStmt> etc.) and for manuscripts (<msDescription>)

For Example

```
<sourceDesc>
 <bibl>
  <title level="a">Enigma</title>,
<title level="j">Punch: or the London
     Charivari</title>, <date when="1914-07-01">July 1,
1914</date>, 147, p. 6</bibl>
</sourceDesc>
```

Association between header and text

By default everything asserted by a header is true of the text to which it is prefixed. This can be over-ridden:

- as when a text header over-rides or amplifies a corpus-header setting
- when model.declarable elements are selected by means of the *@decls* attribute (available on all model.declaring elements)
- using special purpose selection/definition elements e.g. <catRef> and <taxonomy> (see below)

Most components of the encoding description are declarable.

Encoding Description

<encodingDesc> groups notes about the procedures used when the text was encoded, either summarised in prose or within specific elements such as

- <projectDesc>: goals of the project
- <samplingDecl>: sampling principles
- <editorialDecl>: editorial principals, e.g. <correction>, <normalization>, <quotation>, <hyphenation>, <segmentation>, <interpretation>
- <classDecl>: classification system/s used
- <tagsDecl>: specifics about usage of particular elements

The <encodingDesc> can replace the user manual, or facilitate semi-automatic document management, given agreed codes of practice.

<rendition> element

- <rendition>: structured information about appearance in the source document

```
<tagsDecl>
 <rendition xml:id="r-center" scheme="css">text-align:
center;</rendition>
 <rendition xml:id="r-small" scheme="css">font-size:
small;</rendition>
 <rendition xml:id="r-large" scheme="css">font-size:
large;</rendition>
</tagsDecl>
```

<appInfo> element

- <appInfo>: structured information about an application which has edited this TEI file

```
<appInfo>
 <application version="1.7" ident="ImageMarkupTool" no-
tAfter="2008-06-01">
  <label>Image Markup Tool</label>
  <ptr target="#P1"/>
  <ptr target="#P2"/>
 </application>
</appInfo>
```

Profile Description

An extensible rag-bag of descriptions, categorised only as 'non-bibliographic'. Default members of the model.profileDescPart) class include:

- <creation>: information about the origination of the intellectual content of the text, e.g. time and place
- <langUsage>: information about languages, registers, writing systems etc used in the text
- <textDesc> and <textClass>: classifications applied to the text by means of a list of specified criteria or by means of a collection of pointers, respectively
- <particDesc> and <settingDesc>: information about the 'participants', either real or depicted, in the text
- <handList>: information about the hands identified in a manuscript

Classification Methods

<textClass> provides a classification (by domain, medium, topic...) for the whole of a text expressed in one or more of the following ways:

- direct reference to a locally defined category (using <catRef>)
- reference to an externally defined category (using <classCode>)
- documented by <keywords>

Example

```
<textClass>
 <catRef target="#X123"/>
 <classCode scheme="DD12">001.9</classCode>
 <keywords scheme="DD">
  <term>End of the World</term>
  <term>Day of Judgment</term>
  <term>Apocalypse</term>
 </keywords>
</textClass>
<classDecl>
 <taxonomy>
  <category xml:id="X1">
   <catDesc>Homiletic writing</catDesc>
   <category xml:id="X123">
    <catDesc>Day of Judgment</catDesc>
   </category>
  </category>
 </taxonomy>
</classDecl>
```

Detailed characterization of a text

<textDesc> provides a description of a text in terms of its 'Situational parameters'

```
<textDesc n="novel">
 <channel mode="w">print; part issues</channel>
 <constitution type="single"/>
 <derivation type="original"/>
 <domain type="art"/>
 <factuality type="fiction"/>
 <interaction type="none"/>
 <preparedness type="prepared"/>
 <purpose type="entertain" degree="high"/>
 <purpose type="inform" degree="medium"/>
</textDesc>
<!-- These subelements constitute the class
model.textDescPart: redefine that to roll your own. -->
```

Language and character set usage

The <langUsage> element is provided to document usage of languages in the text. Languages are identified by their ISO codes:

```
<langUsage>
 <language ident="en">English</language>
 <language ident="bg-cy">Bulgarian in Cyrillic
characters </language>
 <language ident="bg">Romanized Bulgarian</language>
</langUsage>
```

Revision Description

A list of <change> elements, each with a *@date* and *@who* attributes, indicating significant stages in the evolution of a document. Most recent first.

Example

```
<revisionDesc>
 <change when="2006-08-09" who="#LB">handedits following
newhrdgen.xsl</change>
 <change when="2000-10-11" who="#OUCS">Final manual
corrections for BNC-W</change>
 <change when="2000-10-18" who="#OUCS">Further manual
corrections for BNC-W</change>
 <change when="2000-01-08" who="#OUCS">Manually changed
catdescriptions etc. for
    BNC-W</change>
 <change when="1994-11-30" who="#OUCS">First release for
BNC-1</change>
</revisionDesc>
```

Bibliographies

The TEI provides numerous ways to provide bibliographic citations, from the highly flexible to the highly structured.

Bibliographic Citations

- <bibl> (loosely structured bibliographic citation)
- <biblStruct> (structured bibliographic citation)
- <listBibl> (a list of bibliographic citations such as a bibliography)
- The 'header' module also includes <biblFull> (fully-structured bibliographic citation based on the TEI fileDesc element)

Simple <bibl> Example

```
<p>In Punch there is an brief note which could be
misconstrued as a slur upon Canadians.<note>
   <bibl>
    <title level="a">Men for the Antarctic</title>
    <title level="j">Punch: or the London
Charivari</title>, <biblScope>p.
         6</biblScope>
   </bibl>
 </note> It should not be understood as such.</p>
```

Simple <biblStruct> Example

`Enigma', *Punch: or the London Charivari*, July 1, 1914, 147, pp. 1-20

```
<biblStruct>
 <analytic>
  <title level="a">Enigma</title>
 </analytic>
 <monogr>
  <title level="j">Punch: or the London
Charivari</title>
  <imprint>
   <pubPlace>London</pubPlace>
   <date when="1914-07-01">July 1, 1914</date>
   <biblScope type="vol">147</biblScope>
   <biblScope type="pp">1-20</biblScope>
  </imprint>
 </monogr>
</biblStruct>
```

<biblFull> example

```
<biblFull>
  <titleStmt>
    <title>The Feminist Companion to Literature in English: women
writers from the
       middle ages to the present</title>
    <author>Blain, Virginia</author>
    <author>Clements, Patricia</author>
    <author>Grundy, Isobel</author>
  </titleStmt>
  <editionStmt>
    <edition>UK edition</edition>
  </editionStmt>
  <extent>1231 pp</extent>
  <publicationStmt>
    <publisher>Yale University Press</publisher>
    <pubPlace>New Haven and London</pubPlace>
    <date>1990</date>
  </publicationStmt>
  <sourceDesc>
    <p>No source: this is an original work</p>
  </sourceDesc>
</biblFull>
```

Conclusion

Without proper metadata:

- no one can find your text
- no one knows how or why you made it
- no one knows what it is or what they can do with it

Working with facsimiles

James Cummings and other TEI@Oxford authors

February 2009

Working with facsimiles

Increasingly people want to do not just 'text' editions but text editions with facing page (or otherwise linked) facsimile images. Indeed, some people want to just have images and create and electronic facsimile (perhaps with a view to later eventual transcription). The <facsimile> element is provided to accommodate this desire.

Digital Facsimiles

- <facsimile> contains a representation of some written source in the form of a set of images rather than as transcribed or encoded text
- <surface> defines a written surface in terms of a rectangular coordinate space
  - *@start* points to an element which encodes the starting position of the text
- <zone> defines a rectangular area contained within a <surface> element
- Global *@facs* (facsimile) points directly to an image, or to a part of a facsimile element which corresponds with this element.

Simplest case: 1:1 mapping with *@facs*

If a digital text contains one image per page or column (or similar unit), and no more complex mapping between text and image is envisaged, then the *@facs* attribute may be used to point directly to a graphic resource.

```
<text>
 <body>
  <pb facs="147_13.jpg" n="13"/>
  <div>
   <head>ESSENCE OF PARLIAMENT.</head>
   <p>
<!-- Page 13 continues -->
   </p>
   <pb facs="147_14.jpg" n="14"/>
   <p>Business done.  -- The Commons still...
<!-- Page 14 continues -->
   </p>
  </div>
 </body>
</text>
```

Using *@facs* in conjunction with <facsimile>, <surface>, and <zone>

Using these attributes and elements together enables an editor to

- associate multiple images with each page
- record arbitrary planar coordinates of textual elements on any kind of surface and link such elements to digital facsimile images of them

<facsimile>

The facsimile element is used to represent a digital facsimile. It appears within a TEI document along with, or instead of, the text element introduced in section 5 Default Text Structure. When this module is selected therefore, a legal TEI document may thus comprise any of the following:

- a TEI Header and a text element
- a TEI Header and a facsimile element
- a TEI Header, a facsimile element, and a text element

<facsimile> Example in Context

```
<TEI>
 <teiHeader>
  <fileDesc>
   <titleStmt>
    <title/>
   </titleStmt>
   <publicationStmt>
    <p/>
   </publicationStmt>
   <sourceDesc>
    <p/>
   </sourceDesc>
  </fileDesc>
 </teiHeader>
 <facsimile>
  <graphic url="147_13.jpg"/>
  <graphic url="147_14.jpg"/>
  <graphic url="147_15.jpg"/>
 </facsimile>
 <text>
  <body>
   <p/>
  </body>
 </text>
</TEI>
```

<surface>

The <surface> element may be used to indicate that there are two image files corresponding with the same area of the work:

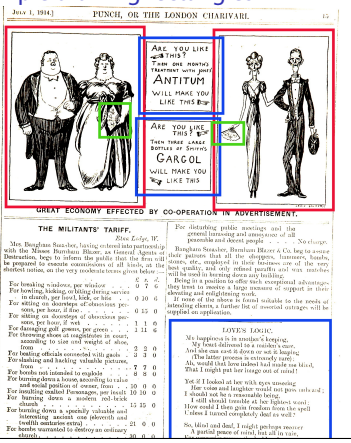
```
<facsimile>
 <surface>
  <graphic url="147_13.jpg"/>
  <graphic url="147_13-huge.jpg"/>
 </surface>
 <graphic url="147_14.jpg"/>
 <graphic url="147_15.jpg"/>
 <graphic url="147_16.jpg"/>
</facsimile>
```

dimensions

The actual dimensions of the object represented are not documented by the surface element; instead, the surface is located within an abstract coordinate space, which is defined by the following attributes, supplied by the att.coordinated class:

- *@ulx* gives the x coordinate value for the upper left corner of a rectangular space
- *@uly* gives the y coordinate value for the upper left corner of a rectangular space.
- *@lrx* gives the x coordinate value for the lower right corner of a rectangular space.
- *@lry* gives the y coordinate value for the lower right corner of a rectangular space.

Example drawing rectangles



<surface> Example

```
<facsimile>
 <surface ulx="0" uly="0" lrx="993" lry="1639">
<!-- ... -->
 </surface>
</facsimile>
```

<zone> in <surface>

To describe the whole image, we will also need to define a zone of interest which represents an area inside this surface. This zone of interest can be defined by a <zone> element, within which we can place the <graphic>:

```
<facsimile>
 <surface ulx="0" uly="0" lrx="993" lry="1639">
  <zone ulx="93" uly="681" lrx="967" lry="1568">
   <graphic url="147_13.jpg"/>
  </zone>
 </surface>
</facsimile>
```

<desc>

The <desc> element may also be used within either <surface> or <zone> to provide some further information about the area being defined.

<desc> Example

```
<facsimile>
 <surface ulx="0" uly="0" lrx="993" lry="1639">
  <desc>Printed page</desc>
  <zone ulx="96" uly="89" lrx="750" lry="657">
   <desc>Cartoon</desc>
   <graphic url="147_13.jpg"/>
  </zone>
  <zone ulx="95" uly="681" lrx="990" lry="1568">
   <desc>Text section</desc>
   <graphic url="147_13.jpg"/>
  </zone>
 </surface>
</facsimile>
```

More uses for <zone>

In addition to acting as a container for <graphic> elements, <zone> elements may also be used to select parts of each surface for analytical purposes.

```
<facsimile>
 <surface ulx="0" uly="0" lrx="993" lry="1639">
  <desc>main text</desc>
  <zone ulx="393" uly="681" lrx="967" lry="890">
   <desc>Cartoon</desc>
   <graphic url="147_13"/>
  </zone>
  <zone ulx="507" uly="596" lrx="704" lry="768">
   <desc>Artist's signature</desc>
  </zone>
 </surface>
</facsimile>
```

Linking transcription and facsimile

1. give each relevant part of the facsimile an identifier
2. using the *@facs* attribute, point from the transcription into the <facsimile>

Linking transcription and facsimile: <facsimile>

```
<facsimile>
 <surface ulx="0" uly="0" lrx="993" lry="1639">
  <zone xml:id="cartoon-
facs" ulx="96" uly="89" lrx="950" lry="657">
   <desc>Cartoon</desc>
   <graphic url="147_10.jpg"/>
  </zone>
  <zone xml:id="text-
facs" ulx="93" uly="681" lrx="967" lry="1568">
   <desc>Text section</desc>
   <graphic url="147_10.jpg"/>
  </zone>
 </surface>
</facsimile>
```

Linking transcription and facsimile: text

```
<body>
 <pb n="p10"/>
 <div facs="#cartoon-facs">
  <figure>
<!-- cartoon info here -->
  </figure>
 </div>
 <div facs="#text-facs">
  <head>THE WALKERS</head>
  <lg>
   <l>There were eight pretty walkers...</l>
  </lg>
 </div>
</body>
```

Pointing from <facsimile> to transcription with *@start* attribute

It is also possible to point in the other direction, from a <surface> or <zone> to the corresponding text. This is the function of the *@start* attribute, which supplies the identifier of the element containing the transcribed text found within the <surface> or <zone> concerned.

@start attribute example: <facsimile>

```
<facsimile>
 <surface start="#p10" ulx="0" uly="0" lrx="993" lry="1639">
  <zone xml:id="cartoon-facs-
2" ulx="96" uly="89" lrx="950" lry="657">
   <desc>Cartoon</desc>
   <graphic url="147_10.jpg"/>
  </zone>
 </surface>
</facsimile>
```

@start attribute example: text

```
<body>
 <pb xml:id="p10" n="10"/>
 <div>
  <figure>
<!-- cartoon info here -->
  </figure>
 </div>
 <div>
  <head>THE WALKERS</head>
  <lg>
   <l>There were eight pretty walkers...</l>
  </lg>
 </div>
</body>
```

Gravestone Example: <facsimile>

```
<facsimile>
 <surface xml:id="grave" ulx="0" uly="0" lrx="355" lry="678">
  <graphic url="gravestone-cropped.jpg"/>
  <zone ulx="83" uly="223" lrx="272" lry="256" xml:id="line1"/>
  <zone ulx="92" uly="251" lrx="256" lry="282" xml:id="line2"/>
  <zone ulx="21" uly="281" lrx="330" lry="308" xml:id="line3"/>
  <zone ulx="36" uly="306" lrx="320" lry="332" xml:id="line4"/>
  <zone ulx="85" uly="535" lrx="249" lry="556" xml:id="line5"/>
  <zone ulx="97" uly="556" lrx="241" lry="576" xml:id="line6"/>
  <zone ulx="58" uly="577" lrx="281" lry="595" xml:id="line7"/>
  <zone ulx="68" uly="595" lrx="271" lry="613" xml:id="line8"/>
 </surface>
</facsimile>
```

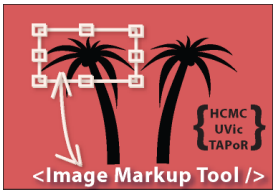
Gravestone Example: text

```
<div facs="#grave">
 <p>Private Moulds' gravestone</p>
 <div>
  <ab>
   <s facs="#line1">12851 PRIVATE</s>
   <lb/>
   <s facs="#line2">H. MOULDS</s>
   <lb/>
   <s facs="#line3">NORTHAMPTONSHIRE REGT.</s>
   <lb/>
   <s facs="#line4">23RD JULY 1916 AGED 21</s>
  </ab>
  <ab>
   <s facs="#line5">LOVING SON OF </s>
   <lb/>
   <s facs="#line6">MRS MOULDS</s>
   <lb/>
   <s facs="#line7">PETERBORO, ENGLAND</s>
   <lb/>
   <s facs="#line8">FOR EVER WITH US</s>
   <lb/>
  </ab>
 </div>
</div>
```

Gravestone Example: rendered

facsimile-demo.html

Introducing the Image Markup Tool



- Written by Martin Holmes at University of Victoria
- Allows image annotation
- Saves its files as pure TEI P5
- http://www.tapor.uvic.ca/~mholmes/image_markup/
- Main window tutorial
- Annotation window tutorial

Loading and Using Images

Loading Images:
- **File / New:** starts new image
- **File / Import:** option of keeping existing annotations and applying to new image

Using Images:
- **Zooming:** scale toolbar box, or Control + mousewheel
- **Scrolling:** mousewheel (up/down); Shift + mousewheel (left-right), or scrollbars

Defining and Editing Categories

- IMT makes you sort your annotations into categories (but which and how many are up to you!)
- You use the 'category manager' to create categories where you can define:
  - an xml:id for identification
  - a title/description of the category
  - a colour for the annotation lines on the image
  - a shape: rectangle (default and only one used in web view), ellipse, cross or spiral

Adding and Deleting Annotation Areas

- Click on 'Annotations / New annotation' (or button), and a resizable shape will appear on the image
- You can resize and move the area with the mouse
- The area, when selected, will have a white bounding box -- otherwise it will be the colour of its category
- The shape depends on the defined shape in the category you've chosen
- To delete, select it and then click on 'Annotations / Delete annotation' or button -- the associated annotations will also be deleted

Adding Annotation Text

- When an annotation is added, it has the last category you used
- You can change this to any other category
- You can provide an annotation title (stored as tei:head)
- You can provide some text (any tei:div content, tei:p default)
- You can change the generated xml:id with the ID button

Hiding and Showing Annotations

- You can hide and show annotations in the annotations window
- You can select either individual annotations to display
- Or you can choose not to display a whole category of annotations
- This is useful when working with lots of overlapping annotations

Creating a "Web View"

- As an example IMT produces a 'web view'
- This includes a web page, with slightly shrunk image, where your annotations appear as mouse-over outlines with boxes displaying the annotations
- Double-clicking on an annotation area produces a zoomed in version of the image for that area
- This is really just a proofing tool, but has been used by some as the basis for annotated facsimiles

IMT Schemas

It is important quickly to note that IMT XML files are pure TEI P5 XML. The RelaxNG schema is available with the installed program, as is the TEI ODD file that Roma uses to generate this schema.

Live Demo

If we have time at this point I'll do a short live demo of the IMT to introduce you to the kinds of things you'll be asked to do in your exercise.

Names, People, and Places

Arianna Ciula and TEI@Oxford authors

February 2009

What's in a name?

- We've already met <name> and <rs> for any form of name or referring string.
- The namesdates module also provides specialisations of these: <persName>, <placeName>, and <orgName>
- Each can be further decomposed
- They can also be associated with a named entity

Personal Names

For example...

- <persName> (personal name) a proper noun or proper-noun phrase referring to a person ... equivalent to <name type="person">
- <surname> a family (inherited) name
- <forename> a forename, given or baptismal name
- <roleName> a name component indicating a particular role or position in society
- <addName> (additional name) an additional name component such as a nickname, epithet, or alias, or any other descriptive phrase used within a personal name
- <nameLink> a connecting phrase or link used within a name but not regarded as part of it

```
<persName xml:lang="xh-tw">
 <forename>□□</forename>
 <forename>□□</forename>
 <surname>□□□□</surname>□
<roleName>
   <placeName>□□</placeName>□□□□□
 </roleName>
</persName>
```

Names as referents (1)

In a text we might find the same person referred to on different occasions in any number of different ways:

```
 ...<persName>Clara
Schumann</persName>.... <persName>Clara</persName> ....
<persName>Frau
Schumann</persName>
```

All of these names refer to the same entity

We can use an attribute on any naming element to specify which entity is being referenced:

- *@key* if we are supplying an externally-defined code for the entity
- *@ref* if we are pointing to a definition of the entity

Names as referents (2)

For example:-

```
 ...<persName ref="#CS">Clara Schumann</persName>....
<persName ref="#CS">Clara</persName> ....
<persName key="CS123">Frau Schumann</persName>
<!-- ... elsewhere -->
<person xml:id="CS">
 <persName xml:lang="de">
  <forename type="first">Clara</forename>
  <forename type="middle">Josephine</forename>
  <surname type="maiden">Wieck</surname>
  <surname type="married">Schumann</surname>
 </persName>
</person>
```

The thing itself (1)

TEI provides special-purpose elements for maintaining structured information about named entities (as well as their names):

- <person>, <place>, <event>
- may be grouped into <listPerson>, <listPlace>, <listEvent>
- relationships can also be modelled, explicitly using <relation> or implicitly by context

```
<person xml:id="VM1893">
 <persName xml:lang="ru">Владимир Владимирович Маяковский</persName>
 <persName xml:lang="fr">Wladimir Maïakowski</persName>
 <birth when="1893-07-19">7 July (OS) 1893,
<placeName ref="#BGDT" xml:lang="en">Baghdati, Georgia</placeName>
 </birth>
 <death when="1930-04-14"/>
 <occupation>Poet and playwright, among the foremost representatives of
early-20th
    century Russian Futurism.</occupation>
<!-- ... -->
</person>
```

Traits, states, and events

The scope of elements one might record for a named entity is *large*. The TEI provides three generic elements, and some specific ones. We identify three main classes of information:

- characteristics or traits which do not, by and large, change over time
- characteristics or states which hold true only at a specific time
- events or incidents which may lead to a change of state or, less frequently, trait

For a person, typical traits are such things as `<faith>`, `<sex>`, `<socEcStatus>`; typical states are such things as `<occupation>`, `<residence>`, `<education>`; typical events are such things as `<birth>` and `<death>`.

Personal Relationships

- `<relationGrp>` (relation group) provides information about relationships identified amongst people, places, and organizations
- `<relation>` (relationship) describes any kind of relationship or linkage amongst a specified group of participants
  - *@name* supplies a name for the kind of relationship of which this is an instance
  - *@active* identifies the 'active' participants in a non-mutual relationship, or all the participants in a mutual one
  - *@mutual* supplies a list of participants amongst all of whom the relationship holds equally
  - *@passive* identifies the 'passive' participants in a non-mutual relationship

Example

```
<person xml:id="jsbach">
 <persName>Johann Sebastian Bach</persName>
</person>
<person xml:id="cdbach">
 <persName>Catharina Dorothea Bach</persName>
</person>
<person xml:id="ghbach">
 <persName>Gottfried Heinrich Bach</persName>
</person>
<!--....-->
<relationGrp type="children" subtype="first-marriage">
 <relation name="parent" active="#jsbach" pas-
sive="#cdbach"/>
<!--....-->
</relationGrp>
<relationGrp type="children" subtype="second-marriage">
 <relation name="parent" active="#jsbach" pas-
sive="#ghbach"/>
<!--....-->
</relationGrp>
```

Other kinds of entity

`<org>`: a named collection of people regarded as a single unit, such as a business, institution, or tribe.
`<place>`: a named location of any kind (including mythological and non-terrestrial places)
These can be grouped in the same way (using `<listOrg>` or `<listPlace>`), and also have states, traits, and events.

Places

- Places can be identified solely in terms of geographical features or locations, e.g.

```
<place>
 <placeName>
  <geogFeat>mount</geogFeat>
  <geogName>Sinai</geogName>
 </placeName>
</place>
```

- More usually, they are identified in geo-political terms, using
  - adminstrative units such as `<bloc>`, `<country>`, `<region>`, `<settlement>`, `<district>`
  - physical location using `<geo>` and `<offset>`
- Note that all these things are traits -- they may change over time

For example: Taipei region

```
<place xml:id="BGDT">
 <placeName xml:lang="zh-tw">□□□□</placeName>
 <placeName xml:lang="en">Shilin</placeName>
 <location type="geopolitical">
  <country>Taiwan</country>
  <region>Taipei</region>
 </location>
 <location type="physical">
  <geo>25.0866 121.5254</geo>
 </location>
</place>
```

Places can be nested (unlike people)

```
<place xml:id="LT">
 <country>Lithuania</country>
 <country xml:lang="lt">Lietuva</country>
 <place xml:id="LT-VN">
  <settlement>Vilnius</settlement>
 </place>
 <place xml:id="LT-KA">
  <settlement>Kaunas</settlement>
 </place>
</place>
```

Sources

Responsibility and uncertainty about the sources can be asserted by using attributes from the att.editLike class:

```
<org xml:id="MXY" type="tribe" resp="#herodotus">
 <orgName>The Maxyans</orgName>
 <country>Libya</country>
 <desc>According to Herodotus, they were a west Libyan
tribe who said
    that they were descended from the men of Troy.</desc>
</org>
```

XML technologies for accessing TEI texts

James Cummings, Sebastan Rahtz and other TEI@Oxford authors

February 2009

What more do you need to know?

You've created some TEI XML documents. What processing technologies are now available to you?

- XPath
- XSLT
- XML Query (XQuery)

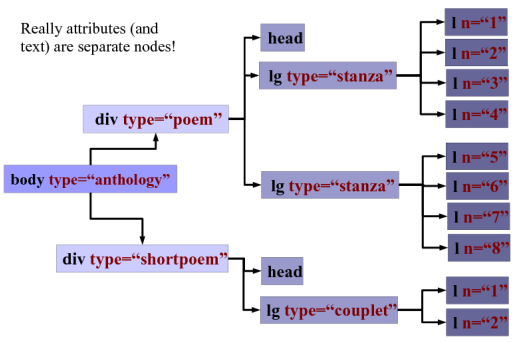
What is XPath?

- A standardized syntax for identifying and accessing parts of an XML document
- A library of standard functions
- A W3C Standard
- A major component of XQuery, XSLT, and almost every XML processing system

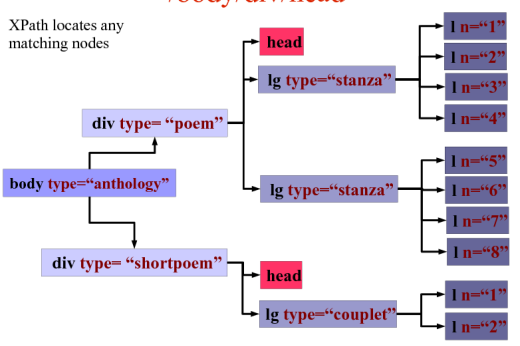
Example text

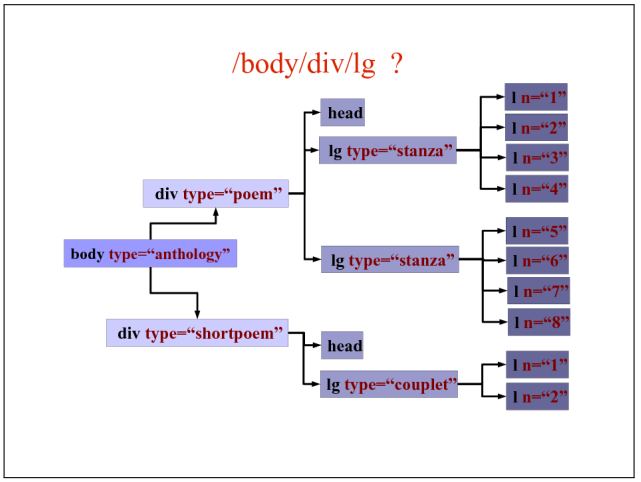
```
<body n="anthology">
 <div type="poem">
  <head>The SICK ROSE </head>
  <lg type="stanza">
   <l n="1">O Rose thou art sick.</l>
   <l n="2">The invisible worm,</l>
   <l n="3">That flies in the night </l>
   <l n="4">In the howling storm:</l>
  </lg>
  <lg type="stanza">
   <l n="5">Has found out thy bed </l>
   <l n="6">Of crimson joy:</l>
   <l n="7">And his dark secret love </l>
   <l n="8">Does thy life destroy.</l>
  </lg>
 </div>
</body>
```

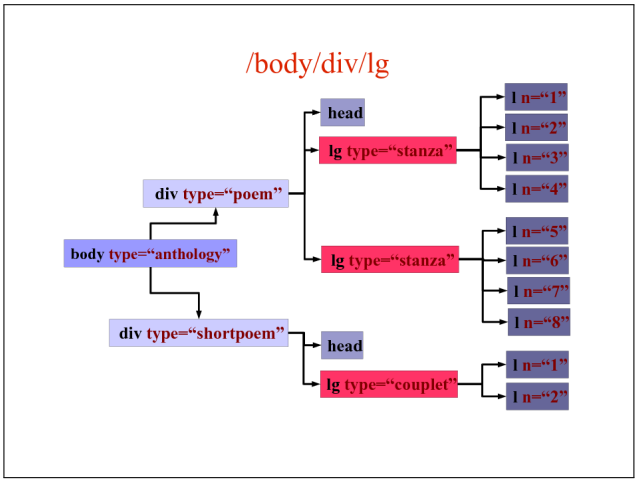
XML Structure

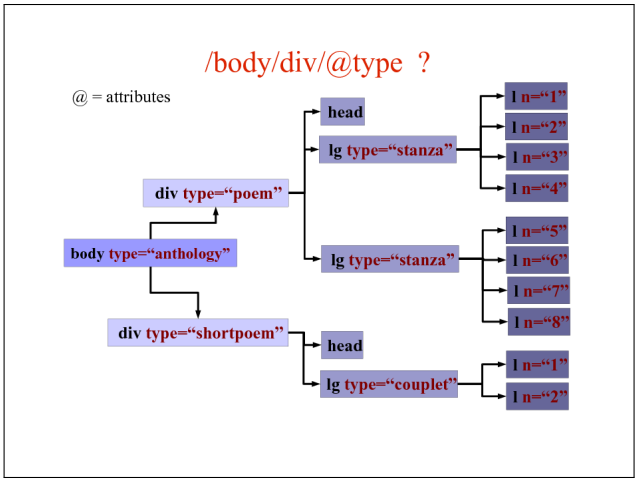


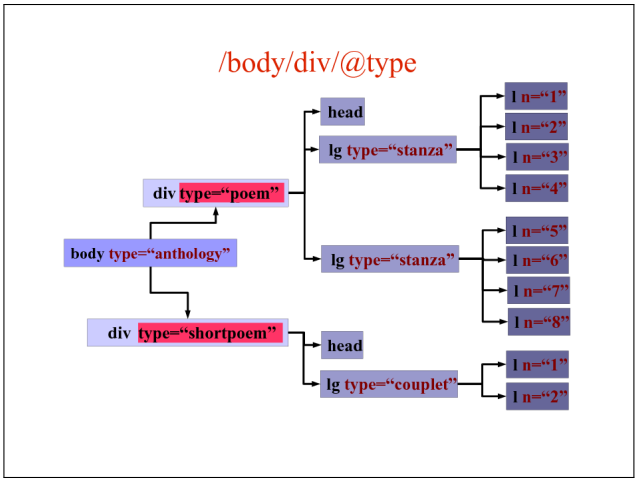
/body/div/head

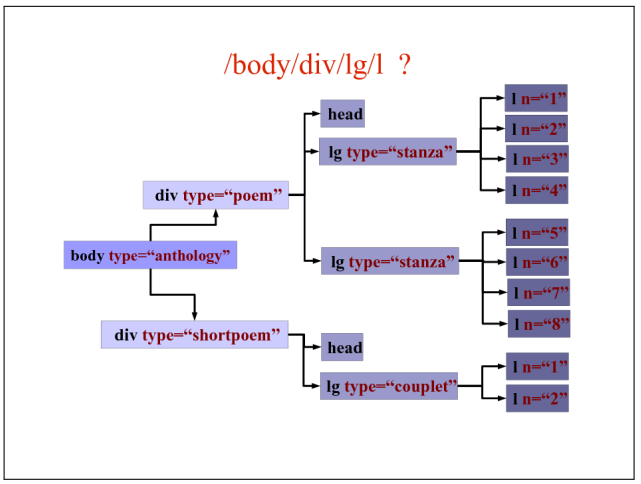


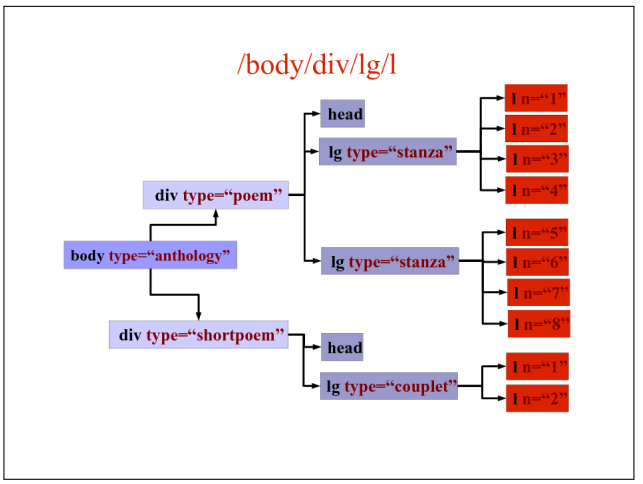

/body/div/lg ?
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

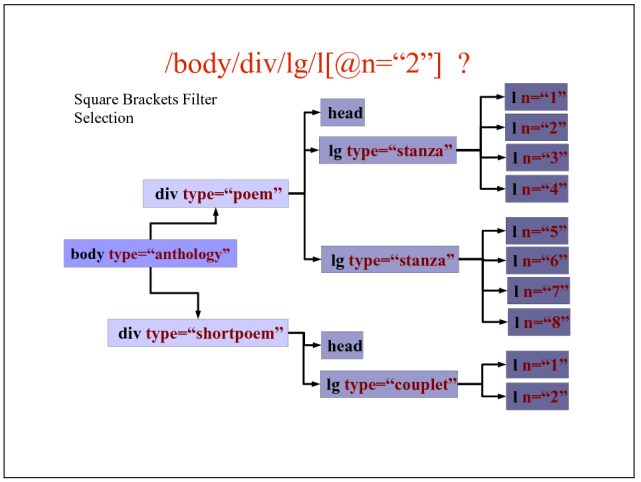

/body/div/lg
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

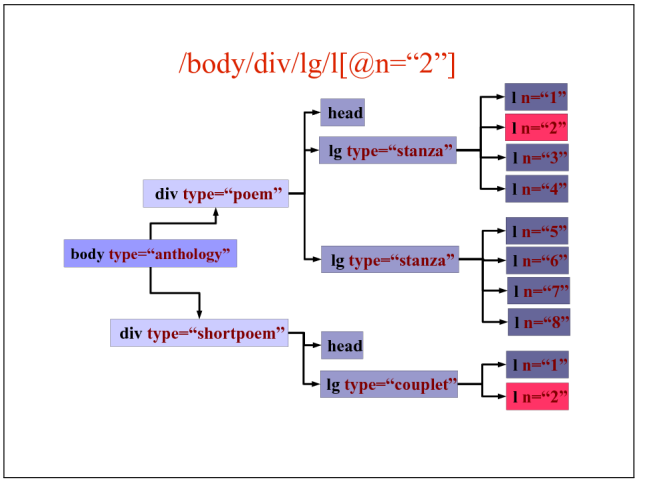

/body/div/@type ?
@ = attributes
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

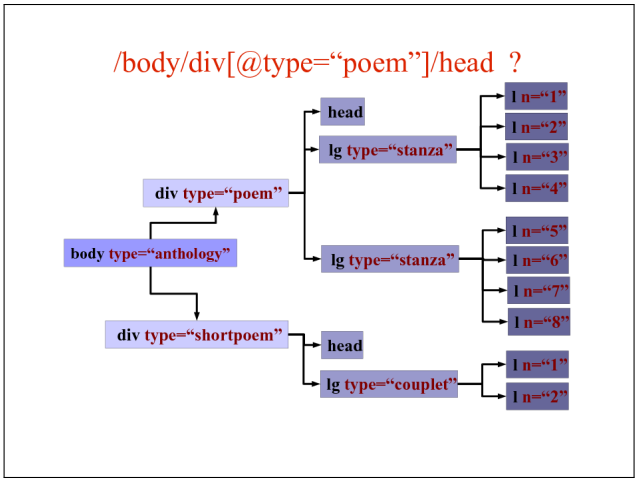

/body/div/@type
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

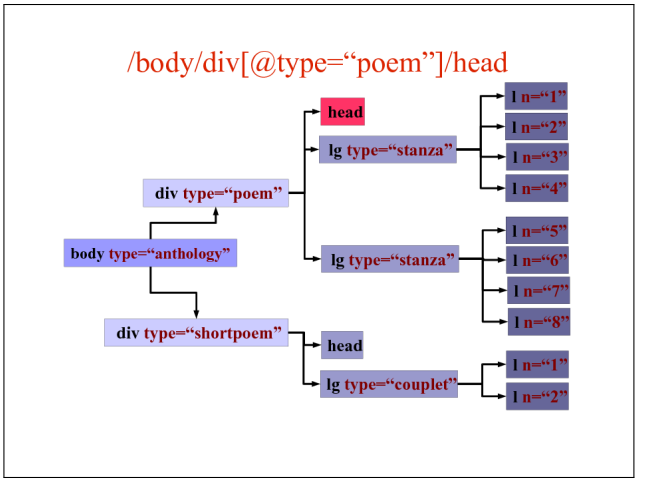

/body/div/lg/l ?
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

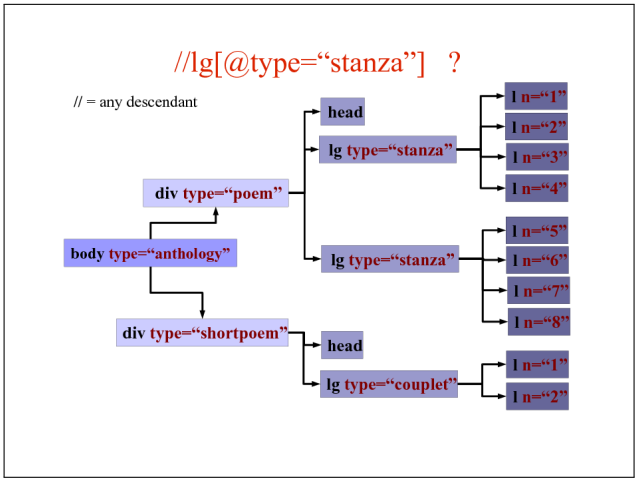

/body/div/lg/l
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

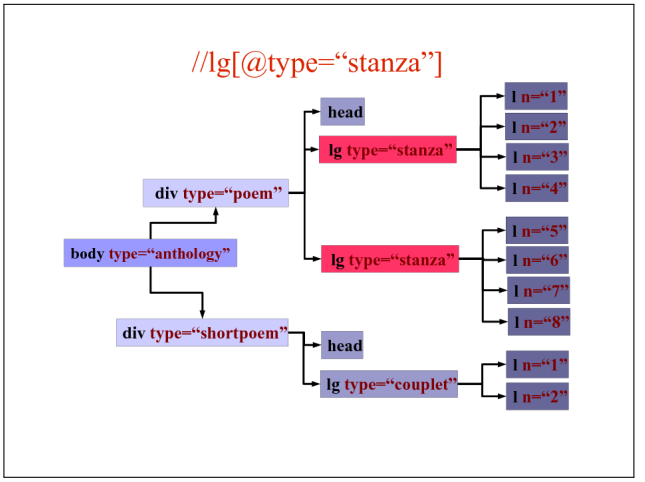

/body/div/lg/l[@n="2"] ?
Square Brackets Filter
Selection
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

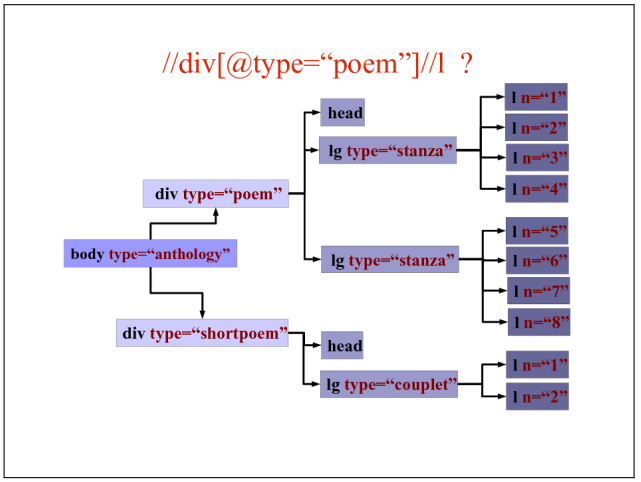

/body/div/lg/l[@n="2"]
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

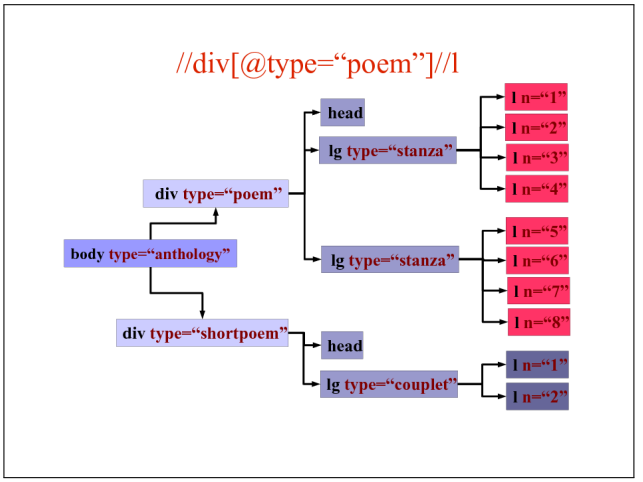

/body/div[@type="poem"]/head ?
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

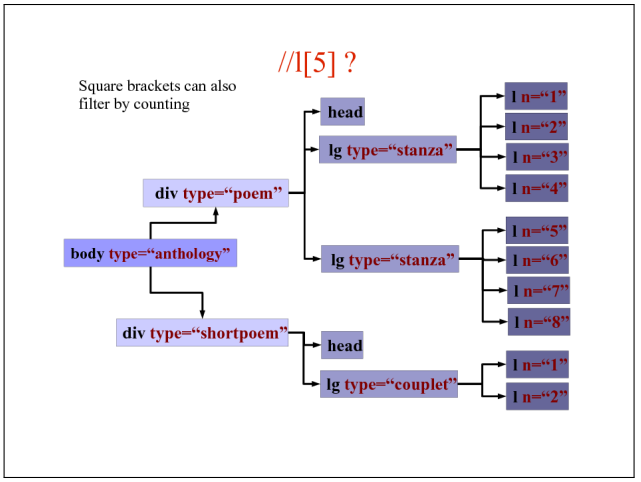

/body/div[@type="poem"]/head
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

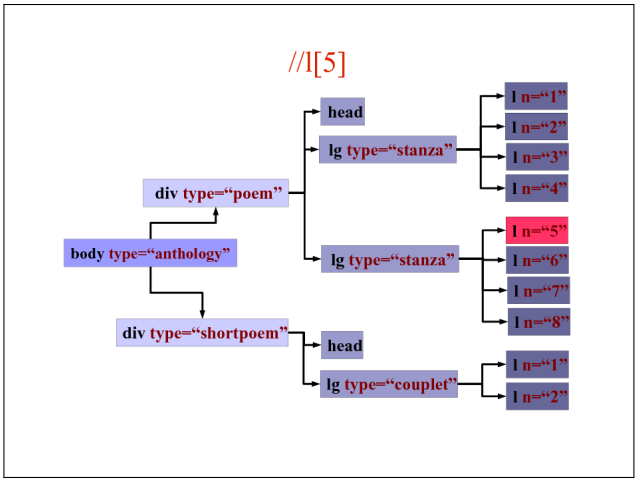

//lg[@type="stanza"] ?
// = any descendant
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

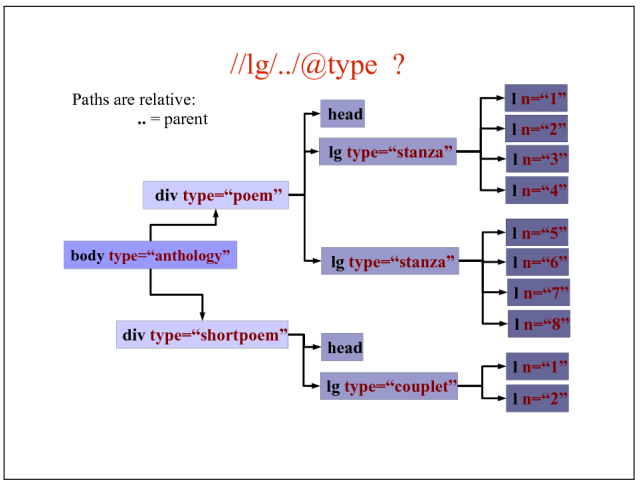

//lg[@type="stanza"]
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

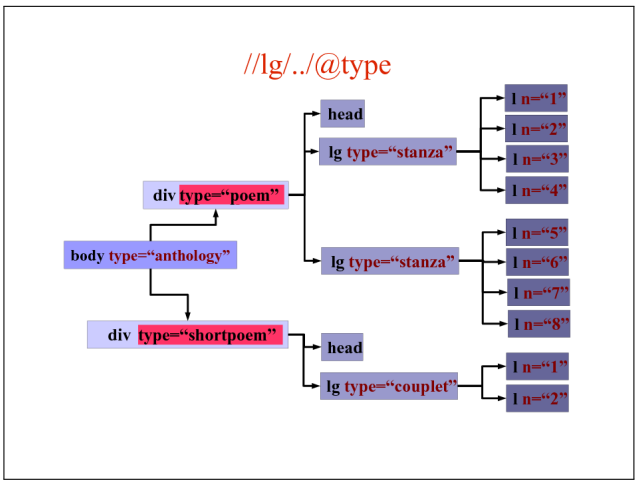

//div[@type="poem"]//l ?
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

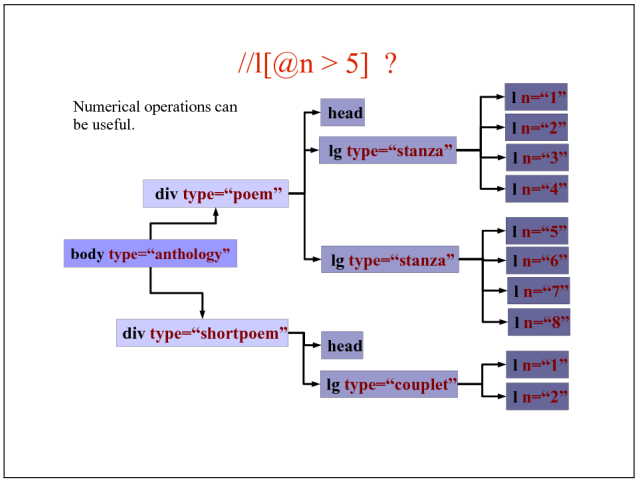

//div[@type="poem"]//l
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

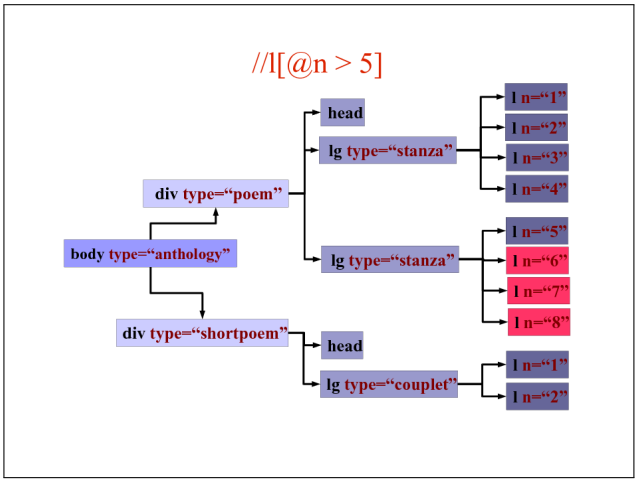

//l[5] ?
Square brackets can also filter by counting
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

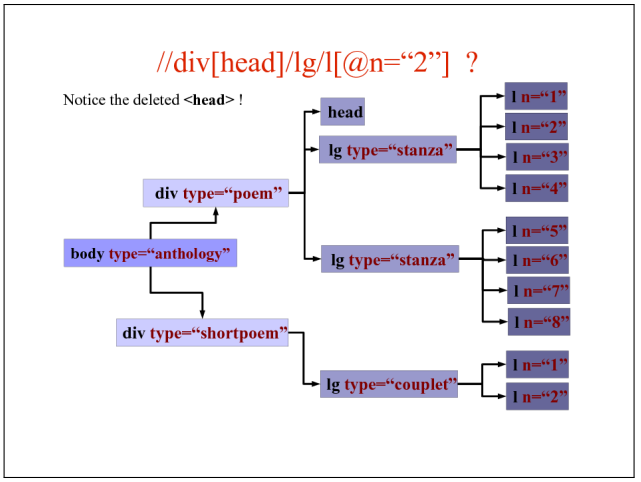

//l[5]
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

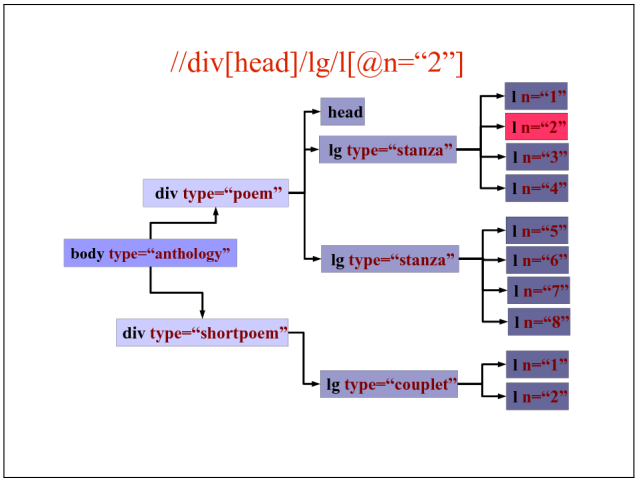

//lg/../@type ?
Paths are relative:
.. = parent
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

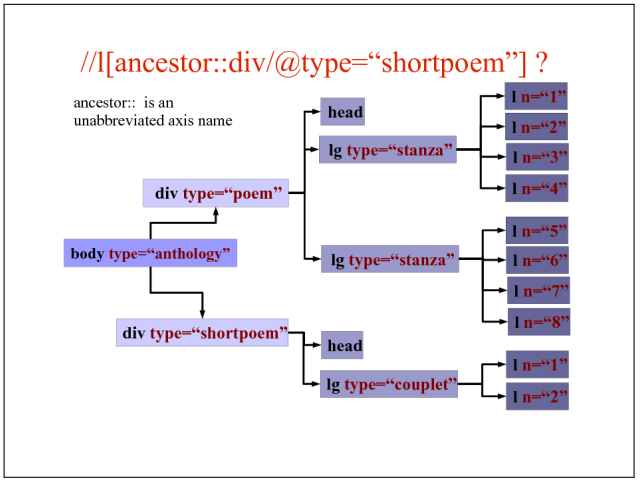

//lg/../@type
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

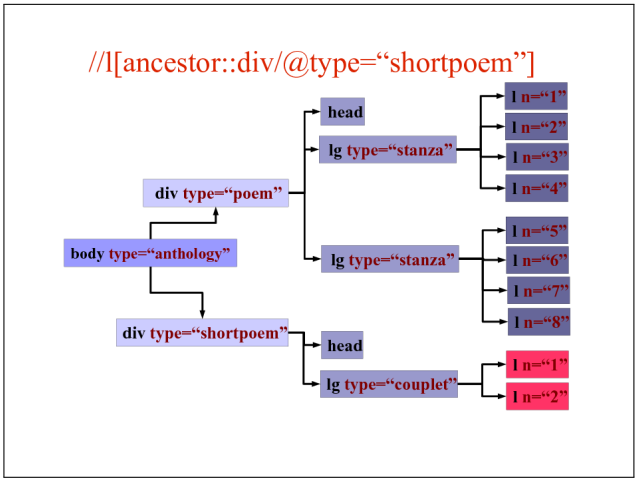

//l[@n > 5] ?
Numerical operations can be useful.
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"


//l[@n > 5]
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"


//div[head]/lg/l[@n="2"] ?
Notice the deleted <head> !
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"


//div[head]/lg/l[@n="2"]
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"


//l[ancestor::div/@type="shortpoem"] ?
ancestor:: is an unabbreviated axis name
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"


//l[ancestor::div/@type="shortpoem"]
body type="anthology"
div type="poem"
div type="shortpoem"
head
lg type="stanza"
lg type="stanza"
head
lg type="couplet"
l n="1"
l n="2"
l n="3"
l n="4"
l n="5"
l n="6"
l n="7"
l n="8"
l n="1"
l n="2"

XPath in summary

- A location path results in a node-set
- Paths can be absolute (`/div/lg[1]/l`)
- Paths can be relative (`l/../../head`)
- Formal Syntax: (`axisname::nodetest[predicate]`)
- For example:`child::div[contains(head, 'ROSE')]`

XPath: Abbreviated Syntax

- nothing is the same as `child::`, so `lg` is short for `child::lg`
- @ is the same as `attribute::`, so `@type` is short for `attribute::type`
- . is the same as `self::`, so `./head` is short for `self::node()/child::head`
- .. is the same as `parent::`, so `../lg` is short for `parent::node()/child::lg`
- // is the same as `descendant-or-self::`, so `div//l` is short for `child::div/descendant-or-self::node()/child::l`

XPath also has built-in functions

These include...

node-set functions for example, to find the number of nodes in a node-set, select nodes by position, name, namespace, etc.

string functions for example, to concatenate string values, do string matching, find substrings, perform one-to-one character translation, etc.

eXtensible StyLesheeTs

The XSLT language is

- expressed in XML; uses namespaces to distinguish output from instructions
- a Turing-complete functional programing language
- reads and writes XML trees
- designed to generate XSL FO, but now widely used to generate HTML or other forms of XML

What do you mean, "transformation"?

Take this

```
<recipe>
<title>Pasta for beginners</title>
<ingredients><item>Pasta</item>
<item>Grated cheese</item>
</ingredients>
<cook>Cook the pasta and mix with the cheese</cook>
</recipe>
```

and make this

```
 <html>
<h1>Pasta for beginners</h1>
<p>Ingredients: Pasta Grated cheese</p>
<p>Cook the pasta and mix with the cheese</p>
</html>
```

How do you express that in XSL?

```
 <xsl:stylesheet
xmlns:xsl='http://www.w3.org/1999/XSL/Transform'
version="1.0">
<xsl:template match="recipe">
<html>
<h1><xsl:value-of select="title"/></h1>
<p>Ingredients:
<xsl:apply-templates
select = "ingredients/item"/>
</p>
<p><xsl:value-of select="cook"/></p>
</html>
</xsl:template>
</xsl:stylesheet>
```

XSL Summary

The core techniques:

- template rules for nodes in the incoming XSL
- taking material from other nodes
- processing nodes several times in different modes
- variables and functions
- choosing, sorting, numbering
- different types of output

What is XQuery?

- a domain-specific method for accessing and manipulating XML
- designed for querying XML
- built upon XPath
- analagous to SQL (but for XML rather than for relational data)
- a W3C recommendation

XQuery: the core techniques

path expressions return a nodeset
element constructors return a new element
FLWOR expressions analogous to SQL Select statement
list expressions operations on lists or sets of values
conditional expressions traditional if then else construction
qualified expressions boolean operations over lists or sets of values
datatype expressions test datatypes of values

XQUERY is a complete programming language

Using XQuery

- The language itself is usually hidden from the end user (e.g. behind an HTML form)
- Implementations such as eXist may be embedded in web publishing framework such as Apache...
- .. but eXist can also be accessed e.g. from a Java program and via HTTP/REST, XML-RPC, SOAP, WebDAV etc.
- eXist provides XUpdate: a means of adding, changing, or deleting nodes from an XML database

Accessing XML: Conclusions

- XML is a set of hierachical nodes that can be identified with XPath
- XSLT is an easy way to transform these into different forms
- XQuery can enable you to perform sophisticated queries on a large database of texts
- ...We haven't put things into TEI XML for the fun of it, but to be able to get things out of it or do things to it. These things may be the text in multiple formats, lists of names, statistical information, linguistic analysis, or additional layers of annotation

5 Exercises

5.1 Exercise: getting started with Oxygen

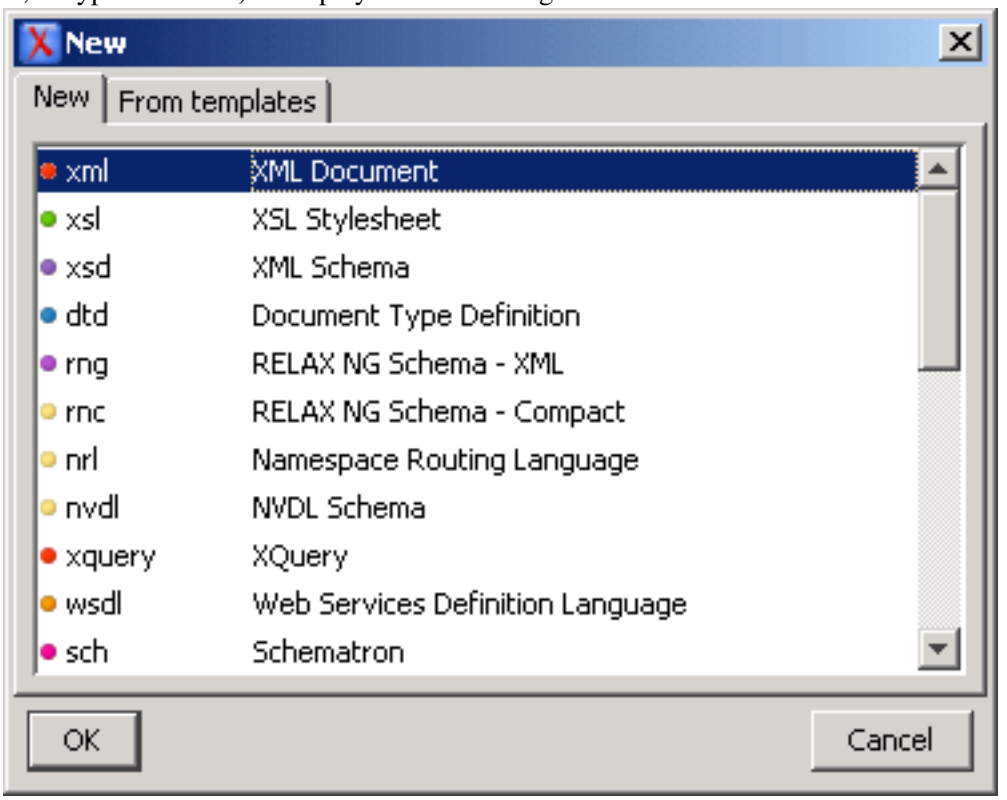
5.1.1 First steps with Oxygen

In this first exercise you will learn how to use Oxygen to

- create a new XML document
- add markup to a document
- keep your document well-formed
- display and edit your document without seeing tags

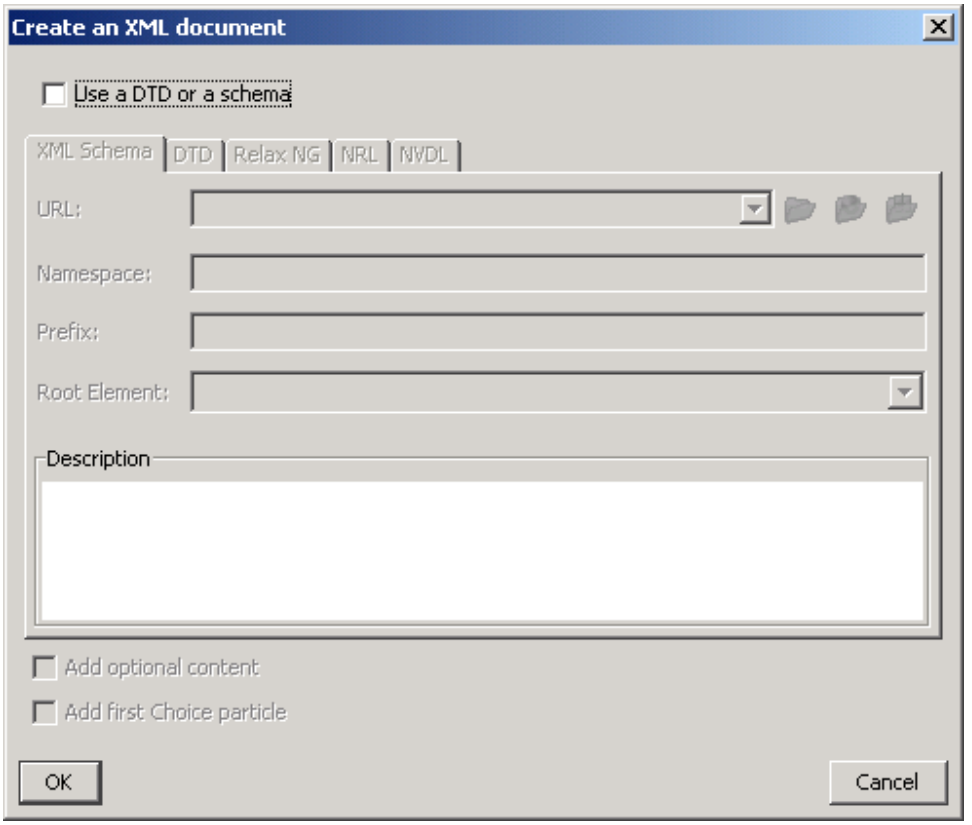
5.1.2 Creating a new document

Start up Oxygen, and click the New Document icon at top left (or select New Document from the File menu, or type CTRL-N.) to display the New dialog.



Select the default XML option, and press OK.

In the "Create an XML Document" dialog, *uncheck* the tick box Use a DTD or a schema, since we don't yet have a schema.



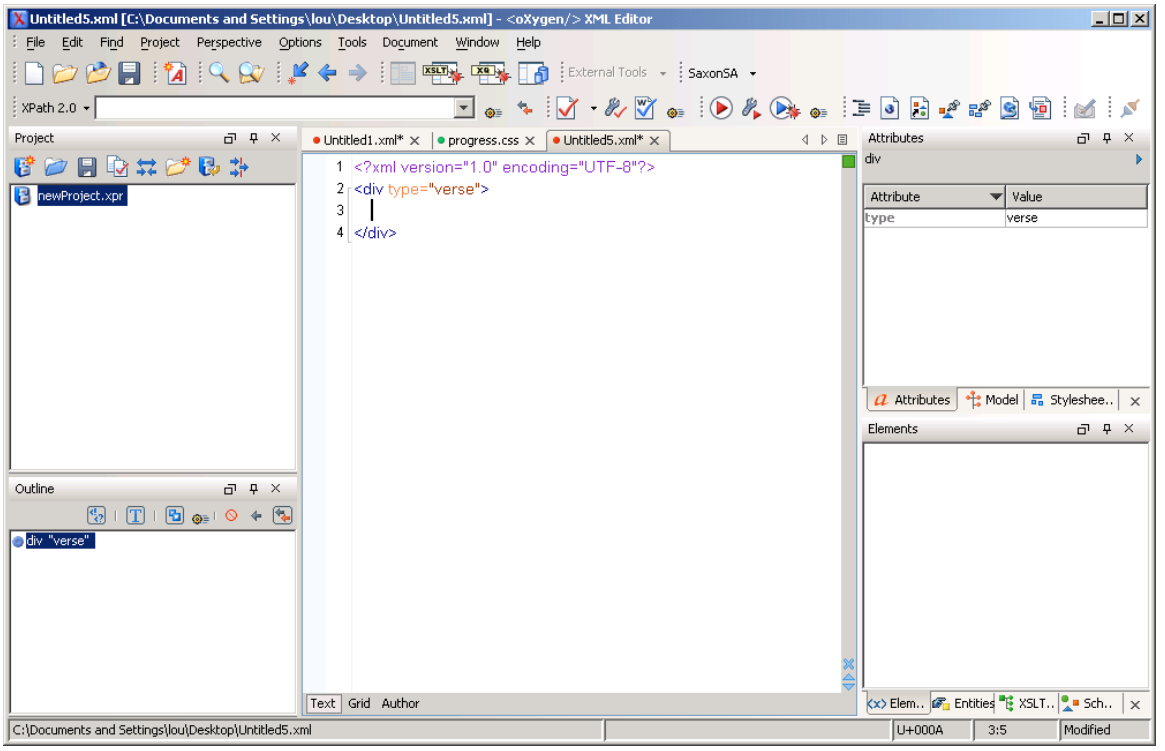
Press OK to continue.

Oxygen opens a blank document, containing just the XML declaration at the start. As you start typing in XML, notice how hard Oxygen tries to make your document well-formed.

- type in `<div type="` and pause
- notice that Oxygen supplies the closing quote for you
- continue typing `verse` and move the cursor *after* the closing quote (use an arrow key)
- Type the closing > and note that Oxygen immediately supplies the closing `</div>` for you. Press the RETURN key

5.1.3 Adding text to a document

Your document should now look like this:

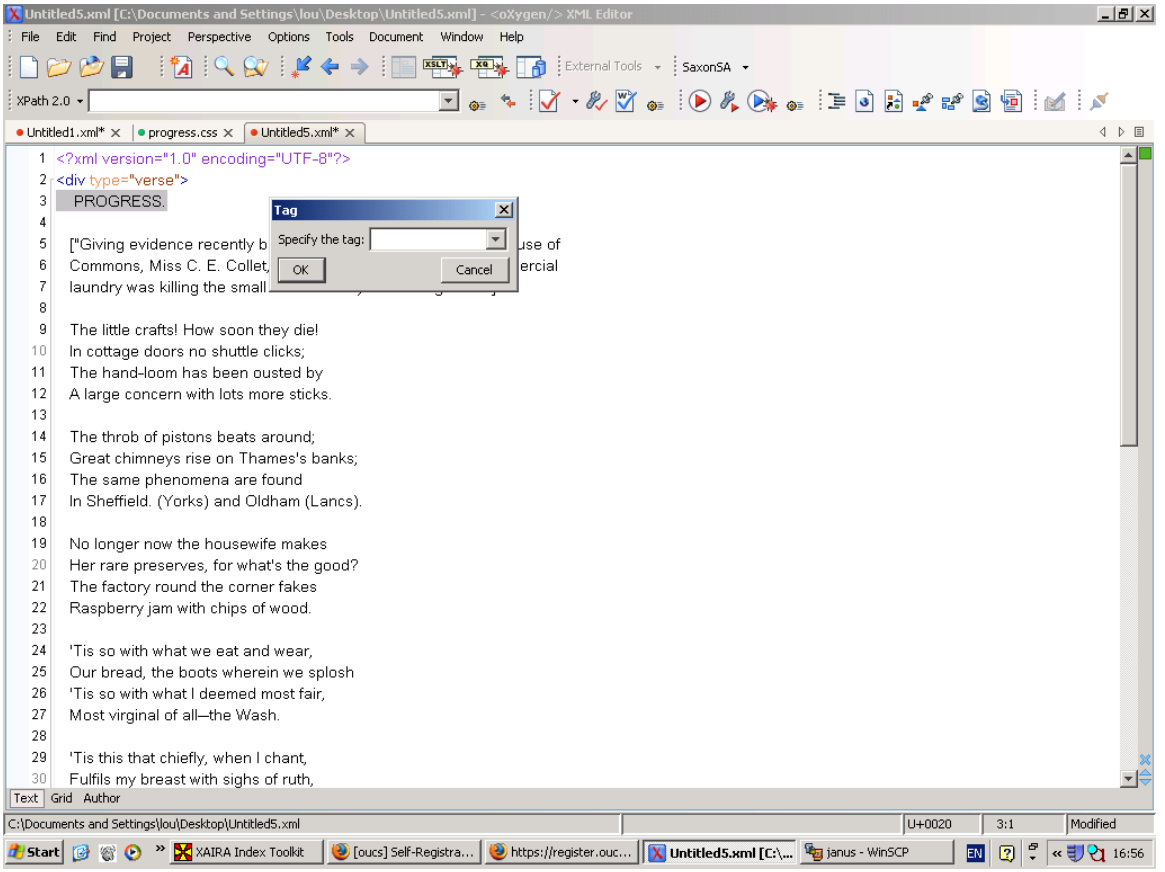


- Select File from the Document menu
- Select Insert File from the submenu that opens.
- Navigate to the file progress.txt in your working directory (or download it) and insert it into your document.

This is a plain text version of the poem at the start of our sample Punch issue. In the rest of this exercise we will add some minimal tagging.

5.1.4 Tagging bits of a document

- With the mouse, select the word 'PROGRESS' at the start of the text.
- Press CTRL-E (or select XML Refactoring/Surround with Tag from the Document menu).
- Type the name of the tag you want to use into the popup: this is a heading, so type `head` and press OK



Repeat the process for the whole of the paragraph in square brackets: this time, tag it as a <p>.

5.1.5 Tagging the poem

We will tag the poem proper using <lg> to enclose each stanza, and <l> to enclose each <l>.

- Try typing the sequence <lg><l> at the start of the first line. Delete the unwanted </l» and </lg> tags inserted by Oxygen.
- Put the cursor at the end of the first *stanza* (after the word 'sticks') and type <
- A small menu opens, showing that you can enter a closing-tag (to close the <l> element) here. Select it.
- Type another < and the same menu appears: but note that this time the closing-tag to be inserted is a </lg>. Select this, and the document becomes valid – no red lines visible.
- What will happen if you repeat the process, do you think? Try it and see!

5.1.6 Tagging the poem properly

Although well-formed our tagging is not very honest. We have a single <l> element containing several lines, and lots of stanzas which are not tagged as anything at all. If we were validating this document against a schema we'd be in trouble.

- put the cursor at the end of the first line (after 'die!')
- select XML Refactoring/Split Element from the Document menu.

This closes the current <l> and immediately opens a new one, so that our document remains valid. We just need to repeat this process for each line. We could do that by repeating what we just did. Or, more simply perhaps, we could add the XML Document Refactoring toolbar, which would provide a single button to do the job.

We'll do something even simpler. Use the mouse to select the sequence of characters you just inserted (`</l><l>`); copy it with CTRL-C; move the cursor to the end of the next line; paste it with CTRL-V. Repeat this for each line, except the last one of course.

5.1.7 Tagging another way

Some people just don't like tags. Fortunately, Oxygen also has a 'tag free' editing mode: it works by displaying parts of the text which are tagged in different ways in different styles. We specify the style for each tag by means of a stylesheet associated with the document.

- Select XML Document/Associate XSLT/CSS Stylesheet from the Document menu (or click the appropriate button if you can find it). In the Associate dialog, navigate to the file progress.css in your working directory (or download it) and select it. Observe that a new processing instruction is added at the start of your document.
- At the bottom of the editing window, you see a choice of Mode displays: Text (the default), Grid, and Author. Select the last. Observe that the display changes and a new Menu option (Author) is now available. Select this and observe the effect of the various command options on it.
- Choose `No Tags` for the next part of this exercise.

5.1.8 Quick splitting

- Use the mouse to select the the rest of the poem (from the word 'The' to the word 'shrink') excluding the name 'Evoe' at the end.
- Use CTRL-E to tag this all as a single `<lg>`.
- Select the same stratch of text again and tag it as a single `<l>` in the same way.
- Now place the cursor at the end of the first line of the second stanza (after 'around'). Press RETURN. A menu offers you the choice of splitting the `<l>` element. Press RETURN again.
- Move on to the next line (after 'banks') and repeat. Repeat for each subsequent line in the poem.
- To split the stanzas, put the cursor at the end of the last line of the stanza, and use the right arrow key to move it between the invisible tags (or switch back to Text mode). You should then be able to split the stanzas in exactly the same way.
- You can also experiment with other ways of splitting the text, of course. If you get into a mess, use CTRL-Z to undo the last change you made. Remember you can switch between Author and Text views as much as you like.
- Finally, tag the name of the author of the poem ('Evoe') using the `<signed>` element.

5.2 Making your own schema with Roma

5.2.1 Before you start

In this exercise, you will use **Roma**, a web tool available from the TEI web site and usable with any web browser: Firefox or Internet Explorer or Opera for example. Once you have created your schema, you will also need an XML-aware editor: oXygen in our case.

Our goal is to make a schema which we can use to mark up the sample page. We don't need all of of TEI Lite, much less the full TEI, but we do need bits of various modules. We'll also have to tinker with some of the modules, to make our schema more helpful with daily editing.

5.2.2 Making your own schema

1. Open the Roma application, by pointing your favourite web browser at `http://www.tei-c.org/Roma/`

2. The Roma start screen allows you to create a new customization, or to upload an existing customization for further work. We will start from scratch, which means ticking the first radio button ("Build schema (Create a new customisation by adding elements and modules to the smallest recommended schema)"). Press the `Submit` button at bottom left of the screen to continue.

The next and subsequent screens show you a row of tabs for acting on your customization (`Save`, `Customize`, `New`, and `Help`), and a row of tabs for each of the major stages or tasks making up a customization (`Modules`, `Add elements`, `Change classes`, `Language`, `Schema`, and `Documentation`. We won't explore all of these in this exercise. By default the `Customize your Customization` screen is displayed. This allows you to specify a file name and other details for the schema, and also to change the interface language if you wish. For now, accept the defaults. Go to the `Modules` tab to proceed.

The modules screen shows two lists: on the left are all available TEI modules; on the right are the modules currently selected for your schema. You can add modules from the list on the left, and remove modules from the list on the right, by clicking the appropriate word next to the module you wish to operate on.

1. For this exercise, we will need the following extra module:

   - `figures`

   Click the word `add` next to the name of the module.

2. The modules chosen contain many more elements than we need, so we will now remove some of them, simplying the view in the XML editor. Click the name of a module in the *List of selected modules* (the right-most column) to see a list of the elements this module defines.

Each element listed has a name, a radio button indicating whether it is to be included or excluded, a tag name, a description, and a link to a further screen where its attributes are specified. You can toggle inclusion or exclusion of all elements in the list by clicking the appropriate column heading. You can click on `Exclude` to remove all elements from the module.

Now work down the list clicking the radio button to restore or add the elements needed for this exercise. Remember to press the `Submit` button when you have finished with each module. Press the `Modules` or `back` links to go back to the list of modules.

from the core module delete `<add>`, `<analytic>`, `<cb>`, `<cit>`, `<corr>`, `<del>`, `<distinct>`, `<expan>`, `<gap>`, `<gloss>`, `<headItem>`, `<headLabel>`, `<index>`, `<meeting>`, `<mentioned>`, `<milestone>`, `<orig>`, `<pb>`, `<postBox>`, `<postCode>`, `<reg>`, `<series>`, `<sic>`, `<street>`, `<time>`, and `<unclear>`

from the textstructure module delete `<argument>`, `<byline>`, `<closer>`, `<dateline>`, `<epigraph>`, `<imprimatur>`, `<opener>`, `<salute>`, `<titlePage>`, and `<trailer>`.

from the figures module delete `<cell>`, `<formula>`, `<row>` and `<table>`.

We are now ready to generate a schema. Click the *Schema* tab, and then press Submit, taking the default option of a RELAXNG compact schema. Your browser will ask whether you want to save or open the generated file: you should save it into your Desktop directory (the default). Look at the result, if you feel strong, or experiment with other options of the web application.

Complete this stage by going to the *Save* tab and saving your work. Do not close the web browser, we'll use it again shortly.

5.2.3 Using your schema in oXygen

You can use oXygen and the file you made earlier to check that you've made your schema correctly. Proceed as follows:

- Open the file you made earlier in oXygen.
- Go to the menu `Document`, then `XML Document` and then `Associate Schema`. Choose the `RELAXNG` tab, and locate your schema file (using the middle folder icon on the right to browse).
- If all goes well, oXygen will insert some code to mark the schema location, and attempt to validate the file. Try inserting some new elements, and you should see a much-reduced collection.

5.2.4 Enhancing your schema

Now let us return to the Roma, and make the schema more constrained, to make sure we get just what we want in our documents. The example we have chosen is to constrain the allowed values of the `type` attribute on `<div>`, and to make it compulsory.

Go back to Roma. If you have closed the browser, you can restart Roma and load the session you saved earlier. Go to the *Modules* tab and click on `textstructure` in the right-hand column. Find `<div>` and click on `Change attributes` on the right-hand side. This will show you all the attributes of `<div>`. Click on `type`, and you will be able to change its properties:

- Change the `Is it optional` radio button to make it compulsory
- Change the radio button for `Closed list?` to make it a closed list
- In the box for `List of values`, type

```
cartoon,verse,drama
```

  (ie a list of possible values, separated by commas).

- Click on Submit Query

Now save the schema as before and reload the file in oXygen. There should be a validation error, because the last `<div>` in the Punch file has no type attribute.

5.2.5 Documenting your schema with Roma

One of the major benefits of Roma, is that after you have customized your schema it can produce two ways of documenting the changes you have made. One of these is the ODD file that you saved earlier (and can save again if you want) which indicates how your schema differs from full TEI, what modules you have included, what elements you may have added and changed, amongst other things. This is a good file to keep with your generated schema, in case you need to generate a new schema with additional elements or constrain it further. It also means that others can generate the schema for your documents in different schema languages if needed.

You may wish to open the ODD file you saved earlier and see how this format works.

You don't need to use the web version of Roma to create ODD files, you can just author them by hand if you are feeling geeky. There is a command-line Roma script which you can use to generate any of the outputs which the web version produces.

However, since the TEI Guidelines themselves are made up of ODD documents, Roma also allows you to generate a set documentation for your particular customization of the TEI. You don't need to worry about creating the documentation for the existing elements you have included, Roma already knows about them. If you've added new elements you are able to provide descriptions and information about the element at that point, and Roma will use that. Moreover, if you've changed the names of the

elements or their descriptions (perhaps for reasons of internationalization) it will use the new names you have provided in the documentation it generates.

You may wish to generate this from the `Documentation` tab.

This concludes the brief exercises on **Roma** for customizing the TEI schema. If you have time you may wish to experiment with making some other customizations to your schema! Some ideas are below.

5.2.6 Other things to try with Roma

1. How do you go about renaming an existing element? What happens in the ODD when you do?
2. How do you add a new element? What namespace does it end up in when you do? How can you control this?
3. Once you have an ODD, you can generate project-specific documentation. Where would you put this prose in the ODD file? Try generating some test documentation with additions you have made.
4. Experiment with starting with different exemplar customization Roma offers. Which do you think would be best for your project?
5. Save the ODD for the TEI ALL Plus schema – how does this add in schemas from other namespaces?
6. How do you go about changing the content model of an element? What implications does this have for your schema? How is it expressed in ODD?
7. Try modifying an ODD 'manually' in oXygen to include/exclude other elements and then submitting it to Roma to generate a schema.

5.3 Exercise: editing a TEI file: metadata

From our previous exercise, we have got a valid TEI document, with the minimum required components. Now let us look at adding some more metadata.

- Replace the contents of `<publicationStmt>` with

```
<idno type="gutenberg">24357</idno>
<availability>
 <p>This text is freely available for re-use under US and UK law,
consult
   your local legal restrictions if elsewhere.</p>
</availability>
```

  to indicate the original catalogue number, and copying conditions.

- Improve the `<sourceDesc>` by adding a `<bibl>` explaining where the text came from:

```
<sourceDesc>
 <bibl>
  <title level="a">Enigma</title>, <title level="j">Punch:  or the
London
     Charivari</title>, <date when="1914-07-01">July 1, 1914</date>,
147, p.  6</bibl>
</sourceDesc>
```

- Add a <profileDesc> after <fileDesc> and specify which languages you can see. at least:

```
<langUsage>
 <language ident="en">English</language>
</langUsage>
```

- We can also have a text classification inside the <profileDesc>, something like

```
<textClass>
 <keywords scheme="DD">
  <term>English Humour</term>
  <term>Magazine</term>
 </keywords>
</textClass>
```

- Finally, look at adding a <revisionDesc> after <fileDesc>, maybe

```
<revisionDesc>
 <change when="2009-03-35"/>
</revisionDesc>
```

  to indicate when you edited the text

Whatever you do, make sure its still valid TEI!

5.4 Exercise: editing a TEI file: names and dates

We have a basic TEI document, but there is a lot more we could do it. There are people, places, and dates mentioned in the prose. Let us try to mark those up.

- For each name you can recognize (it is often marked in bold or italic) put it into a <name> element; distinguish between people and places by using the type attribute on <name>. eg

```
<p>The scene opens at a party given by
<name type="person">Potiphar</name> in <name type="place">Venice</name>.</p>
```

- One of the paragraphs talks about an event of 1873. This needs marking with a <date>.
- The characters Joseph and Potiphar appear often. Make a new <listPerson> element in the header inside <sourceDesc> and create <person> elements for Joseph and Potiphar, giving them xml:id attributes. Link the occurrences of these names to the person using the ref attribute. eg <person xml:id="JO"> is pointed to by

```
<name ref="#JO">Joseph</name>
```

Whatever you do, make sure its still valid TEI!

5.5 Exercise: using the Image Markup Tool

5.5.1 About the Image Markup Tool

The 'Image Markup Tool', or more correctly *The UVic Image Markup Tool Project* (hereafter IMT), is a project which produces a tool for the annotation of images. The IMT allows researchers to annotate portions of images and categorise those annotations. The underlying storage format for these annotations is a pure subset of TEI P5 XML. The IMT is primarily the creation of Martin Holmes at the University of Victoria, British Columbia, Canada. For more information please see its website at: http://www.tapor.uvic.ca/~mholmes/image_markup/.

The goals of the IMT are:

- be simple for novices to use
- load and display a wide variety of different image formats
- allow the user to specify arbitrary rectangles on the image, and associate them with annotations
- allow such rectangles to overlap if the user wishes
- provide mechanisms for bringing overlapped rectangles to the front easily
- allow annotations to be grouped into categories, distinguished by name and appearance, to allow different classes of annotation
- require no significant knowledge of XML or TEI
- allow the use of complex XML code if the user wishes
- save data in an XML file which conforms to a TEI P5-based schema
- reload data from its own files
- generate a simple "Web view" of the data for display in a browser
- come packaged with an installer, Help file, and basic tutorial
- *be simple for novices to use!*

This exercise is designed to be fairly free-form, and you should not feel that you have to follow it too closely as long as you undertake the basic activities. The point of it is to gain some familiarity with a free (and open source) off-the-shelf tool that produces TEI XML for one specific purpose.

5.5.2 First Steps

1. **Load the IMT:** When it starts should see two windows, the IMT itself and its 'annotation window'. Arrange these on the screen by resizing them as you desire.
2. **Load an Image:** To load an image go to `File / Load an image` or press `Control + l` or the appropriate toolbar button. Note that there is a difference between loading an image and loading a project. From the network storage of data files we have provided load the the image '147_15.jpg'. This is a page from Punch magazine.
3. **Scale the image:** The image will have loaded at 100% zoom. This is likely to be an annoying way to view the image. Using the dropdown combo box on the left-hand side of the toolbar, scale the image to such a size as is appropriate for your monitor. Perhaps try 25% to start with, but adjust as you see fit.
4. Give the whole project a title in the Title textbox just under the toolbar. Perhaps something like `Page 15, of the July 1, 1914, issue of "Punch, or the London Charivari"` might be appropriate.

5.5.3 Create Some Annotation Categories

You should create some annotation categories before you start annotating your image. To do this select `Categories / Add a new category` from the menu or select the appropriate toolbar button. I would suggest you might want to add three categories for now:

Structural A category to mark the locations of major structural elements

Interpretative A category to mark locations with annotated interpretations

Textual A category to mark the locations and /or provide transcriptions of distinct textual elements

For each of these, you will want to make the ID something sensible sounding, fill in the description, and choose a different colour from the dropdown menu. Using the default rectangular shape is probably best.

5.5.4 Add Some Annotation Areas

1. After you've made some annotation categories, you should make some annotations in them. Do this by selecting `Annotations / New annotation` from the menu or the appropriate toolbar button.

2. For each annotation move the box which appears in the centre of the image to the appropriate position and resize it to fit around the portion of the image that you are annotating.

3. Using the 'Structural' category, create an annotation area that surrounds the header area of the page (containing the date, journal title, and page number). Give this annotation a title, perhaps `Page Header` and put a short description of this area in between the paragraph tags in the Annotation text box.

4. Provide similar annotations in the structural category for:

   - the printed page as a whole
   - the section containing the cartoon
   - the main textual section
   - any other structural components you feel inclined to mark

Note: Your annotation areas can overlap! And to help in displaying them you can control which areas are being displayed. You may wish to experiment with this before you add too many more annotation areas.

5.5.5 Look At What You've Done!

1. Save your file! Select `File / Save` from the menu to save your file. I suggest a filename like 'p15-imt.xml'.

2. Open this file in oXygen and examine it.

3. Note the metadata in the `<teiHeader>`, specifically the use of the `<rendition>` elements, and the `<appInfo>` element.

4. Scroll down further and see how the `<facsimile>` element has been used.

5. Even further down and you will see some `<div>` elements which contain the content of the annotations you have made.

5.5.6 Add More Annotations

Add some more annotations, this time use the 'Interpretative' annotation category to say what you think about various sections of the page image. For example, you may wish to:

- describe the individuals in the two cartoon panels
- explain some of the items in the Militants' Tariff table
- say what you really think about the themes of the Love's Logic poem
- provide your own interpretations as to possible meanings of the two text boxes in the cartoon

Add at least a few annotations in this category before moving on to the next section.

5.5.7 Re-examine What You've Done!

Save your file and re-examine the resulting XML file to see how the new annotations have been added in. The more annotations you add, the more `<div>` elements there will be in the body section and `<zone>` elements in the facsimile section.

5.5.8 Generate a 'Web View'

Now that you've added a number of annotations, it might be nice to see them in a rendered format.

1. After saving your file again, select `File / Save Web view` from the menu
2. Choose a name for your html file, perhaps 'p15-imt.html'
3. Open this file in a web browser and have a play. Specifically note that:
   - the different sections of the image are highlighted in your chosen colour when you hover over them
   - when you click on them once your annotation appears
   - when you double-click on an annotated section, a larger version of it appears (single-click again to remove)
   - you can navigate via the 'Annotations' box, and your annotations appear as you click on different items in the list
4. You may also wish to look at the other files you have created in this folder. If you named your filename 'p15-imt.html' then you should at least have the following files:

   p15-imt.css - A Cascading Style Sheet for displaying the web view

   p15-imt.html - The HTML file for your web view

   p15-imt.js - The javascript code used in displaying the web view

   p15-imt_wv.xml - A copy of the XML project file containing your annotations

   p15-imt_wv.xsl - The XSLT stylesheet used to transform the XML above into the HTML above

   p15-imt_wv_147_15.jpg - A shrunk down JPEG version of your image for display in the web view

   p15-imt_wv_zoom_147_15.jpg The full-sized JPEG version of your image for zooming display in the web view

IMT web view is designed as an example transformation. One can take the XML, javavscript, CSS and XSLT and further customise them directly to create a web interface to your image(s) that blends in with your website.

5.5.9 Add a Textual Transcription

If you were using the IMT for real, to create a multi-page annotated edition, you might cut-and-paste the IMT markup into your existing XML file. That is really outside the scope of this exercise.

However, if you want, you can easily do the reverse:

1. add an annotation area using the 'Textual' annotation category around the poem "Love's Logic"
2. open a copy of `1914_July_1.xml` in oXygen
3. find the transcribed text of the poem (try searching for 'LOGIC') and copy-and-paste the entire `<div type='verse'>` into the annotation text box, replacing the paragraph that was there. Save the annotation.

5.5.10 Re-examine What You've Done, Again!

Save your project file again, as above, open up the p15-imt.xml file in oXygen and see how the poem you pasted in has been recorded in the file.

5.5.11 Re-Generate a 'Web View'

Regenerated the 'web view' as above, reload the file in your web browser, and select the annotation you have just created to see the affect this has.

5.5.12 Further Work

Other things you may wish to explore or think about if you have time, or at a later date, include:

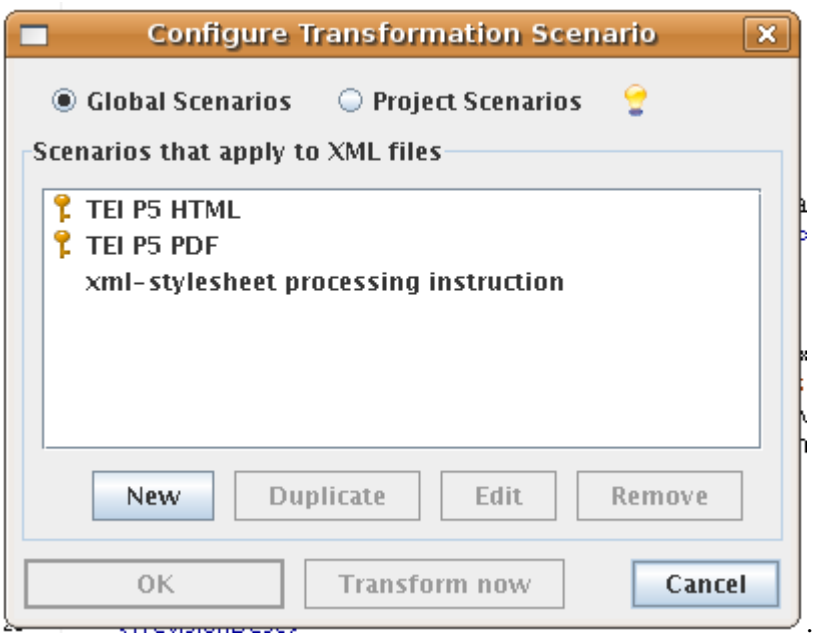
1. How do you delete an annotation?
2. How do you move an annotation from one category to another?
3. What happens to annotations when you delete their category?
4. What happens to annotations when you import a different resolution image of the same page?
5. What sections of a saved IMT XML file do you need to cut and paste into your exist TEI XML file to have that file work with IMT?
6. How many annotations are too many to be useful for one category? How many categories are too many to be useful for one image?
7. IMT can be run entirely off a usb flash key ... how might that be useful?
8. How can you create thumbnail images of all or part of the image with IMT? Why might you want to?
9. How does the IMT schema differ from the full tei_all schema?
10. How could you use IMT with your materials?

5.6 Exercise: using XSLT in oXygen

5.6.1 First steps

For this purpose, use the finished Punch document for July 1st 1914, in the directory `Punch/147/1914_July_1` on your network drive. Load it into oXygen and check that it is well-formed.

Transformations are run in oXygen by pressing the icon like this: . The first time you do this, you should be offered a choice of built-in options:

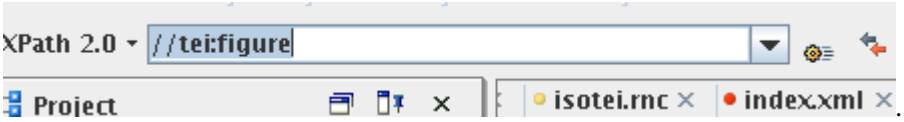


From these, choose 'TEI P5 HTML' and click on `Transform now`.

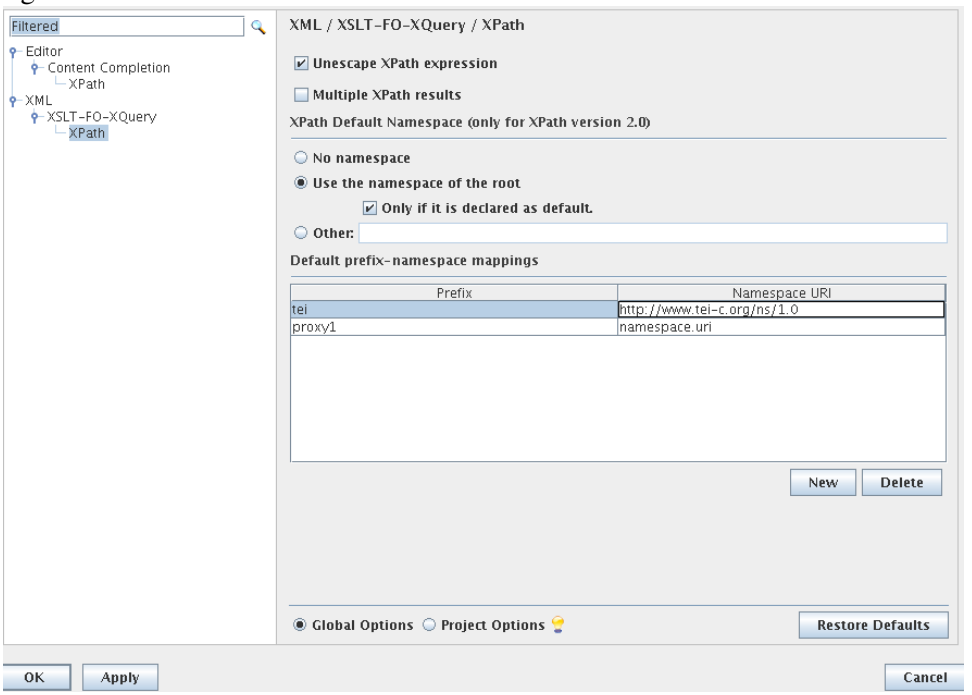
The transformation is run, and the default web browser is fired up with the resulting web page. Check that this looks like a plausible rendition, and that links work etc.

5.6.2 Working with XPath

You can explore the Punch file (or other XML file of your choice!) by typing XPath expressions into the box at the top left:

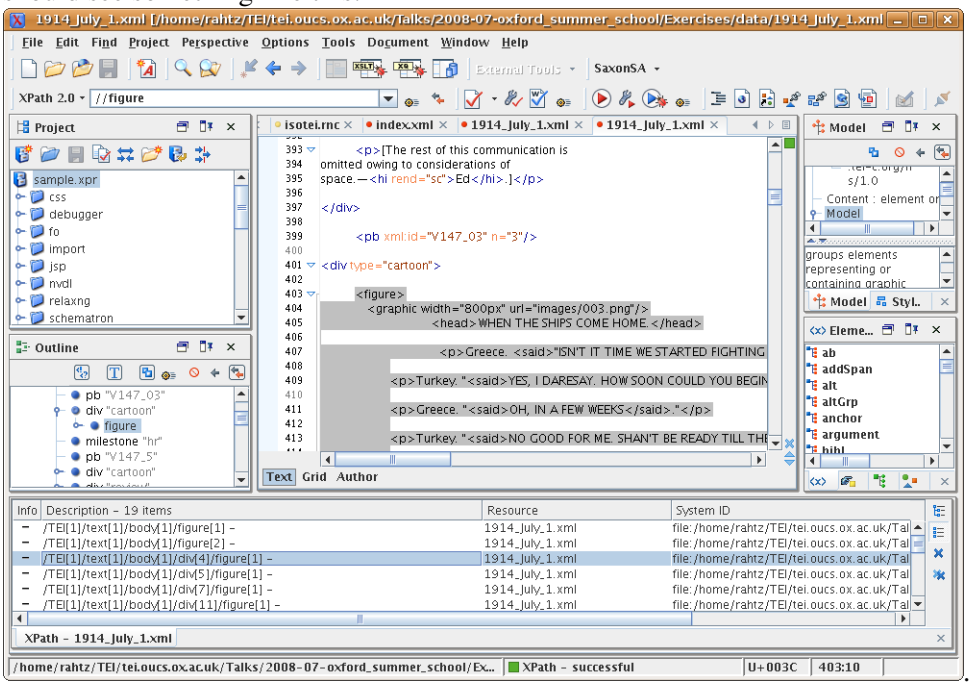


If you use XPath 2.0 (the default), you can put in expressions such as `//figure` without a namespace prefix. If you use XPath 1.0, you have to first set up the namespace by pressing the small icon to the right of the XPath box:



and linking the prefix `tei` with the namespace `http://www.tei-c.org/ns/1.0`. It is probably easier for this exercise to stay with XPath 2.0 and use unprefixed element names.

Type `//figure` into the box, and select one of the results listed in the window at the bottom. You should see something like this:



Try to locate:

- All the `head` elements inside `figure` elements
- All the **name** elements which do **not** have a rend attribute.
- The foreign words (marked as `<foreign>`)
- The sections (`<div>`) which more than 5 paragraphs
- Paragraphs which contain the word 'Dogs'

5.6.3 Trying the PDF rendering

- Load the file `stalky.xml` into oXygen, and choose the TEI P5 to PDF transform scenario (press the 'Configure Transfomation Scenario' icon,). If all goes well, your browser wil load a PDF rendering in due course.
- Now set the parameter `Institution` to 'Oxford Summer School' and rerun the transformation. See the difference?
- More dramatically, change `columnCount` to have the value 2, and see what happens then.
- Set `parIndent` to '0em' and `parSkip` to '2pt'
- Finally, change `pageWidth` to '1755mm', change `columnCount` back to 1, run the transform, and check that the page width is lessened.

5.6.4 Going further with parameters of the HTML

Try some of these changes to the HTML rendering of Punch, by setting parameters, and check the results:

- Set `autoToc` to 'false'
- Set `numberHeadings` to false
- Set `pageLayout` to 'CSS'
- Set `numberParagraphs` to 'true'

5.6.5 Feeling strong?

Download the TEI OpenOffice filters by following links from `http://tei.sf.net` (or go direct to `http://sourceforge.net/project/showfiles.php?group_id=106328&package_id=197808&release_id=587974`). That will be a `.jar` file for TEI P5. Do not open that, but save it on disk. Start OpenOffice, make a new text document, then go to `Tools/XML Filter Settings` and choose `Open Package`. Point it at the `.jar` file you downloaded, close that dialogue, go back to the text document and try 'Save As ... TEI P5'.