Verse, Drama, and Dictionaries

James Cummings

April 2009
The TEI provides more elements to talk encode verse and metrical information than you might first suspect.
• `<l>` (a line of verse)
• `<lg>` (a line group such as stanza or paragraph)

```xml
<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="i10">And Farnaby Fullerton Rigby.</l>
</lg>
```
Many poems consist only of ungrouped lines. This short poem by Emily Dickinson is a simple case:

```html
<body>
<l>To make a prairie it takes a clover and one bee,
<l>One clover, and a bee,
<l>And revery.
<l>The revery alone will do,
<l>If bees are few.
</body>
```
<body>
  <head>My Alba</head>
  <lg>
    <l>Now that I've wasted</l>
    <l>five years in Manhattan</l>
    <l>life decaying</l>
    <l>talent a blank</l>
  </lg>
  <lg>
    <l>talking disconnected</l>
    <l>patient and mental</l>
    <l>sliderule and number</l>
    <l>machine on a desk</l>
  </lg>
</body>
Sire Thopas was a doghty swayn;
White was his face as payndemayn,
His lippes rede as rose;
His rode is lyk scarlet in grayn,
And I yow telle in good certayn,
He hadde a semely nose.

His heer, his ber was lyk saffroun,
That to his girdel raughte adoun;
In the first year of Freedom's second dawn
Died George the Third; although no tyrant, one
Who shielded tyrants, till each sense withdrawn
Left him nor mental nor external sun:
A better farmer ne'er brushed dew from lawn,
A worse king never left a realm undone!

He died — but left his subjects still behind,
One half as mad — and t'other no less blind.
Shakespearean Sonnet with nested `<lg>`

```xml
<lg>
  <lg type="quatrain">
    <l>My Mistres eyes are nothing like the Sunne, </l>
  </lg>
  <!-- 3 more lines -->
</lg>

<lg type="quatrain">
  <!-- 4 lines -->
</lg>

<lg type="quatrain">
  <!-- 4 lines -->
</lg>

<lg type="couplet">
  <l>And yet by heaven I think my love as rare, </l>
  <l>As any she beli'd with false compare. </l>
</lg>
```
Compare the structure...

<div n="I" type="book">
  <div n="I.1" type="canto">
    <div n="I.1.1" type="stanza">
      <l>A noble knight was pricking on the plain</l>
      <l>Ycladd in mightie armes and silver shielde...</l>
    </div>
  </div>
</div>

<div n="I" type="book">
  <div n="1" type="canto">
    <lg n="I.1.1" type="stanza">
      <l>A Gentle Knight was pricking on the plain</l>
      <l>Y cladd in mightie armes and silver shielde,</l>
    </lg>
  </div>
</div>
In a somer seson, when softe was the sonne,
I shoop me into shroudes as I a sheep were,
In habite as an heremite unholy of werkes,
Went wide in this world wondres to here.
Feet vs Syllables

<l>
  <seg type="foot">
    <seg type="syll">Ar</seg>
    <seg type="syll">ma</seg>
    <seg type="syll">vi</seg>
  </seg>
  <seg type="foot">
    <seg type="syll">rum</seg>
    <seg type="syll">que</seg>
    <seg type="syll">ca</seg>
  </seg>
  <seg type="foot">
    <seg type="syll">no</seg>
    <seg type="syll">Tro</seg>
  </seg>...
</l>
In a somer seson, whan softe was the sonne,
I shoop me into shroudes as I a sheep were,
In habite as an heremite unholy of werkes,
Went wide in this world wondres to here.
Metrical Analysis

When the module for verse is in use, the following additional attributes are available to record information about rhyme and metrical form:

- att.metrical defines a set of attributes which certain elements may use to represent metrical information.
  - @met - (metrical structure, conventional) contains a user-specified encoding for the conventional metrical structure of the element.
  - @real - (metrical structure, realized) contains a user-specified encoding for the actual realization of the conventional metrical structure applicable to the element.
  - @rhyme - (rhyme scheme) specifies the rhyme scheme applicable to a group of verse lines.
Metrical Analysis Example

'Tis hard to say, if greater Want of Skill
Appear in Writing or in Judging ill;
But, of the two, less dang'rous is th'Offence,
To tire our Patience, than mis-lead our Sense:
The @real (for ‘realization’) attribute may optionally be specified to indicate any deviation from the pattern defined by the @met attribute which the encoder wishes to record.

\[ <l \text{ real}="+-| -+| -+| -+| -+" >\text{But, of the two, ...} </l> \]

But since this is so common, maybe just:

\[ <l \text{ met}="+-| -+| -+| -+| -+" >\text{But, of the two, ...} </l> \]
Und frische Nahrung, neues Blut

Saug' ich aus freier Welt;

Wie ist Natur so hold und gut,

Die mich am Busen hält!

Die Welle wieget unsern Kahn

Im Rudertakt hinauf,

Und Berge, wolkig himmelan,

Begegnen unserm Lauf.
@met and @real at the same time

<l n="356"> A needless alexandrine ends the song, </l>
<l n="357" met="-+|-+|-+|-+|-+|-+|+-|++|--"> That, like a wounded snake, drags its slow length along. </l>
The @rhyme attribute is used to specify the rhyme pattern of a verse form. It should not be confused with the <rhyme> element, which is used to mark the actual rhyming word or words:

- <rhyme> - marks the rhyming part of a metrical line.
<lg type="couplet" rhyme="aa">
  <l>Outside in the distance a wildcat did</l>
  <rhyme>growl</rhyme>
  <l>Two riders were approaching and the wind began to</l>
  <rhyme>howl</rhyme>
</lg>
Another <rhyme> example

<lg type="quatrain" rhyme="abab">
  <l> I wander thro' each charter'd 
  <rhyme label="a">street</rhyme>, </l>
  <l> Near where the charter'd Thames does 
  <rhyme label="b">flow</rhyme>, </l>
  <l> And mark in every face I 
  <rhyme label="a">meet</rhyme> 
  </l>
  <l> Marks of weakness, marks of 
  <rhyme label="b">woe</rhyme>. </l>
</lg>
Internal <r rhyme> example

<lg rhyme="ABCCBBA">
<l>The sunlight on the <r rhyme label="A">garden</r> </l>
<l>
<r rhyme label="A">Harden</r>s and grows <r rhyme label="B">cold</r>, </l>
<l>We cannot cage the <r rhyme label="C">minute</r> </l>
<l>Wi<r rhyme label="C">thin it</r>s nets of <r rhyme label="B">gold</r> </l>
<l>When all is <r rhyme label="B">told</r> </l>
<l>We cannot beg for <r rhyme label="A">pardon</r>. </l>
</lg>
Metrical Notation Declaration

- `<metDecl>` (metrical notation declaration) documents the notation used to represent a metrical pattern
  - `@pattern` (regular expression pattern) specifies a regular expression of legal values
- `<metSym>` (metrical notation symbol) documents the intended significance of a particular character(s) within a metrical notation
  - `@value` specifies the character or character sequence being documented
  - `@terminal` specifies whether the symbol is defined in terms of other symbols (terminal is set to false) or in prose (terminal is set to true).
<metDecl> example

<metDecl pattern="[DTIS3A]+">
  <metSym n="dactyl" value="D" terminal="false">-oo</metSym>
  <metSym n="trochee" value="T" terminal="false">-o</metSym>
  <metSym n="iamb" value="I" terminal="false">0-</metSym>
  <metSym n="spondee" value="S" terminal="false">--</metSym>
  <metSym n="tribrach" value="3" terminal="false">000</metSym>
  <metSym n="anapaest" value="A" terminal="false">00-</metSym>
  <metSym value="o">short syllable</metSym>
  <metSym value="-">long syllable</metSym>
</metDecl>
Drama or ‘Performance Texts’ is a module of TEI elements for use when encoding printed dramatic texts, screen plays or radio scripts, and written transcriptions of any form of performance.
The Programme

- `<performance>` contains a section describing how a dramatic piece is to be performed in general or how it was performed on some specific occasion
- `<prologue>` contains the prologue to a drama, typically spoken by an actor out of character
- `<epilogue>` contains the epilogue to a drama, typically spoken by an actor out of character
- `<set>` (setting) contains a description of the setting, time, locale, appearance, etc., of the action of a play
- `<castList>` (cast list) contains a single cast list or dramatis personae.
The action of the play is set in Chicago's Southside, sometime between World War II and the present.
Inside <castItem>

- `<role>` the name of a dramatic role, as given in a cast list
- `<roleDesc>` (role description) describes a character's role in a drama
- `<actor>` Name of an actor appearing within a cast list
<castItem>
  <role>Tom Thumb the Great</role>
  <roleDesc>a little hero with a great soul, something violent in his temper, which is a little abated by his love for Huncamunca</roleDesc>
  <actor>Young Verhuyk</actor>
</castItem>
<castGroup rend="braced">
  <head>friends of Mathias</head>
  <castItem>
    <role>Walter</role>
    <actor>Mr Frank Hall</actor>
  </castItem>
  <castItem>
    <role>Hans</role>
    <actor>Mr F.W. Irish</actor>
  </castItem>
</castGroup>
Acts and Scenes

<act n="2">

<scene n="1">
  <head>Scene One</head>
</scene>

<scene n="2">
  <head>Scene Two</head>
</scene>
</act>
Speeches and Speakers

- `<sp>` (an individual speech in a performance text)
- `<speaker>` (the name of the speaker(s) as given in the performance text)
- `<stage>` (a stage direction of any sort within a dramatic text)
Menaechmus

Responde, adulescens, quaeso, quid nomen tibist?

Peniculus

Etiam derides, quasi nomen non noveris?

Menaechmus

Non edepol ego te, quot sciam, umquam ante hunc diem

Vidi neque novi; ...
Stage Directions

- `<stage>` (stage direction) contains any kind of stage direction
  - `@type` indicates the kind of stage direction
- `<move/>` (movement) marks the actual entrance or exit of one or more characters on stage.
  - `@type` characterizes the movement
  - `@where` specifies the direction of a stage movement
  - `@perf` (performance) identifies the performance(s) in which this movement occurred as specified.
<div n="5" type="scene">
  <stage>Elsinore. A room in the Castle. </stage>
  <stage type="setting">Enter Ophelia, distracted. </stage>
  <sp>
    <speaker>Ophelia</speaker>
    <p>Where is the beauteous Majesty of Denmark?</p>
  </sp>
  <sp>
    <speaker>Queen</speaker>
    <p>How now, Ophelia?</p>
  </sp>
</div>
<sp>
  <speaker>Greece. </speaker>
  <said> ISN'T IT TIME WE STARTED FIGHTING AGAIN? </said>
</sp>
<sp>
  <speaker>Turkey. </speaker>
  <said> YES, I DARESAY. HOW SOON COULD YOU BEGIN? </said>
</sp>
<sp>
  <speaker>Greece. </speaker>
  <said> OH, IN A FEW WEEKS. </said>
</sp>
<sp>
  <speaker>Turkey. </speaker>
  <said> NO GOOD FOR ME. SHAN'T BE READY TILL THE AUTUMN. </said>
</sp>
Other types of performance text

- `<view>` describes the visual context of some part of a screen play in terms of what the spectator sees
- `<camera>` describes a particular camera angle or viewpoint
- `<caption>` contains the text of a caption or other text displayed
- `<sound>` describes a sound effect or musical sequence
  - `@type` categorizes the sound in some respect, e.g. as music, special effect, etc.
  - `@discrete` indicates whether the sound overlaps the surrounding speeches or interrupts them
The TEI defines a module for encoding human-oriented monolingual and multilingual dictionaries, glossaries, and similar documents. These are not just for standalone use, but could be for a wordlist or glossary accompanying a digital edition.
Dictionary Structures

- `<entry>` contains a reasonably well-structured dictionary entry
- `<entryFree>` (unstructured entry) contains a dictionary entry which does not necessarily conform to the constraints imposed by the entry element
- `<superEntry>` groups successive entries for a set of homographs
And other structures like...

- `<hom>` (homograph) groups information relating to one homograph within an entry
- `<sense>` groups together all information relating to one word sense in a dictionary entry, for example definitions, examples, and translation equivalents
Inside these structures

- `<form>` groups all the information on the written and spoken forms
- `<gramGrp>` groups morpho-syntactic information about a lexical item
- `<def>` contains a definition
- `<cit>` contains a cited quotation
- `<usg>` contains usage information
- `<xr>` contains a cross-reference
- `<etym>` encloses the etymological information
- `<re>` contains a related entry
- `<note>` contains a note or annotation.
<entry>
  <form>
    <orth>competitor</orth>
    <hyph>com|peti|tor</hyph>
    <pron>k@m"petit@(r)</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <def>person who competes.</def>
</entry>
Multiple senses

<entry>
  <form>
    <orth>disproof</orth>
    <pron>dIs"pru:f</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
  </gramGrp>
  <sense n="1">
    <def>facts that disprove something.</def>
  </sense>
  <sense n="2">
    <def>the act of disproving.</def>
  </sense>
</entry>
Inside `<form>`

- `<orth>` gives the orthographic form
- `<pron>` contains the pronunciation(s)
- `<hyph>` contains a hyphenated form
- `<syll>` contains the syllabification
- `<stress>` contains the stress pattern
- `<lbl>` contains a label for a form, example, translation, or other piece of information
What? There is more inside `<form>`?

- `<gram>` for grammatical information
- `<gen>` identifies the morphological gender
- `<number>` indicates grammatical number
- `<case>` contains grammatical case
- `<per>` contains the grammatical person (1st, 2nd, 3rd, etc.)
- `<tens>` indicates the grammatical tense
- `<mood>` contains information about the grammatical mood of verbs
- `<iType>` indicates the inflectional class
<form> example

<form>
  <orth>brag</orth>
</form>
<gramGrp>
  <pos>vb</pos>
</gramGrp>
<form type="infl">
  <orth>brags</orth>
  <orth>bragging</orth>
  <orth>bragged</orth>
</form>
Multiple Forms

<entry>
  <form type="abbrev">
    <orth>MTBF</orth>
  </form>
  <form type="full">
    <lbl>abbrev. for</lbl>
    <orth>mean time between failures</orth>
  </form>
</entry>
Another tasty <entry>

<entry>
  <form>
    <orth>rémoulade</orth>
    <pron>Remulad</pron>
  </form>
  <gramGrp>
    <pos>n</pos>
    <gen>f</gen>
  </gramGrp>
  <cit type="translation" xml:lang="en">
    <quote>remoulade</quote>
    <quote>rémoulade</quote>
    <def>dressing containing mustard and herbs</def>
  </cit>
</entry>