

TEI BASIC STRUCTURE AND CORE COMPONENTS

Martina Scholger

Centre for Information Modelling - Austrian Centre for Digital Humanities
University of Graz

Thanks to Lou Burnard, James Cummings, Sebastian Rahtz,
and the rest of the TEI community providing workshop material.

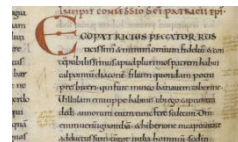
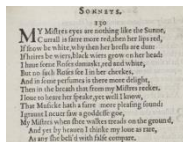
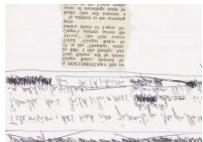


Topic of the session

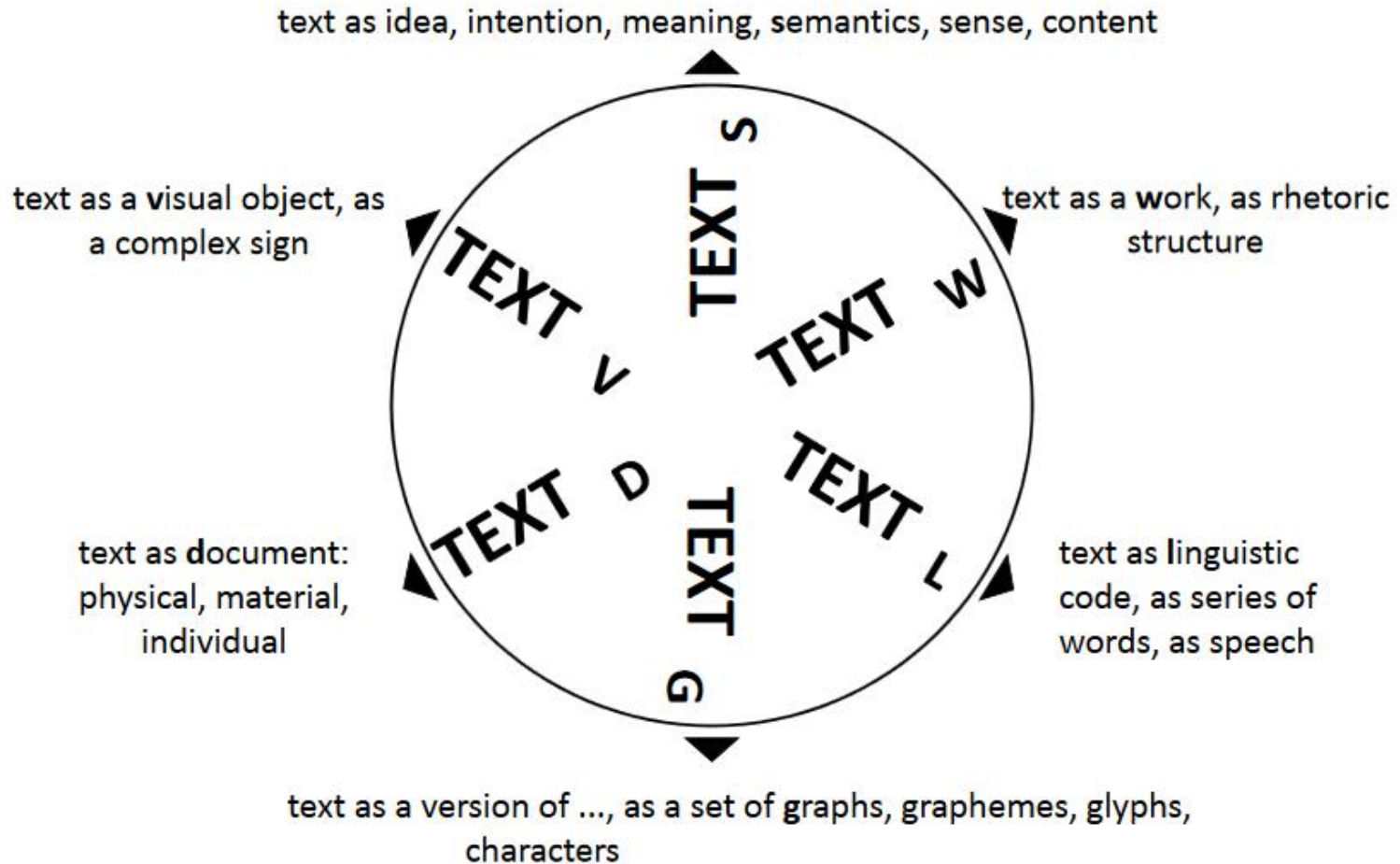
- 4 Default Text Structure (*textstructure*)
- 3 Elements Available in all TEI documents (*core*)

The TEI takes a generalistic approach and should be able to cope with texts

- ... of any size
 - ... language
 - ... complexity
 - ... on all media
 - ... from every time and place
- books, journals, manuscripts, letters, rolls of papyrus, coins, notebooks, postcards, inscription tablets, web pages, etc.



Pluralistic view on text (Sahle)



19th c. print magazine

JULY 1, 1914.]

PUNCH, OR THE LONDON CHARIVARI.

19



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,

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 stoney 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

Dictionaries

dumbledore, dumble-dore ('dʌmb(ə)ldɔə(r)).
local. [f. DUMBLE- + DOR *sb.*¹: see also DRUMBLE-*dore.*] A humble-bee or bumble-bee; also *dial.* a cockchafer.

1787 GROSE *Prov. Gl.*, *Dumble-dore*, an humble, or bumble-bee. 1799 SOUTHEY in *Robberds Mem. W. Taylor* I. 264 Is it not the humble-bee, or what we call the 'dumble dore',—a word whose descriptive droning deserves a place in song? 1837 — *Doctor* IV. *Interch.* xvi. 383 Of Bees, however, let me be likened to a Dumbledore, which Dr. Southey says is the most goodnatured of God's Insects. 1856 MISS YONGE *Daisy Chain* I. xxvi. (1879) 276 Buzzed and hummed over by busy, blacktailed yellow-banded dumbledores. 1863 G. KEARLEY *Links in Chain* iii. 57 In Hampshire these insects [humble bees] are Dumbledors, in other districts Bumble bees, and hummel bees. 1880 *Cornwall Gloss.*, *Dumbledory*, cockchafer.

Table of content



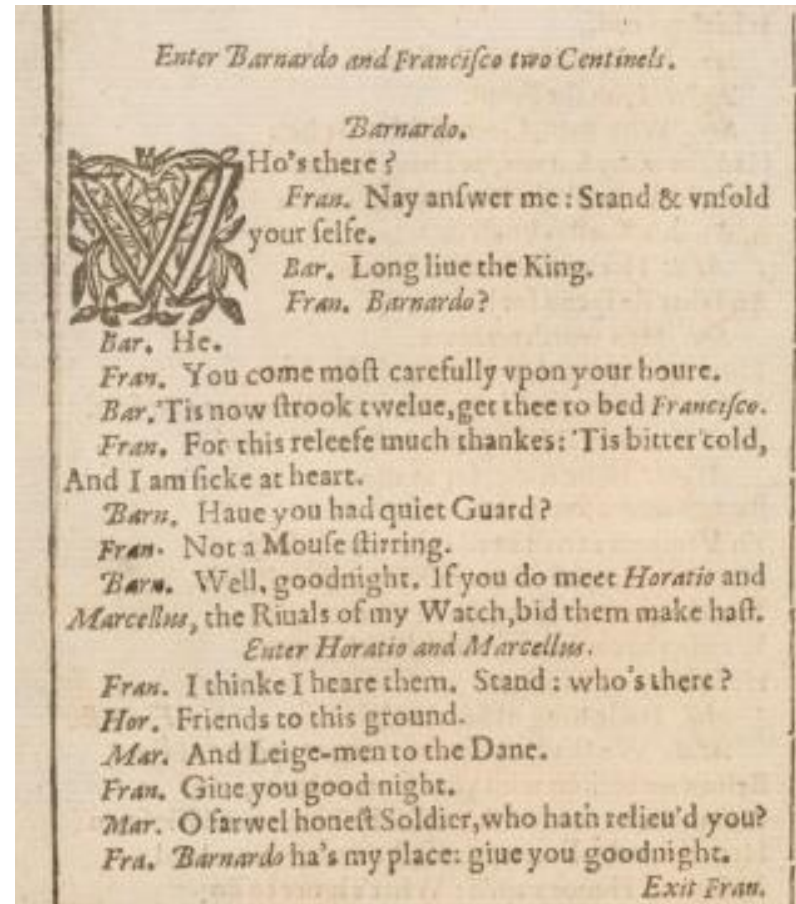
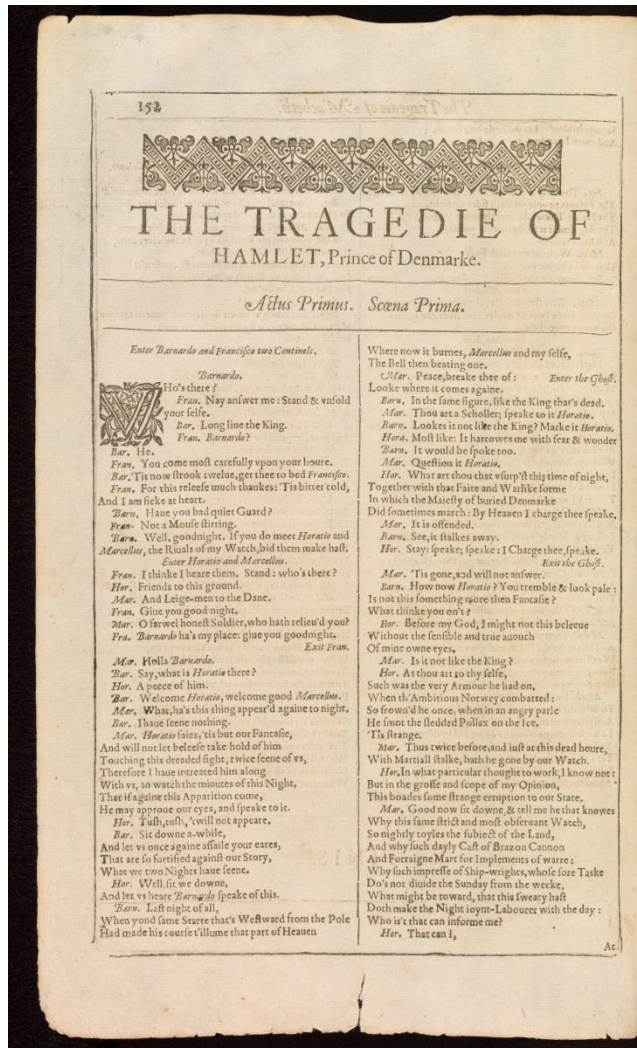
A C A T A L O G V E
of the severall Comedies, Histories, and Tragedies contained in this Volume.

COMEDIES.

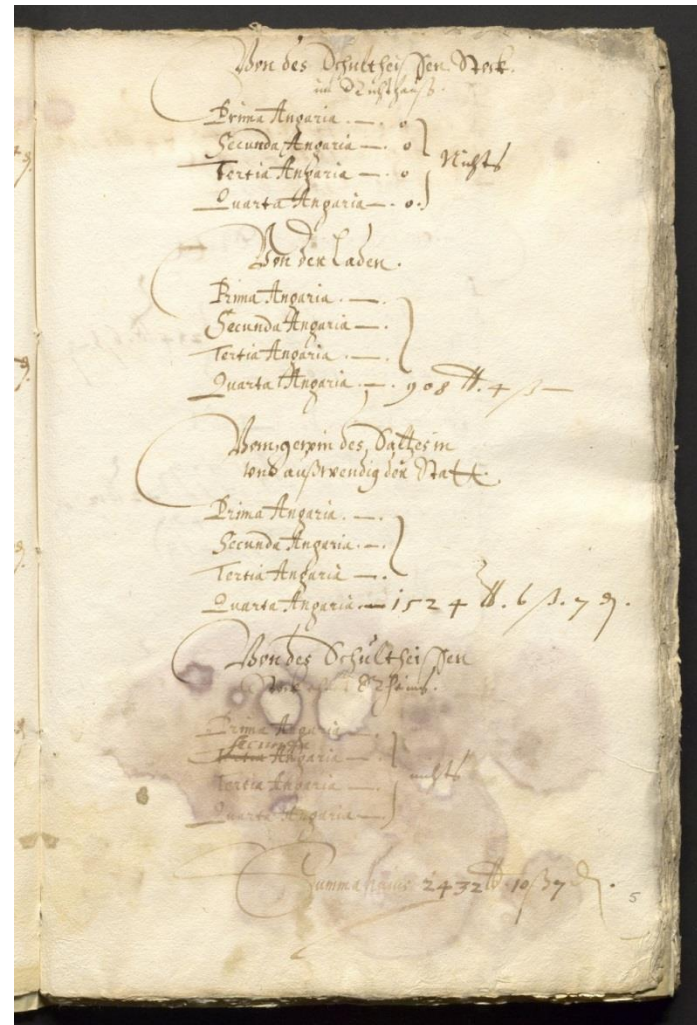
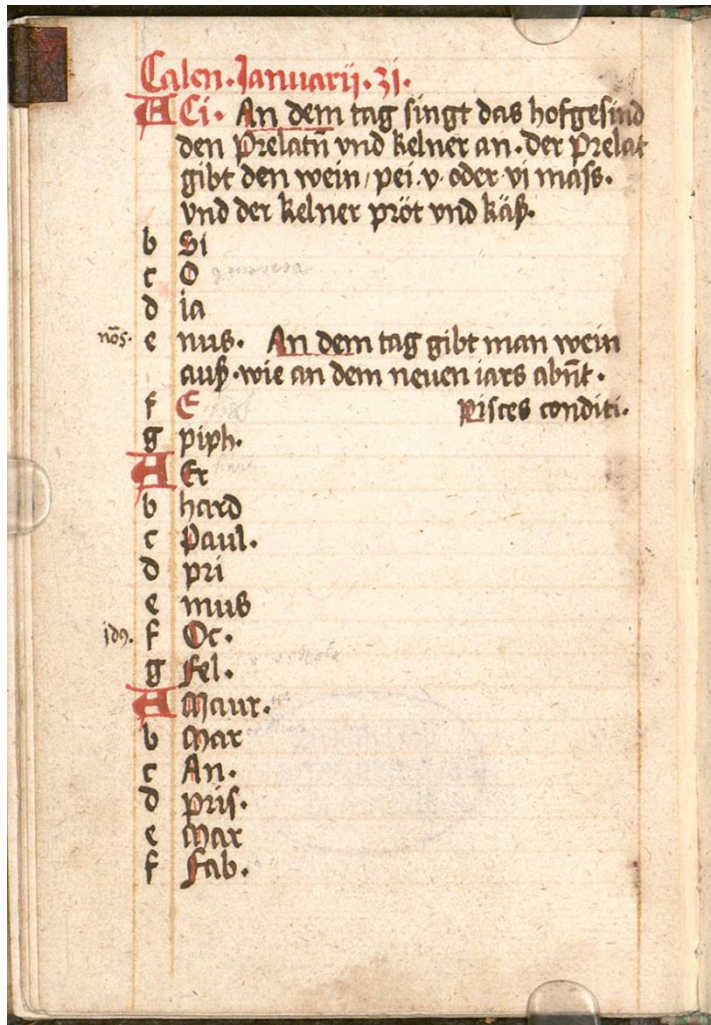
T he Tempest.	Folio 1.	<i>The First part of King Henry the fourth.</i>	46
T he two Gentlemen of Verona.	20	<i>The Second part of K. Henry the fourth.</i>	74
<i>The Merry Wives of Windsor.</i>	38	<i>The Life of King Henry the Fifth.</i>	69
<i>Measure for Measure.</i>	61	<i>The First part of King Henry the Sixth.</i>	96
<i>The Comedy of Errors.</i>	85	<i>The Second part of King Hen. the Sixth.</i>	120
<i>Much ado about Nothing.</i>	101	<i>The Third part of King Henry the Sixth.</i>	147
<i>Loves Labour lost.</i>	122	<i>The Life & Death of Richard the Third.</i>	173
<i>Midsummer Nights Dreame.</i>	145	<i>The Life of King Henry the Eighth.</i>	205

TRAGEDIES.

Dramatic text



Medieval calendar and Account book



Medieval manuscript (St. Patrick Confessio)

aqua
 iam
 luv
 re
 cas
 bar
 ne
 erdo
 qui
 quā
 qual

INCIPIT CONFESSIO SANCTI PATRICKI EPISCOPI.
EGO PATRICKIUS PECCATOR RUS
 ticussim & minimus omnium fidelium & con
 temperabilissimus apud plurimos patrem habui
 calpurnium diaconem. filium quondam pota
 presbiteri. qui fuit in uico banauem tabernic.
 Uillulam eni prope habuit ubi ego captum
 dedi. annorum enim tunc fere sedecim. Om
 enim uerū ignorabam. & hibernone in captiuitate
 adductus sum. circa milia hominum scilicet

EGO PATRICKIUS PECCATOR. rusticissimus & minimus omnium fidelium.
 & contemperabilissimus apud plurimos. Patrem habui calpurnium
 diaconem. quondam filium pota presbiteri. qui fuit in uico banauem tabernic.
 Uillulam eni prope habuit. ubi ego captum dedi. annorum enim tunc
 fere quindecim. De eo eni uerum ignorabam. & hibernone in captiuitate
 adductus sum. circa milia hominum. scilicet merita nostra. quia
 ado recessimus. & precepta eius non custodimus. & sacerdo
 tibus nostris inobedientes fuimus. quoniam salutem admo
 nebant. & dominus induxit super nos iram animationis sue. & dis
 psit nos in gentibus multis. etiam ad uiam terre. Vbi nunc
 paruitas mea uidetur esse inter alienigenas. & ubi dominus aperuit sensus
 cordis. & in misericorditer. ut uel sero rememorare delicta mea.
 & ut conuerterer toto corde ad dominum meum. qui respexit humilitatem
 meam. & miseratus adolescentie mee & ignorantie. & custodit me
 antequam scirem eum. & antequam sciperem uel distinguerem inter bonum
 & malum. & monuit me. & consolatus est me ut pater filium. Vno e
 autem tacere non possum. neque expecto quidem tanta beneficia & tanta

EGO PATRICKIUS PECCATOR. rusticissimus & minimus omnium
 fidelium. & contemperabilissimus apud plurimos.
 Patrem habuit. calpurnium diaconum filium condi
 pota presbiteri qui fuit in uico banauem tabernic. uillula enim
 prope habuit ubi ego captum dedi. annorum enim tunc fere
 xxi. Dominum uerū ignorabam. & hibernone in captiuitate
 adductus sum cum tot milia hominum scilicet merita nostra. quia
 ado recessimus. & precepta eius non custodimus. & sacerdo
 tibus nostris inobedientes fuimus. quoniam salutem admo
 nebant. & dominus induxit super nos iram animationis sue. & dis
 psit nos in gentibus multis. etiam usque ad ultimum terre. Vbi nunc
 paruitas mea uidetur esse inter alienigenas. & ubi dominus aperuit
 sensum cordis mei. in incredulitate uel sero rememorare
 delicta mea. Et ut conuerterer toto corde ad dominum meum
 qui respexit humilitatem meam. & miseratus adolescentie
 mee. & custodit me. antequam scirem eum. & antequam sciperem
 ut distinguerem inter bonum & malum. & monuit me
 consolatore non possum. neque expecto quidem tanta be
 neficia. & tantam gratiam quam mihi dominus prestare dignatus.
 in captiuitate mea. quia hoc est retributio nostra ut post

Seal and Squeeze

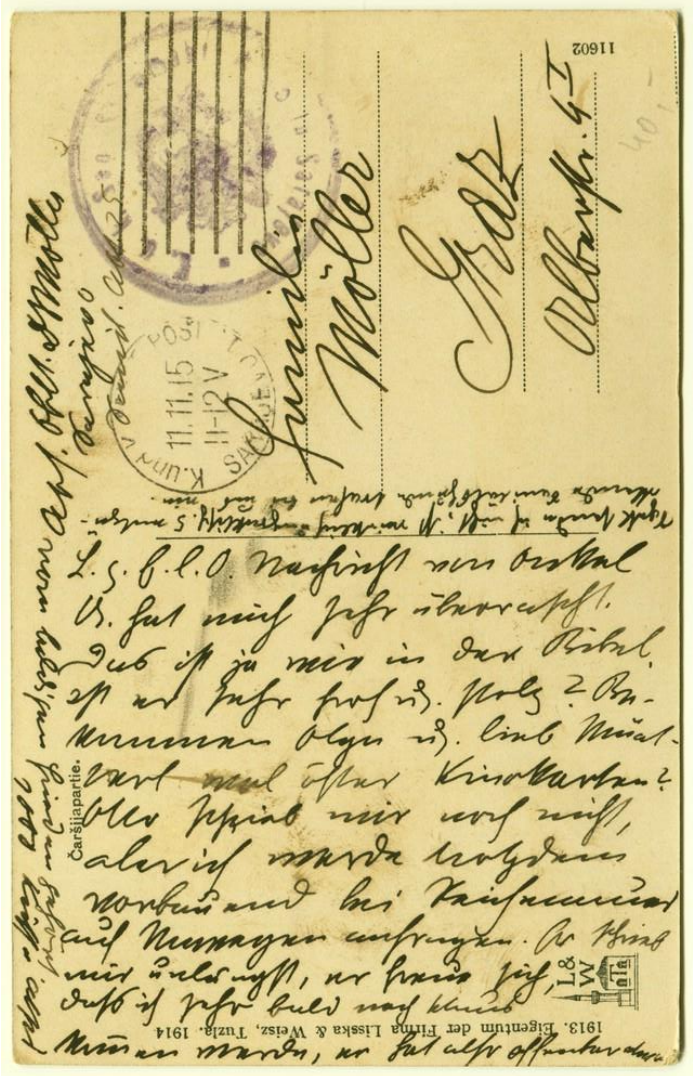


Poems

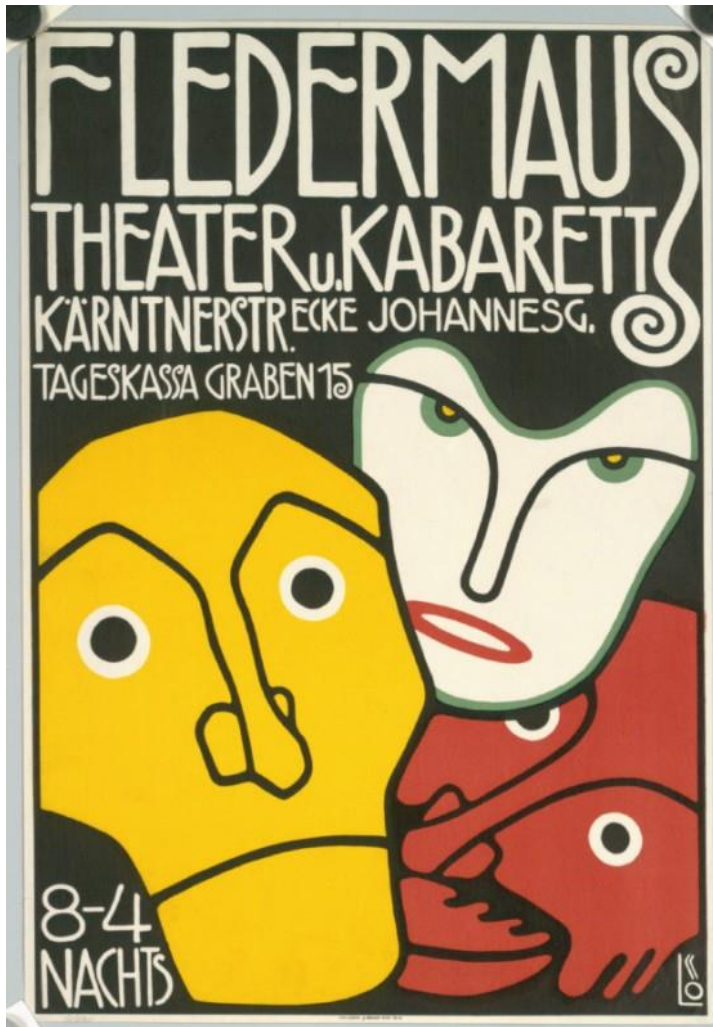
*Cette respectable heronme c'est la
adorable heronme c'est la
vaine heronme c'est la
l'impair
de ton buste
dors sur remon
a. bravis un image*

SONNETS.
130
MY Mistres eyes are nothing like the Sunne,
Currall is farre more red,then her lips red,
If snow be white,why then her brefts are dun:
If haire be wiers,black wiers grow on her head:
I haue scene Roses damask,red and white,
But no such Roses see I in her cheekes,
And in some perfumes is there more delight,
Then in the breath that from my Mistres reekes.
I loue to heare her speake,yet well I know,
That Musicke hath a farre more pleasing found:
Igraunt I neuer saw a goddesse goe,
My Mistres when shee walkes treads on the ground,
And yet by heauen I thinke my loue as rare,
As any she belid with false compare.

Postcard



Poster



Circus Olympicus.

zum
**aller-
 letzten
 Male**

am
Sonntag
 den
10. April
 1888

außerordentliche Vorstellung
Herrn Anton Poletti
 im Gebiete der unterhaltenden Physik
 und der natürlichen Zauberei.

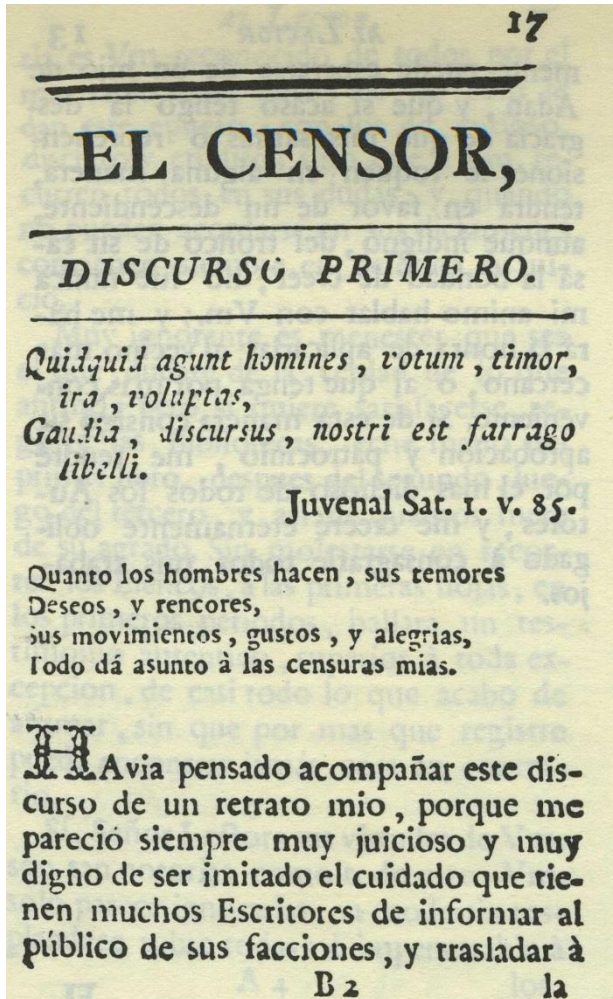
Erste Abtheilung.	Zweite Abtheilung.
1. Die sympathischen Karten.	1. Der Meisterstich.
2. Die Uhr des Sultan's Harun al Raschid.	2. Die Zauber-Uhr.
3. Die Chatonille des Geisterkönigs.	3. Der Brancard Wagen.
4. Die hysterischen Uhren.	4. Hermbold's Blumenpende.
5. Der besauberte Spiegel.	5. Der Heiß-Kaffee.

Non plus ultra!
Die magnetisch-electrische Lotterie. X Die improvisirte
Buchdruckerei.

Preise der Plätze: Erster Platz 20 fr. — Zweiter Platz 10 fr. — Dritter Platz 6 fr. — Das Militär zahlt am dritten Platz 3 fr.

Glocke: Eröffnung um 2 Uhr. — Anfang um 4 Uhr.

Print and Online journal



ride

A review journal for digital editions and resources
published by the IDE

[Home](#)
[About](#)
[Issues](#)
[Tags](#)
[Data](#)
[Reviewers](#)

Go

Home > Issues > Issue 1 > NietzscheSource

Nietzschesource

Nietzschesource, Paolo D'Iorio (ed.), 2009. <http://www.nietzschesource.org/> (Last Accessed: 25.05.2014). Reviewed by Philipp Steinkrüger (KU LEUVEN). philipp.steinkrueger (at) hiw.kuleuven.be

Abstract

This review addresses the digital edition of Nietzsche's works, *Nietzschesource*. The project presents a definite step ahead in the history of editions of the works of the 19th century philosopher and writer and offers the best text available to date. Moreover, it includes a growing archive of digital facsimiles of Nietzsche's manuscripts and thus allows a wide base of scholars to suggest corrections and emendations of the established text. As a digital edition, however, it is in many respects disappointing, for it fails to make use of the great possibilities modern editorial techniques offer, for instance, the possibility to enrich texts with contextual material such as information on persons and places mentioned in the texts.

Review

1 NietzscheSource describes itself as a 'web site devoted to the publication of scholarly content on the work and life of Friedrich Nietzsche'. So far, two editions of Nietzsche's work have been made available on NietzscheSource: the *Digitale Kritische Gesamtausgabe Werke und Briefe* (eKGWB) and the *Digitale Faksimile Gesamtausgabe* (DFGA) of the *Nietzsche Estate*. The eKGWB is based on the latest critical print edition of Nietzsche's works, the *Gesamtausgabe* edited by Giorgio Colli and Mazzino Montinari (Colli-Montinari) and follows both the text as well as the categorisation of Nietzsche's writings established by that edition, which are the following:

- Published works (18 volumes in the eKGWB)
- Private publication (3 volumes)
- Authorised manuscripts (3 documents)
- Posthumous writings (8 documents)

Social

ToC

Abstract

1. Review

2. References

Meta

Last updated: June 6, 2016 at 13:12 pm

Print
 PDF

XML of this review, including formal evaluation data

Factsheet of this project

Citation suggestion

Tags

collected works (1)

Illuminated charter



What is common to all these materials?

- Identification information
e.g. shelfmark, inventory number, page number, titles...
- Divisions and subdivisions
- Pictures, diagrams, some kind of graphical information
- A number of writing modes or registers
e.g. prose, verse, drama...
- With formal structural units
e.g. paragraphs, lists, stanzas, lines, speeches...
- Containing textual distinctions (sometimes signalled by a change in rendition)
e.g. titles, headings, quotes, names...
- Metatextual indications/interventions
e.g. deletions, additions, annotations, revisions...

The TEI core module can cope with this and more phenomena!

TEI document structur(s)

- A TEI document is represented by means of the root element `<TEI>` which contains both, data and metadata
- A sequence of `<TEI>` elements may be combined to form a `<teiCorpus>` element.

The document may be of any size, form and complexity, from a postcard to a multi-volume encyclopedia

Basic <TEI> Structure

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">  
  <teiHeader>  
    <!-- required: the document's metadata -->  
  </teiHeader>  
  <facsimile>  
    <!-- optional: facsimile images -->  
  </facsimile>  
  <sourceDoc>  
    <!-- optional: documentary transcription -->  
  </sourceDoc>  
  <text>  
    <!-- required if neither facsimile nor sourceDoc: the content-  
->  
  </text>  
</TEI>
```

Basic <teiCorpus> Structure

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">  
  <teiHeader>  
    <!-- required -->  
  </teiHeader>  
  <TEI>  
    <!-- required -->  
  </TEI>  
  <!-- more <TEI> elements for each document -->  
</teiCorpus>
```

Each <TEI> element could represent
a collection of encoded texts,
versions of a text,
or samples of a language corpora, etc.

Unitary and composite <text>

A <text> may be

- *Unitary*, forming an organic whole
- *Composite*, consisting of several components which are in some important sense independent of each other

A unitary text contains:

- <front> optional, contains any prefatory matter, found at the start of a document (titlepage, preface, etc.)
- <body> mandatory, contains the whole body of a single text
- <back> optional, back matter containing appendixes following the main part

Unitary <text> forming an organic whole

```
<text>
  <front>
    <!-- optional front matter -->
  </front>
  <body>
    <!-- Either divisions and/or paragraphs -->
    <div n="1">
      <p><!-- paragraph --></p>
    </div>
    <div n="2">
      <p><!-- paragraph --></p>
    </div>
  </body>
  <back>
    <!-- optional back matter -->
  </back>
</text>
```

Composite `<text>`

A composite text contains:

- `<front>` optional, contains any prefatory matter relating to the composite
- `<group>` contains at least one text, grouping together distinct texts
- `<back>` optional, back matter relating to the composite

A group may contain sub-groups represented by nested `<group>` elements.

Composite <text> of several components

```
<text>
  <front>
    <!-- front matter of composite text -->
  </front>
  <group>
    <text n="1">
      <front>
        <!-- front matter of first unitary text -->
      </front>
      <body>
        <!-- body of first unitary text-->
      </body>
      <back>
        <!-- back matter of first unitary text -->
      </back>
    </text>
    <text n="2">
      <body>
        <!-- body of second unitary text -->
      </body>
    </text>
  </group>
  <back>
    <!-- back matter of composite text -->
  </back>
</text>
```

Front Matter

The front matter `<front>` represents distinct sections of a text (e.g. title pages, preface).

- **'preface'** a foreword or preface addressed to the reader
- **'ack'** a declaration of acknowledgement by the author
- **'dedication'** a formal dedication to one or more persons
- **'abstract'** a summary of the content
- **'contents'** a table of contents
- **'frontispiece'** pictorial frontispiece, possibly including a text

Title Pages

For detailed analysis of title pages and other preliminaries of older printed books and manuscripts (selection):

- `<titlePage>` the title page of a text, appearing within the front or back matter
- `<docTitle>` the title of the document
- `<titlePart>` subsection or division of the title of a work
 - `@type` specifies the role: e.g. main, sub, alt, short, desc
- `<docAuthor>` the name of the author
- `<docImprint>` the imprint statement (place, date, publisher)
- `<docDate>` the date of the document

Title Pages

```

<front>
  <titlePage>
    <docTitle>
      <titlePart type="main">Is There a Text in This Class?</titlePart>
      <titlePart type="sub">The Authority of Interpretive
Communities</titlePart>
    </docTitle>
    <docAuthor>Stanley Fish</docAuthor>
    <docImprint>
      <publisher>Harvard University Press</publisher>
      <pubPlace>Cambridge, Massachusetts</pubPlace>
      <pubPlace>London, England</pubPlace>
    </docImprint>
  </titlePage>
</front>

```

*Is There a Text
in This Class?*

The Authority of
Interpretive Communities

Stanley Fish

HARVARD UNIVERSITY PRESS
Cambridge, Massachusetts
London, England

Back Matter

The back matter `<back>` contains any appendixes following the main part of the text.

- 'appendix' an ancillary self-contained section of a work
- 'glossary' a list of terms associated with definition texts (`<list type="gloss">`)
- 'notes' a section where notes are gathered together
- 'bibliogr' list of bibliographic citations (`<listBibl>`)
- 'index' any form of index of the work
- 'colophon' statement describing the physical production of the work

Back Matter

```
<back>
  <div type="index">
    <head>Index</head>
    <list type="index">
      <item>Africa, cause assigned for the barbarous state of the interior
        parts of that continent, <ref target="#P125">125</ref></item>
      <item>Actors, public, paid for the contempt attending their profession,
        <ref target="#P263">263</ref></item>
    </list>
  </div>
</back>
```

Global Attributes

- Some features (potentially) apply to everything, therefore members of the attribute class `att.global` can appear in every TEI element:
 - `@xml:id` provides a unique identifier for any element
 - `@n` provides a number or name for an element (not unique)
 - `@xml:base` provides a base URI reference for resolving relative URIs.
 - `@xml:lang` specifies the language of any element, using an ISO standard code (e.g. ISO 639-1)
 - `@xml:space` specifies how whitespace should be managed by applications
- `@rend`, `@style` and `@rendition` provide ways of specifying the visual appearance (rendition) of any element (`att.global.rendition`)
- `@resp` points to the agency responsible for the intervention or interpretation (`att.global.responsibility`)

Inside the `<body>`: divisions

Hierarchical grouping of text sequences into textual divisions and subdivisions by means of nested `<div>` elements.

- Use of the `@type` attribute to distinguish different kinds of divisions
 - Epic, Bible → `book`
 - Report → `part`, `section`
 - Novel → `chapter`
 - Drama → `acts`, `scenes`
 - Reference book → `sections`
 - Diary → `entries`
 - Newspaper → `issues`
 - ...
- and possibly `@n` to provide a name or number of any kind:

Divisions

```
<body>
  <div type="part" n="1">
    <div type="chapter" n="1">
      <!-- text of part 1, chapter 1 -->
    </div>
    <div type="chapter" n="2">
      <!-- text of part 1, chapter 2 -->
    </div>
  </div>
  <div type="part" n="2">
    <div n="1" type="chapter">
      <!-- text of part 2, chapter 1 -->
    </div>
    <div n="2" type="chapter">
      <!-- text of part 2, chapter 2 -->
    </div>
  </div>
</body>
```

Components of a division

What do divisions contain (apart from other divisions)?

- *Headings*, tagged with `<head>`
- *Prose*, which may be organized as a sequence of paragraphs `<p>`.
- *Poetry*, divided into metrical lines `<l>`, optionally grouped into stanzas `<lg>`.
- *Drama*, divided into speeches `<sp>`, containing an optional speaker label `<speaker>`, followed by a mix of `<p>` or `<l>` elements, optionally mixed up with stage directions `<stage>`.

Original layout information

Within the `<text>` element the logical view is privileged, but the physical view can be encoded as well through empty elements:

- `<pb>` marks the start of a new page
- `<cb>` marks the start of a new column
- `<lb>` marks the start of a new line

Module 3:

Elements Available in all TEI documents

The so called core module describes elements which may appear in any kind of text

- paragraphs
- highlighting, emphasis and quotation
- simple editorial changes
- basic names, numbers, addresses and dates
- links and cross references
- graphics
- lists
- notes
- bibliography

Paragraph

A paragraph is a significant organizational unit for all prose texts

- `<p>` marks paragraphs in prose
- `<p>` can contain all the phrase-level elements in the core module
 - *Phrase-level* elements must be entirely contained within a paragraph
 - *Inter-level* elements can appear either within a paragraph or between paragraphs (e.g. list, bibliographic citations, etc.)
 - *Chunks* (eg. paragraphs, anonymous block)

```

<body>
  <p>
    <lb />Loveinge Contreyman I am bolde of yow as of a ffrende, craveinge
    <lb />yowr helpe with xxxli vppon Mr Bushells &amp; my securytee or Mr Myttons
    <lb />wth me Mr Rosswell is nott come to London as yeate &amp; I have
    <lb />especiall cawse, yow shall ffrende me muche in helpeinge me
    ...
  </p>
</body>

```

Highlighting

Typographic features in order to distinguish passages from its surroundings:

- distinct in some way (e.g. foreign, dialectal, technical, etc.)
- emphatic or stressed when spoken
- not part of the body of the text (e.g. title, head, label, etc.)
- distinct narrative stream (e.g. monologue, commentary, etc.)
- attributed by the narrator to some other agency (e.g. direct speech, quotation, etc.)
- set apart from the text in some other way (e.g. individual names in older texts, editorial corrections or additions, etc.)

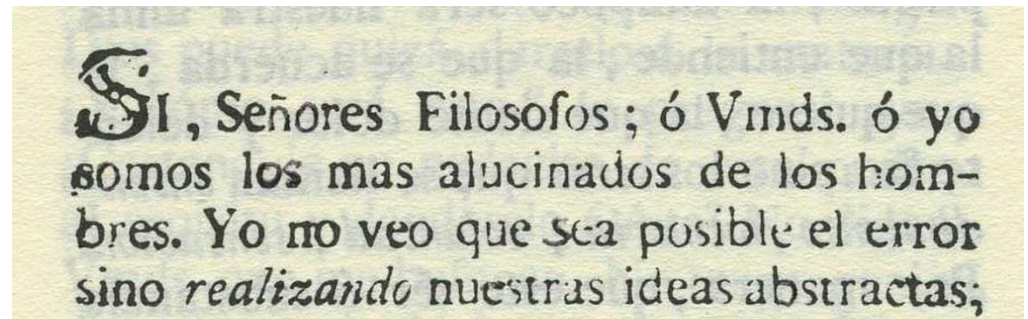
Highlighting

- `<hi>` word or phrase which is graphically distinct from the surrounding text
 - `@rend` specifies the visual appearance; the value is defined by each project
 - `@style`, `@rendition` renditions using external standards, like CSS

`<p>`

`<hi rend="initial">S</hi>I, Señores Filósofos; ó Vmds. ó yo somos los mas alucinados de los hombres. Yo no veo que sea posible el error sino <hi rend="italic">realizando</hi> nuestras ideas abstractas;`

`</p>`



Foreign Phrases

- `<foreign>` word or phrase not written in the same language than the surrounding text
 - `@xml:lang` global attribute to specify the language, using an ISO standard code (e.g. ISO 639-1)

John eats a `<foreign xml:lang="fr">`croissant`</foreign>` every morning.

Emphasis

- `<emph>` words or phrases which are emphasized for linguistic or rhetorical effect
 - `@rend`, `@rendition` and `@style`

`<p>`To Sherlock Holmes she is always
`<emph>`the`</emph>` woman. ... `</p>`

Quotation

The TEI distinguishes a variety of ,distinct' text enclosed in quotation marks:

- `<q>` separated from the surrounding text with quotation marks, e.g. direct speech, technical term, slang etc.
- `<said>` passages thought or spoken aloud
 - `@direct` direct or indirect speech
 - `@aloud` vocalized or signed speech
- `<quote>` passages attributed to an external source
- `<cit>` quotation from some other document, together with a bibliographic reference

Quotation

```
<cit>  
  <quote>  
    <l>To be, or not to be: that is the question</l>  
  </quote>  
  <bibl>  
    <title>Hamlet</title>  
    <author>William Shakespeare</author>  
  </bibl>  
</cit>
```

Simple Editorial Changes

The core module provides some phrase-level elements which may be used to record simple editorial interventions.

- `<choice>` groups alternative encodings for the same point in a text

Abbreviations:

- `<abbr>` abbreviated form
- `<expan>` expanded form
- Errors:
- `<sic>` apparent error
- `<corr>` corrected error

Regularization:

- `<orig>` original form
- `<reg>` regularized form

Abbreviation and expansion

- `<choice>` groups alternative encodings
- `<abbr>` abbreviated form
- `<expan>` expanded form

```
<choice>
  <abbr>Mr. </abbr>
  <expan>Mister</expan>
</choice>
```

```
<choice>
  <abbr>añ</abbr>
  <expan>ante</expan>
</choice>
```

añ	=	ante
ā	=	aut
xp̄s	=	Christus
c̄	=	cum
d̄s	=	deus
diḡ	=	dixit
dñs	=	dominus
ej	=	eius
o	=	ergo
g	=	
.	=	

Emendation and correction

- `<choice>` groups alternative encodings
- `<sic>` apparent error in the copy text
- `<corr>` corrected error

I don't know, Juan. It's so far in the past now — how

`<choice>`

`<sic>`we can`</sic>`

`<corr>`can we`</corr>`

`</choice>` prove or disprove anyone's theories?

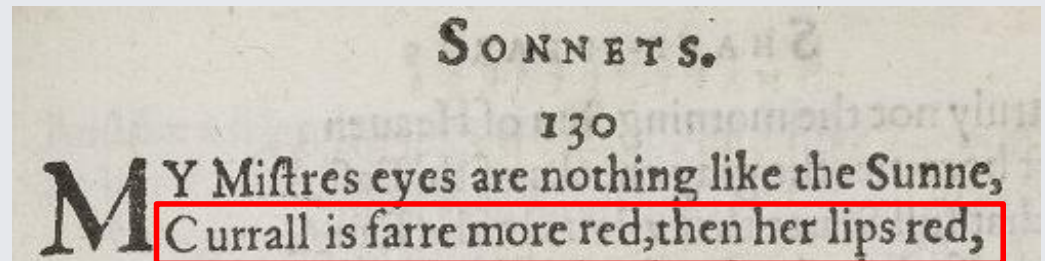
Regularization

- `<choice>` groups alternative encodings
- `<orig>` original form
- `<reg>` regularized form

```

</>
  <choice>
    <orig>Currall</orig>
    <reg>Coral</reg>
  </choice> is
  <choice>
    <orig>farre</orig>
    <reg>far</reg>
  </choice> more red,
  <choice>
    <orig>then</orig>
    <reg>than</reg>
  </choice> her lips red
</l>

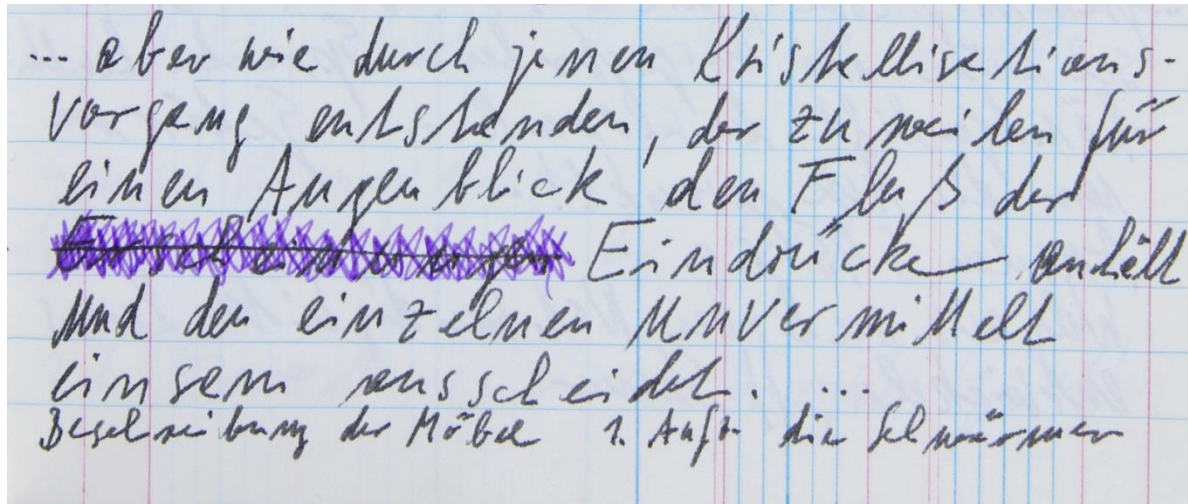
```



Additions, Deletions, and Omissions

- `<add>` addition to the text
 - `` letter, word or phrase marked as deleted in the text
 - `<unclear>` illegible or inaudible passage which cannot be read with confidence
 - `<gap>` indicates a point where material is omitted
-
- *Note: the transcr module (11) provides more elements for the representation of primary sources*

Simple deletion



</>...aber wie durch jenen Kristallisations-</>
 </>vorgang entstanden, der zuweilen für</>
 </>einen Augenblick den Fluß der</>
 </>Erscheinungen Eindrücke enthält</>
 </>...</>

Names

- `<name>` a proper noun or noun phrase
- `<rs>` a string referring to some person, place, object, etc.
- `@type` attribute specifies the type of the name in more detail

```
<p>
```

```
  <q>My dear <name type="person">Mr.  
Bennet</name>,</q> said <rs type="person">his lady</rs>  
to him one day,
```

```
  <q>have you heard that <name type="place">Netherfield  
Park</name> is let at last?</q>
```

```
</p>
```

Addresses

Elements to distinguish postal and electronic addresses

- `<address>` contains a postal address
- `<email>` contains an email address
- `<addrLine>` a non-specific address line
- `<street>` a full street address
- `<postCode>` a postal or zip code
- `<postBox>` a postal box number
- `<name>` can also be used within address



Addresses

```
<address>  
  <name type="city">The Shakespeare Centre</name>  
  <street>Henley Street</street>  
  <postCode>CV37 6QW</postCode>  
  <name type="city">Stratford-upon-Avon</settlement>  
  <name type="country">United Kingdom</country>  
</address>
```

```
<email>info@shakespeare.org.uk</email>
```

Namesdates module for more detailed geographical information

```
<address>  
  <name>The Shakespeare Centre</name>  
  <street>Henley Street</street>  
  <postCode>CV37 6QW</postCode>  
  <settlement>Stratford-upon-Avon</settlement>  
  <country>United Kingdom</country>  
</address>
```

Numbers and Measures

- `<num>` a number of any sort, written in any form
 - `@type` and `@value`
- `<measure>` marks a quantity and/or commodity
 - `@type`, `@unit`, `@quantity`, `@commodity`
- `<measureGrp>` a groups of dimensional specifications

```
<num type="percentage" value="10">ten percent</num>
```

```
<measureGrp type="volume" unit="in">  
  <measure type="height" quantity="14">xiv</measure>  
  <measure type="width" quantity="5">v</measure>  
  <measure type="depth" quantity="10">x</measure>  
</measureGrp>
```

Date & Time

- `<date>` contains a date in any format
 - `@when` contains the regularized form; JJJJ-MM-DD
 - `@calendar` to specify the calendar system
- `<time>` contains a time of day in any format
 - `@when` contains the regularized form: HH:MM:SS

```
<date when="1616">The year 1616</date>
```

```
<date when="1616-04-23">23 April 1616</date>
```

```
<date when="--04">April</date>
```

```
<time when="08:48:00">8:48</time>
```

```
<time when="08:45:00">a quarter to nine</time>
```

Links and Cross References

- `<ptr>` defines a pointer to another location
- `<ref>` defines a reference to another location with an optional linking text
- `@target` taking a URI reference

See `<ptr target="#Section13"/>`.

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

The programm can be found at the

`<ref target="http://digital.humanities.ox.ac.uk/dhoxss/2016">digital.humanities@oxford website</ref>`.

Lists

- `<list>` a sequence of items organized as a list
- `<item>` contains one component of a list
- `<label>` label or heading associated with an item

```
<list>  
  <label>(1)</label>  
  <item>Romeo and Juliet</item>  
  <label>(2)</label>  
  <item>Timon of Athens</item>  
  <label>(3)</label>  
  <item>Julius Caesar</item>  
</list>
```



Notes

Encoding explanatory notes, whether already present in the copy text or supplied by the encoder:

- `<note>` contains a note or annotation
 - `@place` indicates the physical location of the note (e.g. margin, foot, etc.)
 - `@type` to categorize the note (e.g. gloss)

```
<lg>  
<l>The self-same moment I could pray</l>  
<l>And from my neck so free</l>  
<l>The albatross fell off, and sank</l>  
<l>Like lead into the sea.</l>  
<note type="gloss" place="margin">The spell begins to break</note>  
</lg>
```

Index

An index provides access to the content via keywords and phrases which are not necessarily present in the text itself. For converting a *pre-existing index*, nested lists are recommended; *Auto-generated indices* are represented with the help of `<index>`.

- `<index>` marks a location to be indexed
 - `@indexName` to specify which index the entry belongs to
- `<term>` marks a term inside the index element

```
<p>William Shakespeare <index indexName="INDEX-PERSONS">  
  <term>Shakespeare, William</term>  
</index> was born in <index indexName="INDEX-PLACES">  
  <term>Stratford-upon-Avon</term>  
</index> Stratford...</p>
```

Graphics

- `<graphic>` location of an inline graphic, illustration or figure
- `<binaryObject>` binary data embedding graphics or other objects

```
<graphic url="materials/notebook_01-r01.jpg" />
```

Note: the figure module (14) provides `<figure>` and `<figDesc>` for encoding more complex graphical components.

Graphics

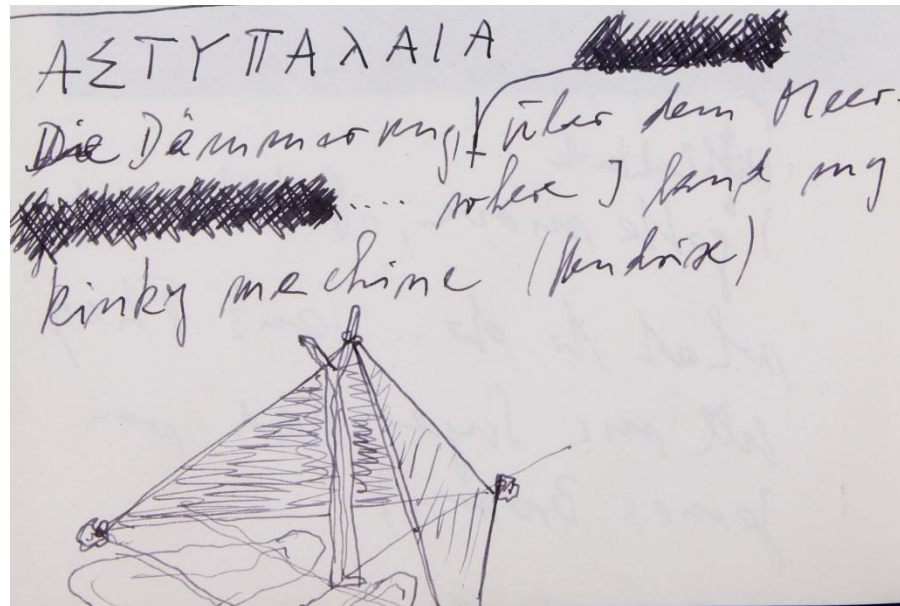
<|> ΑΣΤΥΠΑΛΛΑΙΑ</|>

<|>Die Dämmerung, über dem Meer .. </|>

<|> where I land my </|>

<|>kinky machine (<name type="person">Hendrix</name>)

<graphic url="cubehouse-001.jpg" />



Bibliography

- `<bibl>` a structured or unstructured bibliographic reference
- `<title>`, `<editor>`, `<title>`, `<pubPlace>`, `<publisher>`, `<date>`, etc. for further structuring
- Presentational conventions (punctuation, order, layout, etc.) do not have to be considered; it is a matter of further processing

```
<cit>  
  <quote>Soft morning City!</quote> (<bibl>Joyce, FW</bibl>)  
</cit>
```

```
<bibl xml:id="Joyce_1939">  
  <author>Joyce, James</author>  
  <title>Finnegans Wake</title>  
  <pubPlace>London</pubPlace>  
  <publisher>Faber & Faber</publisher>  
  <date>1939</date>  
</bibl>
```

Core Tags for Verse

- `<lg>` a formal unit (e.g. stanza) containing one or more verse lines
- `<l>` contains a single verse line

```
<lg type="quatrain">  
  <l>Shall I compare thee to a Summer's day?</l>  
  <l>Thou art more louely and more temperate:</l>  
  <l>Rough windes do shake the darling buds of Maie,</l>  
  <l>And Sommers lease hath all too short a date:</l>  
</lg>
```

Note: the verse module (6) provides more elements to represent metrical systems.

Core Tags for Drama

- `<sp>` an individual speech in a performance text, or passages presented as such in prose or verse text
- `<speaker>` provides the name of one or more speakers in a dramatic text
- `<stage>` provides stage directions within a dramatic text

`<stage>`Enter the King, Lord Iohn of Lancaster, Earle of Westmerland, with others.`</stage>`

`<sp>`

`<speaker>`King.`</speaker>`

`<l>`SO shaken as we are, so wan with care,`</l>`

`<l>`Find we a time for frighted peace to pant`</l>`

...

`</sp>`

- *Note: the drama module (7) provides more elements to represent drama texts, screen plays, etc.*

Core Elements

