

<philtag 5> TEI P5: Encoding dictionaries & more

A Summary by Christian Graiger

November 17, 2006

Abstract

Summary of the TEI workshop session with Sebastian Rahtz and Laurent Romary about encoding dictionaries.

1 Motivation

The discussion on how to encode dictionaries was based on the first approach of encoding the entry “Der Lachbaum” taken from the Campe dictionary (see Listing 11). In this approach the entry was encoded by using the following TEI modules *without* changing XML element assignments:

1. `dictionaries` for printed dictionaries,
2. `core` that contains elements common to all TEI documents,
3. `header`, the TEI header,
4. `analysis` containing simple analytic mechanisms,
5. `linking` for linking, segmentation and alignment.

The general difficulties of using a generated schema that contains these modules for encoding a dictionary were outlined and discussed with *Laurent Romary* and *Sebastian Rahtz*. As a result of this discussion a new schema was generated for encoding the entry more suitably. In particular, critical XML elements were assigned to new element classes to adopt the schema for the needs of encoding entries. Along with generating the new schema, the encoding of “Der Lachbaum” was redesigned (see Listing 12, 13).

2 Using ROMA to generate a validators for the TEI

Generating a schema for validating TEI XML documents can be achieved easily by using the ROMA generator [ROM06]. A web browser is all that is necessary:

1. Go to the ROMA URL,
2. select Create a new customization and submit your request.
3. Click the Schema link and select your favoured schema, e.g. Relax NG schema (compact syntax) from the drop-down menu.

4. Submit your request and download the schema.
5. Optionally, click Save Customization for saving the customization that generated the schema above. Submit your request to download the customization.

After downloading the schema, link it to your TEI document and use it for validating the document.

2.1 Understanding ROMA

Steps 1 to 5 describe a very fast approach to obtain a *general* schema for validating TEI documents. However, they don't describe the purpose for which ROMA was designed nor how to use ROMA properly for encoding dictionaries or encoding other documents using TEI P5.

Organisation of TEI elements

The TEI encoding scheme consists of a number of *modules*, each module declares particular XML elements and their attributes for encoding text. XML elements are organised in *model classes*. A class might be considered as a set of elements for encoding particular characteristics of text. [TEI05]

The declaration of an element in a module includes the assignment of the element to one or more classes. Element declarations can be excluded from a module and cannot be assigned to another module. However, elements can be assigned to other classes. By doing so, modules and their *contents* might be combined freely to form a schema for your requirements. [TEI05]

Equivalent to XML elements and model classes, attribute declarations for elements are assigned to *attribute classes*.

Example. For encoding a dictionary the module `dictionaries` is required which declares necessary XML elements such as `<entry>` or `<form>`. The `<entry>` element itself is assigned to the model class `model.entryLike` which has the additional members `<entryFree>` and `<superEntry>`. As content of `<entry>`, the elements `<hom>`, `<sense>` and elements from `model.entryParts.top` are defined.

Representation of TEI elements in ROMA

ROMA provides two possibilities to obtain information about XML elements defined in a module:

1. Select Modules and click on a particular module in the list of modules. You will see a list of all elements declared in this module. By selecting an element, all information about the element is displayed.
2. The other option is to search an element in the TEI *database*. An input field is provided at the bottom of each page in ROMA. Type the element you are searching for and click the Search TEI database button. A new page will display the information.

Information about an element is divided into four parts: the first part Description provides information about the element that is comprehensible to the user. The next

parts Module classes and Attribute classes respectively display assignments of the element or its attributes to classes. The last part Contents provides detailed information about the contents of the element in *Relax NG* notation.

2.2 Customising schemas in ROMA

To generate a schema suitable for your needs, select Generate a new customisation on the start page of ROMA and submit your request. Click the Modules link and select required modules by clicking the add or remove link which is displayed before the name of each module.

The same mechanism is provided for elements. If you click the modules' name, you will get a list of all elements declared in this module. Elements might be *included* or *excluded* from the declaration by using the radio buttons next to an elements' name. Any changes you have made will be assigned if you click the Submit button at the end of the element list.

Changing element assignments

Click an elements' name in the list of elements to display detailed information about the element. By using the check boxes you might change assignments of the element to module or attribute classes.

Example. Select the element `<head>` in the module `core` from the list of elements. If you want to use `<head>` within an `<entry>` element, enable the checkbox referring to `model.entryParts.top` in **Module classes** and submit your request. As the content of `<entry>` is assigned to `model.entryParts.top`, the element `<head>` can be used now within `<entry>` declarations.

Changing attributes to an element

In a modules' list of elements you might use the link Change attributes to *add* new attributes to an element. Also, you might *include* or *exclude* attributes. If you click an attribute's name, a more detailed information about the attribute is provided. You might declare the attribute as optional or required, or assign particular values to the attribute. After submitting your request, any changes made to an attribute will be assigned to your customization.

Handling customizations

Once your adaptations are complete, save your schema as described in step 4 above. It is highly recommended to save your customization as well (step 5). The customization will be stored as an ODD XML file, where the abbreviation ODD stands for *One Document Does it all* [TEI06]. If you want to change adaptations at a later time, just upload your customization by selecting the radio button Upload an existing customization and your customization file in the file upload field at the start page of ROMA. After uploading the customization, change it to your needs and finally save the newly generated schema and customization.

3 Encoding the entry “Der Lachbaum”

3.1 Encoding of the TEI document

Basically, changes in the encoding of “Der Lachbaum” were made in the TEI header description, the *lemma* encoding including inflected forms (<form> element) and in <sense> tags.

Changes in the TEI header

Listing 1 shows the original header information provided in the first encoding approach.

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Campe - Exemplarisches Wörterbuch</title>
10    <author>Christian Graiger, Mirjam Blümm</author>
    </titleStmt>
    <publicationStmt>
      <p>Campe Wörterbuch, Beispiel der Kodierung bishin zur Wortebene</p>
    </publicationStmt>
15    <sourceDesc>
      <p>Wörterbucheintrag "Der Lachbaum" aus Campe L-R, 1809</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

Listing 1: Original TEI header

In listing 2 a more suitable header information is given that describes the state of the document exactly: To the <publicationStmt> element the elements <availability> and <authority> were added. Also <revisionDesc> was introduced for marking up changes in the encoding.

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Campe - Exemplarisches Wörterbuch</title>
8    <author>Christian Graiger, Mirjam Blümm</author>
    </titleStmt>
    <publicationStmt>
      <availability><p>nicht nutzbar</p></availability>
      <authority>KfE UW</authority>
13    </publicationStmt>
    <sourceDesc>
      <p>Wörterbucheintrag "Der Lachbaum" aus Campe L-R, 1809</p>
    </sourceDesc>
18    </fileDesc>
    <revisionDesc>
      <change who="#LR">Der Lachbaum 3rd approach</change>
      <change who="#WW">Texterfassungsform Datum</change>
      <change who="#WW">Erstkodierung Datum</change>
      <change who="#WW">Textkontrolle Datum</change>
23    <change who="#WW">Übernahme in Bandkodierung</change>
    </revisionDesc>
  </teiHeader>
```

Listing 2: New TEI header

Changes for encoding the dictionaries structure

Note that the structure in Listing 3 was used for encoding an entry in a *single* TEI document. Therefore the element <div> was attributed by *type* and *xml:id*.

```

<text>
  <body>
    <div type="dictionary" xml:id="lachbaum" xml:lang="de" rend="unmarked">
      <entry>

```

Listing 3: Original dictionary structure

For encoding more entries in a TEI document, a structure as shown in Listing 4 was suggested: `<div>` elements should be used for grouping particular blocks in a dictionary, e.g. entries starting with the common first letter. Such `<div>` elements might include `<head>` elements which label the common first letter itself. Each `<entry>` within a `<div>` element must be attributed by `xml:id`.

```

<text>
  <body>
    <div n="L" type="groupOfEntriesGroupedByCharacter"
      xml:id="letter-l">
      <head n="L"/>
34    <entry xml:id="lachbaum">

```

Listing 4: New dictionary structure

Changes in encoding the lemma

Listing 5 shows the original encoding of the lemma. The `<form>` element was rearranged to encode the lemma in a more sophisticated way (Listing 6):

- The `<form>` element with attribute `type` and value `"lemma"` was introduced for encoding the *entire* lemma. Its child element `<gramGrp>` was used for determining the gender of the lemma, marked up by the `<gen>` element.
- the `<gram>` element which encoded the article used in the lemma was replaced by another `<form>` element with attribute `type` and its value `"determiner"`. Within this element, `<orth>` was used to encode the article itself.
- Analogically the headword was encoded, using a `<form>` element with attribute `type` and value `"headword"`.
- `<gramGrp>` was replaced by two `<form>` elements with attribute `type` and value `"inflected"` to encode the inflected forms of the headword, one for genitive singular, the other for nominative plural. For both elements `<gramGrp>` and `<orth>` were used to encode the inflection details.

```

39    <form type="headword">
      <gram type="art_m">Der</gram>
      <orth>Lachbaum<c type=",$", ">,</c></orth>
    </form>
    <gramGrp>
      <number value="sing"/>
      <case value="gen">
        <oVar type="gen"><oRef/>-es</oVar>
        <c type=",$", ">,</c>
44    </case>
      <number value="plur">Mz.</number>
      <case value="nom">
        <oVar type="pl"><oRef/>-bäume</oVar>
        <c type=",$", ">,</c>
49    </case>
    </gramGrp>
    <usg type="dom">im Forstwesen<c type=",$", ">.</c></usg>

```

Listing 5: Original encoding of the lemma

```

42 <form type="lemma">
    <gramGrp>
        <gen value="m"/>
    </gramGrp>
    <form type="determiner">
        <orth>Der</orth>
    </form>
    <form type="headword">
47     <orth>Lachbaum</orth><c type="$,"></c>
    </form>
    <form type="inflected">
        <orth>
52         <oVar><oRef/>-es</oVar>
        </orth><c type="$,"></c>
        <gramGrp>
            <case value="G"/>
        </gramGrp>
57    </form>
    <form type="inflected">
        <gramGrp>
            <number value="p">Mz.</number>
        </gramGrp>
62    <orth>
        <oVar type="pl"><oRef/>-bäume</oVar>
        </orth><c type="$,"></c>
    </form>
    <usg type="dom">im Forstwesen<c type="$. "></c></usg>

```

Listing 6: New encoding of the lemma including inflected forms

Changes in encoding the semantic description

Listing 7 shows the original encoding of the semantical description and its parts. Different parts are encoded using `<C>` elements. These parts were replaced by `<head>` elements as in listing 8.

```

    <sense>
        <sense n="1" level="1" rend="numbered">
            <c type="ordnungselement">1</c>
            <def>Ein an sei-<lb/>ner Borke aufgehauener Nadelholzbaum<c
56             type="$,"
            ></c> aus welchem man Harz<lb/> rinnen läßt<c type="$
            . "></c></def>
        </sense>
        <sense n="2" level="1" rend="numbered">
            <c type="ordnungselement">2</c>
            <def>ein Baum<c type="$,"></c> welchen man auf der Grenze
61             des
            Wal-<lb/>des neben den Grenzsteinen stehen läßt und auf
            zwei Seiten
            mit drei<lb/> Hieben bezeichnet<c type="$,"></c> um den
            Grenzstein
            desto besser finden zu können<c type="$. "></c><lb/></def>
            <dictScrap>
                <lbl>nach Popowitsch besser</lbl>
            </dictScrap>
66            <xr type="syn">
                <ref target="#Laßbaum">der Laßbaum<c type="$,"></c></ref>
            </xr>
            <def>weil man ihn stehen läßt<c type="$. "></c><lb/></def>
71            <def>Dann überhaupt ein jeder Grenzbaum<c type="$,"></c> in
            welchen ein
            Merkmal zur<lb/> Auszeichnung eingehauen worden ist<c
            type="$. "></c>
            </def>
        </sense>
        <xr type="syn">
76            <ref target="#Mahlbaum">der Mahlbaum</ref>
            <c type="$,"></c>
            <ref target="#Markbaum">Markbaum</ref>

```

```

      <c type="$,">,</c>
      <lb/>
81      <ref target="#Grenzbaum">Grenzbaum</ref>
      <c type="$,">,</c>
      <lbl>auch</lbl>
      <ref target="#Lochbaum">Lochbaum</ref>
      <lbl>und</lbl>
86      <ref target="#Lachterbaum">Lachterbaum</ref>
      <c type="$.">.</c><lb/>
    </xr>
  </sense>

```

Listing 7: Original encoding of the description

Additionally `<name>` was assigned for marking up persons and names instead of `<dictScrap>`. The attributes `rend` were used to mark up particular text formatting. Also the attribute `type` was introduced for `<def>` elements to provide a more detailed information about definitions in an entry.

```

<sense>
  <sense n="1" rend="numbered">
    <head>1</head>
    <def type="paraphrase">Ein an sei-<lb/>ner Borke aufgehauener
71   Nadelholzbaum<c type="pun">,</c> aus welchem man
      Harz<lb/> rinnen läßt<c type="pun">,</c></def>
  </sense>
  <sense n="2" rend="numbered">
    <head>2</head>
76   <def type="paraphrase">ein Baum<c type="pun">,</c> welchen man auf
      der
      Grenze des Wal-<lb/>des neben den Grenzsteinen stehen
      läßt und auf zwei Seiten mit drei<lb/> Hieben
      bezeichnet<c type="pun">,</c> um den Grenzstein desto
      besser finden zu können<c type="pun">,</c><lb/>
81   nach <name type="person" rend="spaced">Popowitsch</name>
      besser
      <xr>
        <ref target="#Laßbaum">der Laßbaum</ref><c type="pun">,</c>
      </xr> weil man ihn stehen läßt<c type="pun">.</c><lb/>
86   Dann überhaupt ein jeder Grenzbaum<c type="pun">,</c>
      in welchen ein Merkmahl zur<lb/> Auszeichnung
      eingehauen worden ist<c type="pun">,</c>
    </def>
  </sense>
91   <xr type="syn">
    <ref target="#Mahlbaum" rend="0,5+">der Mahlbaum</ref>
    <c type="$,">,</c>
    <ref target="#Markbaum" rend="0,5+">Markbaum</ref>
    <c type="$,">,</c>
96   <lb/>
    <ref target="#Grenzbaum" rend="0,5+">Grenzbaum</ref>
    <c type="$,">,</c>
    <lbl>auch</lbl>
    <ref target="#Lochbaum" rend="0,5+">Lochbaum</ref>
101  <lbl>und</lbl>
    <ref target="#Lachterbaum" rend="0,5+">Lachterbaum</ref>
    <c type="$.">.</c>
    <lb/>
  </xr>
106 </sense>

```

Listing 8: New encoding of the description

Additional changes

There were some more changes for encoding an entry: First, `<revisionDescription>` was assigned to be allowed within `<entry>` elements (listing 9). Second, `<ref>` was also assigned to `<entry>` elements to enable linking. This assignment, however, was

removed later. Instead, the `<xr>` element was reassigned to provide linking functionality. Last, the module linking was removed from the customization completely.

```
<revisionDesc>
  <change who="#LR">Der Lachbaum 2nd approach</change>
</revisionDesc>
```

Listing 9: revision description in an entry

3.2 Adaptation of the Relax NG schema

The customization of the newly generated schema is given in listing 14. It covers the deletion of several unnecessary elements in the modules `analysis`, `core` and `textstructure`.

In module header the element `<revisionDesc>` was assigned to class `model.entryParts.top`. The elements `<cit>` and `<head>` from module `core` were also assigned to this class.

Following assignments changed in module dictionaries:

- `<gramGrp>` was added to class `model.formPart`,
- `model.global` was assigned to `<xr>`,
- `<lbl>` was added to class `model.entryParts.top`,
- `<def>` was complemented by the new attribute `type`.

The element `<def>` was changed by overwriting its *content* directly:

```
<content xmlns:rng="http://relaxng.org/ns/structure/1.0"
  xmlns="http://www.tei-c.org/ns/1.0">
  <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
    name="macro.paraContent"/>
5  <rng:attribute xmlns:rng="http://relaxng.org/ns/structure/1.0" name="type">
  <rng:data xmlns:rng="http://relaxng.org/ns/structure/1.0" type="QName"/>
  </rng:attribute>
</content>
```

Listing 10: content of `<def>`

A Listings for encoding the dictionary entry "Der Lachbaum"

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.tei-c.org/ns/1.0
5  ../CampeExamples/dictionary2.xsd"
  xmlns="http://www.tei-c.org/ns/1.0" xml:lang="de">
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>Campe - Exemplarisches Wörterbuch</title>
10      <author>Christian Graiger, Mirjam Blümm</author>
      </titleStmt>
      <publicationStmt>
        <p>Campe Wörterbuch, Beispiel der Kodierung bishin zur Wortebene</p>
      </publicationStmt>
15      <sourceDesc>
        <p>Wörterbucheintrag "Der Lachbaum" aus Campe L-R, 1809</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
```



```

20 <!--
Der Lachbaum, -es, Mz. -bäume, im Forstwesen. 1) Ein an sei-
ner Borke aufgehauener Nadelholzbaum, aus welchem man Harz
rinnen läßt; 2) ein Baum, welchen man auf der Grenze des Wal-
des neben den Grenzsteinen stehen läßt und auf zwei Seiten mit drei
25 Hieben bezeichnet, um den Grenzstein desto besser finden zu können;
nach P o p o w i t s c h besser der Laßbaum, weil man ihn stehen läßt.
Dann überhaupt ein jeder Grenzbaum, in welchen ein Merkmahl zur
Auszeichnung eingehauen worden ist; der Mahlbaum, Markbaum,
Grenzbaum, auch Lochbaum und Lachterbaum.
30 -->
<text>
<body>
<div type="dictionary" xml:id="lachbaum" xml:lang="de" rend="unmarked">
<entry>
35 <form type="headword">
<gram type="art_m">Der</gram>
<orth>Lachbaum<c type="$,"></c></orth>
</form>
<gramGrp>
40 <number value="sing"/>
<case value="gen">
<oVar type="gen"><oRef/>-es</oVar>
<c type="$,"></c>
</case>
45 <number value="plur">Mz.</number>
<case value="nom">
<oVar type="pl"><oRef/>-bäume</oVar>
<c type="$,"></c>
</case>
50 </gramGrp>
<usg type="dom">im Forstwesen<c type="$. "></c></usg>
<sense>
<sense n="1" level="1" rend="numbered">
<c type="ordnungselement">1</c>
55 <def>Ein an sei-<lb/>ner Borke aufgehauener Nadelholzbaum<c
type="$,"
></c> aus welchem man Harz<lb/> rinnen läßt<c type="$
."></c></def>
</sense>
<sense n="2" level="1" rend="numbered">
<c type="ordnungselement">2</c>
60 <def>ein Baum<c type="$,"></c> welchen man auf der Grenze
des
Wal-<lb/>des neben den Grenzsteinen stehen läßt und auf
zwei Seiten
mit drei<lb/> Hieben bezeichnet<c type="$,"></c> um den
Grenzstein
desto besser finden zu können<c type="$. "></c><lb/></def>
<dictScrap>
65 <lbl>nach Popowitsch besser</lbl>
</dictScrap>
<xr type="syn">
<ref target="#Laßbaum">der Laßbaum<c type="$,"></c></ref>
</xr>
70 <def>weil man ihn stehen läßt<c type="$. "></c><lb/></def>
<def>Dann überhaupt ein jeder Grenzbaum<c type="$,"></c> in
welchen ein
Merkmahl zur<lb/> Auszeichnung eingehauen worden ist<c
type="$. "></c>
</def>
</sense>
75 <xr type="syn">
<ref target="#Mahlbaum">der Mahlbaum</ref>
<c type="$,"></c>
<ref target="#Markbaum">Markbaum</ref>
<c type="$,"></c>
80 <lb/>
<ref target="#Grenzbaum">Grenzbaum</ref>
<c type="$,"></c>
<lbl>auch</lbl>
<ref target="#Lochbaum">Lochbaum</ref>
85 <lbl>und</lbl>
<ref target="#Lachterbaum">Lachterbaum</ref>

```

```

        <c type="$. ">.</c><lb/>
        </xr>
        </sense>
90    </entry>
        </div>
    </body>
</text>
</TEI>

```

Listing 11: "Der Lachbaum" First approach

```

1  <?xml version="1.0" encoding="UTF-8"?>
  <?oxygen RNGSchema="campe_dictionary.rnc" type="compact"?>
  <TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="de">
    <teiHeader>
      <fileDesc>
6     <titleStmnt>
        <title>Campe - Exemplarisches Wörterbuch</title>
        <author>Christian Graiger, Mirjam Blümm</author>
      </titleStmnt>
      <publicationStmnt>
11    <p>Campe Wörterbuch, Beispiel der Kodierung bishin zur
        Wortebene</p>
      </publicationStmnt>
      <sourceDesc>
16    <p>Wörterbucheintrag "Der Lachbaum" aus Campe L-R, 1809</p>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
  <!--
21  Der Lachbaum, -es, Mz. -bäume, im Forstwesen. 1) Ein an sei-
    ner Borke aufgehauener Nadelholzbaum, aus welchem man Harz
    rinnen läßt; 2) ein Baum, welchen man auf der Grenze des Wal-
    des neben den Grenzsteinen stehen läßt und auf zwei Seiten mit drei
    Hieben bezeichnet, um den Grenzstein desto besser finden zu können;
26  nach P o p o w i t s c h besser der Laßbaum, weil man ihn stehen läßt.
    Dann überhaupt ein jeder Grenzbaum, in welchen ein Merkmahl zur
    Auszeichnung eingehauen worden ist; der Mahlbaum, Markbaum,
    Grenzbaum, auch Lochbaum und Lachterbaum.
  -->
  <text>
31  <body>
    <div n="L" type="groupOfEntriesGroupedByCharacter"
      xml:id="letter-l">
      <head n="L"/>
      <entry xml:id="lachbaum">
36    <revisionDesc>
        <change who="#LR">Der Lachbaum 2nd approach</change>
      </revisionDesc>
      <form type="lemma">
        <gramGrp>
41    <gen value="m"/>
        </gramGrp>
        <form type="determiner">
          <orth>Der</orth>
        </form>
46    <form type="headword">
          <orth>Lachbaum</orth><c type="$. ">.</c>
        </form>
      </form>
      <form type="inflected">
51    <orth>
        <oVar><oRef/>-es</oVar>
      </orth><c type="$. ">.</c>
      <gramGrp>
        <case value="G"/>
56    </gramGrp>
      </form>
      <form type="inflected">
        <gramGrp>
          <number value="p">Mz.</number>
61    </gramGrp>
        <orth>
          <oVar type="pl"><oRef/>-bäume</oVar>

```

```

        </orth><c type="$,"></c>
    </form>
66    <usg type="dom">im Forstwesen<c type="$."></c></usg>
    <sense>
        <sense n="1" rend="numbered">
            <head>1)</head>
            <def type="paraphrase">Ein an sei-<lb/>ner Borke aufgehauener
71    Nadelholzbaum<c type="pun"></c> aus welchem man
            Harz<lb/> rinnen läßt<c type="pun"></c></def>
        </sense>
        <sense n="2" rend="numbered">
            <head>2)</head>
76    <def type="paraphrase">ein Baum<c type="pun"></c> welchen man auf
            der
            Grenze des Wal-<lb/>des neben den Grenzsteinen stehen
            läßt und auf zwei Seiten mit drei<lb/> Hieben
            bezeichnet<c type="pun"></c> um den Grenzstein desto
            besser finden zu können<c type="pun"></c><lb/>
81    nach <name type="person" rend="spaced">Popowitsch</name>
            besser
            <xr>
                <ref target="#Laßbaum">der Laßbaum</ref><c type="pun"></c>
                </xr> weil man ihn stehen läßt<c type="pun"></c><lb/>
86    Dann überhaupt ein jeder Grenzbaum<c type="pun"></c>
            in welchen ein Merkmal zur<lb/> Auszeichnung
            eingehauen worden ist<c type="pun"></c>
            </def>
        </sense>
91    <xr type="syn">
        <ref target="#Mahlbaum" rend="0,5+">der Mahlbaum</ref>
        <c type="$,"></c>
        <ref target="#Markbaum" rend="0,5+">Markbaum</ref>
        <c type="$,"></c>
96    <lb/>
        <ref target="#Grenzbaum" rend="0,5+">Grenzbaum</ref>
        <c type="$,"></c>
        <lbl>auch</lbl>
        <ref target="#Lochbaum" rend="0,5+">Lochbaum</ref>
101    <lbl>und</lbl>
        <ref target="#Lachterbaum" rend="0,5+">Lachterbaum</ref>
        <c type="$."></c>
        <lb/>
        </xr>
106    </sense>
        </entry>
    </div>
</body>
</text>
111 </TEI>

```

Listing 12: “Der Lachbaum” Second approach

```

<?xml version="1.0" encoding="UTF-8"?>
<?oxygen RNGSchema="campe_dictionary.rnc" type="compact"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="de">
4    <teiHeader>
        <fileDesc>
            <titleStmnt>
                <title>Campe - Exemplarisches Wörterbuch</title>
                <author>Christian Graiger, Mirjam Blümm</author>
9            </titleStmnt>
            <publicationStmnt>
                <availability><p>nicht nutzbar</p></availability>
                <authority>KfE UW</authority>
            </publicationStmnt>
14            <sourceDesc>
                <p>Wörterbucheintrag "Der Lachbaum" aus Campe L-R, 1809</p>
            </sourceDesc>
        </fileDesc>
        <revisionDesc>
19            <change who="#LR">Der Lachbaum 3rd approach</change>
            <change who="#WW">Texterfassungsform Datum</change>
            <change who="#WW">Erstkodierung Datum</change>
            <change who="#WW">Textkontrolle Datum</change>

```

```

24     <change who="#WW">Übernahme in Bandkodierung</change>
    </revisionDesc>
</teiHeader>
<!--
    Der Lachbaum, -es, Mz. -bäume, im Forstwesen. 1) Ein an sei-
29   ner Borke aufgehauener Nadelholzbaum, aus welchem man Harz
    rinnen läßt; 2) ein Baum, welchen man auf der Grenze des Wal-
    des neben den Grenzsteinen stehen läßt und auf zwei Seiten mit drei
    Hieben bezeichnet, um den Grenzstein desto besser finden zu können;
    nach P o p o w i t s c h besser der Laßbaum, weil man ihn stehen läßt.
34   Dann überhaupt ein jeder Grenzbaum, in welchen ein Merkmal zur
    Auszeichnung eingehauen worden ist; der Mahlbaum, Markbaum,
    Grenzbaum, auch Lochbaum und Lachterbaum.
-->

<text>
39   <body>
    <div n="L" type="groupOfEntriesGroupedByCharacter"
        xml:id="letter-l">
        <head n="L."/>
        <entry xml:id="lachbaum">
44           <form type="lemma">
                <gramGrp>
                    <gen value="m"/>
                </gramGrp>
49           <form type="determiner">
                <orth>Der</orth>
            </form>
            <form type="headword">
                <orth>Lachbaum</orth><c type="$, "></c>
54           </form>
            </form>
            <form type="inflected">
                <orth>
                    <oVar><oRef/>-es</oVar><!-- ORef markiert Textstellen mit
                        Ersatzdarstellung des Stichworts; das Referenzelement ""
                        bleibt im Text -->
59           </orth><c type="$, "></c>
                <gramGrp>
                    <case value="G"/>
                </gramGrp>
            </form>
64           <form type="inflected">
                <gramGrp>
                    <number value="p">Mz.</number>
                </gramGrp>
                <orth>
                    <oVar type="pl"><oRef/>-bäume</oVar>
69           </orth><c type="$, "></c>
            </form>
            <usg type="dom">im Forstwesen<c type="$. "></c></usg>
            <sense>
74           <sense n="1" rend="numbered">
                <head>1)</head>
                <def type="paraphrase">Ein an sei-<lb/>ner Borke aufgehauener
                    Nadelholzbaum<c type="pun"></c> aus welchem man
                    Harz<lb/> rinnen läßt<c type="$, "></c></def>
79           </sense>
                <sense n="2" rend="numbered">
                    <head>2)</head>
                    <def type="paraphrase">ein Baum<c type="$, "></c> welchen man auf
                    der
84           Grenze des Wal-<lb/>des neben den Grenzsteinen stehen
                    läßt und auf zwei Seiten mit drei<lb /> Hieben
                    bezeichnet<c type="$, "></c> um den Grenzstein desto
                    besser finden zu können<c type="$, "></c><lb/>
                    nach <name type="person" rend="spaced">Popowitsch</name>
                    besser
89           <xr>
                <ref target="#Laßbaum">der Laßbaum</ref><c type="$, "></c>
                </xr> weil man ihn stehen läßt<c type="$. "></c><lb/>
                Dann überhaupt ein jeder Grenzbaum<c type="$, "></c>
                in welchen ein Merkmal zur<lb/> Auszeichnung

```

```

94         eingehauen worden ist<c type="$";></c>
        </def>
    </sense>
    <xr type="syn">
        <ref target="#Mahlbaum" rend="0,5+">der Mahlbaum</ref>
99         <c type="$,"></c>
        <ref target="#Markbaum" rend="0,5+">Markbaum</ref>
        <c type="$,"></c>
        <lb/>
104        <ref target="#Grenzbaum" rend="0,5+">Grenzbaum</ref>
        <c type="$,"></c>
        <lbl>auch</lbl>
        <ref target="#Lochbaum" rend="0,5+">Lochbaum</ref>
        <lbl>und</lbl>
109        <ref target="#Lachterbaum" rend="0,5+">Lachterbaum</ref>
        <c type="$."></c>
        <lb/>
    </xr>
    </sense>
    </entry>
114 </div>
</body>
</text>
</TEI>

```

Listing 13: “Der Lachbaum” Third approach

```

<?xml version="1.0"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="en">
3   <teiHeader>
    <fileDesc>
        <titleStmt>
            <title>philtag 5 extension for encoding the Campe dictionary</title>
            <author>generated by Roma 2.3</author>
8         </titleStmt>
        <publicationStmt>
            <p>for use by whoever wants it</p>
        </publicationStmt>
        <sourceDesc>
13        <p>created on Saturday 07th October 2006 08:22:19 AM by the form
            at
            http://www.tei-c.org.uk/Roma</p>
        </sourceDesc>
    </fileDesc>
    </teiHeader>
18   <text>
    <body>
        <divGen type="toc"/>
        <schemaSpec ident="campe_dictionary" xml:lang="en" prefix="">
            <moduleRef key="core"/>
23            <moduleRef key="tei"/>
            <moduleRef key="header"/>
            <moduleRef key="textstructure"/>
            <moduleRef key="dictionaries"/>
            <moduleRef key="analysis"/>
28            <elementSpec module="analysis" ident="cl" mode="delete"/>
            <elementSpec module="analysis" ident="interp" mode="delete"/>
            <elementSpec module="analysis" ident="interpGrp" mode="delete"/>
            <elementSpec module="analysis" ident="m" mode="delete"/>
            <elementSpec module="analysis" ident="phr" mode="delete"/>
33            <elementSpec module="analysis" ident="span" mode="delete"/>
            <elementSpec module="analysis" ident="spanGrp" mode="delete"/>
            <elementSpec module="core" ident="add" mode="delete"/>
            <elementSpec module="core" ident="addrLine" mode="delete"/>
            <elementSpec module="core" ident="address" mode="delete"/>
38            <elementSpec module="core" ident="altIdent" mode="delete"/>
            <elementSpec module="core" ident="analytic" mode="delete"/>
            <elementSpec module="core" ident="binaryObject" mode="delete"/>
            <elementSpec module="core" ident="choice" mode="delete"/>
43            <elementSpec module="core" ident="corr" mode="delete"/>
            <elementSpec module="core" ident="dateRange" mode="delete"/>
            <elementSpec module="core" ident="del" mode="delete"/>
            <elementSpec module="core" ident="desc" mode="delete"/>

```

```

48 <elementSpec module="core" ident="distinct" mode="delete"/>
<elementSpec module="core" ident="divGen" mode="delete"/>
<elementSpec module="core" ident="emph" mode="delete"/>
<elementSpec module="core" ident="equiv" mode="delete"/>
<elementSpec module="core" ident="expan" mode="delete"/>
<elementSpec module="core" ident="foreign" mode="delete"/>
53 <elementSpec module="core" ident="gap" mode="delete"/>
<elementSpec module="core" ident="gloss" mode="delete"/>
<elementSpec module="core" ident="graphic" mode="delete"/>
<elementSpec module="core" ident="headItem" mode="delete"/>
<elementSpec module="core" ident="headLabel" mode="delete"/>
<elementSpec module="core" ident="hi" mode="delete"/>
58 <elementSpec module="core" ident="imprint" mode="delete"/>
<elementSpec module="core" ident="index" mode="delete"/>
<elementSpec module="core" ident="item" mode="delete"/>
<elementSpec module="core" ident="l" mode="delete"/>
<elementSpec module="core" ident="label" mode="delete"/>
63 <elementSpec module="core" ident="lg" mode="delete"/>
<elementSpec module="core" ident="list" mode="delete"/>
<elementSpec module="core" ident="listBibl" mode="delete"/>
<elementSpec module="core" ident="measure" mode="delete"/>
<elementSpec module="core" ident="meeting" mode="delete"/>
68 <elementSpec module="core" ident="milestone" mode="delete"/>
<elementSpec module="core" ident="monogr" mode="delete"/>
<elementSpec module="core" ident="num" mode="delete"/>
<elementSpec module="core" ident="orig" mode="delete"/>
<elementSpec module="core" ident="postBox" mode="delete"/>
73 <elementSpec module="core" ident="postCode" mode="delete"/>
<elementSpec module="core" ident="pubPlace" mode="delete"/>
<elementSpec module="core" ident="reg" mode="delete"/>
<elementSpec module="core" ident="resp" mode="delete"/>
<elementSpec module="core" ident="respStmt" mode="delete"/>
78 <elementSpec module="core" ident="rs" mode="delete"/>
<elementSpec module="core" ident="series" mode="delete"/>
<elementSpec module="core" ident="sic" mode="delete"/>
<elementSpec module="core" ident="soCalled" mode="delete"/>
<elementSpec module="core" ident="sp" mode="delete"/>
83 <elementSpec module="core" ident="speaker" mode="delete"/>
<elementSpec module="core" ident="stage" mode="delete"/>
<elementSpec module="core" ident="street" mode="delete"/>
<elementSpec module="core" ident="term" mode="delete"/>
<elementSpec module="core" ident="time" mode="delete"/>
88 <elementSpec module="core" ident="timeRange" mode="delete"/>
<elementSpec module="core" ident="unclear" mode="delete"/>
<elementSpec module="textstructure" ident="argument" mode="delete
"/>
<elementSpec module="textstructure" ident="byline" mode="delete"/>
93 <elementSpec module="textstructure" ident="closer" mode="delete"/>
<elementSpec module="textstructure" ident="dateline" mode="delete
"/>
<elementSpec module="textstructure" ident="div0" mode="delete"/>
<elementSpec module="textstructure" ident="div1" mode="delete"/>
<elementSpec module="textstructure" ident="div2" mode="delete"/>
<elementSpec module="textstructure" ident="div3" mode="delete"/>
98 <elementSpec module="textstructure" ident="div4" mode="delete"/>
<elementSpec module="textstructure" ident="div5" mode="delete"/>
<elementSpec module="textstructure" ident="div6" mode="delete"/>
<elementSpec module="textstructure" ident="div7" mode="delete"/>
<elementSpec module="textstructure" ident="docAuthor" mode="delete
"/>
103 <elementSpec module="textstructure" ident="docDate" mode="delete
"/>
<elementSpec module="textstructure" ident="docEdition" mode="
delete"/>
<elementSpec module="textstructure" ident="docImprint" mode="
delete"/>
<elementSpec module="textstructure" ident="docTitle" mode="delete
"/>
<elementSpec module="textstructure" ident="epigraph" mode="delete
"/>
108 <elementSpec module="textstructure" ident="group" mode="delete"/>
<elementSpec module="textstructure" ident="imprimatur" mode="
delete"/>
<elementSpec module="textstructure" ident="opener" mode="delete"/>

```

```

113 <elementSpec module="textstructure" ident="salute" mode="delete"/>
<elementSpec module="textstructure" ident="signed" mode="delete"/>
<elementSpec module="textstructure" ident="trailer" mode="delete
"/>
<elementSpec ident="revisionDesc" module="header" mode="change">
  <classes>
    <memberOf key="model.entryParts.top"/>
  </classes>
118 </elementSpec>
<elementSpec ident="gramGrp" module="dictionaries" mode="change">
  <classes>
    <memberOf key="model.entryParts"/>
    <memberOf key="model.entryParts.top"/>
123 <memberOf key="model.formPart"/>
    <memberOf key="model.gramPart"/>
    <memberOf key="att.lexicographic"/>
  </classes>
</elementSpec>
128 <elementSpec ident="xr" module="dictionaries" mode="change">
  <classes>
    <memberOf key="model.entryParts"/>
    <memberOf key="model.entryParts.top"/>
    <memberOf key="model.global"/>
133 <memberOf key="att.lexicographic"/>
  </classes>
</elementSpec>
<elementSpec ident="def" module="dictionaries" mode="change">
  <content xmlns:rng="http://relaxng.org/ns/structure/1.0"
138 xmlns="http://www.tei-c.org/ns/1.0">
    <rng:ref xmlns:rng="http://relaxng.org/ns/structure/1.0"
      name="macro.paraContent"/>
    <rng:attribute xmlns:rng="http://relaxng.org/ns/structure
/1.0" name="type">
      <rng:data xmlns:rng="http://relaxng.org/ns/structure
/1.0" type="QName"/>
143 </rng:attribute>
  </content>
</elementSpec>
<elementSpec ident="lbl" module="dictionaries" mode="change">
148 <classes>
  <memberOf key="model.entryParts"/>
  <memberOf key="model.entryParts.top"/>
  <memberOf key="model.formPart"/>
  <memberOf key="model.gramPart"/>
  <memberOf key="att.lexicographic"/>
153 </classes>
</elementSpec>
<elementSpec ident="head" module="core" mode="change">
  <classes>
158 <memberOf key="model.divWrapper"/>
  <memberOf key="model.entryParts.top"/>
  <memberOf key="model.pLike"/>
  <memberOf key="model.pLike.front"/>
  </classes>
</elementSpec>
163 <elementSpec ident="cit" module="core" mode="change">
  <classes>
    <memberOf key="model.entryParts.top"/>
    <memberOf key="model.qLike"/>
  </classes>
168 </elementSpec>
</schemaSpec>
</body>
</text>
</TEI>

```

Listing 14: Customization for encoding dictionaries

References

[ROM06] Sebastian Rahtz, Arno Mittelbach: ROMA: *generating validators for the TEI*, <http://tei.oucs.ox.ac.uk/Roma/>.

- [TEI05] C.M. Sperberg-McQueen, Lou Burnard: *The TEI Guidelines*, revised 30/04/2006.
- [TEI06] Text Encoding Initiative: *EDW88: The TEI P5 ODD How To*, <http://www.tei-c.org/Drafts/edw88.xml?style=printable>.