



Masterclass “Digital Scholarly Edition”  
Saarbrücken, 19.–23.2.2024



# TEI: Embedded Transcription

**Bernhard Bauer (Universität Graz)**

# TEI - Document



- At least two elements:
  - `<teiHeader>` in which metadata is recorded
  - One of the following elements to represent the text:
    - `<text>` for “normal” transcriptions
    - `<facsimile>` to represent the text with images
    - `<sourceDoc>` for embedded transcriptions

# Representation of Primary Sources (TEI-Guidelines 11)



„A digitized source document may contain nothing more than page images and a small amount of metadata. It may also contain an encoded transcription of the pages represented, which may either be ‘embedded’ within a <sourceDoc> element or supplied in parallel with a <facsimile>.”

# Parallel Transcription vs. Embedded Transcription



- Parallel Transcription:
  - The transcription is a text in its own right and structured independently in a document. It is supplied in parallel with a <facsimile>
- Embedded Transcription:
  - „*An embedded transcription* is one in which words and other written traces are encoded as subcomponents of elements representing the physical surfaces carrying them rather than independently of them.”

# Embedded Transcription (TEI-Guidelines 11.2.2.)



## ○ Available elements:

- <sourceDoc> contains a transcription or other representation of a single source document potentially forming part of a dossier *génétique* or collection of sources.
- <surface> defines a written surface as a two-dimensional coordinate space, optionally grouping one or more graphic representations of that space, zones of interest within that space, and transcriptions of the writing within them.
- <zone> defines any two-dimensional area within a <surface> element.
- <line> contains the transcription of a topographic line in the source document
- <seg> (arbitrary segment) represents any segmentation of text below the 'chunk' level.

# <sourceDoc>



- “This element may be used as an alternative to facsimile for TEI documents containing only page images, or for documents containing both images and transcriptions. Transcriptions may be provided within the surface elements making up a source document, in parallel with them as part of a text element, or in both places if the encoder wishes to distinguish these two modes of transcription.”

# <sourceDoc>



- <sourceDoc> can host the <graphic>-elements

```
17 <sourceDoc>
18   <surfaceGrp n="leaf1">
19     <surface facts="page1.png">
20       <zone>All the writing on page 1</zone>
21     </surface>
22     <surface>
23       <graphic url="page2-highRes.png"/>
24       <graphic url="page2-lowRes.png"/>
25       <zone>
26         <line>A line of writing on page 2</line>
27         <line>Another line of writing on page 2</line>
28       </zone>
29     </surface>
30   </surfaceGrp>
31 </sourceDoc>
```

# <surface>



- „<surface> defines a written surface as a two-dimensional coordinate space, optionally grouping one or more graphic representations of that space, zones of interest within that space, and transcriptions of the writing within them.”

```
17 ▾ <sourceDoc>
18 ▾   <surface ulx="0" uly="0" lrx="200" lry="300">
19     <graphic url="Beatles-49r.png"/>
20   </surface>
21 </sourceDoc>
--
```



# <zone>



- „<zone> defines any two-dimensional area within a surface element..”
- The position of every zone for a given surface is always defined by reference to the coordinate system defined for that surface.
- A graphic element contained by a zone represents the whole of the zone.
- A zone may be of any shape. The attribute points may be used to define a polygonal zone, using the coordinate system defined by its parent surface.
- A zone is always a closed polygon. Repeating the initial coordinate at the end of the sequence is optional. To encode an unclosed path, use the <path> element.

```
17 ▾ <sourceDoc>
18 ▾   <surface ulx="14.54" uly="16.14" lrx="0"
19     lry="0">
20     <graphic url="stones.jpg"/>
21 ▾   <zone points="4.6,6.3 5.25,5.85 6.2,6.6 8.19222,7.4125 9.89222,6.5875 10.9422,6.1375
22     11.4422,6.7125 8.21722,8.3125 6.2,7.65" />
23   </surface>
24 </sourceDoc>
```

# <line>



- “<u>line</u> contains the transcription of a topographic line in the source document”
- “This element should be used only to mark up writing which is topographically organized as a series of lines, horizontal or vertical. It should not be used to mark lines of verse (for which use l) nor to mark linebreaks within text which has been encoded using structural elements such as p (for which use lb).”

```
18 <surface>
19   <zone>
20     <line>I am he as you are he as you are me</line>
21     <line>And we are all together</line>
22     <line>See how they run like pigs from a gun</line>
23     <line>See how they fly</line>
24     <line>I'm crying</line>
25   </zone>
26 </zone>
27   <zone>
28     <line>Goo goo g'joob</line>
29   </zone>
</surface>
```

# How to represent lines



Embedded Transcription	"Normal" Transcription	Lines in poetry
<pre> &lt;surface&gt;   &lt;zone&gt;     &lt;line&gt;A Limerick&lt;/line&gt;     &lt;line&gt;A dozen, a gross, and a score&lt;/line&gt;     &lt;line&gt;Plus three time the square root of four&lt;/line&gt;     &lt;line&gt;Divided by seven&lt;/line&gt;     &lt;line&gt;Plus five times eleven&lt;/line&gt;     &lt;line&gt;Is nine squared and not a bit more&lt;/line&gt;   &lt;/zone&gt; &lt;/surface&gt; </pre>	<pre> &lt;div&gt;   &lt;pb/&gt;   &lt;p&gt;     &lt;lb/&gt;A Limerick     &lt;lb/&gt;A dozen, a gross, and a score     &lt;lb/&gt;Plus three time the square root of four     &lt;lb/&gt;Divided by seven     &lt;lb/&gt;Plus five times eleven     &lt;lb/&gt;Is nine squared and not a bit more   &lt;/p&gt; &lt;/div&gt; </pre>	<pre> &lt;div&gt;   &lt;lg&gt;     &lt;head&gt;A Limerick&lt;/head&gt;     &lt;l&gt;A dozen, a gross, and a score&lt;/l&gt;     &lt;l&gt;Plus three time the square root of four&lt;/l&gt;     &lt;l&gt;Divided by seven&lt;/l&gt;     &lt;l&gt;Plus five times eleven&lt;/l&gt;     &lt;l&gt;Is nine squared and not a bit more&lt;/l&gt;   &lt;/lg&gt; &lt;/div&gt; </pre>

# (Some) Transcription Elements available for Embedded Transcriptions



- <unclear>  
“contains a word, phrase, or passage which cannot be transcribed with certainty because it is illegible or inaudible in the source”
- <damage>  
“(damage) contains an area of damage to the text witness”
- <gap>  
“(gap) indicates a point where material has been omitted in a transcription, whether for editorial reasons described in the TEI header, as part of sampling practice, or because the material is illegible, invisible, or inaudible.”

# (Some) Transcription Elements



## ○ <del>

“(deletion) contains a letter, word, or passage deleted, marked as deleted, or otherwise indicated as superfluous or spurious in the copy text by an author, scribe, or a previous annotator or corrector.”

## ○ <add>

“(addition) contains letters, words, or phrases inserted in the source text by an author, scribe, or a previous annotator or corrector.”

# (Some) Transcription Elements



## ○ <choice>

“groups a number of alternative encodings for the same point in a text”

- <orig>

“contains a reading which is marked as following the original, rather than being normalized or corrected”

- <reg>

“(regularization) contains a reading which has been regularized or normalized in some sense”

- <abbr>

“(abbreviation) contains an abbreviation of any sort”

- <expn>

“(expansion) contains the expansion of an abbreviation”



# Example: Wellcome Library, Misc. Cookery Books MS 18.22 (1746 – c. 1850)



To make a calves foot pye 24  
Take yr calves feet after they are scalded and boyled and put ym  
into a good crust with a little butter and reasens and currans and bake it so  
and when it is enough cutt of ye lit and put into it a glass of white wine  
mixt with som butter and suger and a little vinegar and ye yolk of an egg  
and made very hott over ye fire  
To make ye ~~sweet~~ <sup>sweet</sup> oylment a most excellent and  
approved good one for paine in ye lining or a crush  
Take sage two one pound bay leaves and wormewood one pound mellilet  
herbes and flowers of camomile spike roemany Red roeis ! John wort and  
march mallons of each a good handfull chop all ye herbes as small as can be

<https://wellcomecollection.org/works/favfamjv?c=0&m=0&s=0&cv=48&z=-0.3899%2C-0.1193%2C1.8624%2C1.1699>



# Example: Wellcome Library, Misc. Cookery Books



To make a calves foot pye 24

Take yr calves feet after they are scalded and boyled and put ym into a good crust with a little butter and redons and currans and bake it so and when it is enough cutt of ye lit and put into it a glass of white wine mixt with som butter and suger and a little vineger and ye yolk of an egg and made very hott over ye fire

To make ye <sup>sweet</sup> ~~good~~ oyntment a most excellent and approved good one for paine in ye lining or a crush

Take Sage two one pound Bay leaves and wormwood one pound mellilet hearbes and flowers of camanile spike rosmary Red rosis ! John wort and march mallows of each a good handfull chop all ye hearbes as small as can be

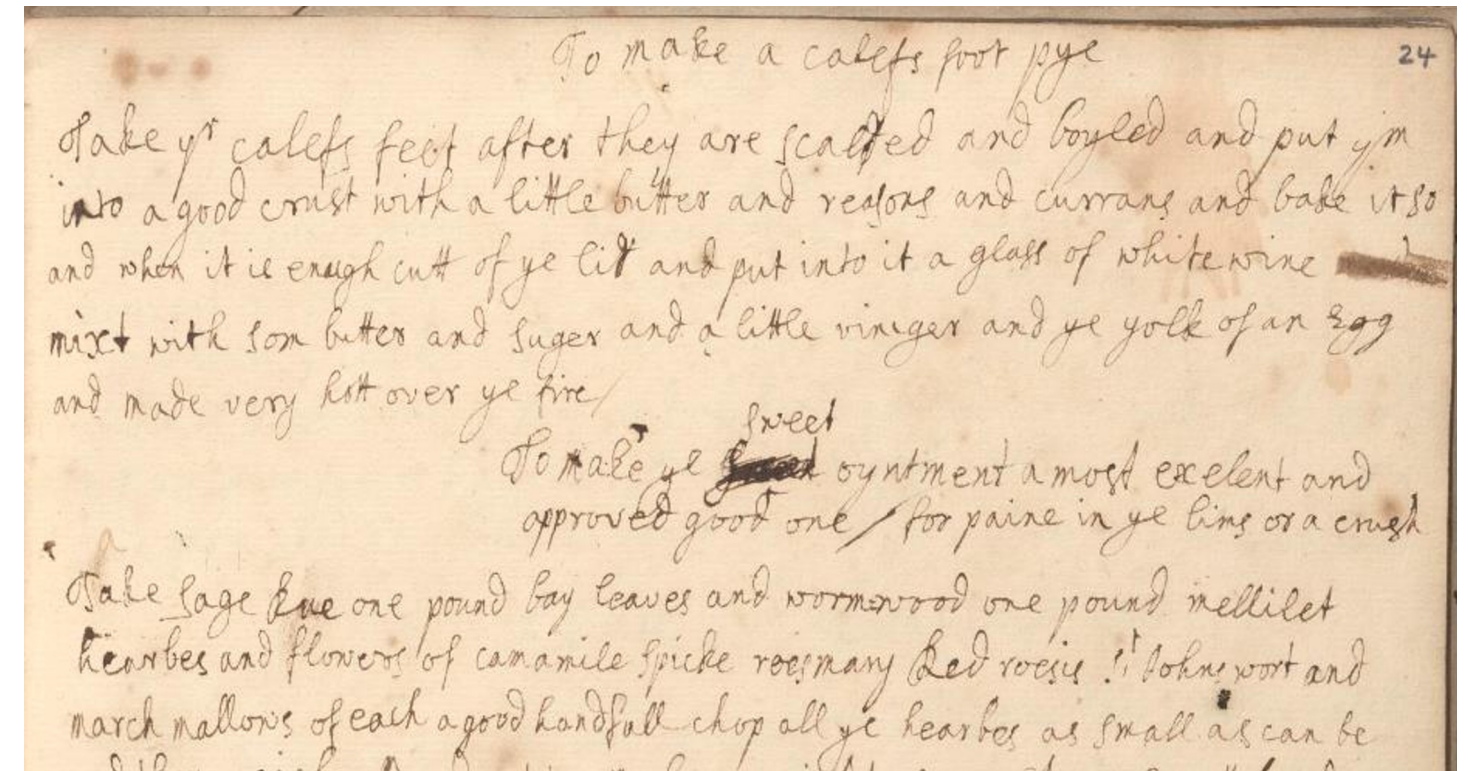
<https://wellcomecollection.org/works/favfamjv#c=0&m=0&s=0&cv=48&z=-0.3899%2C-0.1193%2C1.8624%2C1.1699>



# Example: Wellcome Library, Misc. Cookery Books



```
19 ▾ <sourceDoc>
20 ▾   <surface>
21     <graphic url="page24.png" />
22 ▾     <zone type="page-number">
23       <line>24</line>
24     </zone>
25 ▾     <zone type="heading">
26       <line>To make a calefs foot pye</line>
27     </zone>
28 ▾     <zone type="main-text">
29       <line>Take yr calefs feet after they are scalded and boyled and put ym</line>
30       <line>into a good crust with a little butter ...</line>
31     </zone>
32   </surface>
33 </sourceDoc>
```





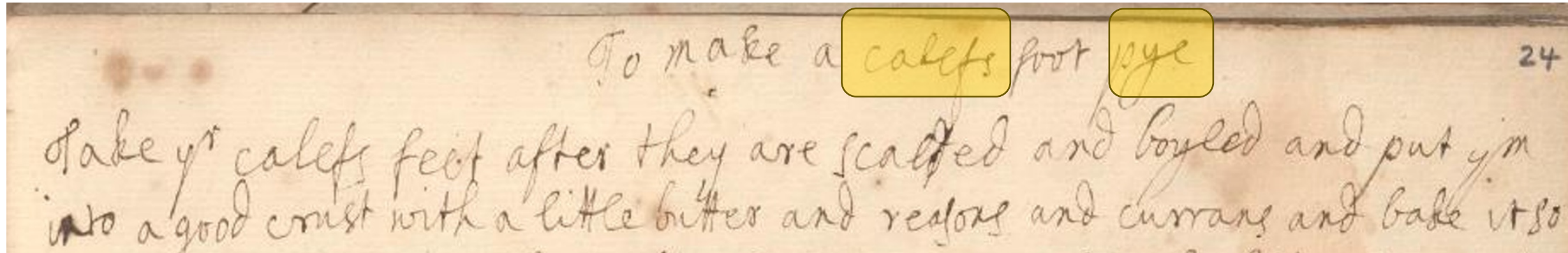
# Example: Wellcome Library, Misc. Cookery Books



To make a calves foot pye 24  
Take yr calves feet after they are scalded and boyled and put ym  
into a good crust with a little butter and redons and currans and bake it so  
and when it is enough cutt of ye lit and put into it a glass of white wine  
mixt with som butter and suger and a little vineger and ye yolk of an egg  
and made very hott over ye fire  
To make ye ~~good~~ sweet oylment a most excellent and  
approved good one for paine in ye lining or a crush  
Take sage one pound bay leaves and wormwood one pound mellilet  
herbes and flowers of camomile spicke rosmary Red rosis ! John wort and  
march mallows of each a good handfull chop all ye herbes as small as can be

<https://wellcomecollection.org/works/favfamjv#c=0&m=0&s=0&cv=48&z=-0.3899%2C-0.1193%2C1.8624%2C1.1699>

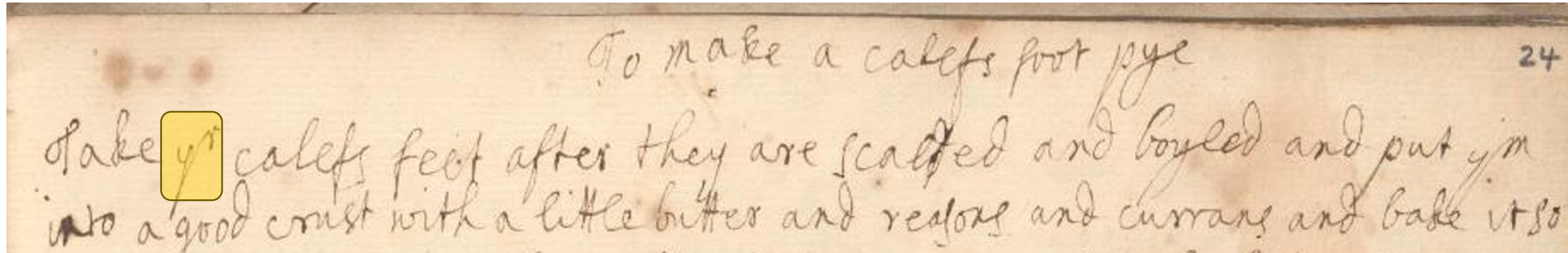
# <orig> / <reg>



```
<zone type="heading">  
  <line>To make a <choice>  
    <orig>calefs</orig>  
    <reg>calf's</reg>  
  </choice> foot  
  <choice>  
    <orig>pye</orig>  
    <reg>pie</reg>  
  </choice>  
</line>  
</zone>
```

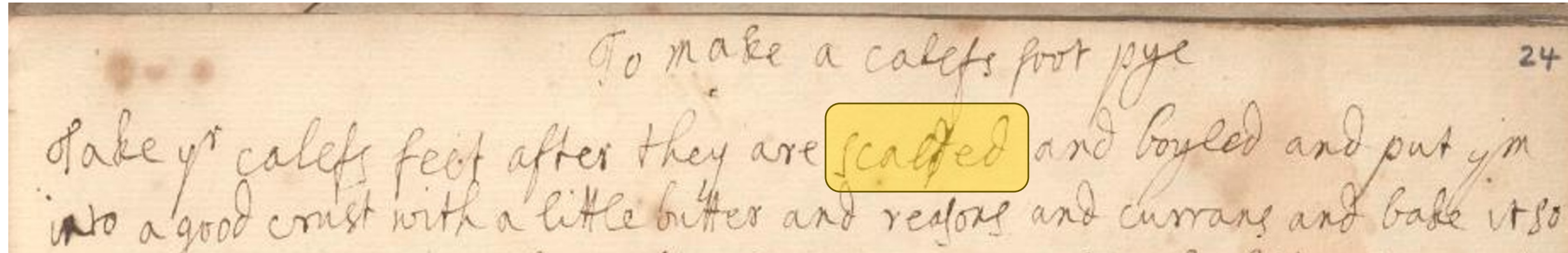


# <abbr> / <expan>



```
<zone type="main-text">  
  <line>Take  
    <choice>  
      <abbr>yr</abbr>  
      <expan>your</expan>  
    </choice>  
  calefs feet after they are
```

<unclear>



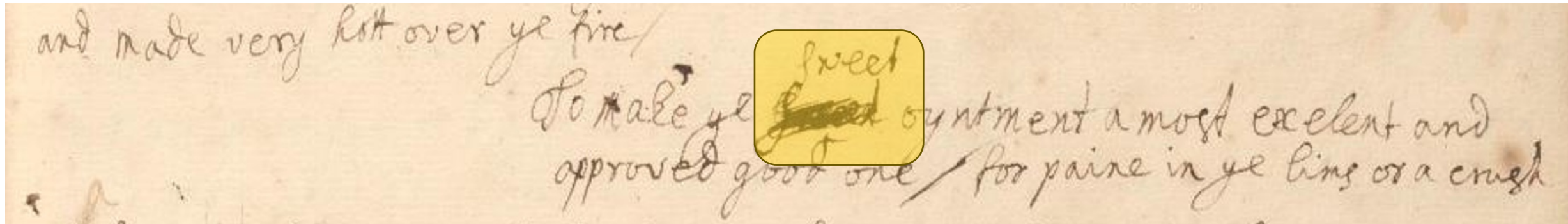
<zone type="main-text">

<line>Take yr calefs feet after they are <unclear reason="illegible">scalded</unclear> and boyled and put ym</line>

<line>into a good crust with a little butter ...</line>

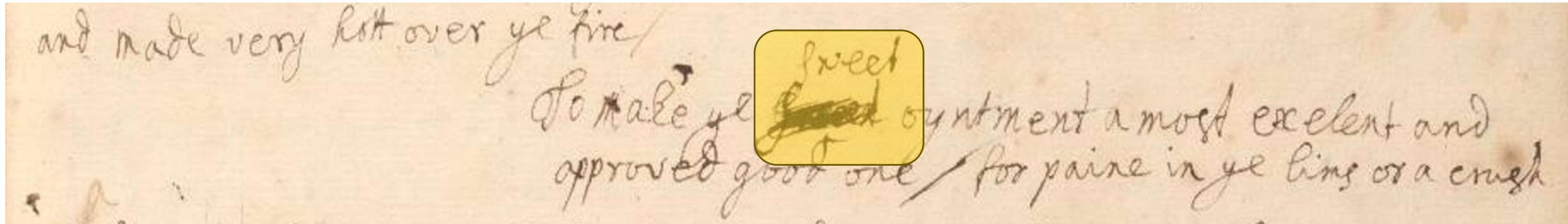
</zone>

<del>



```
<line>To make ye  
  <del>s<gap reason="cancelled"  
    quantity="3" unit="chars"/>t  
  </del> ...  
</line>
```

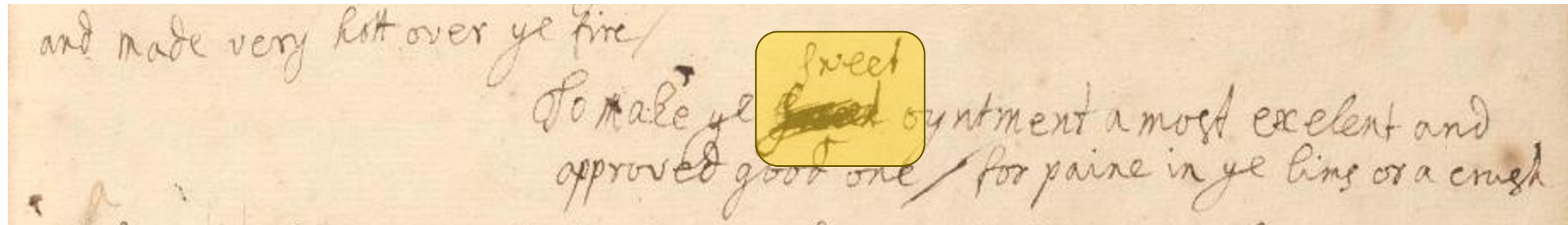
<add>



<line>To make ye <add place="above">sweet</add> oyntment a most ... </line>



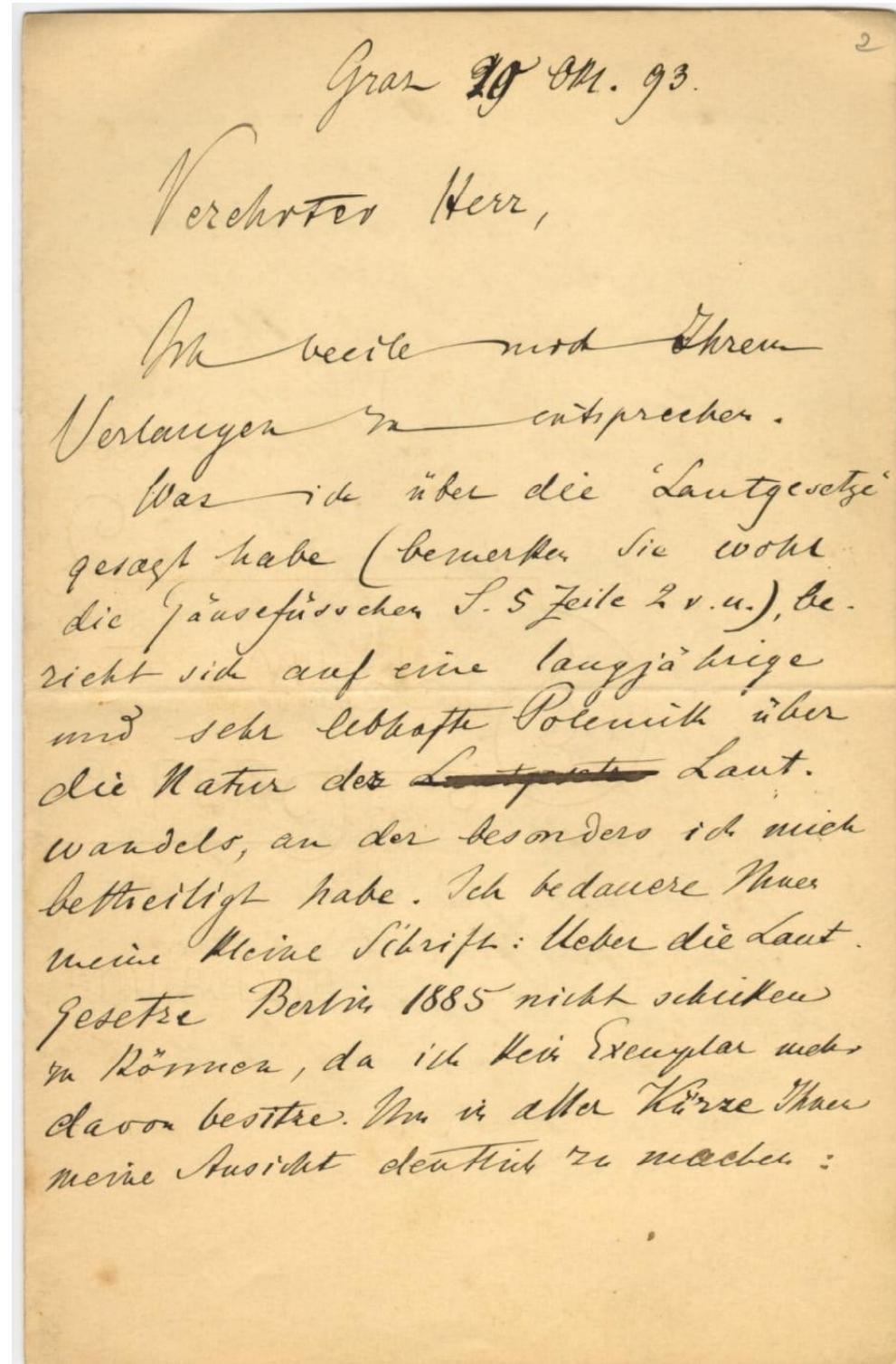
# Combining <del> and <add>



```
<line>To make ye  
<del>s<gap reason="cancelled" quantity="3" unit="chars"/>t</del>  
<add place="above">sweet</add>  
oyntment a most ... </line>
```



# Hands-on Training



Try an embedded transcription by yourself:

- Use either your own material or
- yesterday's transcription example ([https://drive.google.com/file/d/1a5ifjpJKsJMxLuWsSEAbijqxxhLJ2pflz/view?usp=drive\\_link](https://drive.google.com/file/d/1a5ifjpJKsJMxLuWsSEAbijqxxhLJ2pflz/view?usp=drive_link))
- Set up your own TEI-document and use the respective elements