

Encoding pre-modern writing systems

Paolo Monella

CARMEN Workshop, Rome 22 June 2022



SAPIENZA
UNIVERSITÀ DI ROMA

1. Issues

2. Tables of signs

3. Epigraphy

ERC PAGES



European Research Council

Established by the European Commission

ERC PAGES (AdG 2019 n° 882588)

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 882588)

1. Issues

Issues: pre-modern writing systems

- ① Graphemes/allographs (graphematic system)
- ② Brevigraphs (one sign: p/pro)
- ③ Abbreviation combinations (more signs: AUGG/Augusti, utra'q;/utramque)

Context

- Medieval manuscripts
- All pre-modern documents

Question: how much of this applies to Epigraphs?

- Allographs
- Nature of abbreviations

2. Tables of signs

Source files

- Edition of Romualdus Salernitanus, Chronicon
- See www.paolomonella.it/carmen2022 → Materials → Tables of signs
- Key files to download:
 - XML transcription
 - X2: source code - original
 - Tables of signs
 - T3: Table of **graphemes** for MS A, including **brevigraphs** - edited
 - T5: Table of **abbreviation combinations** for MS A, including **brevigraphs** - edited

Graphemes/allographs

EpiDoc

See EpiDoc version: 9.3 → Regularization:

```
<choice>
  <reg>f</reg><orig>s</orig>ua
</choice>
```

With table

fua

Table of allographs

Allographs	Graphemes	Notes
f	s	
s	s	

Brevigraphs

EpiDoc

See CARMEN materials → 2. Text transcription → Symbols (video, slides):

```
<expan>
  <abbr>
    <am>p</am>
  </abbr>
  <ex>pro</ex>
</expan>
```

With table + markup

```
<abbr>p</abbr>
```

With table only

p*ficio*

Brevigraphs

Table of graphemes

Graphemes	Interpretive	Type	Notes
þ	pro	Brevigraph	

Abbreviation combinations

EpiDoc

See CARMEN mat. → 2. Text transcr. → Abbreviations (video, slides)

```
<expa>
```

```
    utra
```

```
    <abbr>
```

```
        <am>' </am>
```

```
    </abbr>
```

```
    <ex>m</ex>
```

```
q
```

```
    <abbr>
```

```
        <am>; </am>
```

```
    </abbr>
```

```
    <ex>ue</ex>
```

```
</expa>
```

Abbreviation combinations

With table + markup

See EpiDoc slide Unexpanded abbreviation:

utr

```
<abbr>a'</abbr>
```

```
<abbr>q;</abbr>
```

With table only

utra'q;

Abbreviation combinations

Table of graphemes

Graphemes	Interpretive	Type	Notes
'		Abbreviation mark	Written over...
;		Abbreviation mark	It looks like...

Table of abbreviation combinations

Graphemes	Interpretive	Notes
q;	que	Vowel+' = vowel+m
[aeiouy]'	[aeiouy]m	Vowel+'=vowel + m

Abbreviation combinations (whole words)

EpiDoc

See CARMEN mat. → 2. Text transcr. → Abbreviations (video, slides)

```
<expn>
```

```
  <abbr>aug</abbr>
```

```
  <ex>ustus</ex>
```

```
</expn>
```

With table + markup

See EpiDoc slide Unexpanded abbreviation:

```
<abbr>aug</abbr>
```

With table only (word distinction needed!)

et aug feliciter

Abbreviation combinations (whole words)

Table of abbreviation combinations

Graphemes	Interpretive	Notes
<aug>	Augustus	

Abbreviations can always be expanded in transcription

Diplomatic

generationes uero iuxta utra3q9 editione3
 numero dece3, que uniu2sali est deleta diluuio,
 si

<choice>

<abbr>c0</abbr>

<expn>cu</expn>

</choice>

t
 prima3

Interpretive

generationes uero iuxta utramque editionem
 numero decem, que uniuersali est deleta diluuio, *sicut* primam

3. Epigraphy

General issues of the “table” approach

- Rationale
 - Writing is systematic (mostly)
 - Including abbreviations (mostly)
- Each table → a graphic system
- Scope
 - 1 MS → 1 system → 1 table

Saussure

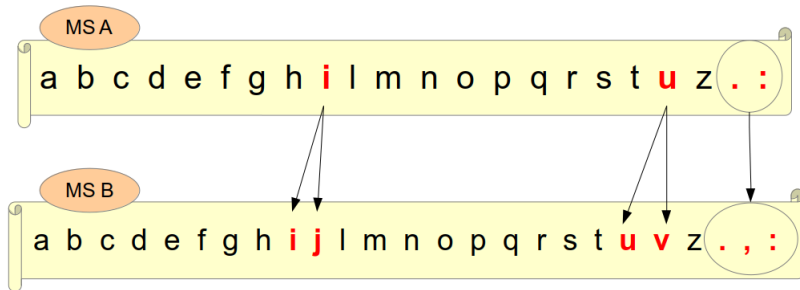


Figura 1: Signs are defined within a system

Epigraphy

- The “table” approach works **if** the sytem is “systematic”
 - ① The corpus is **wide** → we can populate the table
 - ② Abbreviation conventions are **consistent**
 - ③ Practically: how recurrent/consistent are common abbreviations?
- Inscriptions
 - Is this approach **scalable** to a definite **(sub-)corpus** of inscriptions?
 - Are **abbreviation** conventions (brevigraphs, combinations) frequent and consistent **enough** in specific sub-corpora?
 - How about **allographs** in inscriptions?

License



Figura 2: Creative Commons Attribution-ShareAlike 4.0 International License