Paolo Monella

La tabella dei segni
nell'edizione scientifica digitale:
un approccio saussuriano

The table of signs
in the scholarly digital edition:
a Saussurian approach

paolo.monella@unipa.it
http://www1.unipa.it/paolo.monella/calabriatabella2019/index.html

Philosophy of Language and Digital Humanities
University of Calabria (Rende, CS, Italy)
Saussurian tradition session
May 8, 2019

CC BY SA

Outline

- Digital encoding of handritten writing systems
  - Saussure
  - Table of signs (of all graphemes in a witness)
    - Possible implementations
    - Processing

Saussure

Ferdinand De Saussure

Relational nature of signs *within* a semiotic system

Saussure

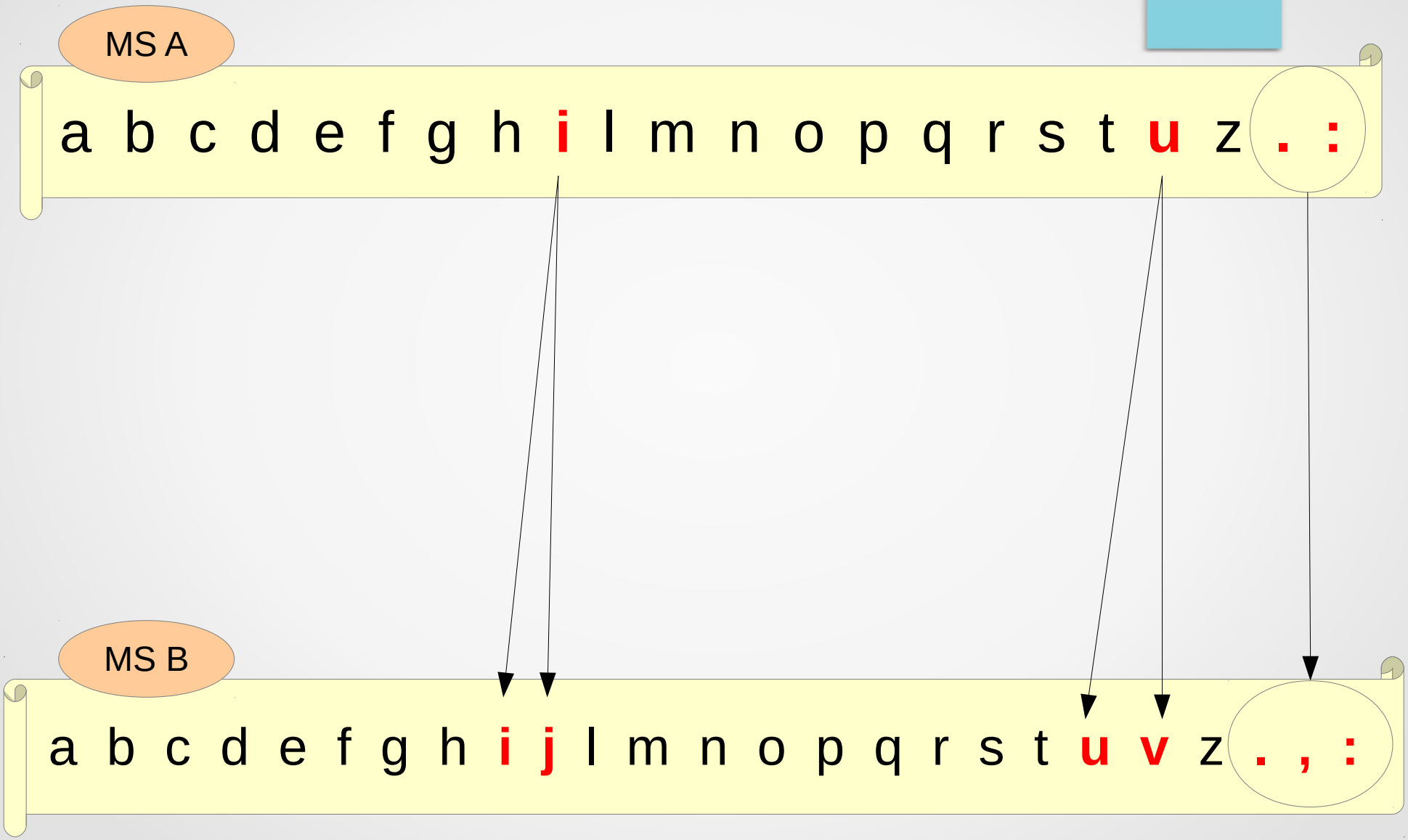
MS A

a b c d e f g h **i** l m n o p q r s t **u** z **. :**

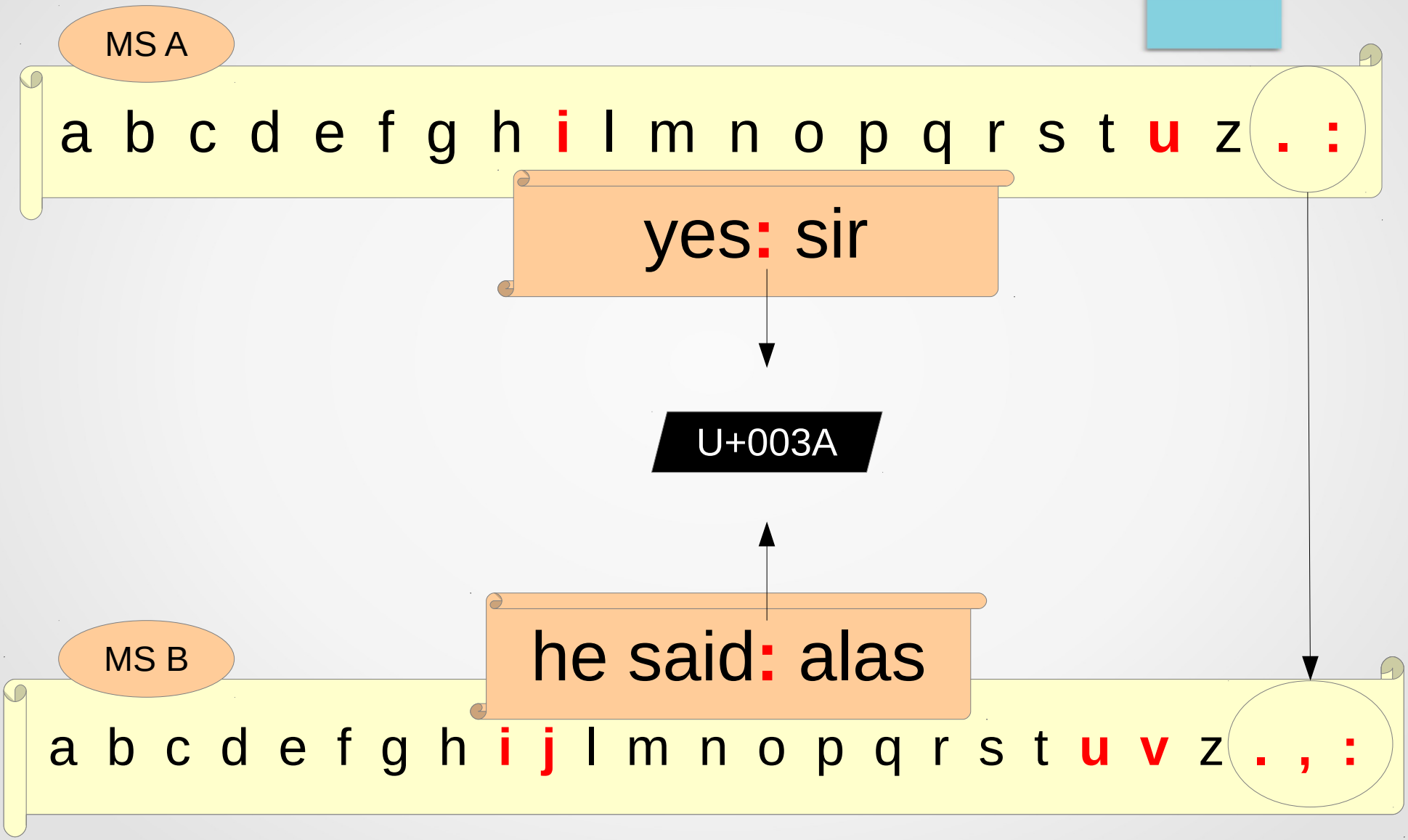
MS B

a b c d e f g h **i j** l m n o p q r s t **u v** z **. , :**

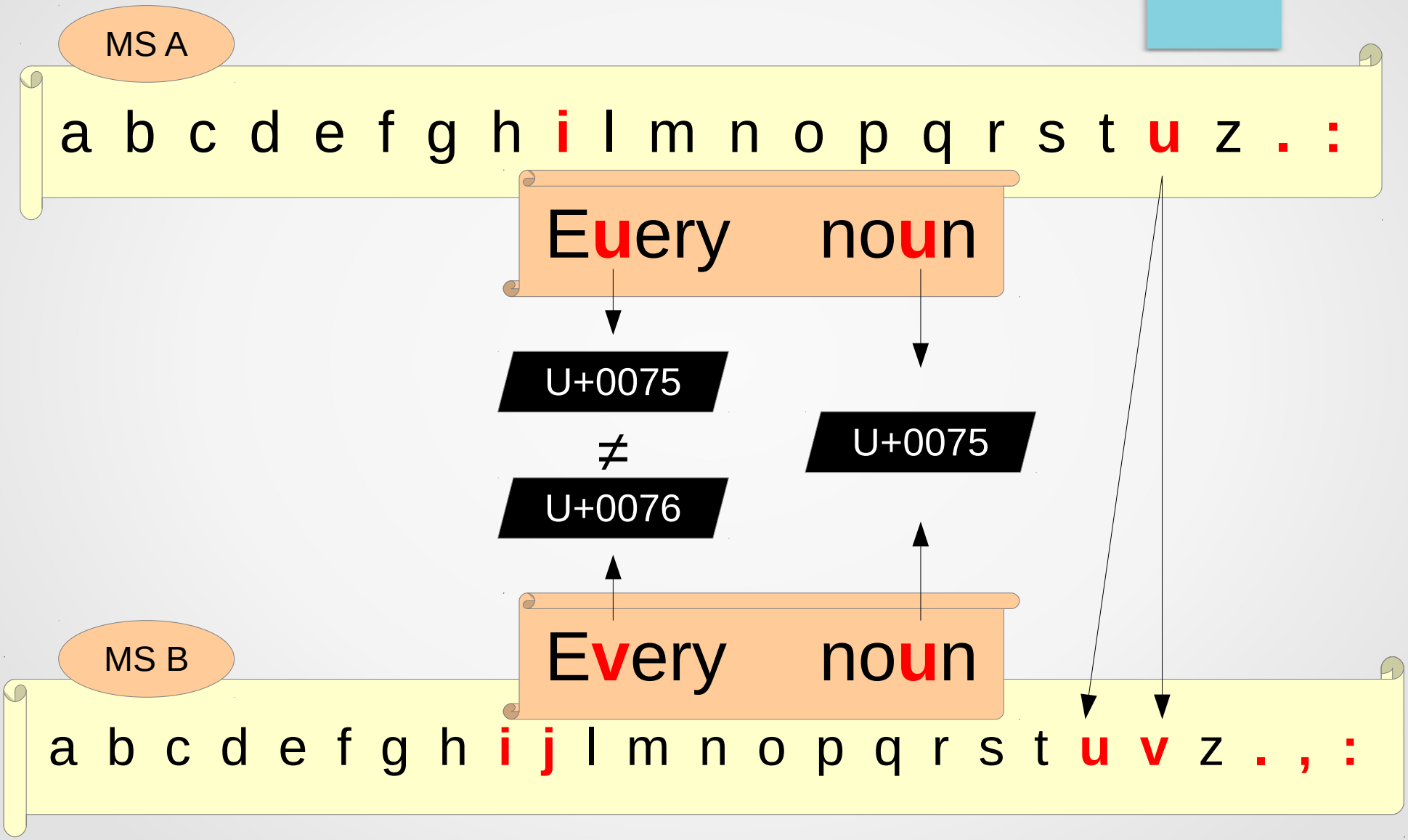
Saussure



Saussure

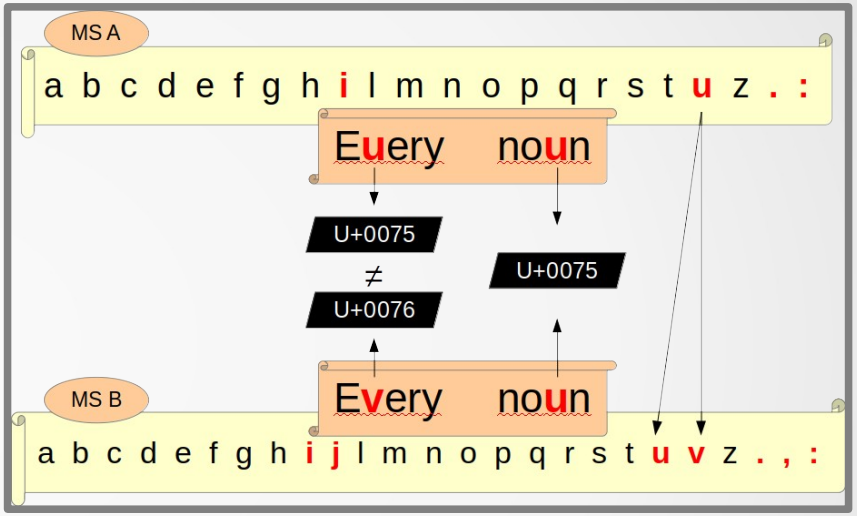

MS A
a b c d e f g h i l m n o p q r s t u z . :
yes: sir
U+003A
he said: alas
MS B
a b c d e f g h i j l m n o p q r s t u v z . , :

Saussure


MS A
a b c d e f g h i l m n o p q r s t u z . :
Euery noun
U+0075
≠
U+0076
U+0075
Every noun
MS B
a b c d e f g h i j l m n o p q r s t u v z . , :

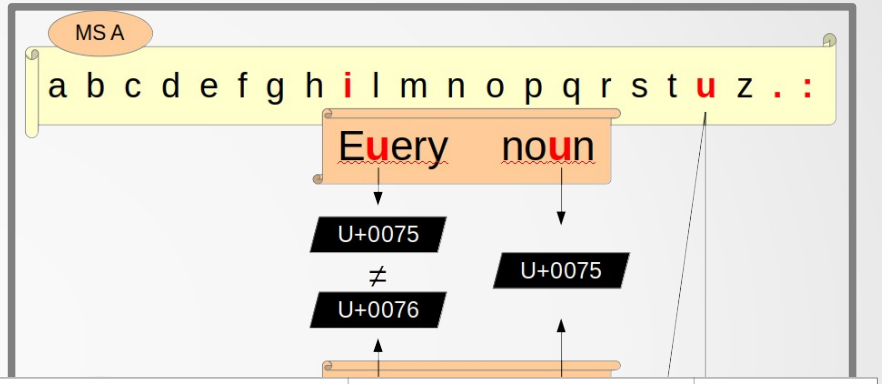
T. Orlandi: “table of signs”

- Tito Orlandi, *Informatica testuale. Teoria e prassi*, Laterza, 2010



T. Orlandi: “table of signs”

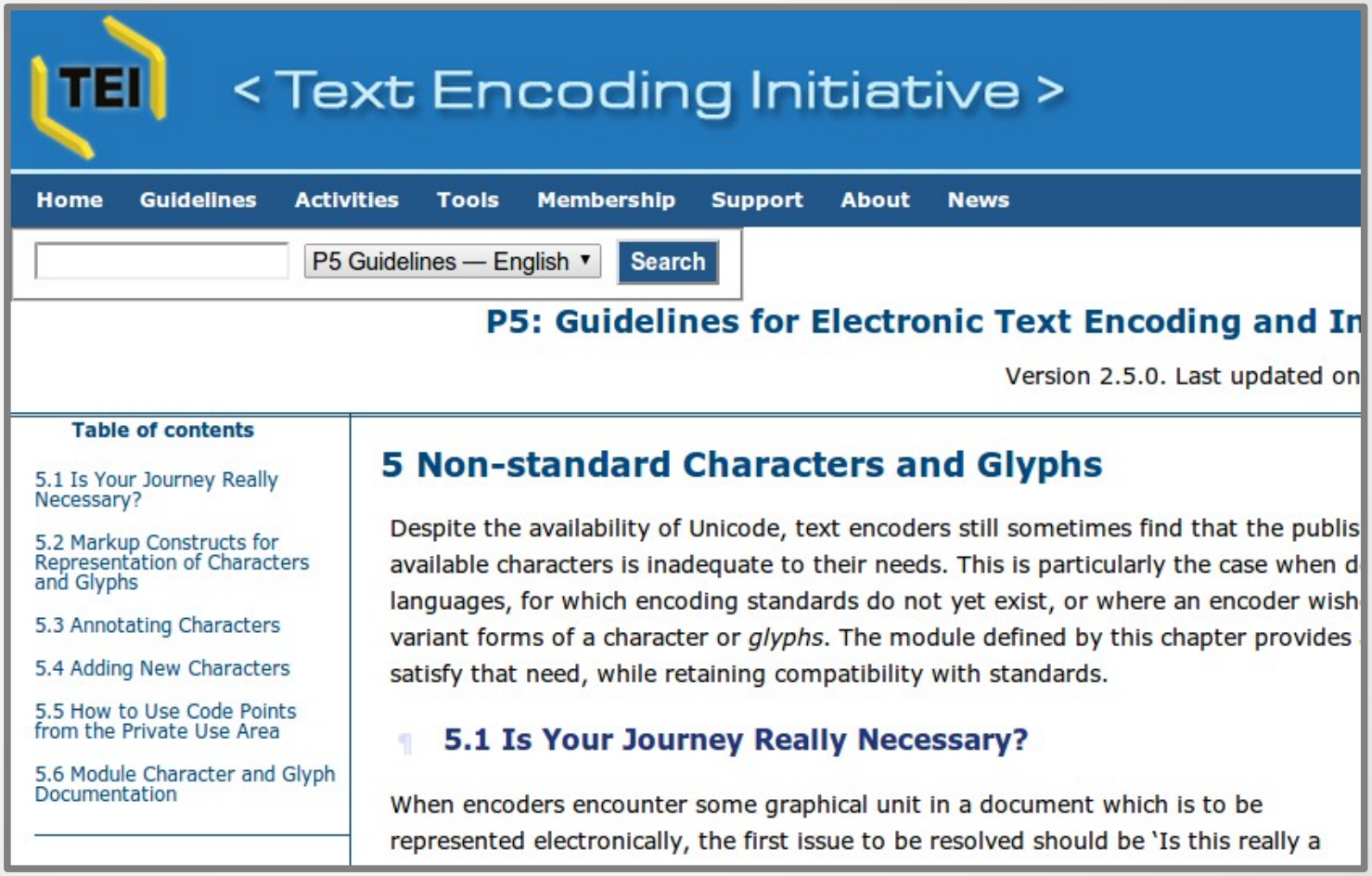
- Tito Orlandi, *Informatica testuale. Teoria e prassi*, Laterza, 2010



Graphemes	Alphabemes	Grapheme Visualization	Type	Notes
a	a	a	Alphabetic	
b	b	b	Alphabetic	
c	c	c	Alphabetic	
d	d	d	Alphabetic	
e	e	e	Alphabetic	
æ	ae	ę	Alphabetic	An 'e'
f	f	f	Alphabetic	

TEI XML → *Unicode*

- Unicode (a “u” is a “u”)



TEI XML → *Unicode*

- Unicode (a “u” is a “u”)

TEI XML → *Unicode*

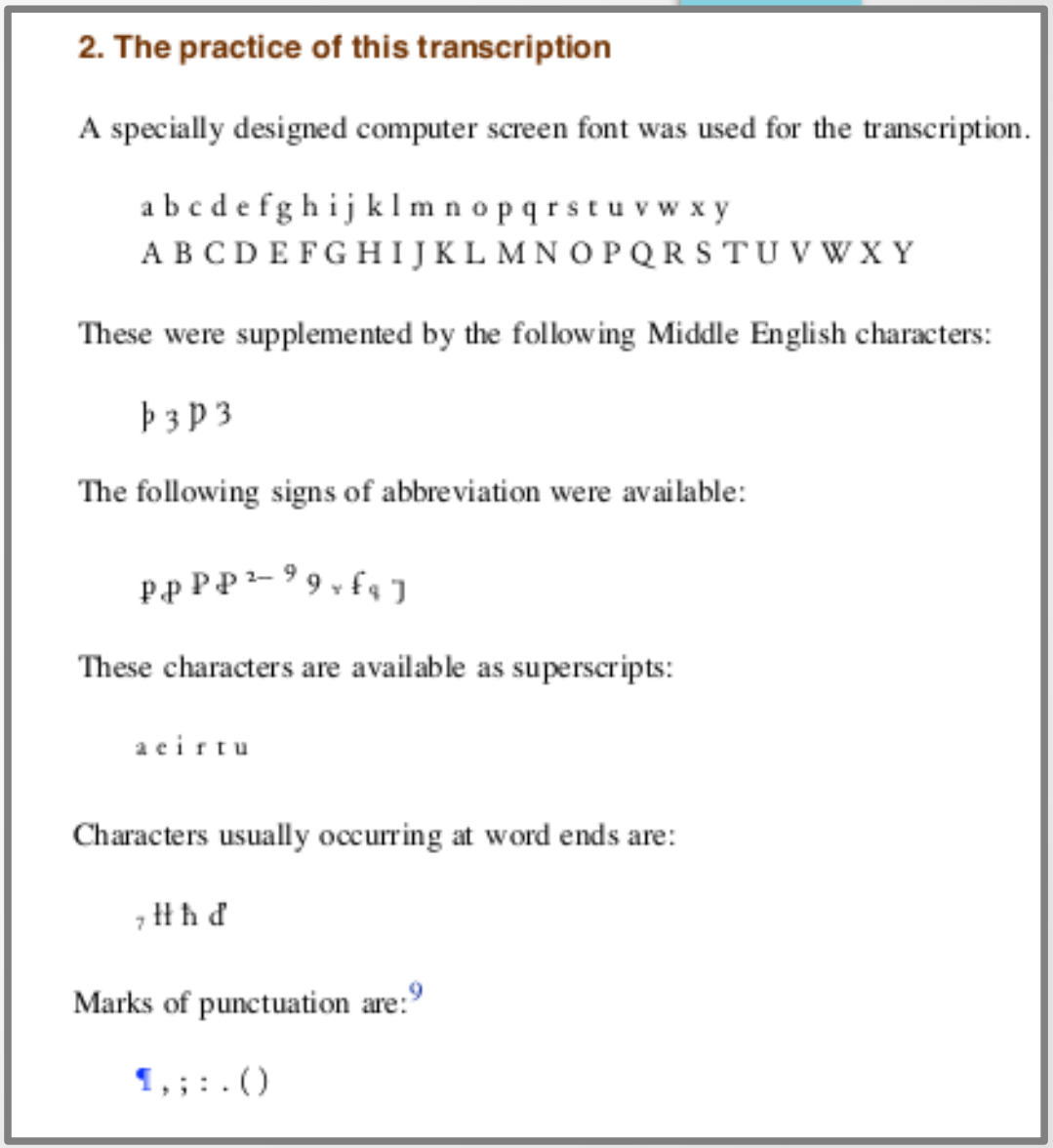
- Unicode
  (a “u” is a “u”)
- Corpus-wide
  normalization

TEI XML → *Unicode*

- Unicode
  (a “u” is a “u”)
- Corpus-wide
  normalization
  (Canterbury
  Tales)

TEI XML → *Unicode*

- Unicode
  (a “u” is a “u”)
- Corpus-wide normalization (Canterbury Tales)
- Documentation



2. The practice of this transcription

A specially designed computer screen font was used for the transcription.

a b c d e f g h i j k l m n o p q r s t u v w x y
A B C D E F G H I J K L M N O P Q R S T U V W X Y

These were supplemented by the following Middle English characters:

þ ȝ Þ Ȝ

The following signs of abbreviation were available:

ꝑ ꝓ Ꝑ Ꝓ ꝛ– ꝰ ꝯ ꝟ ꝗ ꝉ

These characters are available as superscripts:

a e i r t u

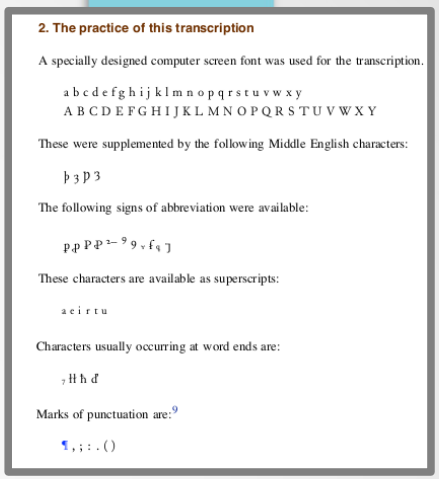
Characters usually occurring at word ends are:

ꞁ ꝉ ħ ꝺ

Marks of punctuation are:[9]

¶ , ; : . ( )

TEI XML → *Unicode*



2. The practice of this transcription

A specially designed computer screen font was used for the transcription.

a b c d e f g h i j k l m n o p q r s t u v w x y
A B C D E F G H I J K L M N O P Q R S T U V W X Y

These were supplemented by the following Middle English characters:

þ ȝ Þ Ȝ

The following signs of abbreviation were available:

ꝑ ꝓ Ꝑ Ꝓ ꝛ ꝰ ꝯ ꝰ ꝭ ꝗ ꝙ

These characters are available as superscripts:

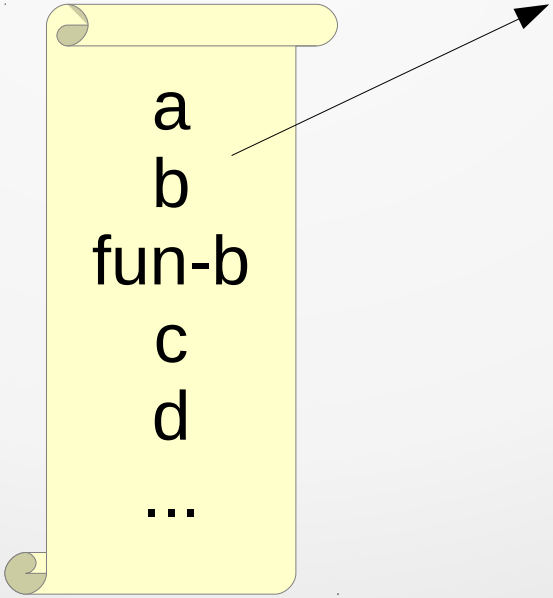
a e i r t u

Characters usually occurring at word ends are:

ꝫ ꝫ ꝫ ꝫ

Marks of punctuation are:[9]

¶ , ; : . ( )



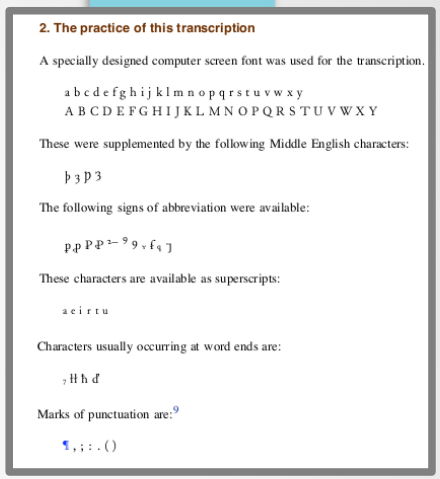
TEI XML → *Unicode*

<charDecl>

~~a~~

~~b~~

<glyph id="fun-b">



2. The practice of this transcription

A specially designed computer screen font was used for the transcription.

a b c d e f g h i j k l m n o p q r s t u v w x y
A B C D E F G H I J K L M N O P Q R S T U V W X Y

These were supplemented by the following Middle English characters:

þ ȝ Þ Ȝ

The following signs of abbreviation were available:

ꝑ ꝓ Ꝑ ꝓ ꝛ ꝰ ꝯ ꝰ ꝭ ꝗ ꝙ

These characters are available as superscripts:

a e i r t u

Characters usually occurring at word ends are:

ꝛ ꝉ ꝉ ꝺ

Marks of punctuation are:[9]

¶ , ; : . ( )

a
b
fun-b
c
d
...

Reforming TEI? <*charDecl*>

<charDecl>
~~a~~
~~b~~
<glyph id="fun-b">

Guide-lines

<charDecl>
<char xml:id="a">
<char xml:id="b">
<glyph id="fun-b">

a
b
fun-b
c
d
...

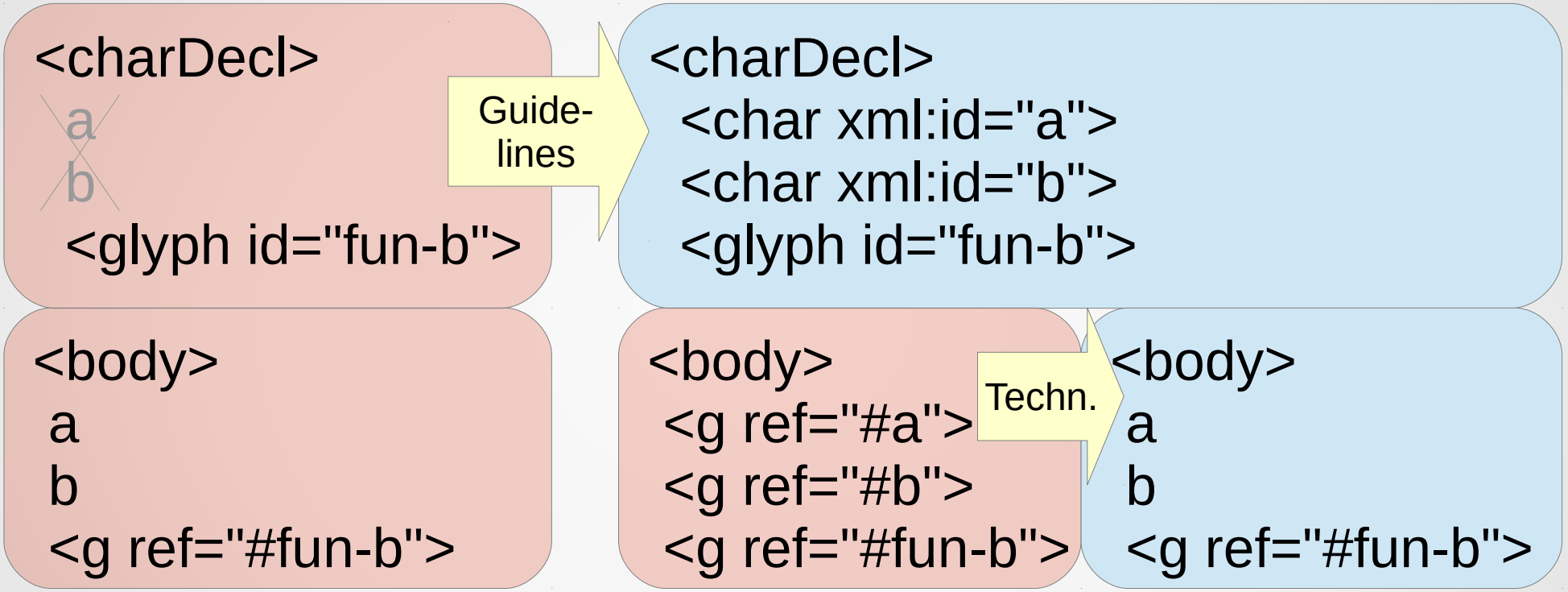
Reforming TEI? *Mapping*

```
<charDecl>
  a
  b
  <glyph id="fun-b">
```

Guide-lines

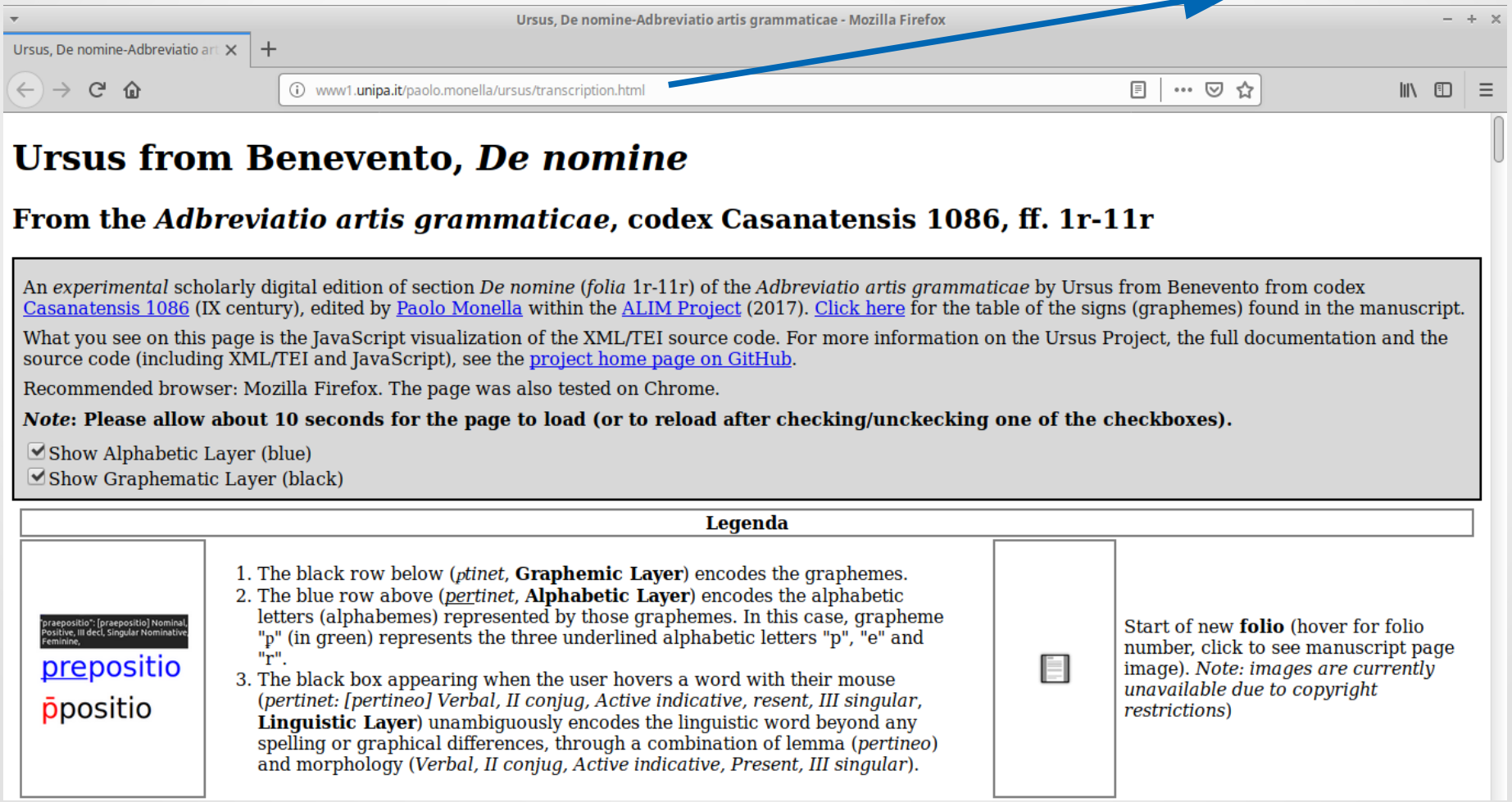
```
<charDecl>
  <char xml:id= "v">
    <charName>Premodern Latin
      uncial lowercase v</charName>
    <charProp>
      <localName>Expression
      </localName>
        <value>U+0076</value>
      <localName>Content
      </localName>
        <value>v</value>
    <charProp>
    <mapping type="Unicode">v
    </mapping>
    <graphic url="v.jpg"/
  </char>
<charDecl>
```

Reforming TEI? *No <g>*



Ursus edition

http://www.unipa.it/paolo.monella/ursus



GToS: *Graphematic Table of Signs*

Grapheme	Alphabeme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs depending on the ligature. For example, allograph e2.png is often found in the common ligature "&" for "et".	e1.png e2.png	1	
æ	ae	e	Alphabetic	This is an "e caudatum": an "e" grapheme with a tail on the bottom left or left side. It is often (but not always) used to represent the diphthong "ae", which is often represented by a simple "e" grapheme elsewhere in the manuscript. It has one allograph (see image ecaudatum1.png) with the tail underneath the body of the grapheme, turning right, and a second allograph (ecaudatum2.png) with the tail on the left of the body (mostly in ligatures).	ecaudatum1.png ecaudatum2.png	1	1r.a.9

http://www1.unipa.it/paolo.monella/ursus/GToS.csv
https://github.com/paolomonella/ursus/blob/master/GToS.csv

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs	e1.png e2.png	1	
æ	ae	e	Alphabetic	This is an "e caudatum": an "e" grap	ecaudatum1.png e	1	1r.a.9
f	f	f	Alphabetic		f.png	1	
g	g	g	Alphabetic		g.png	1	1r.a.7
h	h	h	Alphabetic		h.png	1	
i	i	i	Alphabetic		i.png	1	
k	k	k	Alphabetic	First found in 5r.a.9 when translitera	k.png	1	5r.a.9
l	l	l	Alphabetic		l.png	1	
m	m	m	Alphabetic		m.png	1	
n	n	n	Alphabetic		n.png	1	
o	o	o	Alphabetic		o.png	1	
p	p	p	Alphabetic		p.png	1	
q	q	q	Alphabetic		q.png	1	
r	r	r	Alphabetic		r.png	1	

http://www1.unipa.it/paolo.monella/ursus/GToS.csv
https://github.com/paolomonella/ursus/blob/master/GToS.csv

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs	e1.png e2.png	1	
æ	ae	e	Alphabetic	This is an "e caudatum": an "e" grap	ecaudatum1.png e	1	1r.a.9
f	f	f	Alphabetic		f.png	1	
g	g	g	Alphabetic		g.png	1	1r.a.7
h	h	h	Alphabetic		h.png	1	
i	i	i	Alphabetic		i.png	1	
k	k	k	Alphabetic	First found in 5r.a.9 when translitera	k.png	1	5r.a.9
l	l	l	Alphabetic		l.png	1	
m	m	m	Alphabetic		m.png	1	
n	n	n	Alphabetic		n.png	1	
o	o	o	Alphabetic		o.png	1	
p	p	p	Alphabetic		p.png	1	
q	q	q	Alphabetic		q.png	1	
r	r	r	Alphabetic		r.png	1	

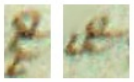
http://www1.unipa.it/paolo.monella/ursus/GToS.csv
https://github.com/paolomonella/ursus/blob/master/GToS.csv

GToS: *Graphematic Table of Signs*

Graphemic table of signs - Mozilla Firefox

Graphemic table of signs

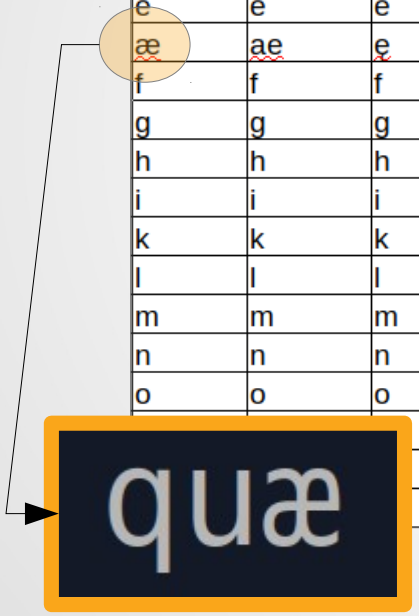
file:///home/ilbuonme/siti/paolo.monella/ursus/gtos.html

Grapheme	Alphabeme(s)	Grapheme visualization	Type	Notes	Image(s)
a	a	a	Alphabetic		
b	b	b	Alphabetic		
c	c	c	Alphabetic		
d	d	d	Alphabetic		
e	e	e	Alphabetic	The grapheme has many allographs depending on the ligature. For example, allograph e2.png is often found in the common ligature "&" for "et".	
æ	ae	ę	Alphabetic	This is an "e caudatum": an "e" grapheme with a tail on the bottom left or left side. It is often (but not always) used to represent the diphthong "ae", which is often represented by a simple "e" grapheme elsewhere in the manuscript. It has one allograph (see image ecaudatum1.png) with the tail underneath the body of the grapheme, turning right, and a second allograph (ecaudatum2.png) with the tail on the left of the	

http://www1.unipa.it/paolo.monella/ursus/GToS.csv
https://github.com/paolomonella/ursus/blob/master/GToS.csv

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs	e1.png e2.png	1	
æ	ae	e	Alphabetic	This is an "e caudatum": an "e" grap	ecaudatum1.png e	1	1r.a.9
f	f	f	Alphabetic		f.png	1	
g	g	g	Alphabetic		g.png	1	1r.a.7
h	h	h	Alphabetic		h.png	1	
i	i	i	Alphabetic		i.png	1	
k	k	k	Alphabetic	First found in 5r.a.9 when translitera	k.png	1	5r.a.9
l	l	l	Alphabetic		l.png	1	
m	m	m	Alphabetic		m.png	1	
n	n	n	Alphabetic		n.png	1	
o	o	o	Alphabetic		o.png	1	
			Alphabetic		p.png	1	
			Alphabetic		q.png	1	
			Alphabetic		r.png	1	



http://www1.unipa.it/paolo.monella/ursus/GToS.csv
https://github.com/paolomonella/ursus/blob/master/GToS.csv

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
a	a	a	Alphabetic
b	b	b	Alphabetic
c	c	c	Alphabetic
d	d	d	Alphabetic
e	e	e	Alphabetic
æ	ae	ę	Alphabetic
f	f	f	Alphabetic
g	g	g	Alphabetic
h	h	h	Alphabetic
i	i	i	Alphabetic

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
k	k	k	Alphabetic
l	l	l	Alphabetic
m	m	m	Alphabetic
n	n	n	Alphabetic
o	o	o	Alphabetic
p	p	p	Alphabetic
q	q	q	Alphabetic
r	r	r	Alphabetic
s	s	s	Alphabetic
t	t	t	Alphabetic

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
u	u	u	Alphabetic
x	x	x	Alphabetic
y	y	y	Alphabetic
z	z	z	Alphabetic
			Space
.		.	Punctuation
.		.	Punctuation
|		,	Punctuation
~		~	Punctuation
:		:	Punctuation

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
ł		;	Punctuation
ï		¨ ;	Punctuation
¦ ¦		: ;	Punctuation
8		◌̂	Abbreviation mark
2		◌̀	Abbreviation mark
7		◌̄	Abbreviation mark

GToS: *Graphematic Table of Signs*

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
–		◌̄	Abbreviation mark
;		;	Abbreviation mark
ų	u	◌̆	Abbreviation mark
þ	per	ꝑ	Brevigraph
ŋ	pro	ꝓ	Brevigraph
¢	qui	ꝗ	Brevigraph
÷	est	⁘	Logograph

GToS processing: *Grapheme visualization*

```
function graph(myString) {
    // This function inputs a string of characters representing graphemes
    // in the XML/TEI source file and returns a string of characters
    // that will be used to display the graphemes in the browser
    for (var igr = 0; igr < gtos.length; igr++) {
        // gtos[igr][0] (value of first column of GToS.csv):
        //  character representing the grapheme in the XML/TEI file
        // gtos[igr][1] (value of second column of GToS.csv):
        //  character representing the alphabeme
        if (gtos[igr][0] != '' && gtos[igr][2] != '') {
            myString = myString.replace(gtos[igr][0], gtos[igr][2]);
        }
    }
    return myString;
}
```

GToS processing: *Grapheme visualization*

```
function graph(myString) {
    for (var igr = 0; igr < gtos.length; igr++) {
        if (gtos[igr][0] != '' && gtos[igr][2] != '') {
            myString = myString.replace(
            gtos[igr][0],
            gtos[igr][2]);
        }
    }
    return myString;
}
```

GToS processing: *Grapheme visualization*

quæ

```
function graph(myString) {
    for (var igr = 0; igr < gtos.length; igr++) {
        if (gtos[igr][0] != '' && gtos[igr][2] != '') {
            myString = myString.replace(
            gtos[igr][0],
            gtos[igr][2]);
        }
    }
    return myString;
}
```

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
æ	ae	ę	Alphabetic

GToS processing: *Grapheme visualization*

quæ

```
function graph(myString) {
    for (var igr = 0; igr < gtos.length; igr++) {
        if (gtos[igr][0] != '' && gtos[igr][2] != '') {
            myString = myString.replace(
            gtos[igr][0],
            gtos[igr][2]);
        }
    }
    return myString;
}
```

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
æ	ae	ę	Alphabetic

GToS processing: *Grapheme visualization*

quæ

quae

quę

```
function graph(myString) {
    for (var igr = 0; igr < gtos.length; igr++) {
        if (gtos[igr][0] != '' && gtos[igr][2] != '') {
            myString = myString.replace(
            gtos[igr][0],
            gtos[igr][2]);
        }
    }
    return myString;
}
```

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
æ	ae	ę	Alphabetic

GToS processing: *Alphabeme derivation*

```
function alph(myString) {
    // This function inputs a string of characters representing graphemes
    // (myString) and returns a string of characters representing alphabemes,
    // mapping the graphemes to their standard alphabetical meaning as encode
    // in the first and second column of the "Graphemic Table of Signs" (GToS).
    for (var ialph = 0; ialph < gtos.length; ialph++) {
        // gtos[ialph][0] (value of first column of GToS.csv):
        //  character representing the grapheme in the XML/TEI file
        // gtos[ialph][1] (value of second column of GToS.csv):
        //  character representing the alphabeme
        if ( gtos[ialph][0] != ''
            && gtos[ialph][1] != ''
            && gtos[ialph][0] != gtos[ialph][1] ) {
            myString = myString.replace(gtos[ialph][0], gtos[ialph][1]);
        }
    }
    return myString;
}
```

GToS processing: *Alphabeme derivation*

```
function alph(myString) {
    for (var ialph = 0; ialph < gtos.length; ialph++) {
        if ( gtos[ialph][0] != ''
            && gtos[ialph][1] != ''
            && gtos[ialph][0] != gtos[ialph][1] ) {
            myString = myString.replace(
                gtos[ialph][0],
                gtos[ialph][1]);
        }
    }
    return myString;
}
```

GToS processing: *Alphabeme derivation*

quæ

```
function alph(myString) {
    for (var ialph = 0; ialph < gtos.length; ialph++) {
        if ( gtos[ialph][0] != ''
            && gtos[ialph][1] != ''
            && gtos[ialph][0] != gtos[ialph][1] ) {
            myString = myString.replace(
                gtos[ialph][0],
                gtos[ialph][1]);
        }
    }
    return myString;
}
```

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
æ	ae	ę	Alphabetic

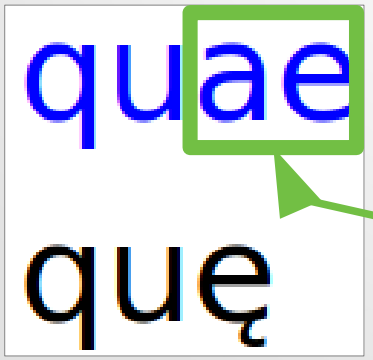
GToS processing: *Alphabeme derivation*

quæ

```
function alph(myString) {
    for (var ialph = 0; ialph < gtos.length; ialph++) {
        if ( gtos[ialph][0] != ''
            && gtos[ialph][1] != ''
            && gtos[ialph][0] != gtos[ialph][1] ) {
            myString = myString.replace(
                gtos[ialph][0],
                gtos[ialph][1]);
        }
    }
    return myString;
}
```

Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
æ	ae	ę	Alphabetic

GToS processing: *Alphabeme derivation*




```
function alph(myString) {
    for (var ialph = 0; ialph < gtos.length; ialph++) {
        if ( gtos[ialph][0] != ''
            && gtos[ialph][1] != ''
            && gtos[ialph][0] != gtos[ialph][1] ) {
            myString = myString.replace(
                gtos[ialph][0],
                gtos[ialph][1]);
        }
    }
    return myString;
}
```

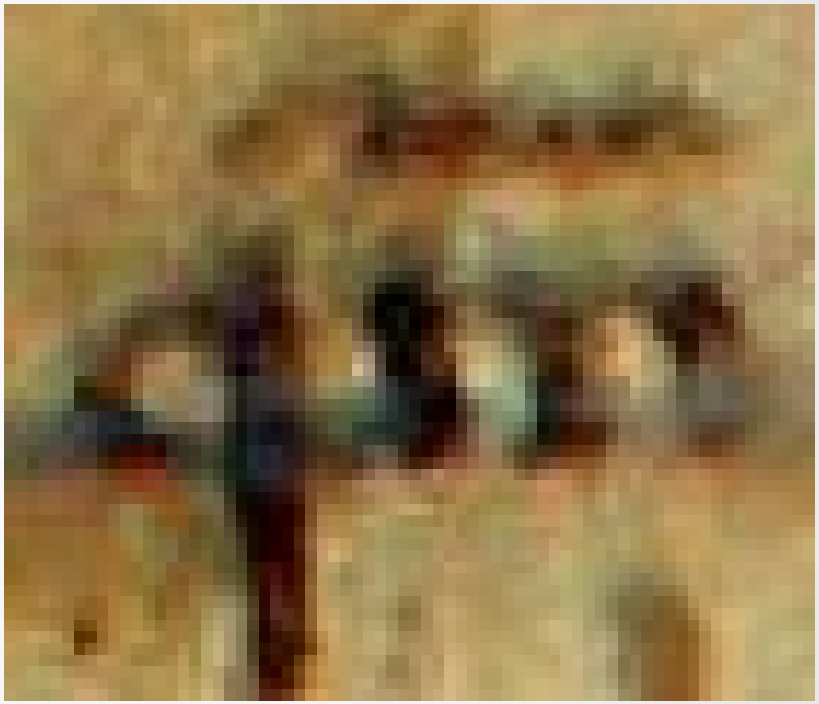
Gra-pheme	Alpha-beme(s)	Grapheme visualiza-tion	Type
æ	ae	ę	Alphabetic

Abbreviations: *Alphabeme encoding*

```
<choice>
    <abbr type="superscription">
        qm<am>̄</am>
    </abbr>
    <expan>
        quoniam
    </expan>
</choice>
```

quoniam

qm̄

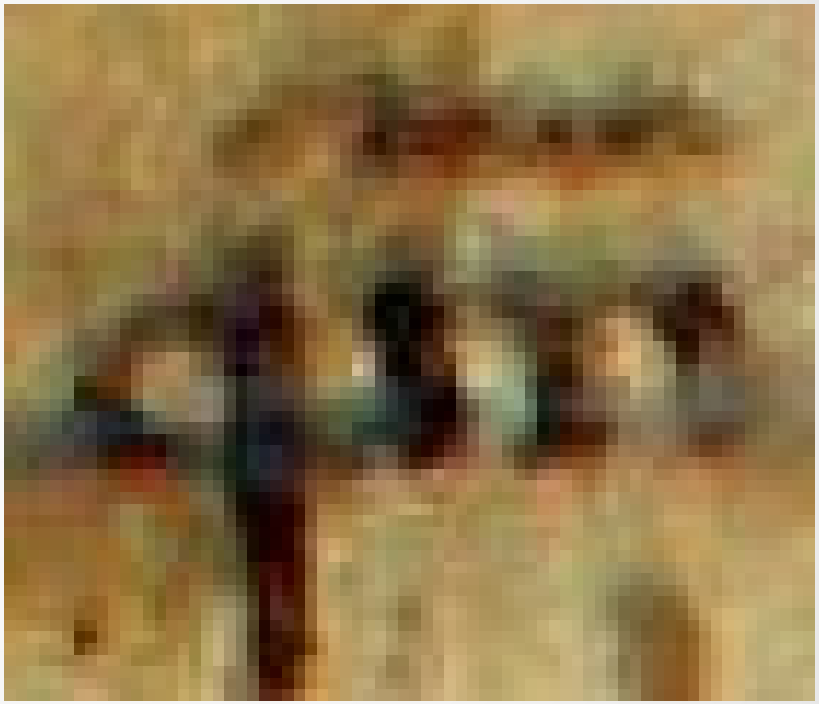


Abbreviations: *Alphabeme encoding*

```
<choice>
    <abbr>qm̄</abbr>
    <expan>quoniam</expan>
</choice>
```

quoniam

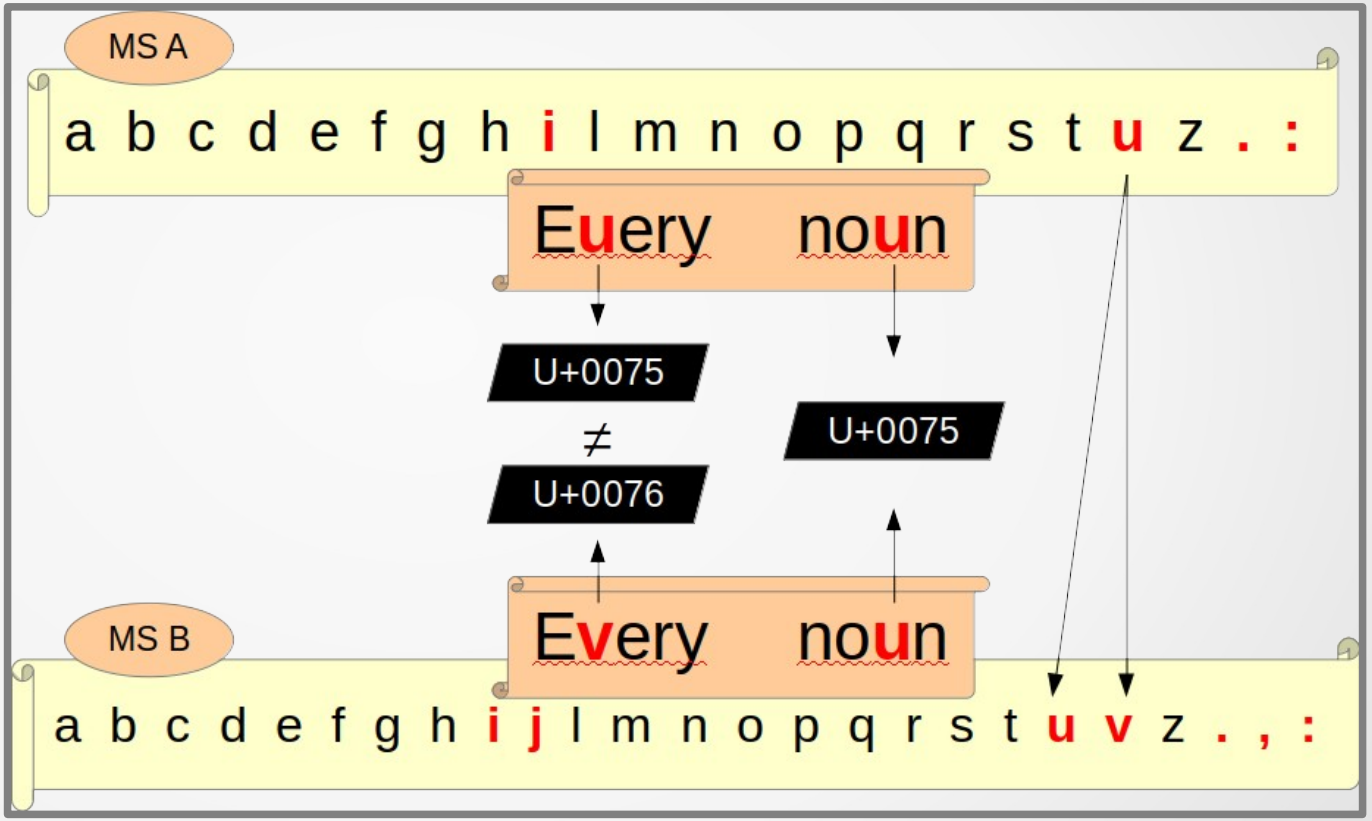
qm̄



Digital encoding of handritten writing systems

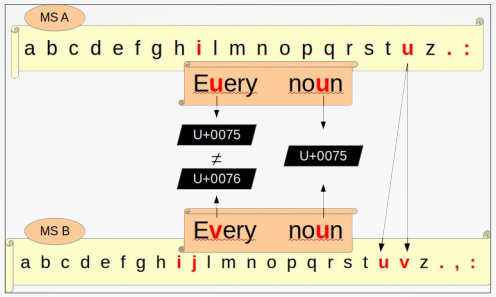
Digital encoding of handritten writing systems

- Saussure



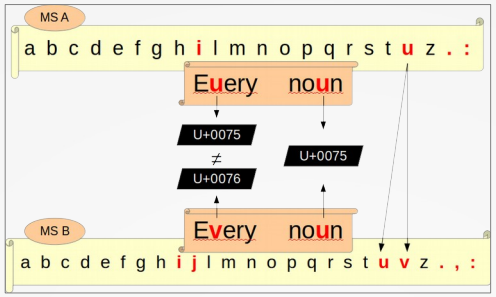
Digital encoding of handritten writing systems

- Saussure



Digital encoding of handritten writing systems

- Saussure

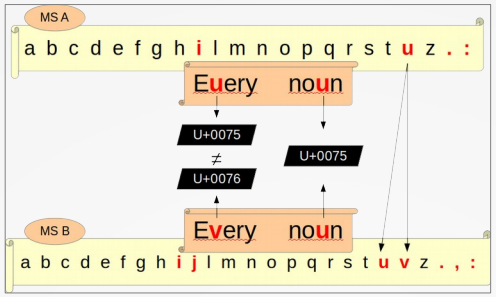


- Table of Signs

Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs	e1.png e2.png	1	
æ	ae	ę	Alphabetic	This is an "e caudatum": an "e" grap	ecaudatum1.png e	1	1r.a.9
f	f	f	Alphabetic		f.png	1	
g	g	g	Alphabetic		g.png	1	1r.a.7
h	h	h	Alphabetic		h.png	1	
i	i	i	Alphabetic		i.png	1	
k	k	k	Alphabetic	First found in 5r.a.9 when translitera	k.png	1	5r.a.9
l	l	l	Alphabetic		l.png	1	
m	m	m	Alphabetic		m.png	1	
n	n	n	Alphabetic		n.png	1	
o	o	o	Alphabetic		o.png	1	
p	p	p	Alphabetic		p.png	1	
q	q	q	Alphabetic		q.png	1	
r	r	r	Alphabetic		r.png	1	

Digital encoding of handritten writing systems

- Saussure

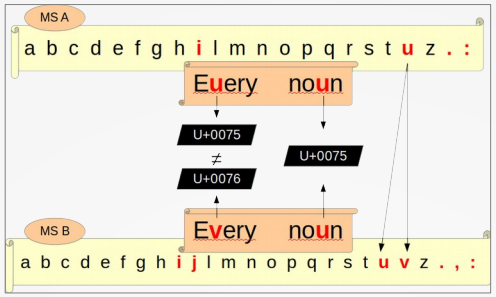


- Table of Signs

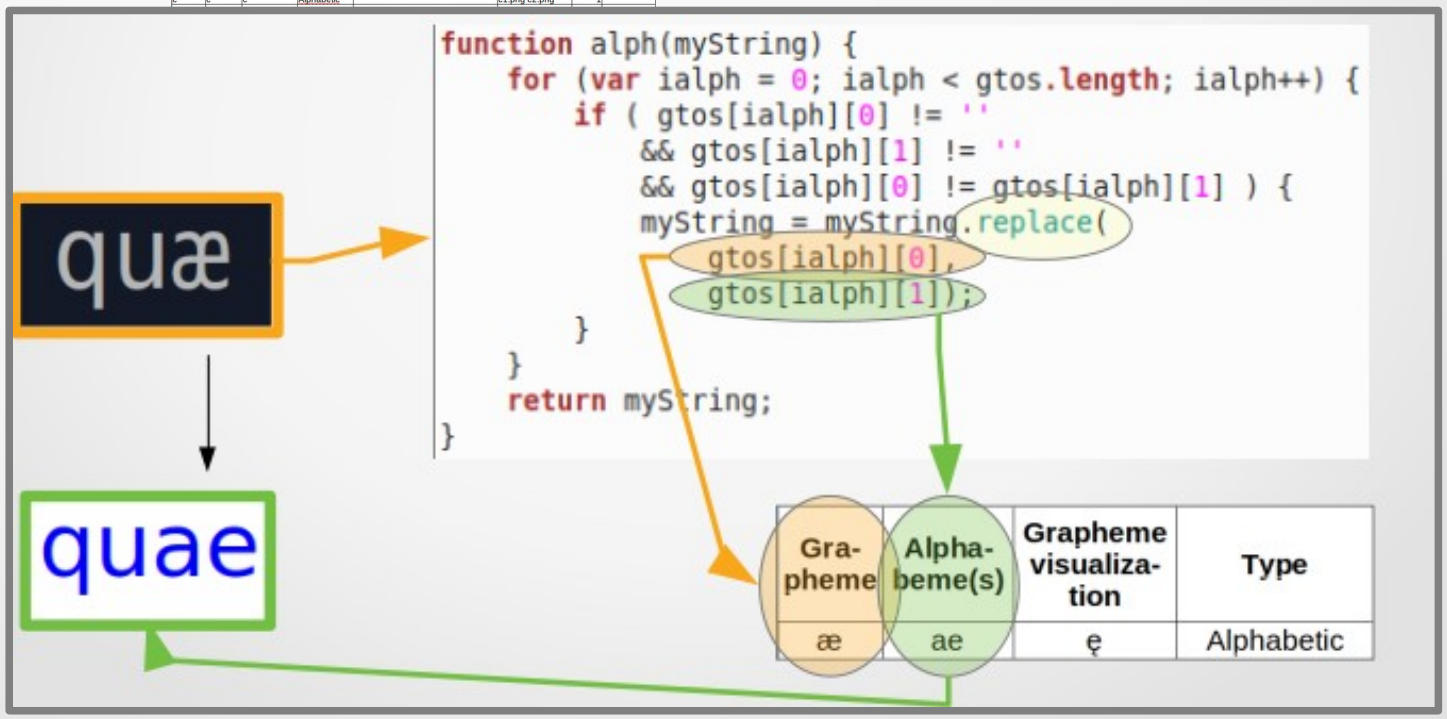
Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs	e1.png e2.png	1	
æ	ae	æ	Alphabetic	This is an "e caudatum": an "e" grap	ecaudatum1.png e	1	1r.a.9
f	f	f	Alphabetic		f.png	1	
g	g	g	Alphabetic		g.png	1	1r.a.7
h	h	h	Alphabetic		h.png	1	
i	i	i	Alphabetic		i.png	1	
k	k	k	Alphabetic	First found in 5r.a.9 when translitera	k.png	1	5r.a.9
l	l	l	Alphabetic		l.png	1	
m	m	m	Alphabetic		m.png	1	
n	n	n	Alphabetic		n.png	1	
o	o	o	Alphabetic		o.png	1	
p	p	p	Alphabetic		p.png	1	
q	q	q	Alphabetic		q.png	1	
r	r	r	Alphabetic		r.png	1	

Digital encoding of handritten writing systems

- Saussure
- Table of Signs
- Processing

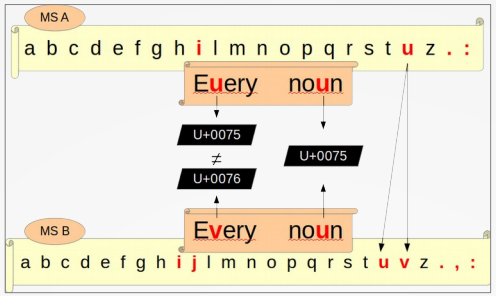


Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	



Digital encoding of handritten writing systems

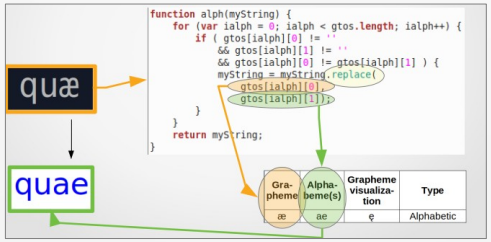
- Saussure



- Table of Signs

Gra-pheme	Alpha-beme(s)	Grapheme visualization	Type	Notes	Image(s)	Found	Found Where
a	a	a	Alphabetic		a.png	1	
b	b	b	Alphabetic		b.png	1	
c	c	c	Alphabetic		c1.png c2.png	1	
d	d	d	Alphabetic		d.png	1	
e	e	e	Alphabetic	The grapheme has many allographs	e1.png e2.png	1	
æ	ae	ę	Alphabetic	This is an "e caudatum": an "e" grap	ecaudatum1.png #	1	1r.a.9
f	f	f	Alphabetic		f.png	1	
g	g	g	Alphabetic		g.png	1	1r.a.7
h	h	h	Alphabetic		h.png	1	
i	i	i	Alphabetic		i.png	1	
k	k	k	Alphabetic	First found in 5r.a.9 when transliterat	k.png	1	5r.a.9
l	l	l	Alphabetic		l.png	1	
m	m	m	Alphabetic		m.png	1	
n	n	n	Alphabetic		n.png	1	
o	o	o	Alphabetic		o.png	1	
p	p	p	Alphabetic		p.png	1	
q	q	q	Alphabetic		q.png	1	
r	r	r	Alphabetic		r.png	1	

- Processing



Paolo Monella

La tabella dei segni
nell'edizione scientifica digitale:
un approccio saussuriano

The table of signs
in the scholarly digital edition:
a Saussurian approach

paolo.monella@unipa.it
http://www1.unipa.it/paolo.monella/calabriatabella2019/index.html

Philosophy of Language and Digital Humanities
University of Calabria (Rende, CS, Italy)
Saussurian tradition session
May 8, 2019

CC BY SA