

Paolo Monella  
paolo.monella@unipa.it

# Fuori dalla torre di Babele: interoperabilità e sistemi grafici pre-moderni



*Out of the Tower of Babel :  
interoperability and pre-modern  
writing systems*

December 4, 2019

[ve]dph

Venice Centre for  
Digital and Public  
Humanities





**In a nutshell**



# In a nutshell

- Digital scholarly editions: diplomatic/normalized
- Hjelmslev's "analysis" as digital modelling for pre-modern writing systems
- Open issues
- Interoperability through modelling



# My focus

- *European Medieval handwriting*
- Pre-Gutenberg
  - Handwriting
  - Early print → imitating handwriting
- Alphabetic writing systems
  - **Latin** script (Italian, English...), Greek, Cyrillic...
  - No Cuneiform, Arabic, Chinese etc.

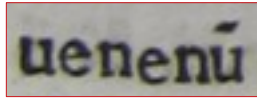
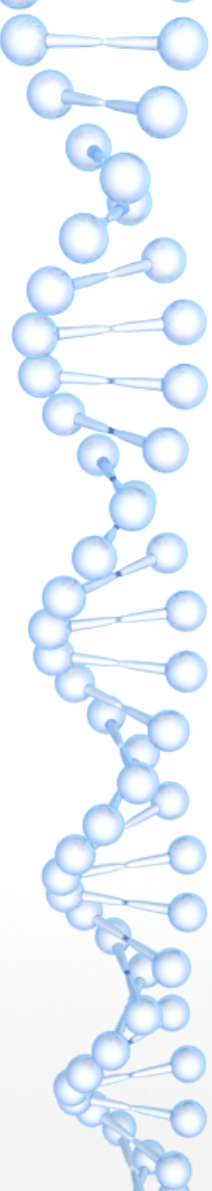


**Diplomatic/normalized**

# Diplomatic/normalized

uenenũ

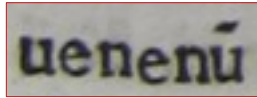
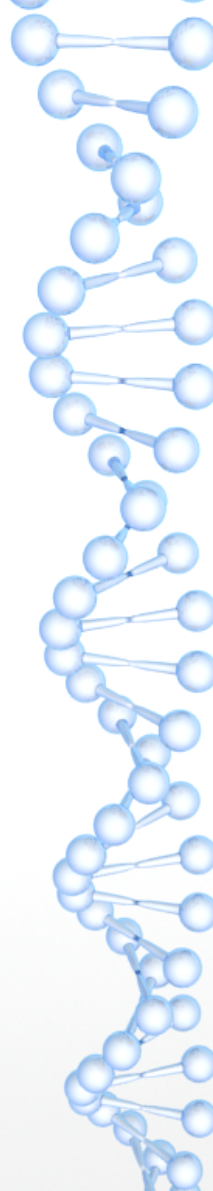
# Diplomatic/normalized



uenenū

- uenenū

# Diplomatic/normalized



- uenenū

Diplomatic

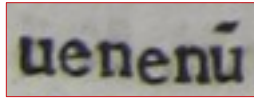
- Historical documentation
- Visualization





# Diplomatic/normalized

- venenum



uenenū

- uenenū

Diplomatic

- Historical documentation
- Visualization

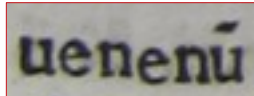


# Diplomatic/normalized

- venenum

Normalized

- Processing
  - Search
  - Collation
  - NLP (lemma, PoS etc.)
  - Statistics (distant reading)...



uenenū

- uenenū

Diplomatic

- Historical documentation
- Visualization

# Diplomatic/normalized

The Digital Vercelli Book **beta**

Magnifier Q HotSpot O TextLink S

DOTR 104v Diplomatic

7 **h**Wæt ic swefna cýrt secgan wylle hæf mege mætte  
8 to midre nihte syððan reord berend reste wunedon.  
9 þuhte me þæt ic ge sawe syllicre treow onlyft  
10 lædan leohte be wunden beama beorhtost call þæt  
11 beacen wæs be goten mid golde gimmar stodon fegere  
12 æt foldan sceatum . swylce þær life wæron uppe  
13 on þam eaxe ge sþanne be heoldon þær engel dryht  
14 nes ealle fegere þurh forð ge sceaft newær ðærhuru  
15 fracodes gealga . æchine þær beheoldon halige galtar  
16 men ofer moldan ƿeall þeor mære ge sceaft :  
17 Syllic wæs se pige beam ƿic pynnū fah for wounded  
18 mid wommū ge seah ic wuldres treow. wædum geweor  
19 ðode wynnū scinan ge gyred mid golde gimmar hæfdon  
20 be wrigene weorðlice wealdes treow : hwæðre ic  
21 þurh þæt gold ongytan meahthe eamra ærgewin  
22 þæt hit ærest ongan swatan on þa fwiðran healf

1:1 27%

7 **h**Wæt ic swefna

<http://vbd.humnet.unipi.it/beta/#104v>

# Diplomatic/normalized

The Digital Vercelli Book beta

Magnifier 🔍 HotSpot ○ TextLink ℹ️

DOTR 104v Interpretative 🔍

Diplomatic  
Interpretative

7 Hwæt ic swefna cyst secgan wylle hæt me gemætte  
8 to midre nihte syðan reordberend reste wunedon.  
9 þuhte me hæt ic gesawe syllicre treow on lyft  
10 lædan leohte bewunden beama beorhtost eall þæt  
11 beacen wæs begoten mid golde gimmas stodon fægere  
12 æt foldan sceatum . swylce þær fife wæron uppe  
13 on þam eaxlegespanne beheoldon þær engel dryht  
14 nes ealle fægere þurh forðgescaft ne wæs ðær huru  
15 fracodes gealga . ac hine þær beheoldon halige gastas  
16 men ofer moldan and eall þeos mære gescaft .  
17 Syllic wæs se sigebeam and ic synnum fah forwunded  
18 mid womnum geseah ic wuldres treow . wædum geweor  
19 ðode wynnum scinan gegyred mid golde gimmas hæfdon  
20 bewigene weorðlice wealdes treow . hwæðre ic  
21 þurh hæt gold ongytan meahthe earmra ærgewin  
22 hæt hit ærest ongan swætan on þa swiðran healfc

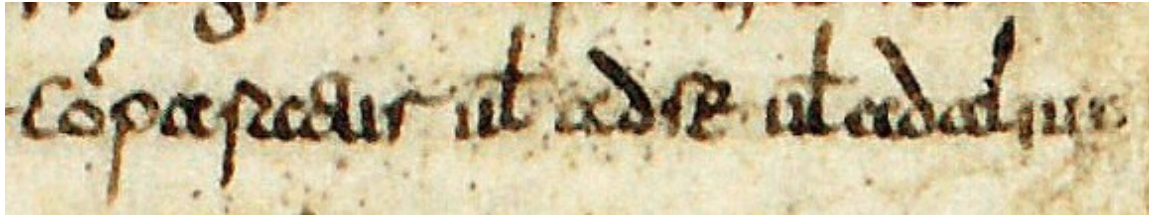
1:1 27%

7 Hwæt ic swefna

<http://vbd.humnet.unipi.it/beta/#104v>

# Diplomatic/normalized

- Comparatur vel ad se vel ad alium



- cōparaEur      uť   adfe   uťadaliu

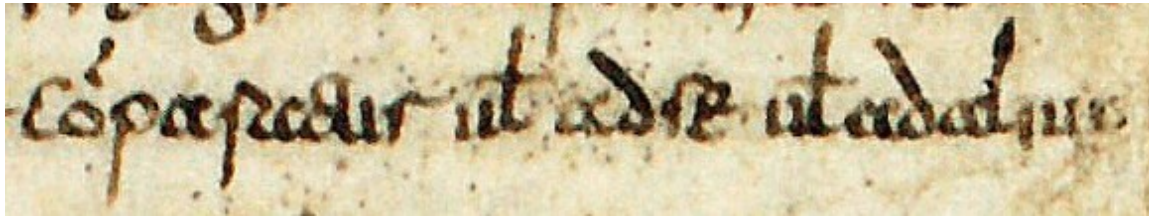
A white speech bubble with a dark blue background. The bubble has a rectangular body and a triangular tail pointing downwards and to the left.

**What's “normal”?**



# What's "normal"?

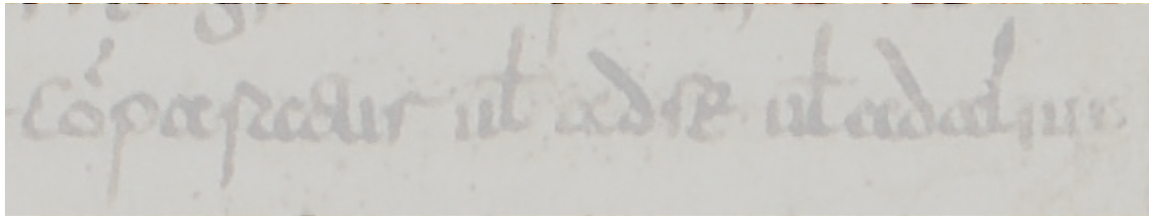
- Comparatur vel ad se vel ad alium



- cōparaEur      uł   adfe   uładaliu

# What's “normal”?

- Comparatur vel ad se vel ad alium



- cōparaEur      uť   adfe   uťadaliu



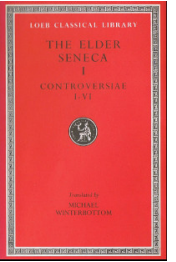


# What's “normal”?

- Comparatur vel ad se vel ad alium

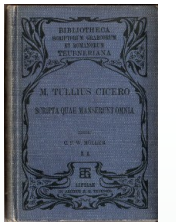
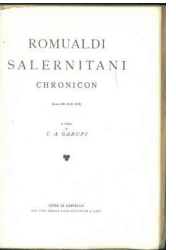
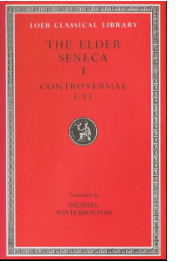
# What's “normal”?

- Vel comparatur ad se vel iam ad Valerium



# What's “normal”?

- Vel comparatur ad se vel **i**am ad Valerium
- Vel comparatur ad se vel **j**am ad Valerium
- **U**el comparatur ad se **u**el iam ad **U**alerium
- **V**el comparatur ad se **u**el iam ad **V**alerium
- **u**el comparatur ad se **u**el iam ad **V**alerium





## What's “normal”?

- te. Ego / te. ego
- u / v / U / V
- i / j
- Ae / æ
- Oe / œ
- bonum est / bonum'st  
/ bonumst
- [punctuation]
- [macron / breve]



# What's “normal”?

- te. Ego / te. ego
- u / v / U / V
- i / j
- Ae / æ
- Oe / œ
- bonum est / bonum'st  
/ bonumst
- [punctuation]
- [macron / breve]

- Processing
  - Search
  - Collation
  - NLP (lemma, PoS etc.)
  - Statistics (distant reading)...

# What's "normal"?

- te. Ego / te. ego
- u / v / U / V
- i / j
- Ae / æ
- Oe / œ
- bonum est / bonum'st  
/ bonumst
- [punctuation]
- [macron / breve]

- Processing
  - Search
  - **Collation**
  - NLP (lemma, PoS etc.)
  - Statistics (distant reading)...

"pertinet": [pertineo] Verbal, II conjug,  
Active indicative, Present, III singular

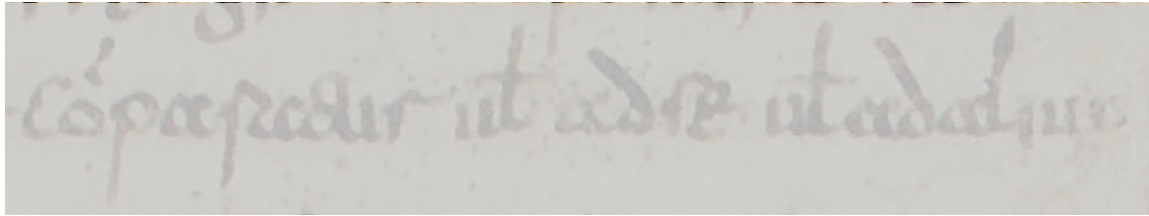
pertinet

p<sup>̄</sup>tin<sup>̄</sup>et

- My Ursus edition (IX Century)  
[www.unipa.it/paolo.monella/ursus](http://www.unipa.it/paolo.monella/ursus)

# What's “normal”?

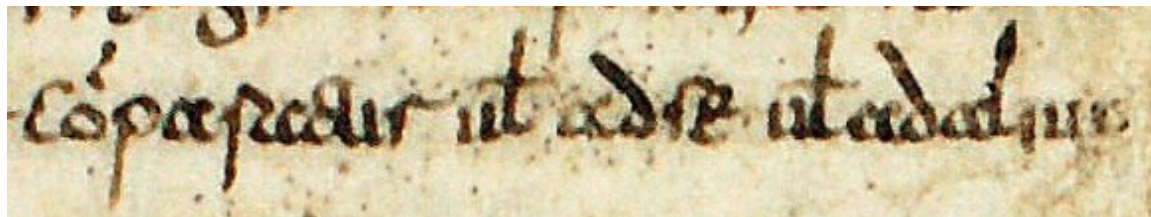
- Comparatur vel ad se vel ad alium



- cōparaEur      uť   adfe   uťadaliuM

## “Diplomatic”...

- Comparatur vel ad se vel ad alium



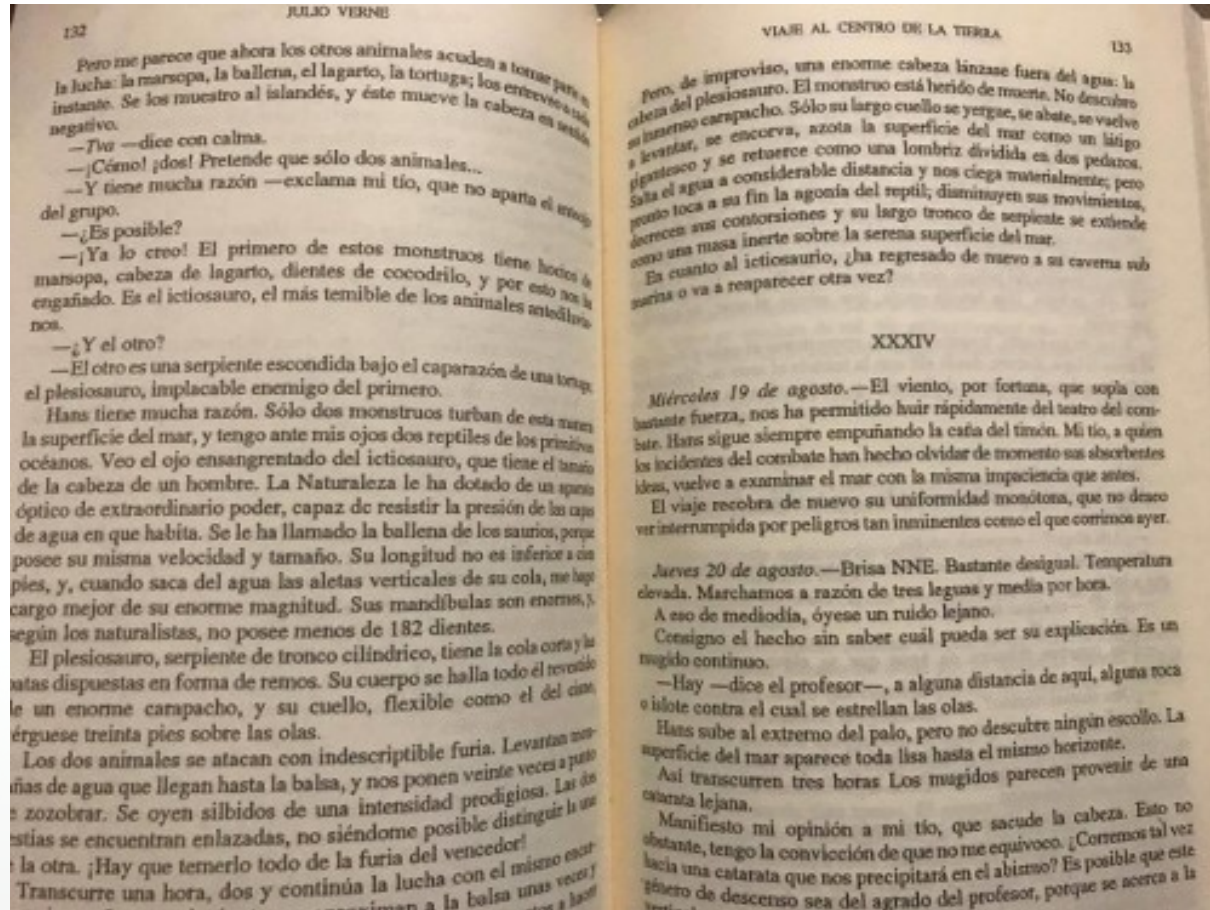
- cōparaEur      uť   adfe   uťadaliu



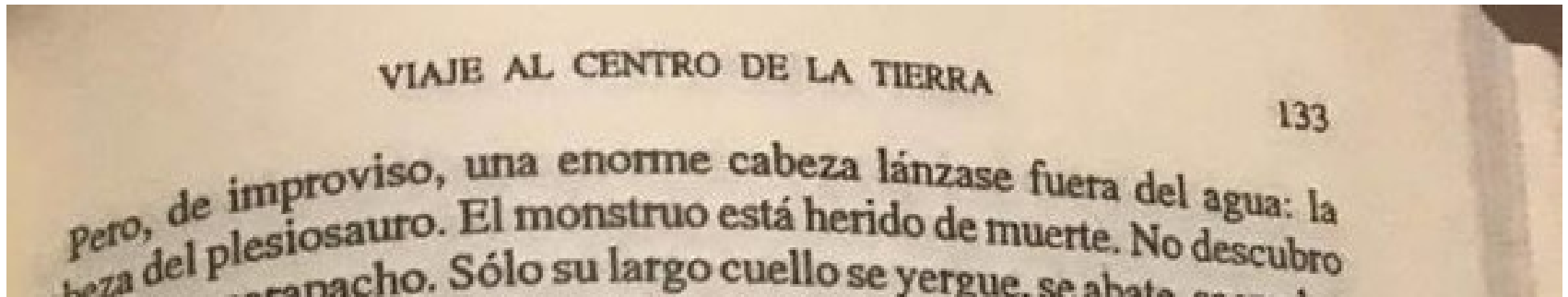
A white speech bubble with a dark blue background. The bubble has a rectangular top and a triangular tail pointing downwards and to the left. The text is centered within the rectangular part of the bubble.

**Modelling  
“Gutenberghian”  
writing systems**

# Modelling “Gutenberghian” writing systems



# Modelling “Gutenberghian” writing systems

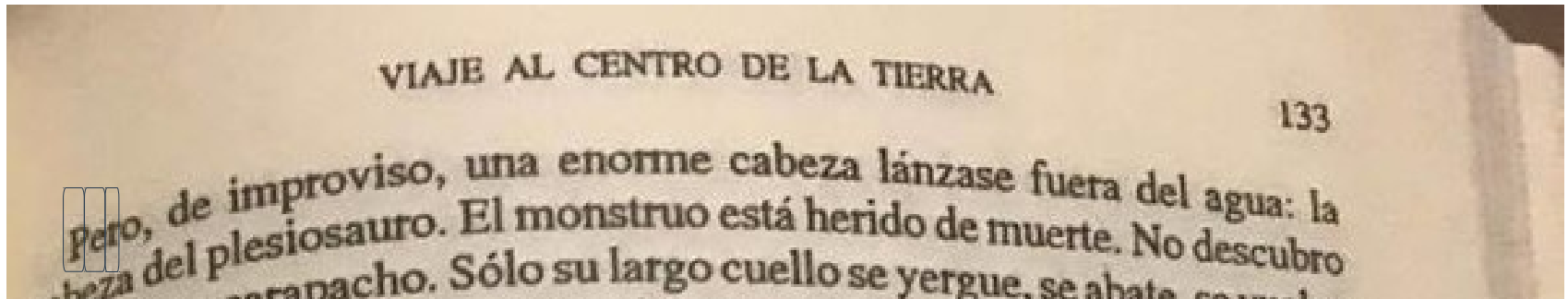


*Digital modelling*

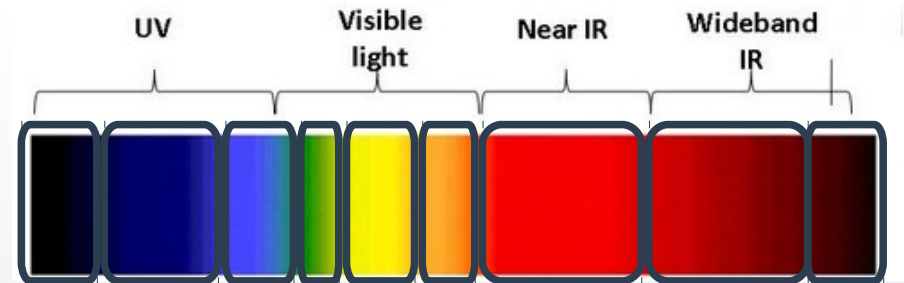


# Modelling “Gutenberghian” writing systems

- Pero, de improviso, una enorme cabeza...



*Digital modelling*



# “Gutenberghian” principles


- □□□ Pero, de improvviso, una enorme cabeza...

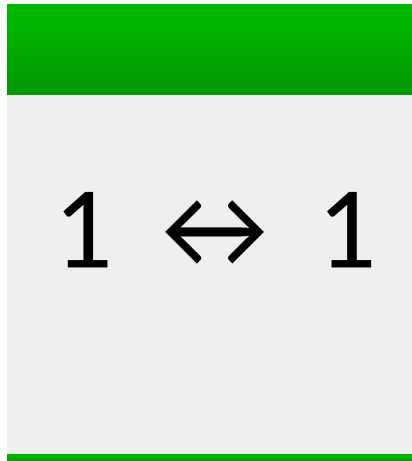
1 ↔ 1

1 = 1

1, 2, 3...

# “Gutenberghian” principles

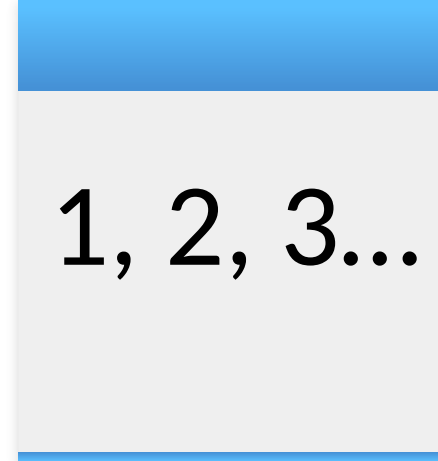
-  Pero, de improviso, una enorme cabeza...



1 mark,  
1 letter



All marks  
are (born)  
equal



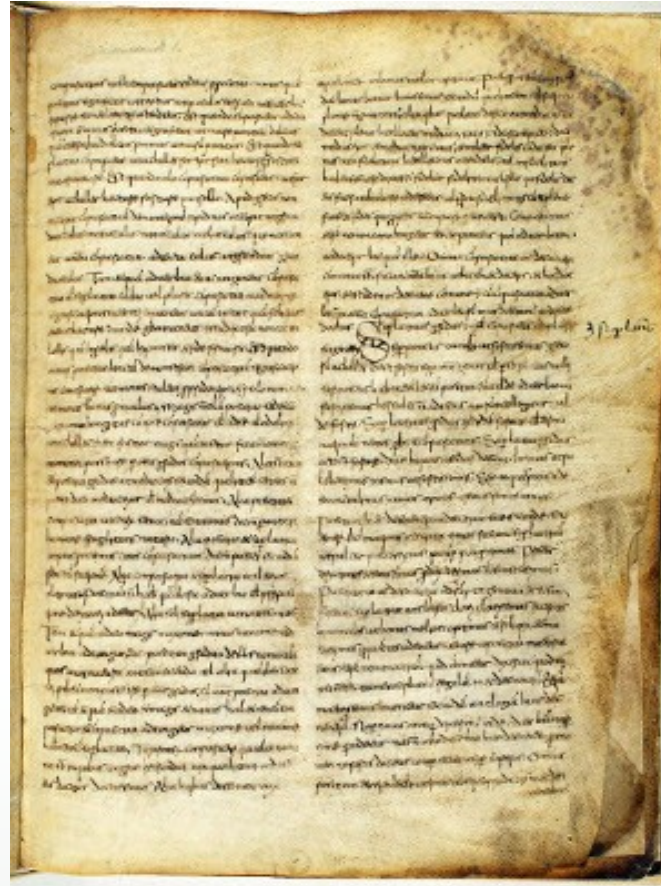
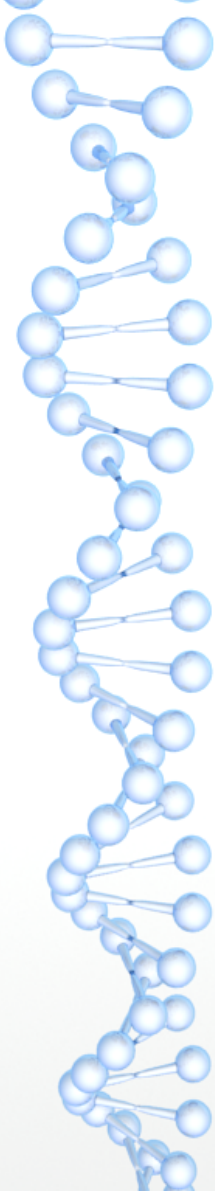
Sequence  
(one dimension)

A white speech bubble with a dark blue background. The bubble has a rectangular body and a triangular tail pointing towards the bottom-left corner.

**Modelling  
pre-modern  
writing systems**



# Modelling pre-modern writing systems







# Modelling pre-modern writing systems

3 Gutenberghian principles

$1 \leftrightarrow 1$

1 mark,  
1 letter

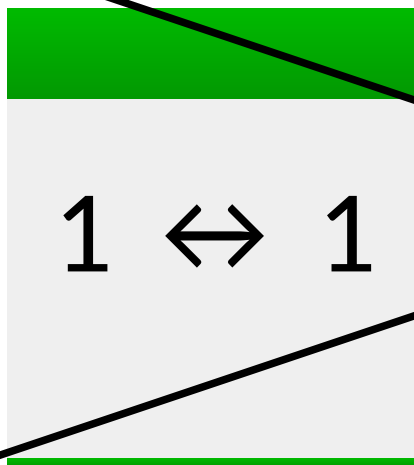
$1 = 1$

All marks  
are (born)  
equal

1, 2, 3...

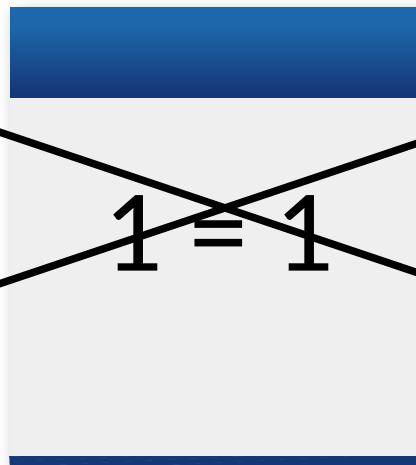
Sequence  
(one dimension)

# Modelling pre-modern writing systems



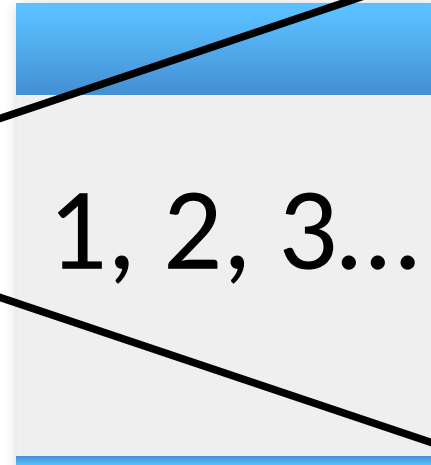
pertinet

pertinet



plerosque

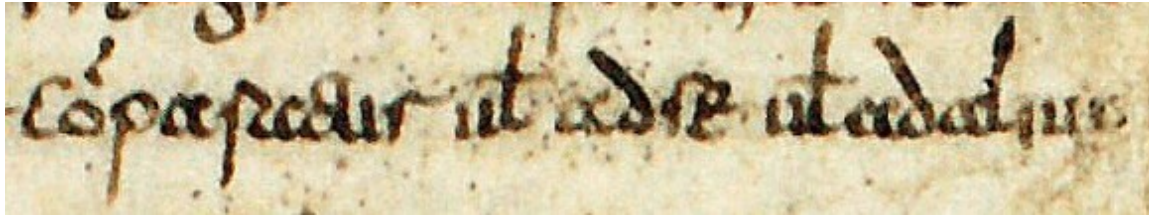
plerosq;



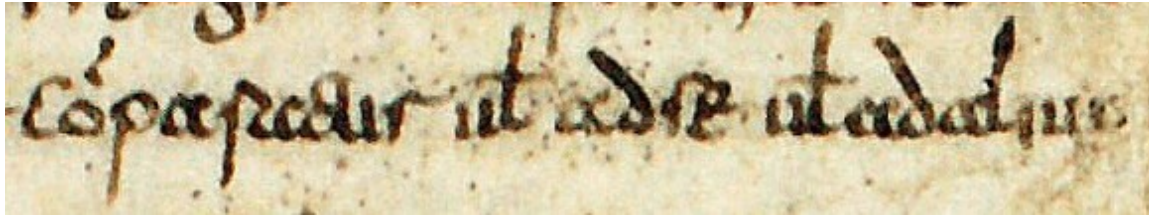
ordinem

ordin<sup>2</sup>e

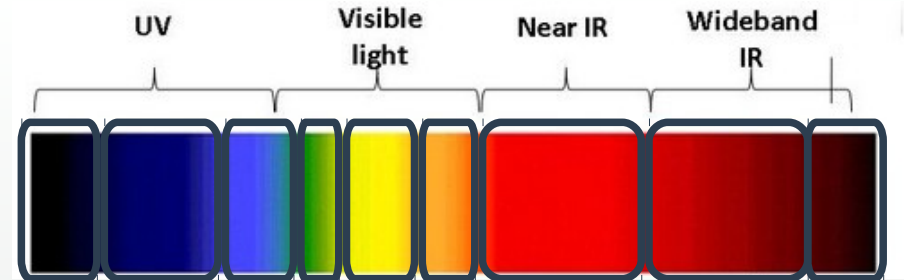
# Modelling pre-modern writing systems



# Modelling pre-modern writing systems

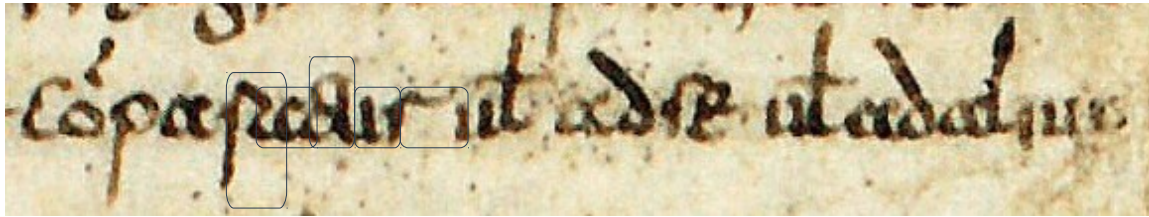


*Digital* modelling



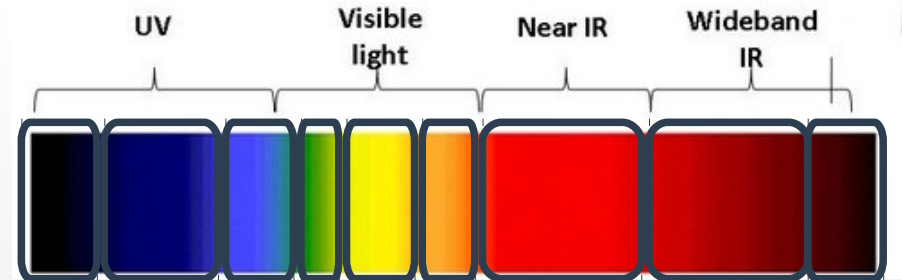
# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium



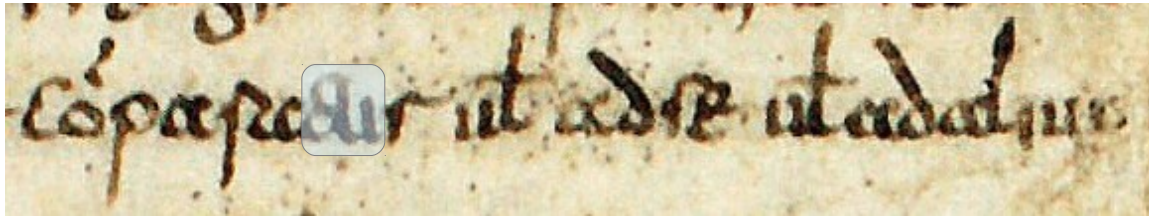
- cōparaſetur ut ad ſe ut ad alium

*Digital modelling*



# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium

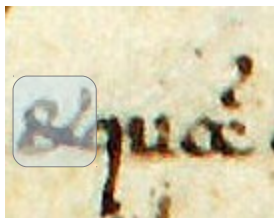


- cōparaEur ut adfe utadalium

Ligatures

# Modelling pre-modern writing systems

- et quam



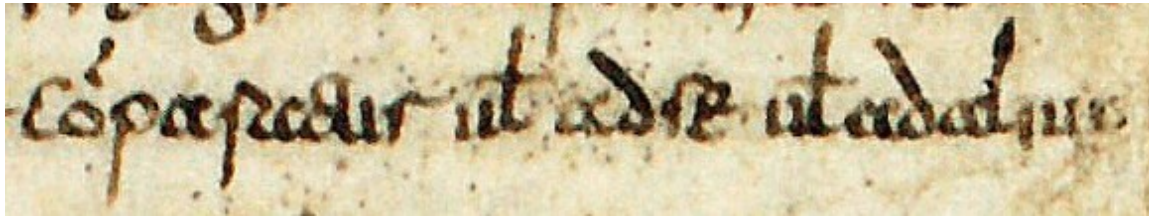
- &quam

Ligatures



# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium

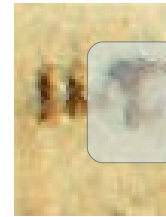
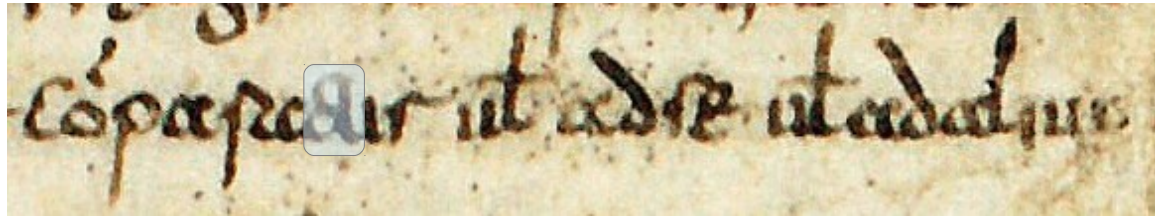


- cōparaEur      uł   adfe   uładaliu



# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium ut

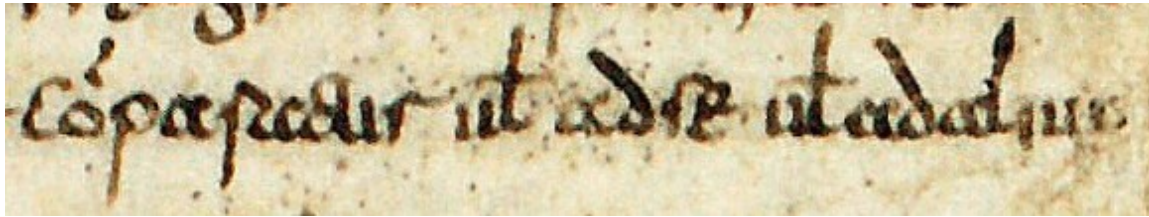


- cōparaEur ut adfe uładadium uo

Allographs

# Modelling pre-modern writing systems

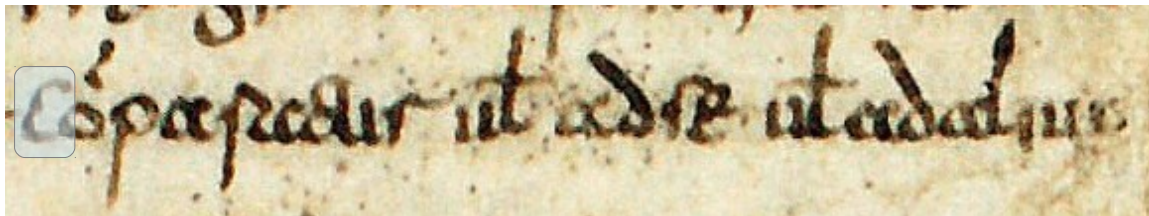
- Comparatur vel ad se vel ad alium



- cōparaEur      uł   adfe   uładaliu

# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium



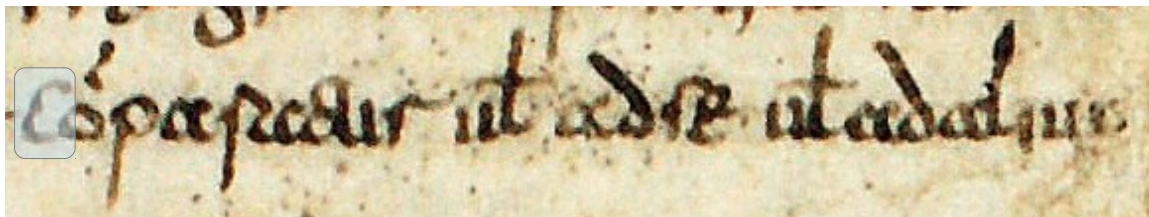
- cōparaEur uł adfe uładaliu

Capitalization?

Allographs

# Modelling pre-modern writing systems

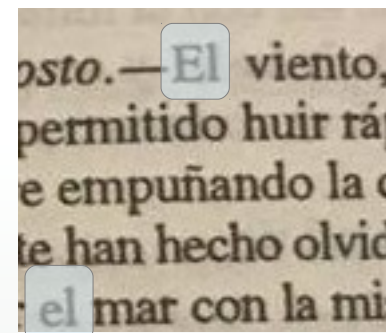
- Comparatur vel ad se vel ad alium



- cōparaEur uł adfe uładaliu

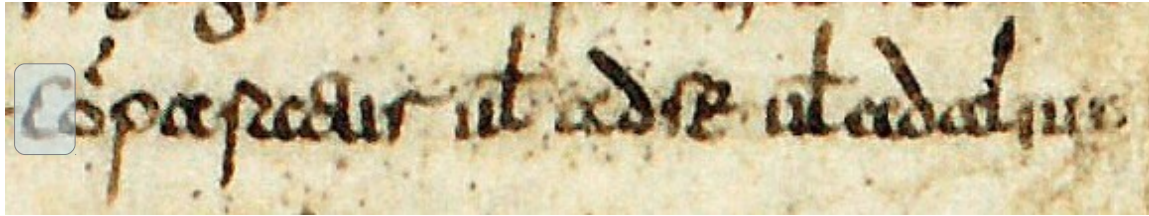
Capitalization?

Allographs



# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium



- cōparaEur uł adfe uładaliu

Capitalization?

Allographs

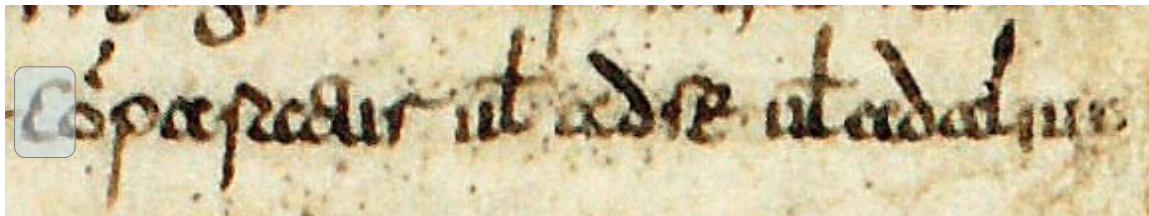
Bari

bari



# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium



- cōparaEur ut adfe utadalium

Capitalization?

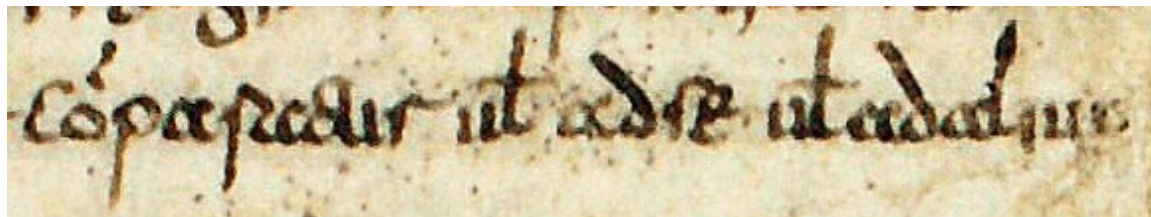
Allographs

ODD

odd

# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium



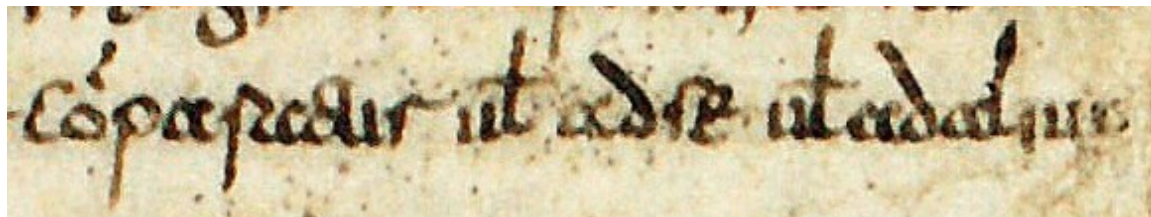
- cōparaEur      uł   adfe   uładaliu

Grapheme set

“Alphabet”

# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium.



- cōparaEur      uł   adfe   uładaliu;

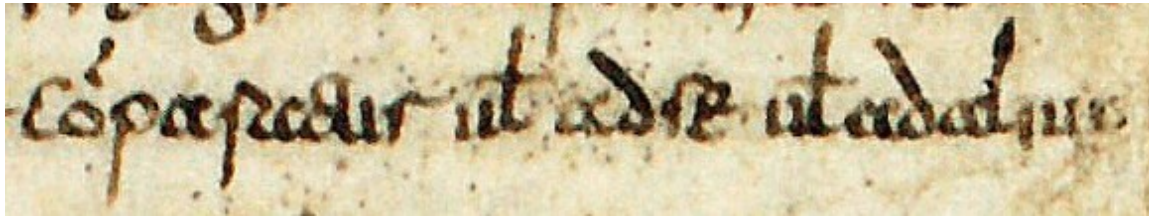
Punctuation

Grapheme set



# Modelling pre-modern writing systems

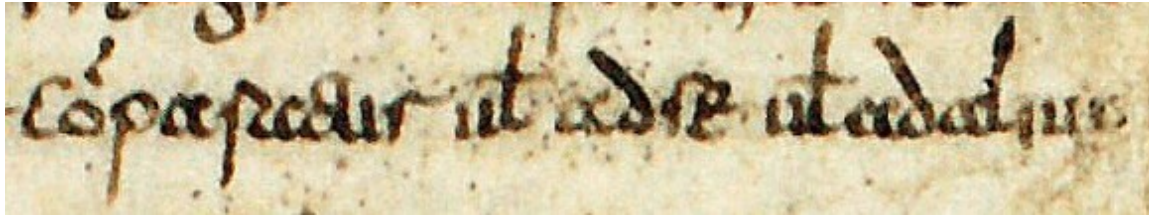
- Comparatur vel ad se vel ad alium



- cōparaEur      uł   adfe   uładaliu

# Modelling pre-modern writing systems

- Comparatur vel ad se vel ad alium



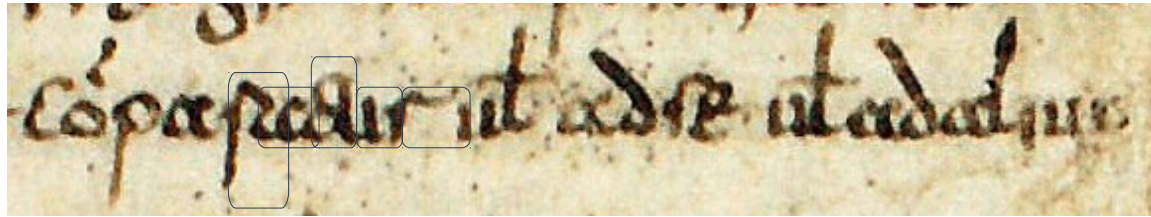
- cōparaEur uł adfe uładadium

Abbreviations

**Hjelmslev's "analysis"  
as digital modelling for  
pre-modern writing systems**

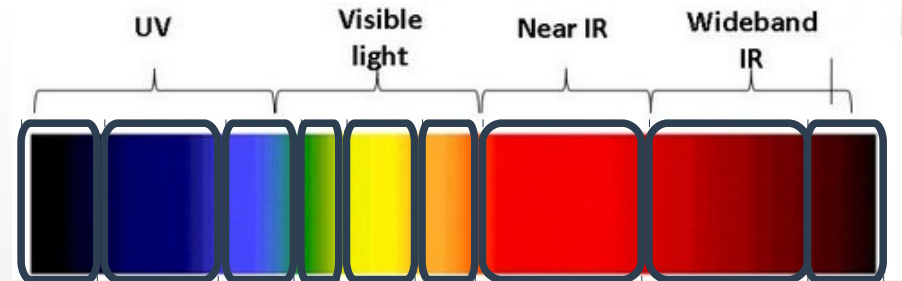
# Modelling

- Comparatur vel ad se vel ad alium



- cōparaEur ut adfe utadalium

*Digital modelling*





# Modelling

- còparaEur uł adfe uładaliu



# Modelling

- còparaEur u† adfe u†adaliu

# System / text

Paradigmatic

Syntagmatic (text, process)

• còparaEur    uʔ    adfe    uʔadaliu

(langue, system)

# System / text

System

<s>

<t>

Text

- còparaEur

uł

adfe

uładaliu

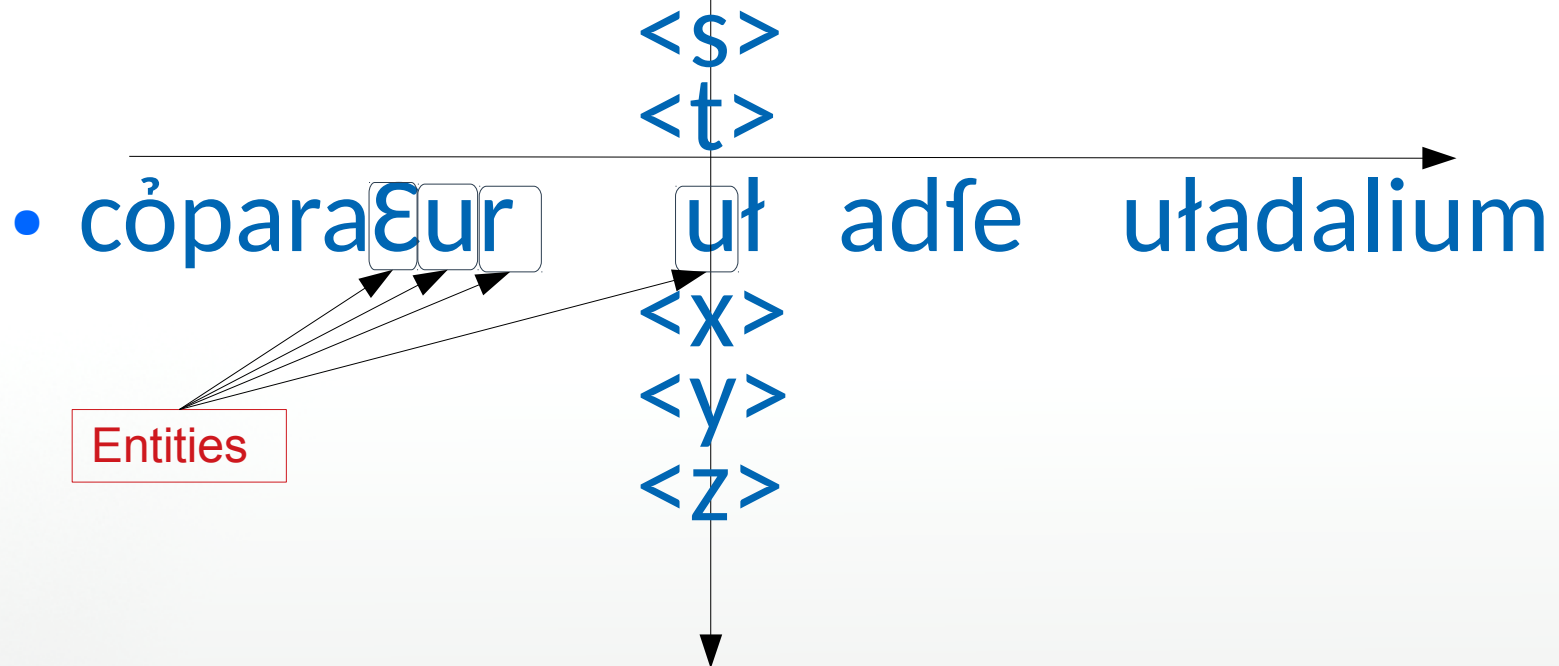
<x>

<y>

<z>



# Modelling/“analysis”: entities



# Modelling/“analysis”: entities

• còparaEur

Entities

<s>

<t>

ut

<x>

<y>

<z>

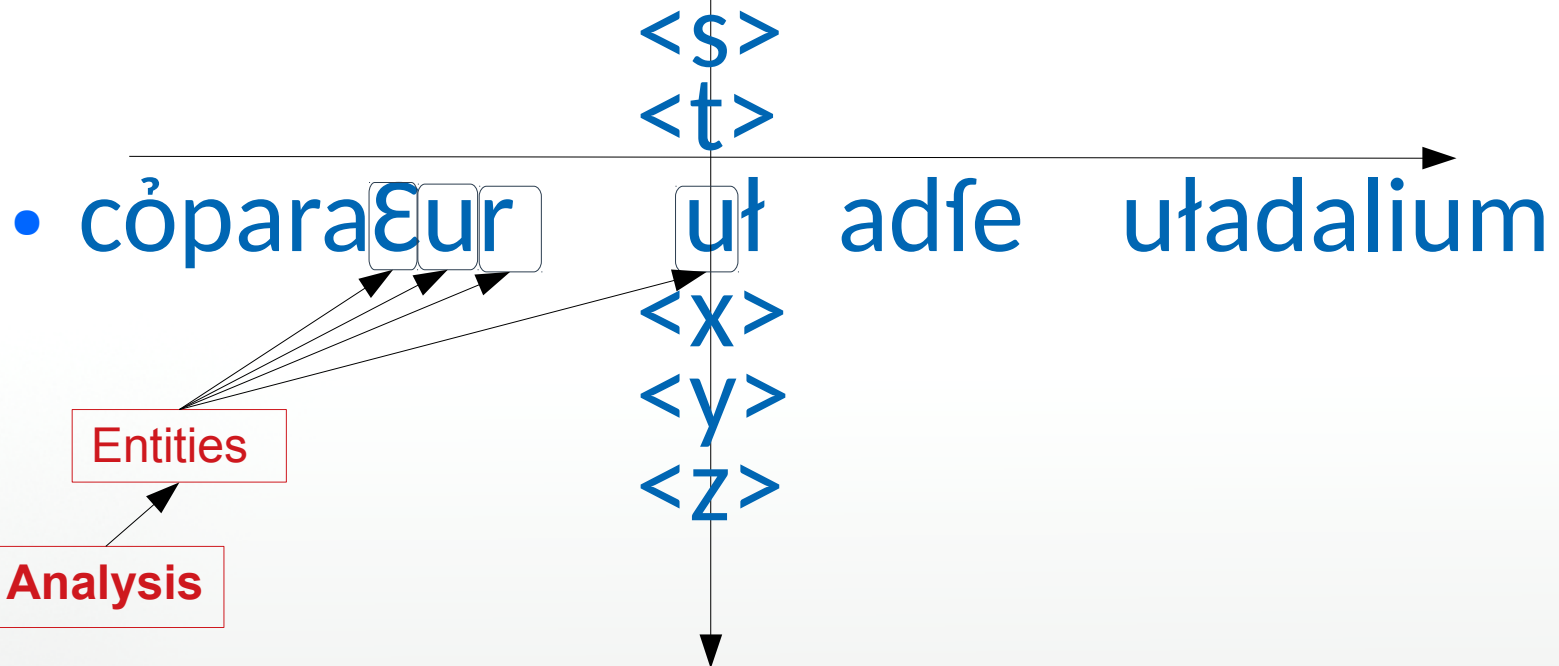
adfe

utadadium

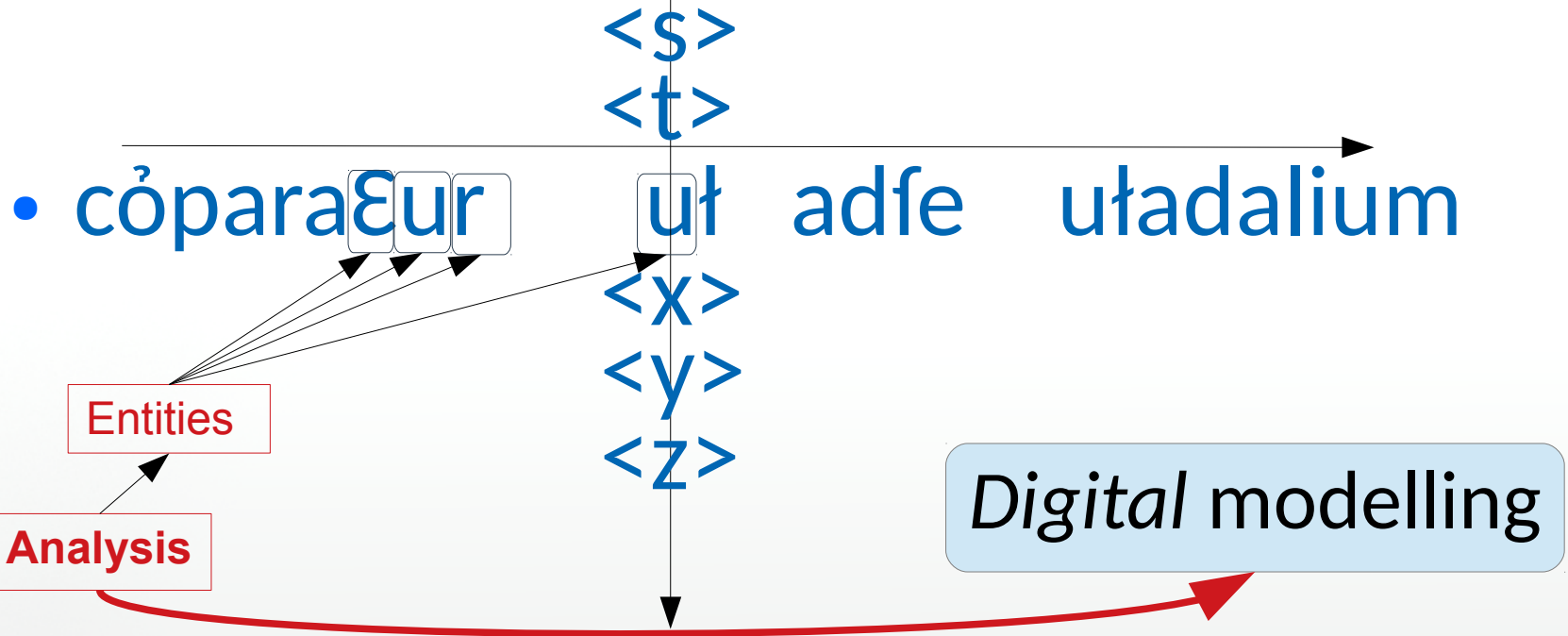


“If [we are] given anything...,  
it is the as yet unanalyzed *text*  
in its undivided  
and absolute integrity”  
 (“**deduction**”, *Prol.* Ch. 4)

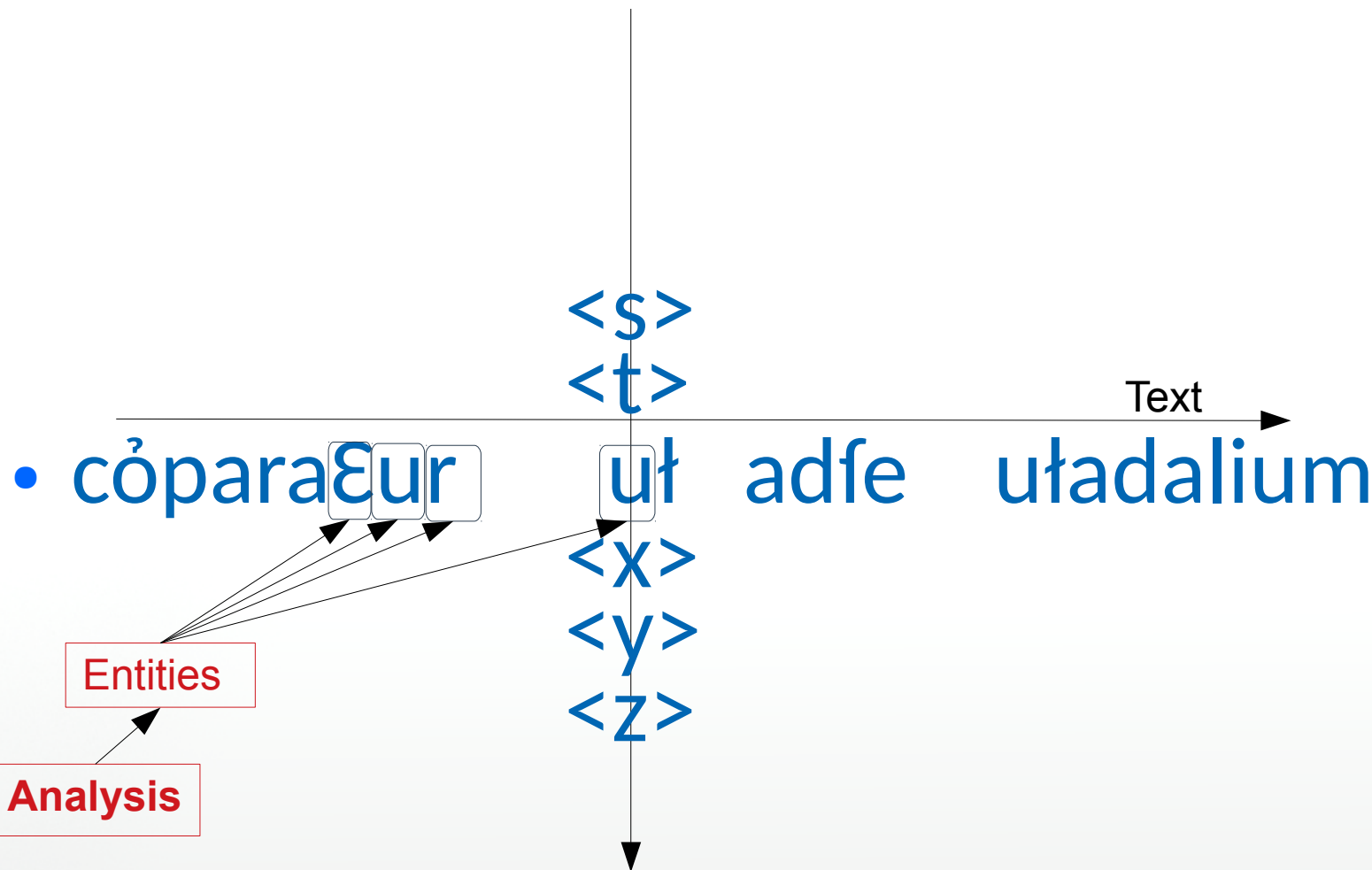
# Modelling/“analysis”: entities



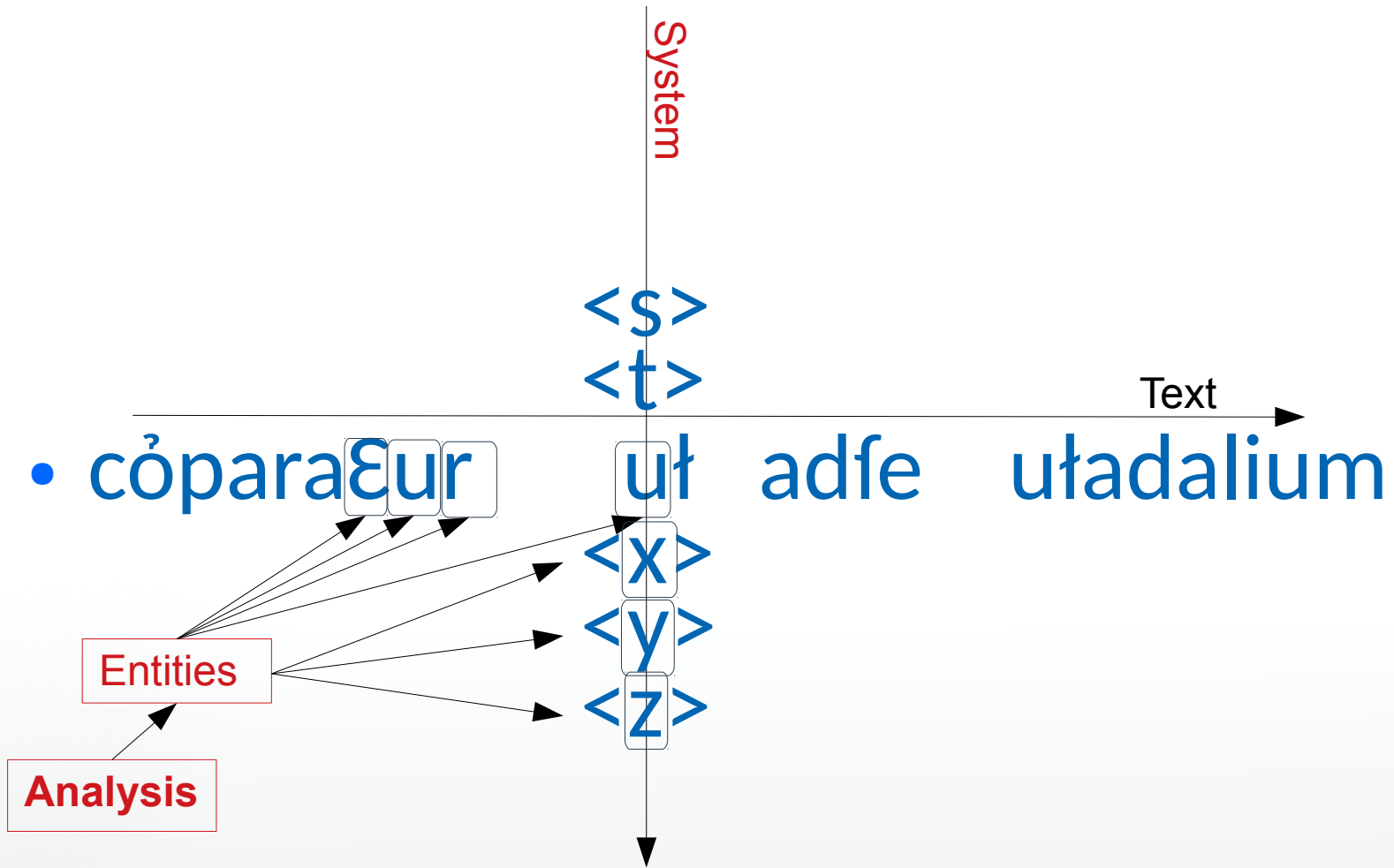
# Modelling/“analysis”: entities



# Modelling/“analysis”: entities

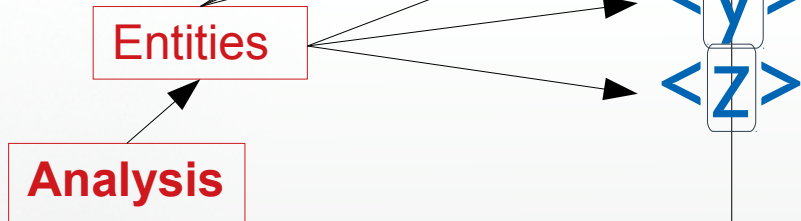


# Modelling/“analysis”: entities



# Graphemes as entities?

• còparaEUR ut adfe utadalium



# Graphemes as entities?

• còparaEur

<s>

<t>

ut

adfe

utadaliu

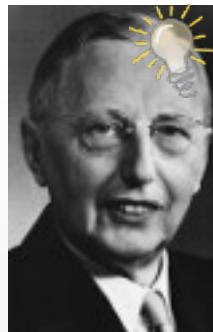
Entities

Analysis

<x>

<y>

<z>



“Thus the same linguistic form may also be manifested in writing... Here is a graphic ‘substance’... Describing the actually present expression... system” (*Prol.* Ch. 21)



# Graphemes as entities?

• còparaξur uł adfe uładalium

Entities

Analysis

<s>

<t>

<x>

<y>

<z>



# Graphemes as entities?

- Digital (discrete)
- Ligatures
- Allographs
  - (Capitalization)
- Grapheme set
  - (Punctuation)
- Abbreviations

• còparaEUR ut adfe utadalium

<s>

<t>

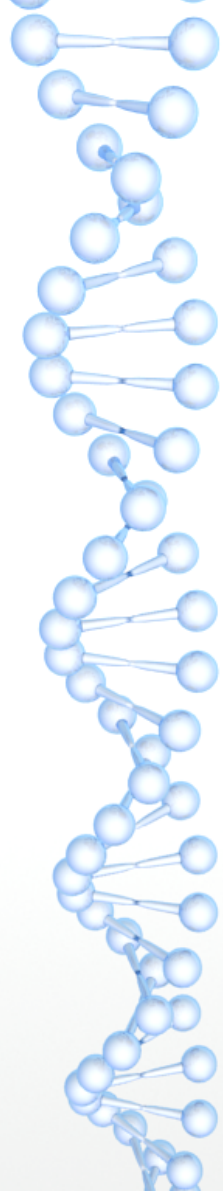
<x>

<y>

<z>

Entities

Analysis



# Graphemes as entities?

- ✓ Digital (discrete)
- Ligatures
- Allographs
  - (Capitalization)
- Grapheme set
  - (Punctuation)
- Abbreviations

• còparaEUR ut adfe utadalium

Entities

Analysis

<s>

<t>

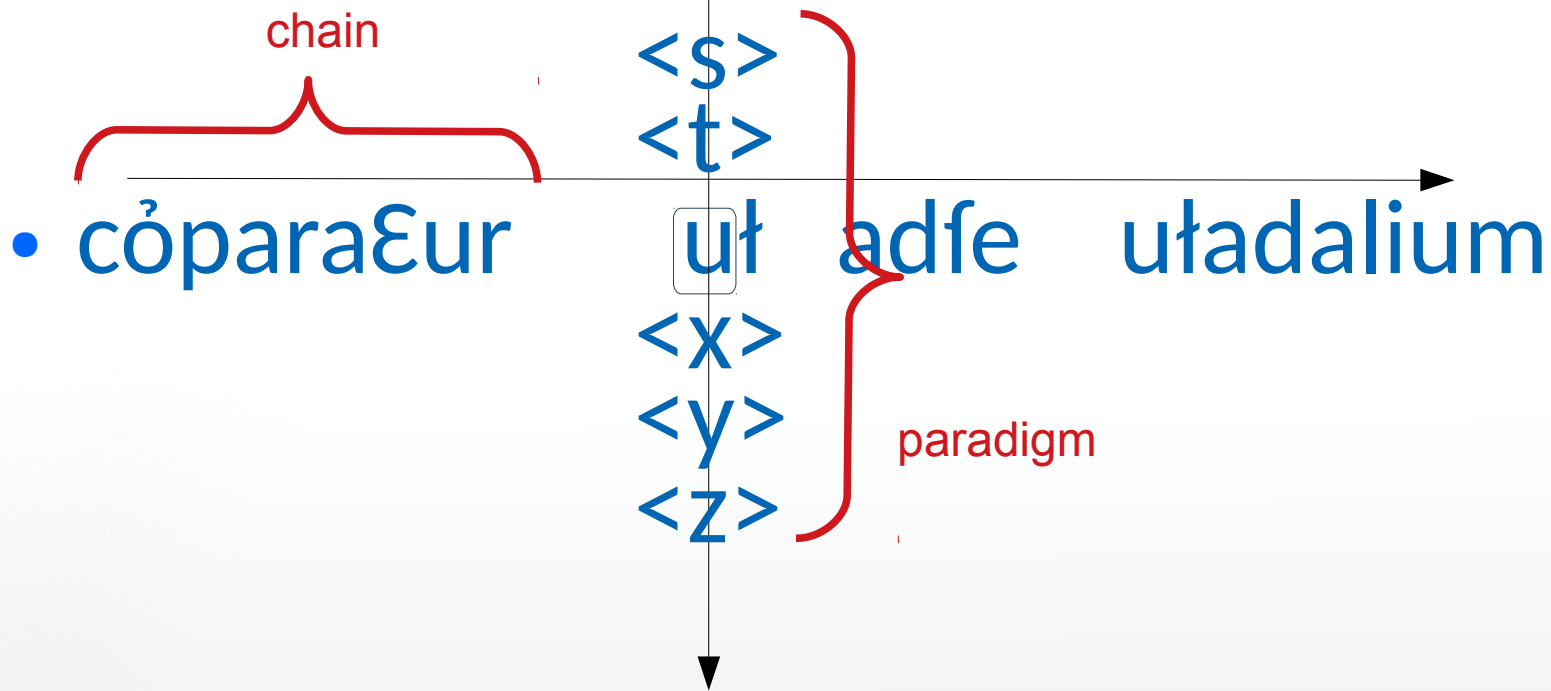
<x>

<y>

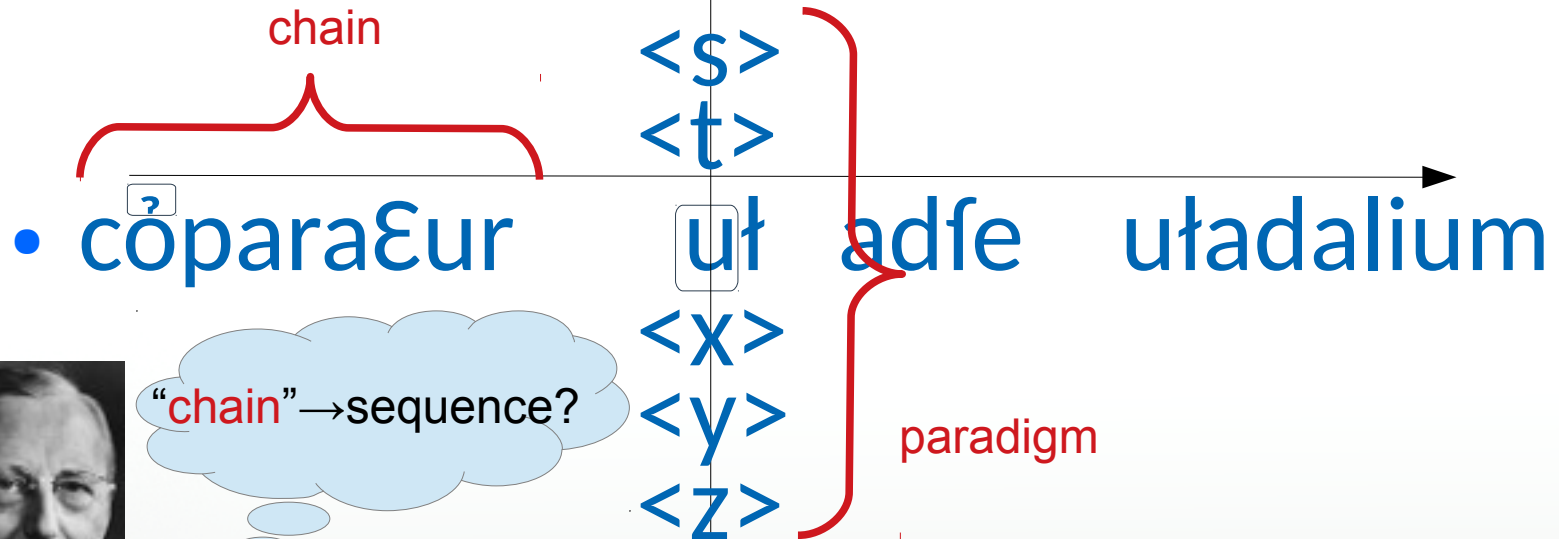
<z>

<z>

# Chains and paradigms

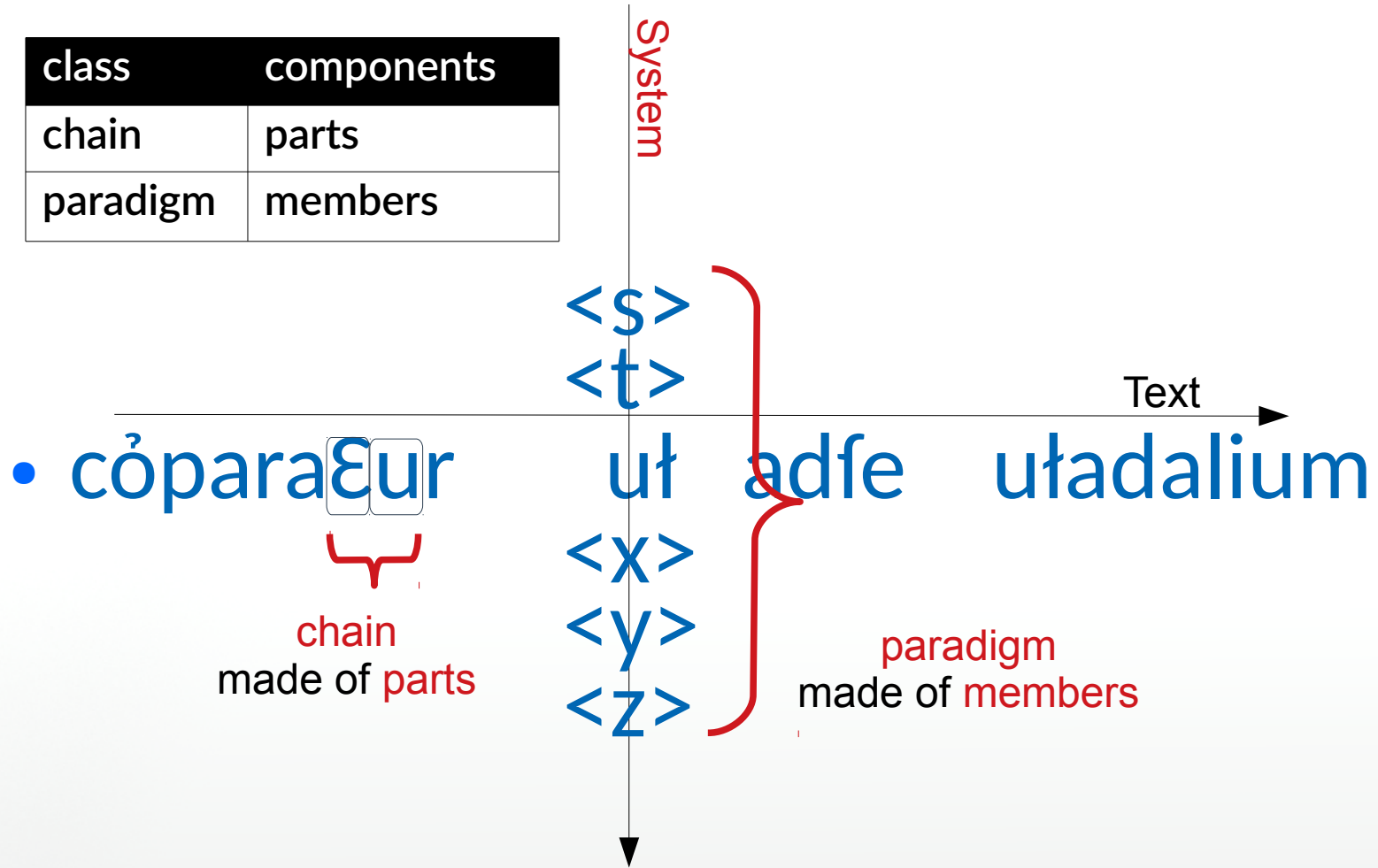


# Chains and paradigms



# Classes: chains and paradigms

class	components
chain	parts
paradigm	members



# Functions



• còparaEur



Function

<s>

<t>

uł

adfe

uładaliu

<x>

<y>

<z>

A white speech bubble with a dark blue background. The bubble has a rectangular body and a triangular tail pointing towards the bottom-left corner.

**Ligatures**



# Ligatures



• còparaEur



Function

“Ligature”  
(entities, parts of a chain)

<s>

<t>

uł

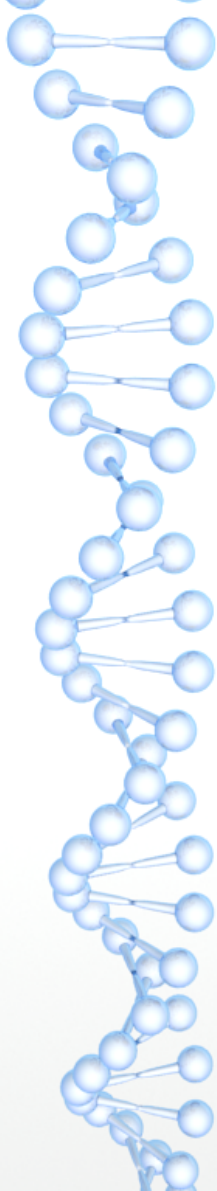
adfe

uładaliu

<x>

<y>

<z>



# Ligatures

Varieties

«σ»

<s>

<t>

• còparaEur

«√»

uł

<x>

<y>

<z>

adfe

uładadium

# Ligatures

Varieties:  
Solidal  
variants

«σ»

<s>

<t>

• còparaur

«√»

uł

<x>

<y>

<z>

adfe

uładaliu

# Ligatures

Varieties:  
Solidal  
variants

- ✓ Digital (discrete)
- Ligatures
- Allographs
  - (Capitalization)
- Grapheme set
  - (Punctuation)
- Abbreviations

• còparaEur

<s>

<t>

uł

adfe

uładaliu

<x>

<y>

<z>

# Ligatures

Varieties:  
Solidal  
variants

- ✓ Digital (discrete)
- ✓ **Ligatures**
  - Allographs
    - (Capitalization)
  - Grapheme set
    - (Punctuation)
  - Abbreviations

• còparaEur

<s>

<t>

uł

adfe

uładaliu

<x>

<y>

<z>



**Graphemes/allographs**

# The commutation test

- còparaEur

<s>

<t>

u

adfe

uɫadaliu

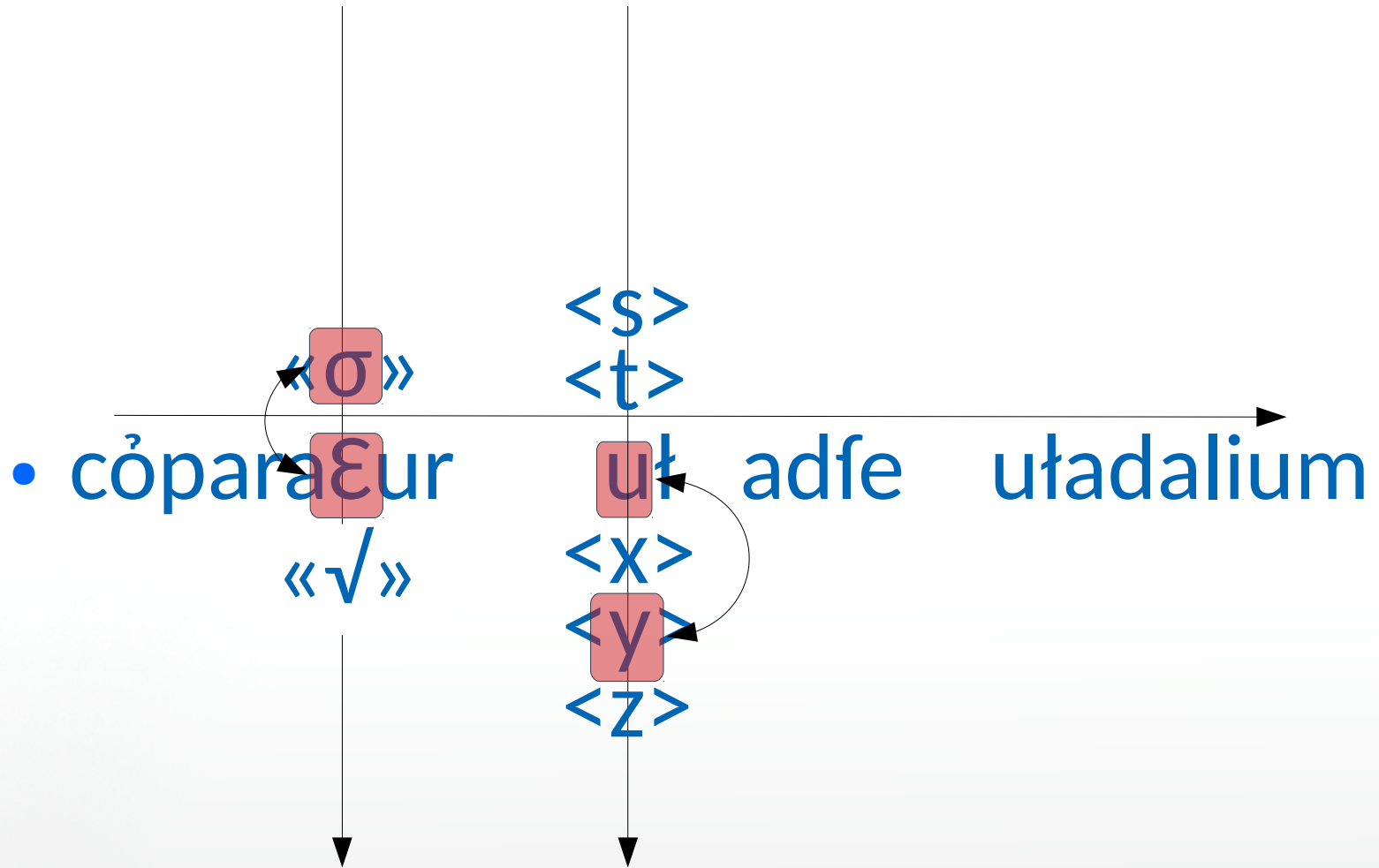
<x>

<y>

<z>



# The commutation test





# The commutation test

• còparaEur

Function

**Substitution:**  
→ No change  
in “meaning”

daEur / daEur

«σ»

«√»

<s>  
<t>

ut

<x>  
<y>

<z>

adfe

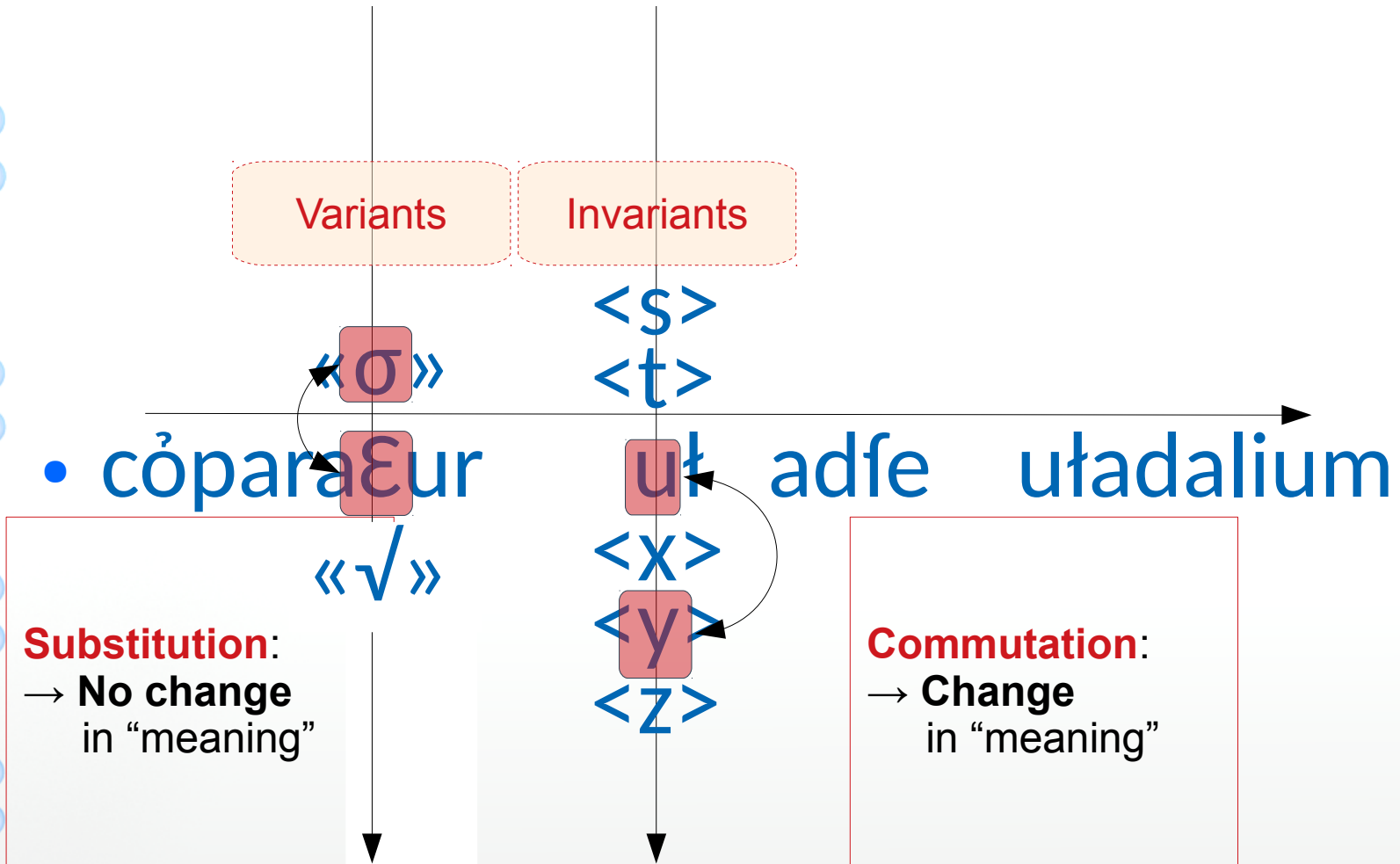
utadaliu

Function

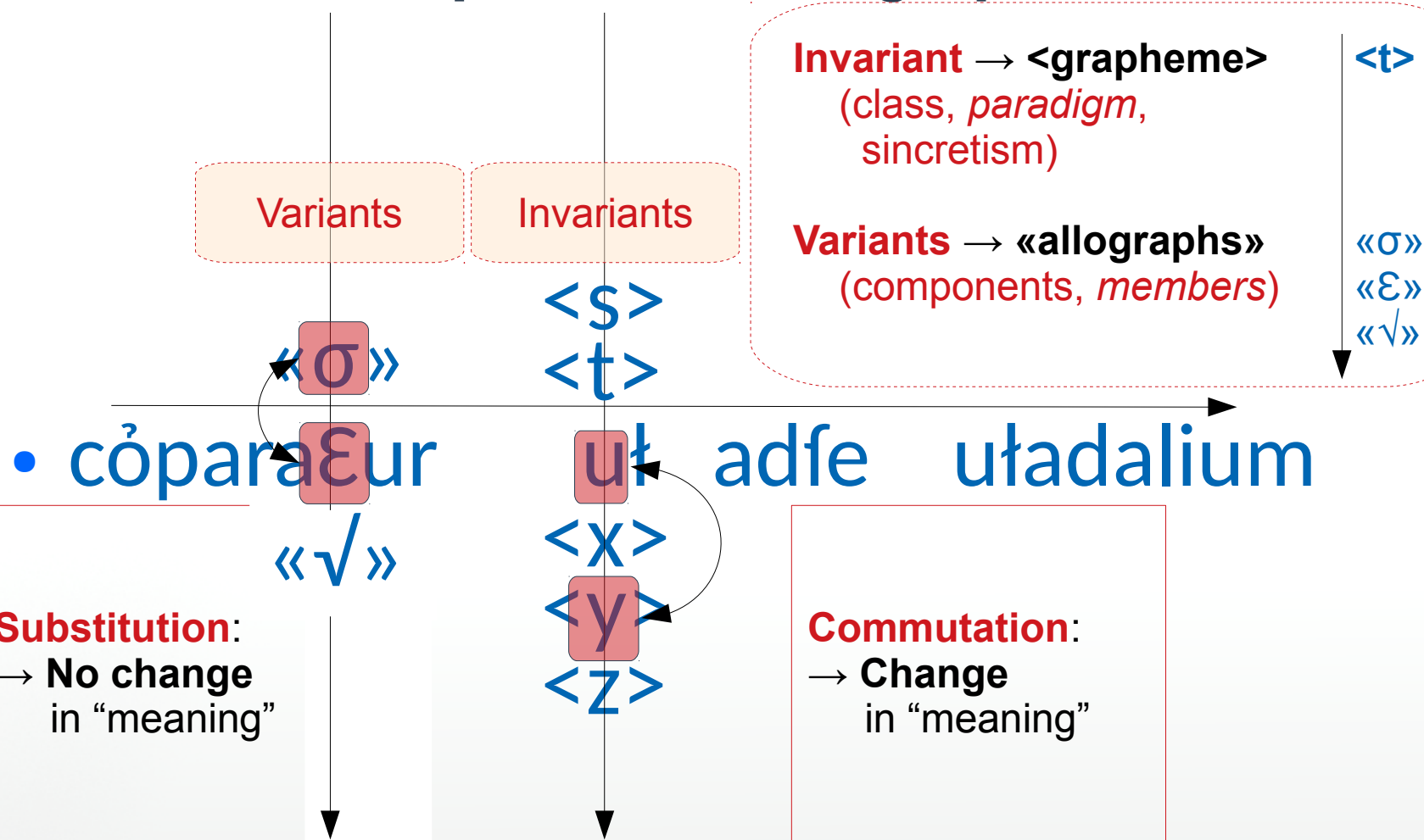
**Commutation:**  
→ Change  
in “meaning”

annus / annys

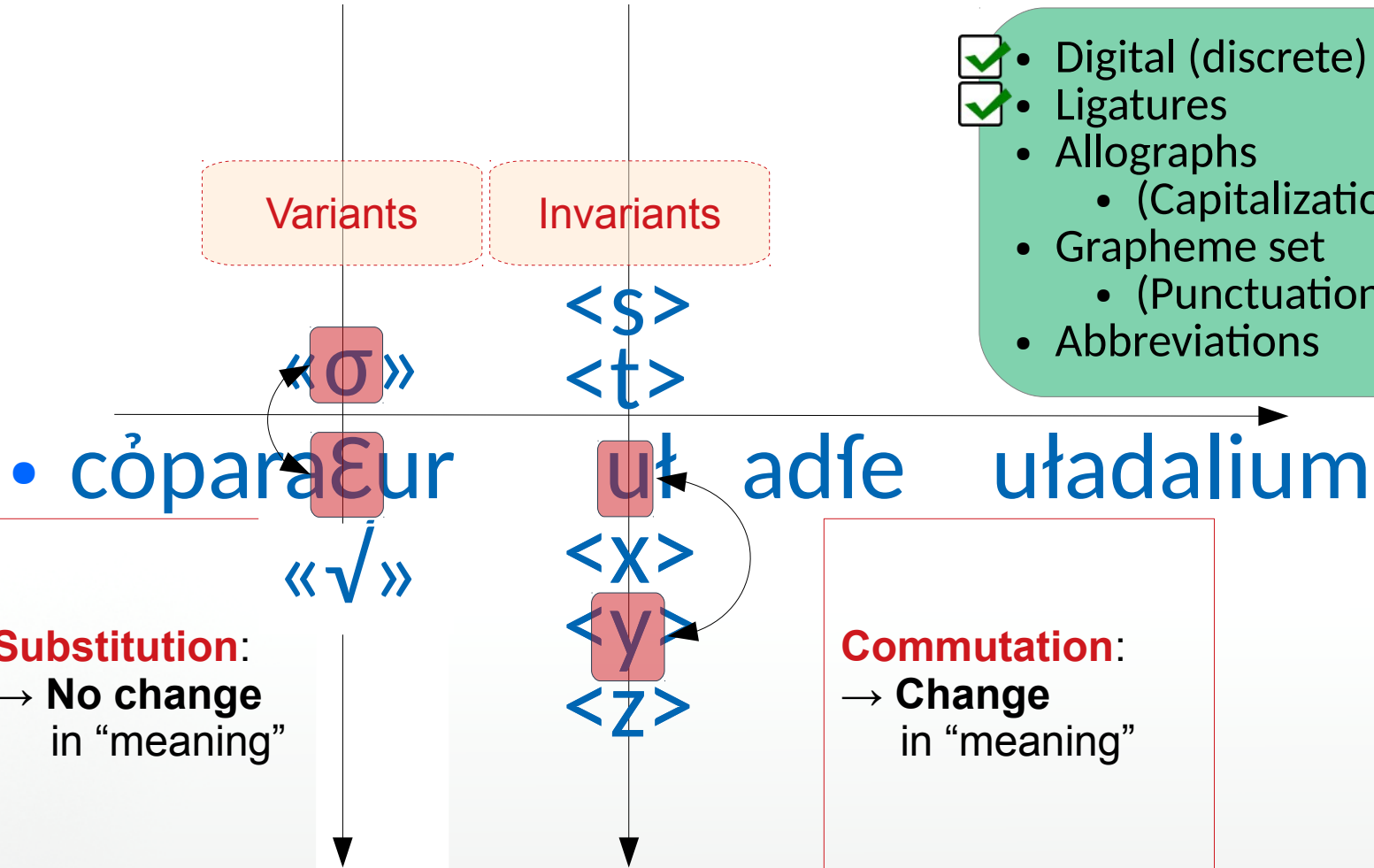
# The commutation test



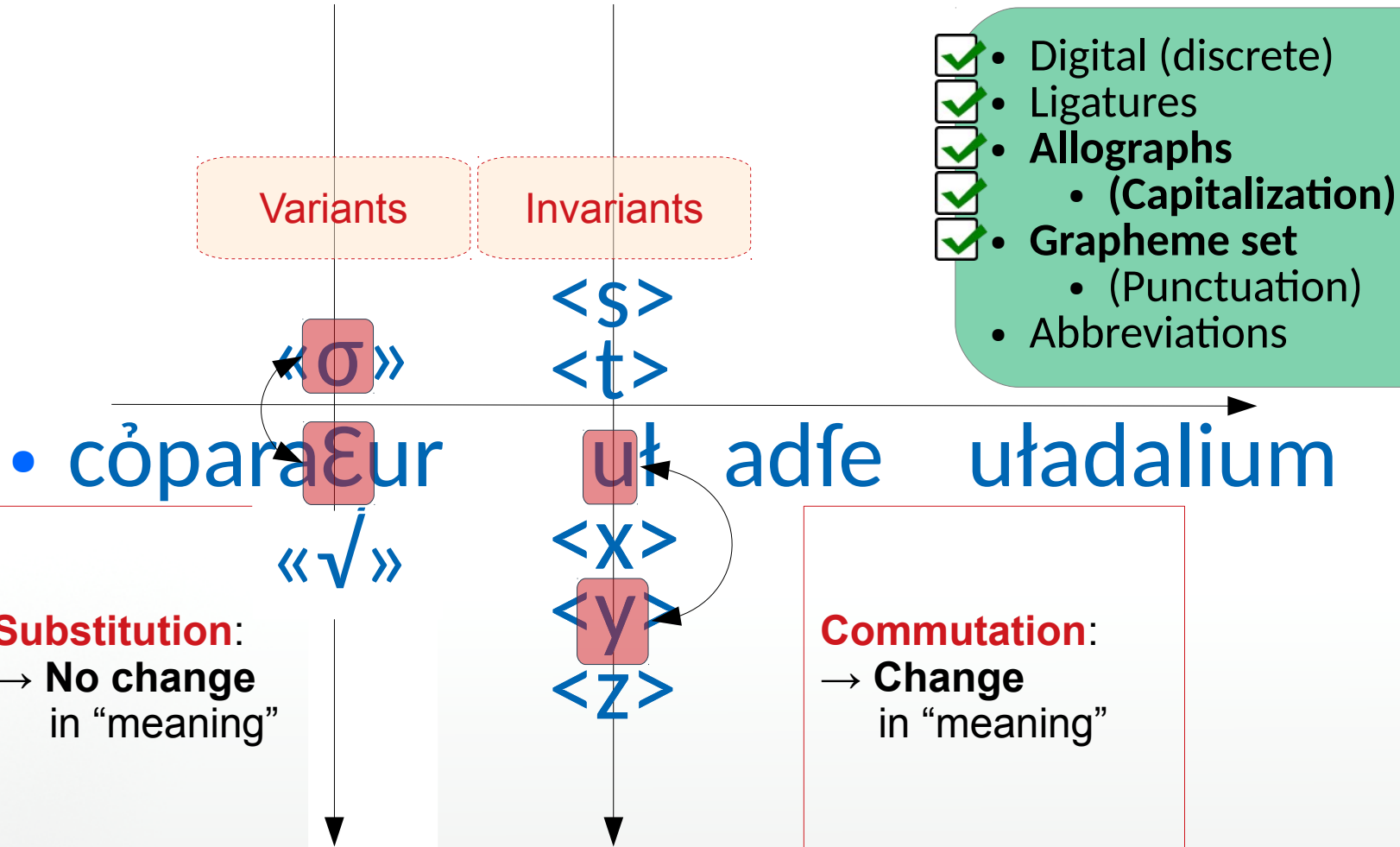
# Graphemes / allographs



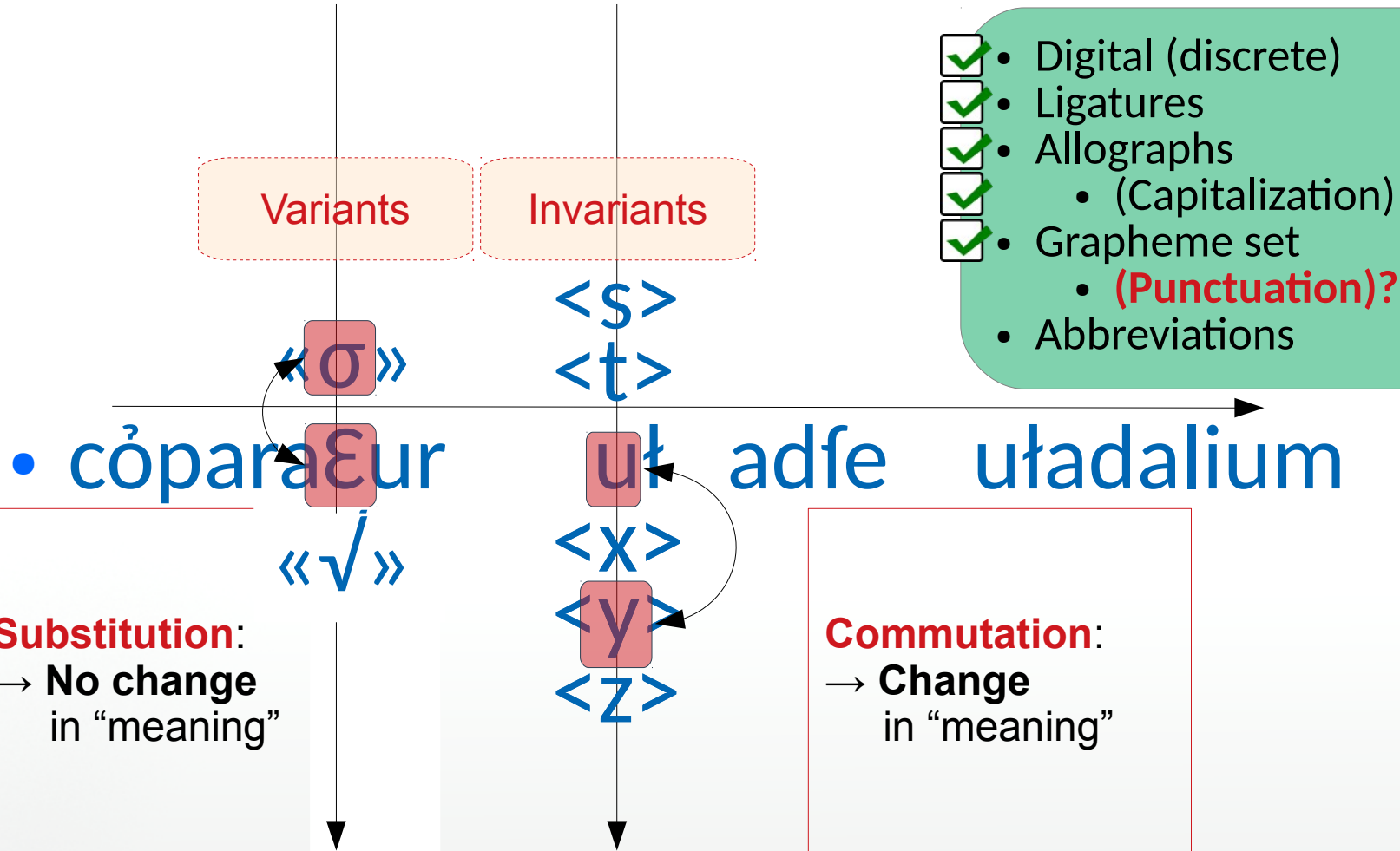
# Graphemes / allographs



# Graphemes / allographs



# Graphemes / allographs





**Punctuation**

# Punctuation

• còparaEur

**Substitution:**  
→ No change  
in “meaning”

«σ»

«√»

<s>  
<t>

uł

<x>

<y>

<z>

adfe

uładadium

**Commutation:**  
→ Change  
in “meaning”



“content-form”  
(Prol. Ch. 13)



# Punctuation



“content-form”  
(*Prolog* Ch. 13)

Including larger  
units, such as  
sentences

• còparaEur

«σ»

«√»

**Substitution:**  
→ No change  
in “meaning”

<s>  
<t>

ut

<x>

<y>

<z>

adfe

utadadium

**Commutation:**  
→ Change  
in “meaning”

# Punctuation



“content-form”  
(*Prol. Ch. 13*)

Including larger  
units, such as  
sentences

Truly I tell you, today you will be with me in paradise

Truly I tell you today, you will be with me in paradise

**Substitution:**  
→ No change  
in “meaning”

**Commutation:**  
→ Change  
in “meaning”



# Punctuation



“content-form”  
(*Prol. Ch. 13*)

Including larger  
units, such as  
sentences

Truly I tell you, today you will be with me in paradise

Truly I tell you today, you will be with me in paradise

**Substitution:**  
→ No change  
in “meaning”

**Commutation:**  
→ Change  
in “meaning”

# Punctuation

• còparaEur

«σ»

«√»

**Substitution:**  
→ No change  
in “meaning”

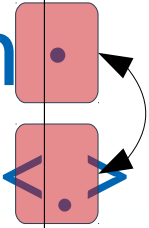
<s>  
<t>

ut

<x>  
<y>  
<z>

adfe

utadaliu



**Commutation:**  
→ Change  
in “meaning”

# Punctuation

- ✓ • Digital (discrete)
- ✓ • Ligatures
- ✓ • Allographs
  - (Capitalization)
- ✓ • Grapheme set
  - (Punctuation)
- ✓ • Abbreviations

• còparaEur

«σ»

«√»

**Substitution:**  
→ No change  
in “meaning”

<s>  
<t>

ut

<x>

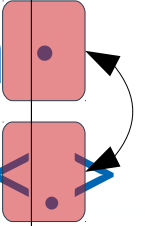
<y>

<z>

adfe

utadaliu

**Commutation:**  
→ Change  
in “meaning”



# Punctuation

- ✓ • Digital (discrete)
- ✓ • Ligatures
- ✓ • Allographs
  - (Capitalization)
- ✓ • Grapheme set
  - **(Punctuation)**
- ✓ • Abbreviations

• còparaEur

«σ»

«√»

**Substitution:**  
→ No change  
in “meaning”

<s>  
<t>

ut

<x>

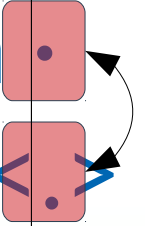
<y>

<z>

adfe

utadaliu

**Commutation:**  
→ Change  
in “meaning”





# **Abbreviations**

# Abbreviations

- 
-  paraEur      uł    adfe      uładaliu



# Abbreviations: one grapheme?

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)

• còparaEur

uł adfe

uładadium



# Abbreviations: one grapheme?

• cōparaEur

uł adfe uładadium

<o>

<ô>

<ō>

<p>

<q>



Principles: **simplicity**,  
**economy**, **reduction**

(*Prol.* Ch. 6)

“lowest possible number  
of elements” (Ch. 13)

# Abbreviations: function

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)

•   coparaEur      uł      adfe      uładadium

Function

(**entities**)

# Abbreviations: function → ligature?

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)
  - 2A. Ligature (**solidal variants** in a **chain**)

•  comparaEur    ut    adfe    utadaliu

Function

“Ligature”  
(**Solidal** variants)



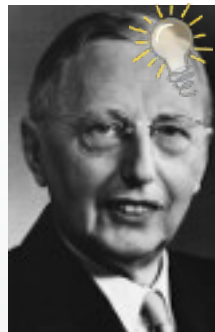
# Abbreviations: function → ligature?

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)
  - 2A. Ligature (**solidal variants** in a **chain**)

•  comparatur ut adfe utadadium

Function

“Ligature”  
(**Solidal** variants)



**Solidal**: both mandatory

# Abbreviations: function → selection?

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)
  - 2A. Ligature (**solidal variants** in a **chain**)
  - 2B. **Selection** (one governs the other)

•  cōparaEur      uł   adfe      uładadium

**Function**

**Selection:**

- one is optional
- one governs the other



# Abbreviations: function → selection?

•  comparatur      ut   adfe   utadaliu

Function

Selection:

- one is optional
- one governs the other



“chain” → sequence?

# Abbreviations: function → complementarity?

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)
  - 2A. Ligature (**solidal variants** in a **chain**)
  - 2B. **Selection** (one governs the other)
  - 2C. **Complementarity** (**interdependence**)

• comparatur ut adfe utadaliu

Function

Complementarity  
(interdependence in a system)



# Abbreviations: function → complementarity?

•   comparatur ut adfe utadaliū

Function

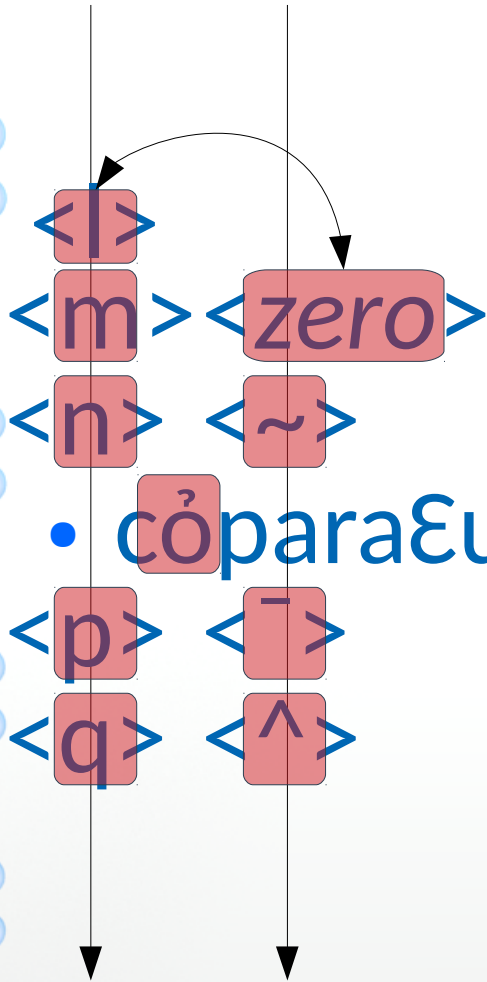
Complementarity  
(interdependence in a system)

Example of  
complementarity

[case + gender + number]

Alt	-	us	nom+masc+sing
Alt	-	orum	gen+m/neu+plur
Alt	-	arum	gen+ fem + plur

# Abbreviations: function $\rightarrow$ complementarity?



ut adfe utadaliu



# Abbreviations

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)
  - 2A. Ligature (**solidal variants** in a **chain**)
  - {** 2B. **Selection** (one governs the other)
  - 2C. **Complementarity** (**interdependence**)

• **c**õparaEur      uł   adfe   uładadium

# Abbreviations

- ✓ • Digital (discrete)
- ✓ • Ligatures
- ✓ • Allographs
  - ✓ • (Capitalization)
- ✓ • Grapheme set
  - ✓ • (Punctuation)
- **Abbreviations**

• cōparaEur      uł   adfe   uładaliu

# Abbreviations

- ☒ • Digital (discrete)
- ☒ • Ligatures
- ☒ • Allographs
  - ☒ • (Capitalization)
- ☒ • Grapheme set
  - ☒ • (Punctuation)
- ☐ • **Abbreviations**

• cōparaEur      uł   adfe   uładaliu

A white speech bubble with a dark blue background. The bubble has a rectangular body and a triangular tail pointing downwards and to the left.

**Open issues**



# Issues: 1. Abbreviations → functions

1. One **entity**
  - Whole abbreviation = grapheme (**invariant**)
2. Two **entities** (**function**)
  - 2A. Ligature (**solidal variants** in a **chain**)
  - {** 2B. **Selection** (one governs the other)
  - 2C. **Complementarity** (**interdependence**)

• **c**ōparaEur      uł    adfe    uładaliu

## Issues: 2. Abbreviations → one:many

- còpara<sup>?</sup>σur



1:1

- còpara<sup>?</sup>Eur








## Issues: 2. Abbreviations → one:many

- **comp**paraσur
- **cō**paraEur

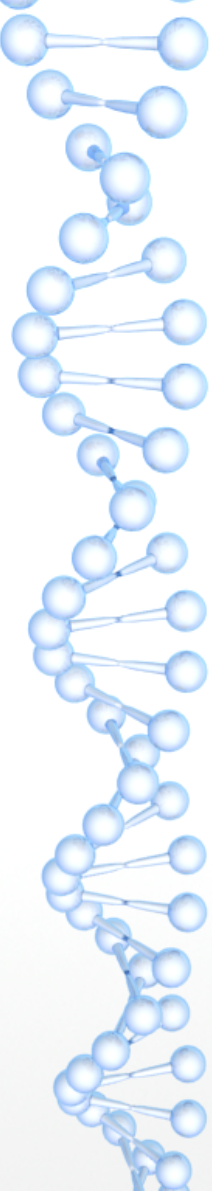
2:2 (2 ≠ 2)



## Issues: 2. Abbreviations → one:many

- **prae**positio
  - **p̄**positio
- 2:4
- 

## Issues: 2. Abbreviations → one:many

- 
- perfecta
  - pfecta
- 1:3
- An arrow points from the 'p' in 'pfecta' to the 'p' in 'perfecta'.

## Issues: 2. Abbreviations → one:many



- perfecta

Alphabemes  
(alphabetical letters)

1:3

- pfecta

Graphemes

# Issues: 3. Ligatures

(syntagmatic)



## Issues: 3. Ligatures



& (U+0026;  
ASCII 38)

# Issues: 3. Ligatures



& (U+0026;  
ASCII 38)

Historical/“etymological”  
considerations

# Issues: 3. Ligatures



& (U+0026;  
ASCII 38)

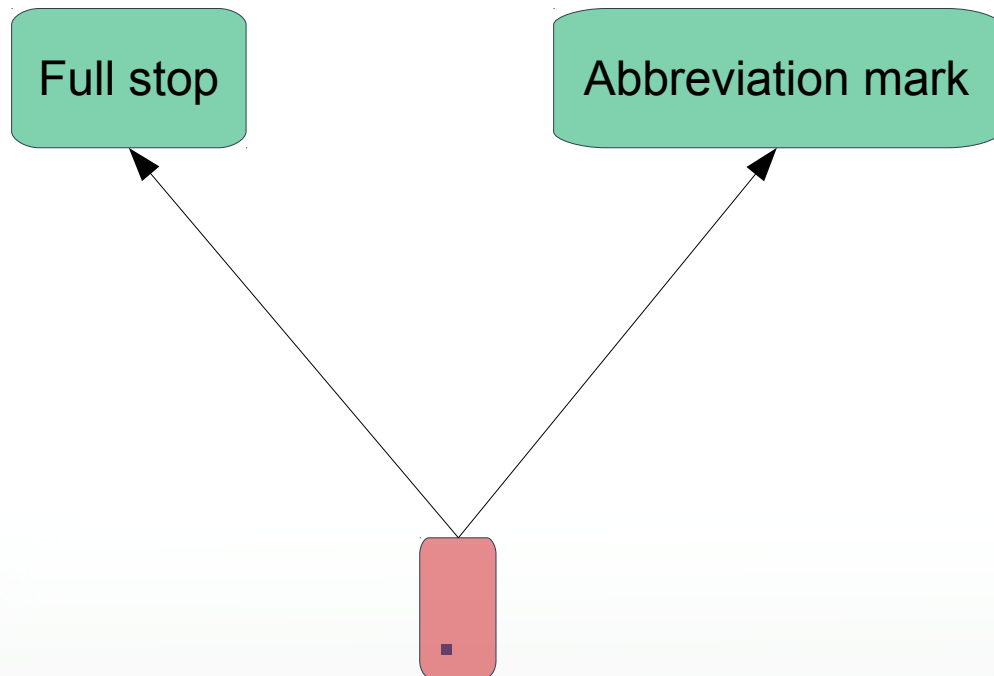
Historical/“etymological”  
considerations





# Issues: 4. Grapheme definition

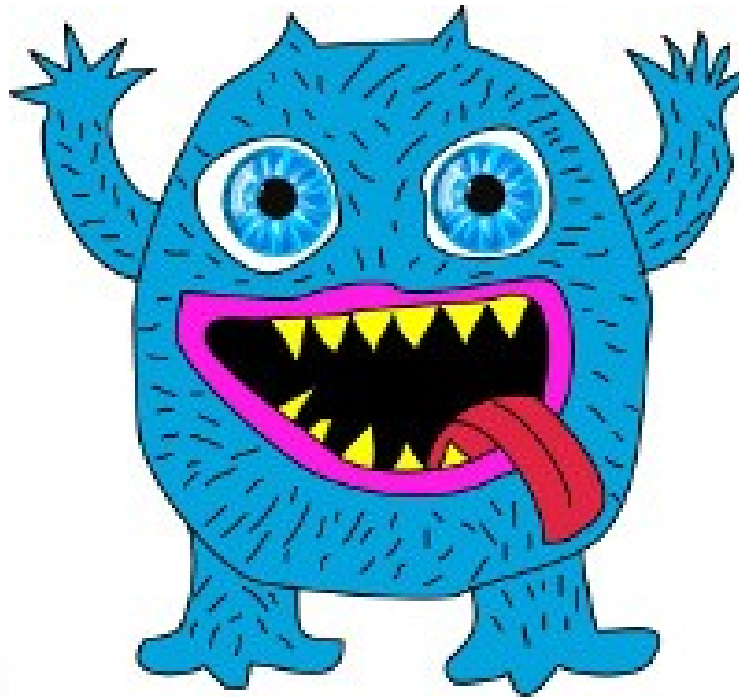
(paradigmatic)





## Issues: 5. Allograph definition (metasemiology)

## Issues: 5. Allograph definition (metasemiology)



# Issues: 5. Allograph definition (metasemiology)

Allographs  
(variants)

« $\sigma$ »

• còparaEur

uł

adfe

uładaliu

« $\sqrt{\quad}$ »



## Issues: 5. Allograph definition (metasemiology)

« $\sigma$ »

« $\sigma$ »

« $\sigma$ »

« $\sigma$ »

« $\sigma$ »

« $\sigma$ »

• còpara $\epsilon$ ur

uł

adfe

uładalium

« $\sqrt{\phantom{x}}$ »

« $\sqrt{\phantom{x}}$ »

« $\sqrt{\phantom{x}}$ »

« $\sqrt{\phantom{x}}$ »



# Open issues

1. Abbreviations → functions
2. Abbreviations → one:many
3. Ligatures
4. Grapheme definition
5. Allograph definition (metasemiology)

A white speech bubble with a dark blue background. The bubble has a rectangular body and a triangular tail pointing downwards and to the left. The text "In the Tower of Babel" is centered within the rectangular body.

**In the Tower of Babel**

# In the Tower of Babel: “diplomatic”

MS A

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



*Ceci n'est pas un linguiste.*

MS B

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



# In the Tower of Babel: “normalized”

OCR from Teubner

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

OCR from Loeb

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



*Ceci n'est pas un linguiste.*

# In the Tower of Babel

- Disposable home-made solutions (project-specific)
  - TEI: theory-agnostic
  - ***Diplomatic***: Modelling pre-modern writing systems
  - ***Normalized***: Normalization models / software



A white speech bubble with a dark blue background. The bubble has a rectangular body and a triangular tail pointing downwards and to the left. The text "Out of the Tower" is centered within the rectangular body.

**Out of the Tower**

# The (long) way out of the Tower



**WANTED**

**DEAD OR ALIVE  
REWARD  
\$4,000**

- Scholarly discussion on **modelling**
- Documentation on project-specific modelling
  - formal (data models, software code, tables)
  - prose
- Shared models
- Reusable **software** libraries



# Is Unicode the short way out (as TEI says)?

- Solution for new digital texts
- Not enough for pre-modern writing systems
  - Ligatures
    - **&** (U+0026; ASCII 38)
    - Have I encoded that it is equivalent to “e + t” in that MS?
  - Allographs
    - **ſ** (U+017F) / **s** (U+0073; ASCII 115)
    - Have I encoded that they are variants of grapheme <s>?
  - Grapheme set
    - **u** (U+0075; ASCII 117)
    - Have I encoded whether it “covers” (or not) <u> *and* <v>?



**In a nutshell**



# In a nutshell

- Digital scholarly editions: diplomatic/normalized
- Hjelmslev's "analysis" as digital modelling for pre-modern writing systems
- Open issues
- Interoperability through modelling