

Sunoikisis Digital Classics 2017

Session 16: Digital Papyrology, Part 2  
May 18, 2017

# Linguistic annotation of Greek and Latin papyri, ostraca and tablets

Marja Vierros (University of Helsinki)

# Treebanking Greek and Latin papyri – WHY?

- To build a lemmatised, morphologically and syntactically annotated corpus of papyrological texts
- Papyrological corpus offers interesting **direct** source for Greek and Latin
  - Historical development of linguistic forms, linguistic variation and language change
  - Sociohistorical linguistics
- The whole existing corpus of documentary papyri is digitally available in papyri.info (DDb**DP**)
  - What about literary and paraliterary papyri?
- Excellent way to deep-read the text and learn papyrological language use
  - And once a text is annotated, it stays in the corpus to benefit the papyrologists and linguists

# DDbDP, Papyrological Navigator, papyri.info

- Over 60.000 texts = ca. 5 million tokens
- Leiden mark-up for editorial information in print editions = TEI EpiDoc XML mark-up in digital form
- Linguistically annotated corpora are based on tokenized plain text; TEI EpiDoc XML data is neither tokenized nor plain

```
<lb n="2"/>τῷ ἀδελφῷ π<unclear>ο</unclear><supplied reason="lost">λλὰ <expan>χ<ex>αίρειν</ex></expan>. καλῶς</supplied>  
<lb n="3"/><choice><reg>ποιήσεις</reg><orig>πυήσις</orig></choice>, ἀδελφε, <unclear>ἐὰ</unclear><supplied reason="lost">ν ἔλθῃ</supplied>  
<lb n="4"/>ἡ <choice><reg>πορεία</reg><orig>πορήε</orig></choice> τῇ νυκτὶ ταύτῃ <add place="above">πέμψας μοι</add>  
<lb n="5"/>τρία <num value="3"/> ζεύγη ἄρτων ἐπὶ οὐκ <choice><reg>ἔ  
<lb n="6" break="no"/>χω</reg><orig>ἐ<lb n="6" break="no"/>χο</orig></choice> ἄρτους καὶ ὅταν ἔλθῃ ἡ <choice><reg>πο  
<lb n="7" break="no"/>ρεία</reg><orig>πο<lb n="7" break="no"/>ρήα</orig></choice> πέμψω <choice><reg>σοι</reg><orig>συ</orig></choice> αὐτά.
```

# Preprocessing in SEMATIA

- Getting rid of the tags but preserving the editorial work
- Dividing the document where handwriting changes: Acts of writing



Two layers of each act of writing

ORIGINAL	Editorial STANDARD
Presents the original forms which the ancient writer wrote and nothing more	Presents supplied text, regularized spellings and expanded abbreviations
Includes <b>dummy markers</b> for the supplied forms in the standard layer in order to keep tokenization identical in both layers	Includes mostly whole text, but can also include some <b>dummy markers</b>

# Layers in Sematia

Original

Supplied text: SU

Lacuna, i.e. gap : G

ΠετενεφSU SU τῷ ἀδελφῷ ποSU SU . SU πυήσις, ἀδελφε, ἐὰSU SU ἡ πορήε τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχο ἄρτους καὶ ὅταν ἔλθῃ ἡ πορήα πέμψω σοι αὐτά. ἀσπάζομε Κολοφονήν. καλῶς πυης, ἀδελφε, ὅταν ἔλθῃ ἡ πορήα ἔτησον τῶν χαλκὸν τον τασκου καὶ πέμψῃς με τὰς ἄλλαA A θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσα.

Πετενεφώτης Οὐαλερίω τῷ ἀδελφῷ πολλὰ χαίρειν. καλῶς ποιήσεις, ἀδελφε, ἐὰν ἔλθῃ ἡ πορεία τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχω ἄρτους καὶ ὅταν ἔλθῃ ἡ πορεία πέμψω σοι αὐτά. ἀσπάζομε Κολοφονήν. καλῶς ποιήσεις, ἀδελφε, ὅταν ἔλθῃ ἡ πορεία αἴτησον τὸν χαλκὸν τῶν τασκου καὶ πέμψῃς μοι τὰς ἄλλας δραχμὰς θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσο.

Standard

Supplied text: as supplied

Lacuna, i.e. gap : G

<lb n="2"/>τῷ ἀδελφῷ π<unclear>o</unclear><supplied reason="lost">λλὰ <expan>χ<ex>αίρειν</ex></expan>. καλῶς</supplied>  
<lb n="3"/><choice><reg>ποιήσεις</reg><orig>πυήσις</orig></choice>, ἀδελφε, <unclear>ἐὰ</unclear><supplied reason="lost">ν ἔλθῃ</supplied>  
<lb n="4"/>ἡ <choice><reg>πορεία</reg><orig>πορήε</orig></choice> τῇ νυκτὶ ταύτῃ <add place="above">πέμψας μοι</add>  
<lb n="5"/>τρία <num value="3"/> ζεύγη ἄρτων ἐπὶ οὐκ <choice><reg>ἔ</reg>  
<lb n="6" break="no"/>χω</reg><orig>ἐ</orig><lb n="6" break="no"/>χο</orig></choice> ἄρτους καὶ ὅταν ἔλθῃ ἡ <choice><reg>πο</reg>  
<lb n="7" break="no"/>ρεία</reg><orig>πο</orig><lb n="7" break="no"/>ρήα</orig></choice> πέμψω <choice><reg>σοι</reg><orig>σου</orig></choice> αὐτά.

# Layers in Sematia

Original

From <choice>: <orig>

ΠετενεφSU SU τῷ ἀδελφῷ ποSU SU . SU πυήσις, ἀδελφε, ἐὰSU SU ἡ πορήε τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχο ἄρτους καὶ ὅταν ἔλθῃ ἡ πορήα πέμψω σὺ αὐτά. ἀσπάζομε Κολοφονήν. καλῶς πυης, ἀδελφε, ὅταν ἔλθῃ ἡ πορήα ἔτησον τῶν χαλκὸν τον τασκου καὶ πέμψῃς με τὰς ἄλλαA A θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσA.

Πετενεφώτης Οὐαλερίω τῷ ἀδελφῷ πολλὰ χαίρειν. καλῶς ποιήσεις, ἀδελφε, ἐὰν ἔλθῃ ἡ πορεία τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχω ἄρτους καὶ ὅταν ἔλθῃ ἡ πορεία πέμψω σοι αὐτά. ἀσπάζομε Κολοφονήν. καλῶς ποιήσεις, ἀδελφε, ὅταν ἔλθῃ ἡ πορεία αἴτησον τὸν χαλκὸν τῶν τασκου καὶ πέμψῃς μοι τὰς ἄλλας δραχμὰς θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσο.

Standard

From <choice>: <reg>

```
<lb n="2"/>τῷ ἀδελφῷ π<unclear>o</unclear><supplied reason="lost">λλὰ <expan>χ<ex>αίρειν</ex></expan>. καλῶς</supplied>  
<lb n="3"/><choice><reg>ποιήσεις</reg><orig>πυήσις</orig></choice> ἀδελφε, <unclear>ἐὰ</unclear><supplied reason="lost">ν ἔλθῃ</supplied>  
<lb n="4"/>ἡ <choice><reg>πορεία</reg><orig>πορήε</orig></choice> τῇ νυκτὶ ταύτῃ <add place="above">πέμψας μοι</add>  
<lb n="5"/>τρία <num value="3"/> ζεύγη ἄρτων ἐπὶ οὐκ <choice><reg>ἔ</reg>  
<lb n="6" break="no"/>χω</reg><orig>ἐ<lb n="6" break="no"/>χο</orig></choice> ἄρτους καὶ ὅταν ἔλθῃ ἡ <choice><reg>πο</reg>  
<lb n="7" break="no"/>ρεία</reg><orig>πο<lb n="7" break="no"/>ρήα</orig></choice> πέμψω <choice><reg>σοι</reg><orig>συ</orig></choice> αὐτά.
```



# Layers in Sematia

Original  
Abbreviations: A

ΠετενεφSU SU τῷ ἀδελφῷ ποSU SU .SU πυήσις, ἀδελφε, ἐὰSU SU ἡ πορήε τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχο ἄρτους καὶ ὅταν ἔλθῃ ἡ πορήα πέμψω σὺ αὐτά. ἀσπάζομε Κολοφονήν. καλῶς πυης, ἀδελφε, ὅταν ἔλθῃ ἡ πορήα ἔτησο ν τῶν χαλκὸν τον τασκου καὶ πέμψης με τὰς ἄλλαA A θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσA.

Πετενεφώτης Οὐαλερίω τῷ ἀδελφῷ πολλὰ χαίρειν. καλῶς ποιήσεις, ἀδελφε, ἐὰν ἔλθῃ ἡ πορεία τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχω ἄρτους καὶ ὅταν ἔλθῃ ἡ πορεία πέμψω σοι αὐτά. ἀσπάζομε Κολοφονήν. καλῶς ποιήσεις, ἀδελφε, ὅταν ἔλθῃ ἡ πορεία αἴτησον τὸν χαλκὸν τῶν τασκου καὶ πέμψης μοι τὰς ἄλλας δραχμὰς θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσο.

Standard  
Abbreviations: expanded

<expan>ἄλλα<ex>ς</ex></expan>

<expan><ex>δραχμὰς</ex></expan>

<lb n="2"/>τῷ ἀδελφῷ π<unclear>ο</unclear><supplied reason="lost">λλὰ<expan>χ<ex>αίρειν</ex></expan>. καλῶς</supplied>  
<lb n="3"/><choice><reg>ποιήσεις</reg><orig>πυήσις</orig></choice>, ἀδελφε, <unclear>ἐὰ</unclear><supplied reason="lost">ν ἔλθῃ</supplied>  
<lb n="4"/>ἡ <choice><reg>πορεία</reg><orig>πορήε</orig></choice> τῇ νυκτὶ ταύτῃ <add place="above">πέμψας μοι</add>  
<lb n="5"/>τρία <num value="3"/> ζεύγη ἄρτων ἐπὶ οὐκ <choice><reg>ἔ</reg>  
<lb n="6" break="no"/>χω</reg><orig>ἐ<lb n="6" break="no"/>χο</orig></choice> ἄρτους καὶ ὅταν ἔλθῃ ἡ <choice><reg>πο</reg>  
<lb n="7" break="no"/>ρεία</reg><orig>πο<lb n="7" break="no"/>ρήα</orig></choice> πέμψω <choice><reg>σοι</reg><orig>συ</orig></choice> αὐτά.

# Layers in Sematia

Original

Uncertain letters: as certain

ΠετενεφSU SU τῷ ἀδελφῷ ποSU SU . SU πυήσις, ἄδελφε, ἐὰSU SU ἡ πορήε τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχο ἄρτους καὶ ὅταν ἔλθῃ ἡ πορήα πέμψω σοι αὐτά. ἀσπάζομε Κολοφονήν. καλῶς πυης, ἄδελφε, ὅταν ἔλθῃ ἡ πορήα ἔτησον τῶν χαλκὸν τον τασκου καὶ πέμψῃς με τὰς ἄλλαA A θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσA.

Πετενεφώτης Οὐαλερίω τῷ ἀδελφῷ πολλὰ χαίρειν. καλῶς ποιήσεις, ἄδελφε, ἐὰν ἔλθῃ ἡ πορεία τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχω ἄρτους καὶ ὅταν ἔλθῃ ἡ πορεία πέμψω σοι αὐτά. ἀσπάζομε Κολοφονήν. καλῶς ποιήσεις, ἄδελφε, ὅταν ἔλθῃ ἡ πορεία αἴτησον τὸν χαλκὸν τῶν τασκου καὶ πέμψῃς μοι τὰς ἄλλας δραχμὰς θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσο.

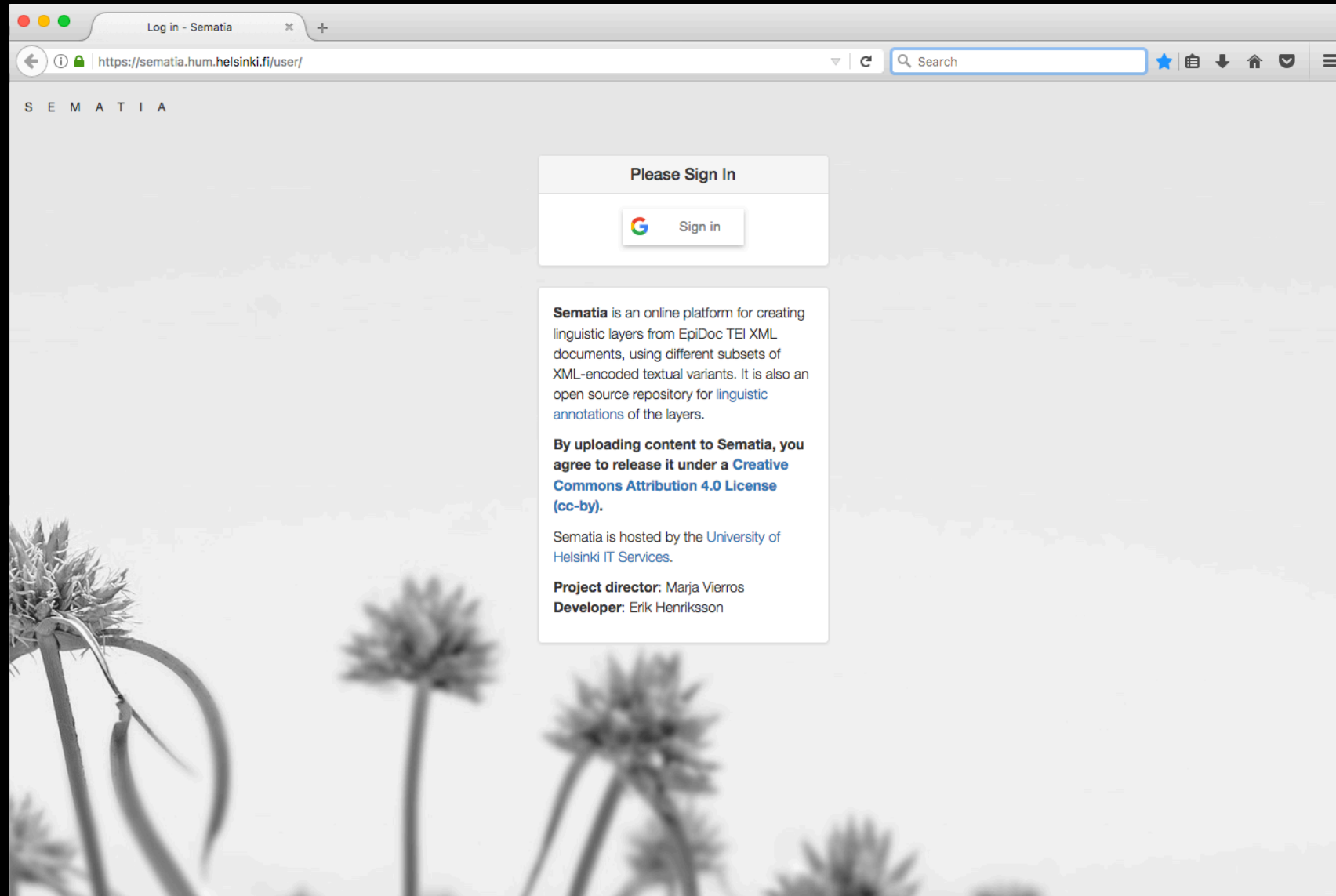
Standard

Uncertain letters: as certain

<lb n="2"/>τῷ ἀδελφῷ π<unclear>o</unclear><supplied reason="lost">λλὰ <expan>χ<ex>αίρειν</ex></expan>. καλῶς</supplied>  
<lb n="3"/><choice><reg>ποιήσεις</reg><orig>πυήσις</orig></choice>, ἄδελφε, <unclear>ἐὰ</unclear><supplied reason="lost">ν ἔλθῃ</supplied>  
<lb n="4"/>ἡ <choice><reg>πορεία</reg><orig>πορήε</orig></choice> τῇ νυκτὶ ταύτῃ <add place="above">πέμψας μοι</add>  
<lb n="5"/>τρία <num value="3"/> ζεύγη ἄρτων ἐπὶ οὐκ <choice><reg>ἔ</reg>  
<lb n="6" break="no"/>χω</reg><orig>ἐ<lb n="6" break="no"/>χο</orig></choice> ἄρτους καὶ ὅταν ἔλθῃ ἡ <choice><reg>πο</reg>  
<lb n="7" break="no"/>ρεία</reg><orig>πο<lb n="7" break="no"/>ρήα</orig></choice> πέμψω <choice><reg>σοι</reg><orig>συ</orig></choice> αὐτά.



# Using Sematia: [www.sematia.hum.helsinki.fi](https://sematia.hum.helsinki.fi)

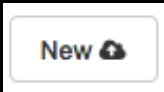


- Google sign-in

# Adding a new text

- Info-button:  
How to Use

Step 1: Click on



Add document

XML Document URL Example: <http://papyri.info/ddbdp/p.petra;1;1/source>

Cancel Import to Sematia

Step 2: Find the text you want to add from papyri.info

DDbDP transcription: o.claud.2.280 **[xml]**

II spc Mons Claudianus

Ἡρακλειδη(\*) καὶ Διονυσίῳ πα-  
νισκος(\*) καὶ Ἑρμῖνος(\*) τοῖς ἀμ-  
φοτέροις πολλὰ χαίρειν). τὸ προσ-  
κύνημα ὑ(\*)μῶν ποιῶμεν(\*)  
5 παρὰ τῇ Τύχῃ τοῦ πραισιδίου.  
ἔγραψες ἡμεῖν τότε ὅτι 'ἔ-

Step 3: click [xml] in the DDbDP transcription title, and the address line gives you the correct URL:



Step 4: Copy and paste the URL in the Add document box and click Import to Sematia

Add document

XML Document URL Example: <http://papyri.info/ddbdp/p.petra;1;1/source>

Cancel Import to Sematia

# Acts of writing and writer&author metadata

Title		Data	
bgu.3.994		<div><div><div>&gt; -113</div><div>&lt; -113</div><div>Pathyris</div><div>24</div></div><div><div>m1</div><div><div>original</div><div>standard</div><div>variation</div></div><div><div>Description (edition)</div><div>Description (custom)</div><div>Professional</div><div>Scibal name</div><div>Scibal title</div><div>Text type</div></div><div><div>Cursive</div><div>Perhaps the "Ammonios-hand", cf. Vierros 2012, 94. (autopsy)</div><div>Professional</div><div>Ammonios; Heliodoros; Sosos</div><div>agoranomos</div><div>Sale contract</div></div></div><div><div>m2</div><div><div>original</div><div>standard</div><div>variation</div></div><div><div>Professional</div><div>Scibal name</div><div>Scibal title</div><div>Text type</div></div><div><div>Professional</div><div>Apollonios</div><div>banker</div><div>Bank and taxation</div></div></div></div>	
c.ep.lat.144		<div><div>&gt; 101</div><div>&lt; 125</div><div>Alexandria (?)</div><div>1</div></div>	
chla.4.249		<div><div>&gt; 101</div><div>&lt; 200</div><div>Philadelphia (Arsinoites)</div><div>1</div></div>	
chr.mitt.233		<div><div>&gt; -111</div><div>&lt; -111</div><div>Pathyris</div><div>12</div></div>	

Edit hand

Handwriting

Description in the edition:

Cursive

Custom description:

Perhaps the "Ammonios-hand", cf. Vierros 2012, 94. (autopsy)

Professional:

Professional

Same hand in other documents:

Actual Writer

Name:

Title:

Trismegistos ID:

Scribal Official

Name:

Ammonios; Heliodoros; Sosos

Title:

agoranomos

Trismegistos ID:

Author

Name:

Title:

Trismegistos ID:

# The layer view

SEMATIA

Type

original

Hand no.

1

Document

o.claud.2.245

Contributors

Lauri Marjamäki

URL

http://papyri.info/ddbdp/o.claud;2;245/source

Elements

add

choice

ex

expan

Layer

Show treebank

Transcription

1 ΠΕΤΕΝΕΦΩΤΗΣ ΟΥΑΛΕΡΙΩ 2 τῷ ἀδελφῷ πολλὰ χαίρειν . καλῶς 3

ποιήσεις πυήσις , ἀδελφε, ἐὰν ἔλθῃ 4 ἢ πορεία πορήε τῇ νυκτὶ ταύτῃ πέμψας μοι

5 τρία ζεύγη ἄρτων ἐπὶ οὐκ ἐ 6 χω ἐ 6 χο ἄρτους καὶ ὅταν ἔλθῃ ἡ

πο 7 ρεία πο 7 ρήα πέμψω σοι σου αὐτά. ἀσπάζομε 8 Κολοφονήν. καλῶς

ποιήσεις πυης , ἀδελ 9 φε, ὅταν ἔλθῃ ἡ


πορεία αἴτῃ 10 σον τὸν πορήα ἔτῃ 10 σο ν τῶν χαλκὸν τῶν τον τασκου

11 καὶ πέμψεις μοι πέμψης με τὰς ἄλλας 12 δραχμὰς θ ἵνα πληρωθῇς ὧν

13 μοι ἔδωκας. ἔρρωσο .

Plain text

ΠΕΤΕΝΕΦSU SU τῷ ἀδελφῷ ποSU SU . SU πυήσις, ἀδελφε, ἐὰSU SU ἡ πορήε τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχο ἄρτους καὶ ὅταν ἔλθῃ ἡ πορήα πέμψω σου αὐτά. ἀσπάζομε Κολοφονήν. καλῶς πυης, ἀδελφε, ὅταν ἔλθῃ ἡ πορήα ἔτῃ σο ν τῶν χαλκὸν τον τασκου καὶ πέμψης με τὰς ἄλλας Α Α θ ἵνα πληρωθῇς ὧν μοι ἔδωκας. ἔρρωσα.

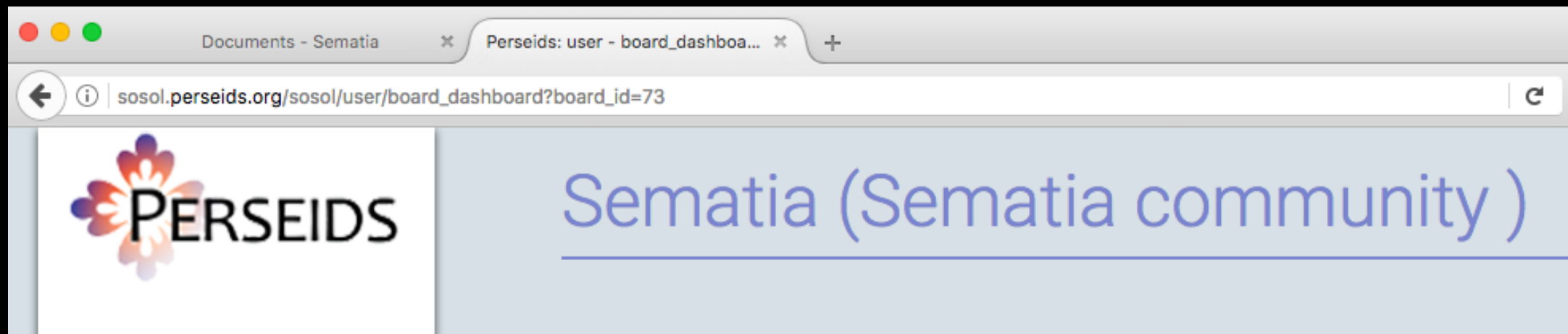


You should export BOTH layers separately to Perseids' Arethusa annotation service for treebanking

If several suggestions for a reading are presented in the standard layer, you can choose the one you think best in the transcription window -> change will appear in the Plain text window. In the original layer you should be very cautious in changing anything.

# External linguistic annotation

- According to the Ancient Greek and Latin Dependency Treebank Guidelines
  - See e.g. Sunoikisis DC 2017, session 3 on how.
- In Perseids/Arctura (<http://sites.tufts.edu/perseids/>)
  - Automatic tokenization
  - Semiautomatic lemmatization and morphological analysis (annotator checks!)
  - Manual syntactic annotation by the annotator
- Submit annotated texts to the Sematia board

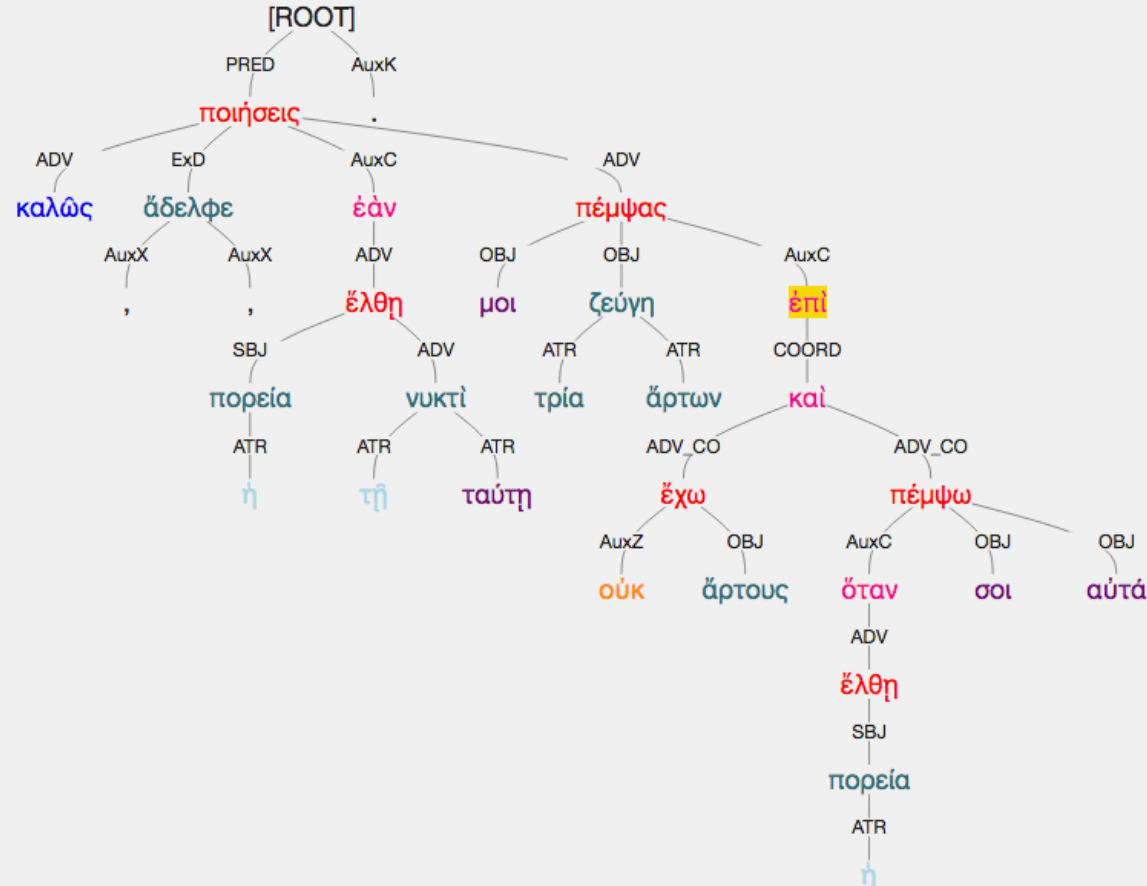




# Standard layer annotation

καλῶς ποιήσεις , ἀδελφε , ἐὰν ἔλθῃ ἢ πορεία τῇ νυκτὶ ταύτῃ πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχω  
ἄρτους καὶ ὅταν ἔλθῃ ἢ πορεία πέμψω σοι αὐτά .

selection none 0 unused highlight unused



morph

relation

aT

selector

history

comments



0 of 30 unused

ἐπὶ 2-18



ἐπὶ

r-----

bsp/morpheus

adp



ἐπεὶ

c-----

document

conj

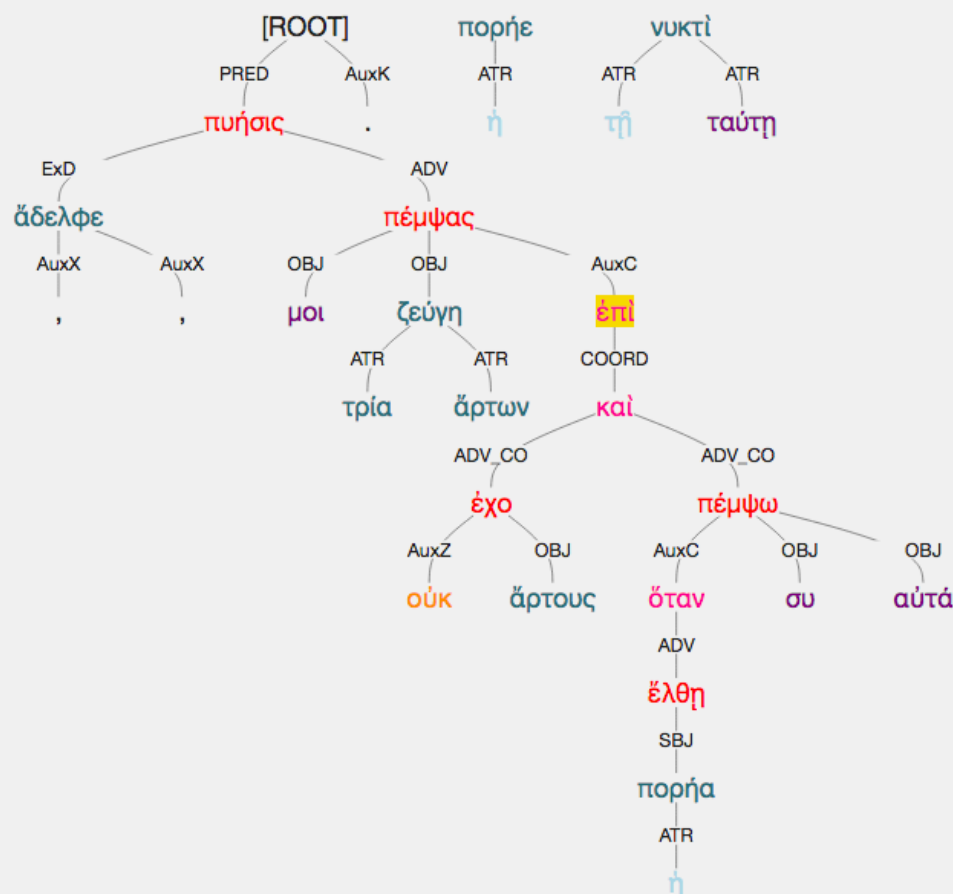
Note that even in standard layer some forms may not be in regularized form!

- Annotator can create new lemmas in the tree view as well as change erroneous forms in the xml in Perseids

# Original layer annotation

SU πυήσις , ἄδελφε , ἐὰν SU SU ἡ πορήε τη νυκτὶ ταύτη πέμψας μοι τρία ζεύγη ἄρτων ἐπὶ οὐκ ἔχο ἄρτους καὶ ὅταν ἔλθῃ ἡ πορήα πέμψω συ αὐτά .

selection none 5 unused highlight unused



morph relation aT selector history comments

ἐπὶ 2-18

3 of 30 unused

<input type="checkbox"/>	ἐπὶ	r-----	bsp/morpheus
	adp		
<input checked="" type="checkbox"/>	ἐπεὶ	c-----	document
	conj		

In the original layer, annotate according to the original forms and only as much as has survived; no supplied or abbreviated case endings etc.

- Result may be a fragmentary tree

# Treebank XML of o.claud.2.245 sentence 2 (original layer)

```
<sentence document_id="https://sematia.hum.helsinki.fi/edit/553" id="2" span="" subdoc="">
  <word form="SU" head="" id="1" lemma="" postag="" relation="" />
  <word form="πυήσις" head="0" id="2" lemma="ποιέω" postag="vZsfia---" relation="PRED" />
  <word form="," head="4" id="3" lemma="punc1" postag="u-----" relation="AuxX" />
  <word form="ἄδελφε" head="2" id="4" lemma="ἄδελφός" postag="n-s---mv-" relation="ExD" />
  <word form="," head="4" id="5" lemma="punc1" postag="u-----" relation="AuxX" />
  <word form="ἐὰνSU" head="" id="6" lemma="" postag="" relation="" />
  <word form="SU" head="" id="7" lemma="" postag="" relation="" />
  <word form="ἦ" head="9" id="8" lemma="ὅ" postag="l-s---fn-" relation="ATR" />
  <word form="πορήε" head="" id="9" lemma="πορεία" postag="n-s---fn-" relation="" />
  <word form="τῇ" head="11" id="10" lemma="ὅ" postag="l-s---fd-" relation="ATR" />
  <word form="νυκτὶ" head="" id="11" lemma="νύξ" postag="n-s---fd-" relation="" />
  <word form="ταύτῃ" head="11" id="12" lemma="οὗτος" postag="p-s---fd-" relation="ATR" />
  <word form="πέμψας" head="2" id="13" lemma="πέμπω" postag="v-sapamn-" relation="ADV" />
  <word form="μοι" head="13" id="14" lemma="ἐγώ" postag="p1s---md-" relation="OBJ" />
  <word form="τρία" head="16" id="15" lemma="τρία" postag="n-p---na-" relation="ATR" />
  <word form="ζεύγη" head="13" id="16" lemma="ζεύγος" postag="n-p---na-" relation="OBJ" />
  <word form="ἄρτων" head="16" id="17" lemma="ἄρτος" postag="n-p---mg-" relation="ATR" />
  <word form="ἐπὶ" head="13" id="18" lemma="ἐπεί" postag="c-----" relation="AuxC" />
  <word form="οὐκ" head="20" id="19" lemma="οὐ" postag="d-----" relation="AuxZ" />
  <word form="ἔχω" head="22" id="20" lemma="ἔχω" postag="v1spia---" relation="ADV_CO" />
  <word form="ἄρτους" head="20" id="21" lemma="ἄρτος" postag="n-p---ma-" relation="OBJ" />
  <word form="καὶ" head="18" id="22" lemma="καί" postag="c-----" relation="COORD" />
  <word form="ὅταν" head="27" id="23" lemma="ὅταν" postag="c-----" relation="AuxC" />
  <word form="ἔλθῃ" head="23" id="24" lemma="ἔρχομαι" postag="v3sasa---" relation="ADV" />
  <word form="ἦ" head="26" id="25" lemma="ὅ" postag="l-s---fn-" relation="ATR" />
  <word form="πορήα" head="24" id="26" lemma="πορεία" postag="n-s---fn-" relation="SBJ" />
  <word form="πέμψω" head="22" id="27" lemma="πέμπω" postag="v1sfia---" relation="ADV_CO" />
  <word form="σου" head="27" id="28" lemma="σύ" postag="p-s---mn-" relation="OBJ" />
  <word form="αὐτά" head="27" id="29" lemma="αὐτός" postag="p-p---na-" relation="OBJ" />
  <word form="." head="0" id="30" lemma="punc1" postag="u-----" relation="AuxK" />
</sentence>
```

Postag:

- 1: part-of-speech
- 2: person
- 3: number
- 4: tense
- 5: mood
- 6: voice
- 7: gender
- 8: case
- 9: degree

o.claud.2.243		> 101	< 200	Tiberiane	1  2	Lauri Marjamäki	
o.claud.2.245		> 101	< 200	Tiberiane	1  2	Lauri Marjamäki	
<div> <div> <div>original</div> <div>standard</div> <div>variation</div> </div> <div> </div> <div> </div> </div> <div> <p>Same hand</p> <p>http://papyri.info/ddbdp/o.claud;2;243, http://papyri.info/ddbdp/o.claud;2;245, http://papyri.info/ddbdp/o.claud;2;246, http://papyri.info/ddbdp/o.claud;2;247, http://papyri.info/ddbdp/o.claud;2;248, http://papyri.info/ddbdp/o.claud;2;249, http://papyri.info/ddbdp/o.claud;2;250, http://papyri.info/ddbdp/o.claud;2;251, http://papyri.info/ddbdp/o.claud;2;252, http://papyri.info/ddbdp/o.claud;2;253, http://papyri.info/ddbdp/o.claud;2;254</p> <p>Writer name Petenephotes</p> <p>Writer TM ID 144388</p> <p>Author name Petenephotes</p> <p>Author TM ID 144388</p> <p>Text type Private letter</p> <p>Addressee Private</p> <p>Addressee name Valerius</p> <p>Addressee TM ID 144389</p> </div>							
o.claud.2.246		> 101	< 200	Tiberiane	1  2	Lauri Marjamäki	
o.claud.2.248		> 101	< 200	Tiberiane	1  2	Lauri Marjamäki	
o.claud.2.249		> 101	< 200	Tiberiane	1  2	Lauri Marjamäki	

# The corpus and queries

- Size still microscopic:
  - ca. 100 texts annotated = 505 sentences = ca. 7000 tokens (one layer count)
  - Growing (new annotators welcome!)
- Querying
  - Exporting layers to external query engines (PML Tree Query, XQuery etc.)
  - Querying in Sematia (beta version exists): layer comparisons!

## Tools

Export Layers

All

Original

Standard

Export Word Frequencies

All

Original

Standard

Hierarchical Treebank Clustering

All

Original

Standard



Original

Word

A >\_

Lemma

A >\_

Relation

>\_

Postag

^.....d

>\_

Standard

Word

A >\_

Lemma

A >\_

Relation

>\_

Postag

^.....g

>\_



Results

Show

50

▼

entries

Search:

S-W	Word	Lemma	Relation	Postag	Hand	Doc.	TM IDs
1-9	Ἀπολλωνίωι Ἀπολλωνίου	Ἀπολλώνιος Ἀπολλώνιος	ADV ADV	a-s---md- a-s---mg-	m1	upz.1.13	35723 0 35723 12661
2-3	Μαρονᾶτι Μαρωνᾶτος	Μαρονᾶς Μαρωνᾶς	ADV ADV	n-s---md- n-s---mg-	m1	o.claud.2.248	144399 0 144399 144400
2-3	Λογγινᾶτι Λογγινᾶτος	Λογγινᾶς Λογγινᾶς	ADV ADV	n-s---md- n-s---mg-	m1	o.claud.2.249	144405 0 144405 144406

Querying  
dative vs.  
genitive  
case

- Create an account in Perseids:
  - <http://sites.tufts.edu/perseids/>
- Annotation Guidelines:
  - Ancient Greek:  
[https://github.com/PerseusDL/treebank\\_data/blob/master/AGDT2/guidelines/Greek\\_guidelines.md](https://github.com/PerseusDL/treebank_data/blob/master/AGDT2/guidelines/Greek_guidelines.md)
  - older version:  
<http://nlp.perseus.tufts.edu/syntax/treebank/greekguidelines.pdf>
  - Latin: <http://nlp.perseus.tufts.edu/syntax/treebank/1.3/docs/guidelines.pdf>