From Annotated Text to Vocabulary Exercises

Sunoikisis Digital Classics, Summer 2019

Thursday June 13, 17:00 - 18:15 CEST
Neven Jovanović, Petar Soldo

University of Zagreb
Faculty of Humanities and Social Sciences
Department of Classical Philology
github.com/SunoikisisDC/SunoikisisDC-2018-2019/wiki/Summer2019-Session11

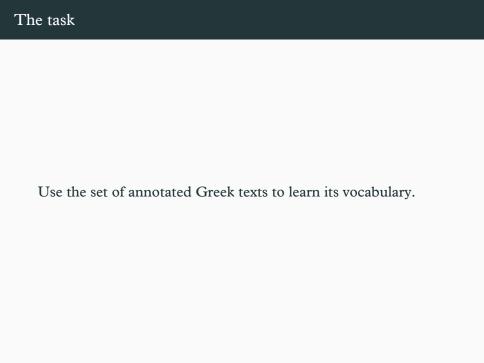# The plan

The plan

The task

Materials and methods

Tools

Creating the database

Creating the exercises

Demonstration (results)

Discussion

The task

Use the set of annotated Greek texts to learn its vocabulary.

Materials and methods

Lemmatize words in a text

Add morphological descriptions of lemmatized words (parts of speech)

Reorder words by form, lemma, part of speech

Regroup words by forms, lemmata, parts of speech (a group is a set of identical forms, identical lemmata etc; from it we get information about frequency of an item in the collection)

Analyze frequencies, retrieve reoccurring forms, lemmata, parts of speech

Select certain forms, lemmata, parts of speech

Retrieve all forms from a text (but not punctuation)

For a form, retrieve lemma: $\pi\rho\acute{\alpha}\xi\varepsilon\iota\varsigma \rightarrow \pi\rho\hat{\alpha}\xi\iota\varsigma$

For a form, retrieve part of speech information:
$\pi\rho\acute{\alpha}\xi\varepsilon\iota\varsigma \rightarrow$ n - p - - - fa -

XML manipulation, common tasks

For a lemma, find whether it occurs in other texts, or in another database (DC Greek Core)

For a lemma, retrieve form in a text

For a lemma, retrieve (Croatian) meaning from a lexical database (DC Greek Core)

Format results for import into Anki, our vocabulary learning
program (a simple three-fields CSV is needed)

Spaced repetition: we tend to remember things more effectively if we spread reviews out over time, instead of studying multiple times in one session.

Make your own cards!

The act of making the card helps you remember the stuff on it.

The way you learn and the things you care about are unique to you.

Tools

Arethusa / Perseids

BaseX / XQuery

Anki

Dickinson College Core Vocabulary

"The DCC Core Vocabulary lists represent the thousand most common words in Latin and the 500 most common words in ancient Greek. They were originally composed in 2012–13 by a team at Dickinson College led by Christopher Francese. The lists are downloadable in various formats and licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License."

Creating the database

Creating the exercises

1. For a word form, give the lemma (standard vocabulary entry):
Q πράξεις, A πρᾶξις

2. For a word form, give the grammatical description:
Q *πράξεις*, A noun, feminine, plural, nominative / accusative / vocative

3. For a lemma (of a word from the DC Greek core list), give the translation:

Q *ποταμός*, A rijeka

Demonstration (results)

Discussion

From Annotated Text to Vocabulary Exercises

Sunoikisis Digital Classics, Summer 2019

Thursday June 13, 17:00 - 18:15 CEST
Neven Jovanović, Petar Soldo

University of Zagreb
Faculty of Humanities and Social Sciences
Department of Classical Philology
github.com/SunoikisisDC/SunoikisisDC-2018-2019/wiki/Summer2019-Session11