

Aim of the SLUW-Project is to produce a study of the **Syntax (and Morphosyntax) of the Luwian language.**

The Principal Investigator is **Dr. Federico Giusfredi**
under the scientific supervision of **Prof. Paola Cotticelli-Kurras.**

Luwian was a language belonging to the Ancient Anatolian branch of Indo-European. It was spoken and written in Anatolia and Syria between the first half of the second millennium to the end of the eight century BCE.



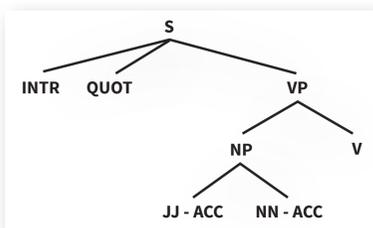
Luwian was written using two different writing systems, the cuneiform script (on clay tablets) and Anatolian hieroglyphic one (on stone, metal or other materials).

In order to work on a statistically significant range of materials, the texts to be examined will belong both to the cuneiform and the hieroglyphic corpora, and all the attested text-types (e.g. rituals, "songs", myths, royal inscriptions, letters) will be taken into consideration.

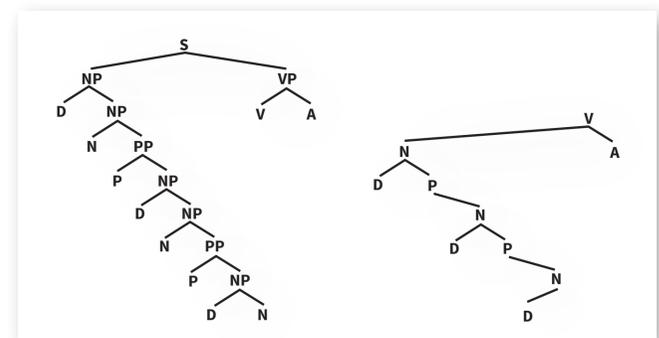
A substantial number of complete texts and single sentences will be analyzed in a Treebank based on an expanded and adapted version of the POS tagging protocol of the Penn Treebank II (for the annotation of constituents). The protocol is expanded and adapted in order to better match the structure of the Anatolian languages and to provide the type of information necessary for the study of syntax and morphosyntax; for instance, tags are employed to mark agreement and alignment.

- **Host Institution:** Università degli Studi di Verona.
- **P.I.:** Federico Giusfredi.
- **Scientific Supervisor:** Paola Cotticelli-Kurras.
- **Timing:** June 2015 - May 2017.

Luwian
Anatolian language (Indo-European)
1800 to 700 BCE
Ca. 700 tokens
Cuneiform Corpus
Hieroglyphic Corpus



Constituency Treebanks help to highlight the S-Structure of sentences, by annotating the single phrases and their specific attributes.



Dependency Treebanks, on the contrary, will not be employed, because they require a complete syntactic theory to start from, which is what the SLUW project aims at producing.

The corpus of Luwian Constituency Trees will be analyzed with queries in order to extrapolate standard and non-standard patterns, that will be assessed and explained on the levels of:

- Morphosyntax (phrase level and constituent analysis)
- Syntax-Semantics interface
- Interference with other languages
- Other social factors

and, where possible:

- Pragmatics of the written text
- Stylistics

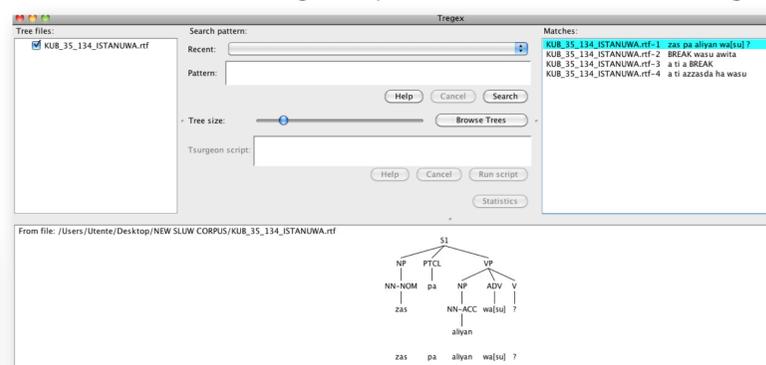
The final results will be organized in a monographic study of the Luwian morpho-syntax at the end of the project; relevant intermediate results will be made available earlier in the timeline.

The first step of the project consists in annotating the morpho-syntax of the selected corpus, using a tagset derived from the one used for the Penn II Treebank, with the addition of specific extensions that will be used to mark the special features of an Anatolian language (e.g.: a large set of clitics and sentence particles, complex marker of possession such as the genitival adjectives, omitted conjunctions).

A short sample of how an annotated Luwian sentence looks like:

```
(S
  (ADV unu)
  (PTCL pa)
  (QUOT wa)
  (VP
    (NP (JJ-ACC sanawinzi) (NN-ACC atutinzi))
    (V-2SI harawani)
  )
)
```

Texts will be browsed using the Open Source software Stanford Tregex:



As soon as a sufficient amount of texts will have been collected, the annotated corpus will be analyzed in order to study the morpho-syntax of Luwian.



The results will be published in papers, in a final monograph as well as in open-data, made accessible on the web (on the project's website <http://luwiansyntax.info>). Workshops will be organized in order to share the results and to promote the circulation of ideas in the scientific community.



Contacts:
federico.giusfredi@univr.it
paola.cotticelli@univr.it

Website:
<http://luwiansyntax.info>