# India 2011 - workshop

Our workshop is meant to jump-start work on a phenomenology of grammatical features and construction types for several Indian languages

as listed under: [workshop](#)

Our focus is on distinguishing features in selected areas of the grammars. Our approach combines morpheme-level digital annotation with AVM-based syntactic-semantic theoretical analysis.

Lars and I would now like to elicit proposals for what phenomena should be focused on. As a point of departure, we list some phenomena which we have seen mentioned in the literature.

(1) Verb morphology:
      Finite and non-finite verb morphology
      Tense-aspect agreement across main and subordinate verb phrases and in multi-verb constructions

(2) Verb constructions
      Compound Verbs (vector versus polar verbs)
      Conjunct Verbs
      Echo Verbs

Lars has written a short [overview](#) over various types of multi-verb constructions and 'complex predicates'.

(3) Syntax of complementizers
(4) Scrambling
(5)  Wh in situ
(6)  Case copying in Telugu
(7)  Reflexives and argument structure alternations
(8)  Honorifics

## Method

We aim at creating  a research corpus using TypeCraft facilities,  partly  for representative collections of examples and partly consisting of coherent text of your choice.

In addition we suggest that all participants agree to record their description of the following short YouTube video called ['Simon's cat'](#) . A similar experiment lead to the [following recordings,](#) documented on TypeCraft.

As for the collections of examples,

(a) we would like to invite the participants to suggest examples of constructions, whether based on occurring text or intuition.  In order to be telling examples have to have a certain length and naturalness, yet since we want to work on the examples for morphological annotation and AVM analysis, they should not be too long either. Again important is that they are telling for the point of interest.

At present we find  [Bengali Constructions](#) represented on the TypeCraft wiki with 37 sentences in the TC database. Telugu is with 585 sentences present in the database . For the latter please go to

Text search and search for Telugu.

(b) we would like to collect seminal articles around the phenomena chosen for focus.
    who can suggest articles?  Which of these articles, which exist electronically - do  need to be
     scanned?

In doing digital text analysis, the issue of script is essential, and will be addressed on the first day, parallel to initiating morphological annotation.

**Keynotes**

As you already know we will also offer keynote talks
A keynote talk will be offered **on Monday** by Prof. Sengupta, on Indic scripts.

I will look at  collaborative corpus creation which to my mind is totally under-estimated in its potentials as a linguistic method and as a means for data harvesting. This will be my introductory note for our workshop.

**On Tuesday,** a discussion on syntactic construction types will parallel the annotation project, and there will be a keynote talk by Prof. K.V. Subbarao on 'The Syntax and Semantics of Non-nominative subject in South Asian languages'.

**On Wednesday** we commence with the issue of AVM representation of constructions, and a  talk on the  integration of such representations into digital linguistic analysis. Prof. L. Hellan gives a keynote talk on this.


**Thursday and Friday** are devoted to pursuing the themes hereby introduced, developing corpora, databases and analyses at various levels. More presentations may be given by the participants.

In the evening on Thursday, Prof. Sengupta gives a public lecture over Brahmi-Derived Scripts