

**Project Part I:
Defining and testing a small grammar**

(Final version due in last week of classes by email attachment to dm@ling.osu.edu)

The final project consists of implementing a grammar for a small text of your choice. More concretely:

- Pick a text consisting of approx. 10 sentences with approx. 100 words. Beware of complicated linguistic phenomena like parentheticals etc.—when in doubt, pick something simpler. While, as usual, you are encouraged to work together, each student should pick a separate text and submit a separate grammar.

By **Thursday, February 28**, send me an email with a copy of the text (or texts if you are thinking of different options) to get it approved for this exercise.

- Think about syntactic analyses for these sentences. Try to capture generalizations where possible. However, your main focus should be to try to license the sentences in your 10 sentence corpus and to exclude related ungrammatical ones.
- Implement and test a grammar incorporating your analyses. Include comments in your grammar on which part does what (but don't overdo it: more than three lines of comment per line of code is too much). Include a test predicate which starts parsing of the examples of your corpus. You should submit a single prolog file containing:

- the commented grammar
- the test suite predicates with your 10 sentences, and
- additional comments, added to the grammar file delimited by `/* ... */`

Please include a comment for each category name (what it stands for, what kind of strings it is supposed to capture) and the same for each rule you specify.

In terms of parser to use, please use the Earley recognizer which you can find in the code directory under `parser/earley/non_atomic/earley.pl` and `earley_trace_verbose.pl`

The grammar format this recognizer expects is exemplified by the small grammar in `parser/earley/non_atomic/grammar.pl`