A minimalist parser for Finnish
A computational model

Mikko Määttä
mikko.maatta@helsinki.fi

Introduction
- This research project started in June 2007.
- The goal of the project is to develop a parser for Finnish that utilizes minimalist ideas and implement the parser as a computational model.

Minimalism
- The main idea of minimalism is to explain linguistic phenomena with only the minimal conceptual apparatus.
- Syntax is viewed as an optimal mapping from the articulatory/perceptual (A/P) cognitive system to the conceptual/intentional (C/I) system.
- Sentences are constructed by the simple operations Merge and Move.

Example 1: Deriving \( \text{DP tuo mies} \) ('that man')
1. The root MIES is merged with the nominalizer n.
2. n is not a free morpheme. Therefore the closest free morpheme, MIES in this case, is moved and attached to it.
3. The determiner TUO is merged with the new complex constituent.

Brattico’s model
- The parser is based on a particular minimalist theory of grammar developed by Pauli Brattico, University of Jyväskylä, Finland.
- In order to reduce computational complexity, the derivation of a sentence proceeds in phases.
  - Certain phrases, e.g. DP and CP, are first assembled as separate units.
  - Later they are combined to form the complete sentence.
  - Within the phases, all morphemes are combined in every possible order.
  - The requirements of the C/I interface (Logical Form) and the A/P interface (Phonetic Form) filter out all but the correct phrase structures.

The parser
- A central hypothesis: The parser and the competence are not separate faculties.
- The task of the parser is to detect the boundaries of the phases in the input sentence.
- This involves a search algorithm that generates hypotheses about the boundaries of phases and tests them by deriving all possible phrase structures from the words in the phase, using the rules of the grammar.

1) Tuo \( \text{DP mies } \) näkee tämän naisen
2) \( \text{DP Tuo mies } \) näkee tämän naisen
3) * \( \text{DP Tuo mies näkee } \) tämän naisen
4) * Tuo \( \text{DP mies näkee } \) tämän naisen

The search algorithm uses heuristics to diminish the search space. For example:
- In every DP there has to be either an overt determiner or a noun.
- In every CP there has to be a finite verb.
- The parser is going to be implemented as an actual computer program.

Further development
- A secondary objective of this project is to use the computational model as a part of a model for actual human language processing.