# Grammar Engineering: Grammar Matrix

21.05.2012

In this exercise, you will build a small HPSG grammar for English based on the Grammar Matrix.

1. Download *grammar04* from `http://www.coli.uni-saarland.de/~yzhang/ge/grammar/grammar04.tgz`.

   Instead of using the online customization interface, you will be starting with *grammar04*, which is essentially a version of Grammar Matrix without language-specific customizations. The types defined in file *matrix.tdl* are provided by the Grammar Matrix, and are considered to be universal to all languages. You should not change anything in *matrix.tdl*, instead, selectively inherit from those universal types, and define your own types in *english.tdl*. If some of the types provided by Matrix do not suit you, redefine them (and their subtypes) in file *english.tdl* with different names. You should also have a look into the file *items*, which contains our test suite.

2. Introduce word categories and lexical types
   Inheriting from the basic types under *lex-item*, introduce lexeme types for nouns, verbs, and determiners. Some of the basic types are already exsiting in Matrix. Inherit from appropriate ones, and specify the necessary extra constraints (e.g. valence) if necessary. For verbs, you need to have at least two types, for transitive and intransitive verbs respectively. You will need to use these types to construct the lexicon in the next step.

3. Build lexicon and morphological rules
   The grammar comes with only one dummy lexical entry. You need to delete it and add lexical entries for words in the test suite (*items*). Try to avoid specifying too much information in the lexicon. Reuse the constraints in types. This allows you to achieve higher generalization

with your grammar. Also build appropriate morphological rules to handle inflections (this should be easy now).

4. Build phrasal rules
   You need to build three phrasal rules for your grammar: *spec-head*, *subj-head* and *head-comp*. But before you carving into *rules.tdl*, make sure you have all appropriate types. It is generally good idea to have more information encoded as types, rather than rules, for the former can always be reused by inheritance. Look at the Matrix type *head-valence-phrase* and its offspring types and decide which ones you need to inherit from. After adding phrase types in *english.tdl*, add corresponding rule entries in *rules.tdl*. Note that the valence phrase types in Grammar Matrix deliberately underspecify word orders. Introduce these extra constraints for your English grammar. Also, you need to introduce agreement constraints to avoid overgeneration. Whenever you introduce new types, always check Matrix to avoid reinventing the wheels.

5. Test and document
   Use the provided test suite *items* to test your grammar. Revise your analysis if your grammar over/under-generates. Document the new types, lexical entries, morphological and phrasal rules you have introduced, together with any problems or questions.