

Language Resources and Evaluation

A GF miniature resource grammar for Tswana: modelling the proper verb

Laurette Pretorius; Laurette Marais; Ansu Berg

Abstract

The Grammatical Framework (GF) not only offers state of the art grammar-based machine translation support between an increasing number of languages through its so-called Resource Grammar Library, but is also fast becoming a de facto framework for developing multilingual controlled natural languages (CNLs). For a natural language to share maximally in the opportunities that GF-based multilingual CNL support presents, it has to have a GF resource grammar. Tswana, an agglutinating Bantu language, spoken in Southern Africa as one of the eleven official languages of South Africa, does not yet have such a grammar. This article reports on the development of a so-called miniature resource grammar, a first step towards a full resource grammar for Tswana. The focus is on the modelling of the Tswana proper verb as it occurs in simple sentences. The (proper) verb is the morphologically most complex word category in Tswana, and therefore constitutes a notable contribution towards the development of a GF resource grammar for Tswana. The computational model is discussed in some detail, implemented and tested on a systematically constructed treebank.