

The goal of this thesis is to create a dependency treebank for Yorùbá, a language with very little pre-existing machine-readable resources. The treebank follows the Universal Dependencies (UD) annotation standard, certain language-specific guidelines for Yorùbá were specified. Known techniques for porting resources from resource-rich languages were tested, in particular projection of annotation across parallel bilingual data.

Manual annotation is not the main focus of this thesis; nevertheless, a small portion of the data was verified manually in order to evaluate the annotation quality. Also, a model was trained on the manual annotation using UDPipe.