

Syntax for "Exotic" languages

- [Developing Universal Dependencies for Wolof](#)
- [Towards a dependency-annotated treebank for Bambara \(Aplonova & Tyers 2018\)](#)

Developing Universal Dependencies for Wolof

The paper can be found [here](#).

Wolof is a Niger-Congo language (however it is Senegambian where Bantu languages are Benue-Congo (citation needed)).

Computational grammar of Wolof in the LFG framework (page 1). This was used to create the first treebank of Wolof (see the ParGramBank paper Sulger et al. 2013).

The dependency treebank created is not the result of automatic conversion of the LFG treebank, though the LFG treebank did serve as a basis for annotation. However, this is because they see significant mapping issues between LFG and UD (though they plan to do this automatic conversion at a later time).

13 noun classes (8 singular, 2 plural, 2 locative and 1 manner). Locus of noun class marking is on the nominal modifiers not the noun.

Determiners encode proximal and distal relations for both the speaker and addressee.

Noun classes in Wolof lack semantic coherence (citing McLaughlin, 1997).

“ Wolof nouns are typically not inflected except for the genitive and possessive case

No adjective category, stative verbs used instead (similar to Swahili though there are still a small set of Adjectives in Swahili).

Towards a dependency-annotated treebank for Bambara (Aplonova & Tyers 2018)

- POS tags automatically converted (using rules), treebank handcrafted
- AS of writing, only 116 sentences with dependency annotations
- Using UD 2.0
- Bambara is predominately isolating
- The Daba analyzer tool was used to create the original Bambara Reference corpus
- Morphological features generated by looking at both the glosses and the morphological breakdown in CBR (the reference corpus).
- compounding and derivation not treated productively so lemmas are not split into compound components
- Original reference corpus did this thing where it had multiple POS tags in cases where the POS was ambiguous. These were resolved using largely manual methods.
- All copulas were annotated as verbs? Weird choice not to have them as aux.
- Topicalization involves resumptive pronouns in Bambara