

# 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

---

Revision #1

Created Wed, Apr 15, 2020 8:51 PM by [kenneth](#)

Updated Wed, Apr 15, 2020 8:58 PM by [kenneth](#)