

Annotation Issues

- relative clauses without overt modifiers
- case with no noun?
- Things to go back and fix in manually annotated corpora
- SWH in UD questions

relative clauses without overt modifiers

Consider this noun phrase: Idadi ya waliofariki (from sentence #6799)

the issue is that waliofariki is a relative clause meaning something like "who have died".

The issue is that there seems to be an elided "watu" which serves as the thing that waliofariki modifies. An alternative, is that waliofariki is a noun derived from fariki. E.g. it's those who have died not who have died.

Currently, I am treating these as nouns derived from verbs. (see # 5260)

case with no noun?

What the heck is up with `mpaka hivi sasa`?

Why do I have an adposition modifying an adverb???

Things to go back and fix in manually annotated corpora

- check that iobj is used correctly. E.g. check that verbs which could have iobj but weren't marked with one, don't have one and check that all use of iobj is appropriate.
 - Follow up: see if HCS has iobj indicated somewhere on nouns. this doesn't seem to be making it through the tagger. NO rules are currently leveraging iobj.
- mark should be used when "prepositions" precede a clause. E.g. after he went outside, after should be connected to went with a mark
- ni is always a copula. kuwa can be sconj. verbal kuwa is a verb. Go back and fix uses of kuwa.
- any cases where there's a hyphenated demonym, use the version generated by the rules and correct it, the tokenizer was fixed to treat these correctly.

Additional things to go back and fix.

gani and ngapi should be changed from ADJ to DET, have their arcs changed to be det and nummod respectively and have PronType=Int added as a morph feature.

ka should be assigned continuative aspect?

a- indefinite tense marker should be assigned some special features.

check that hu- is assigned habitual aspect.

Check that ki- TAM marker is conditional mood.

All infinitive verbs should be given VerbForm=Inf and assigned verbal dependencies. Other

morphological features will also need to be adjusted.

When reduplication happens, the second reduplicant is the head [source](#).

Progress

- Longer sentence sample:
 - ka has been assigned continuative aspect (1 examples).
 - hu has been assigned habitual aspect (no examples).
 - ki has been assigned conditional mood. (5 examples).
 - ni has been assigned copula.
 - Verbal kuwa has been assigned verb status.
 - no hyphenated demonyms left to change.
 - Prepositions with verbal heads are now using the deprel mark (1 change)
 - gani fixed (3 changes)
 - ngapi fixed (no changes)
- Shorter sentence sample:

SWH in UD questions

should a derived noun like waliokusanyika be a Vnoun?

should statives be

what is kuna/hakuna? is it a compound (kuwa + na) or is it related to kuna from arabic? should it be a pleonastic or a copula?