A New Perspective on the Swahili Dual-Complementizer System

Aron Finholt, University of Kansas, finhola@ku.edu

Research on languages which employ two distinct complementizers to introduce a finite embedded clause has been profoundly influential in our understanding of the clausal left periphery. A number of open questions still remain, notably whether the choice of complementizer is purely a result of syntactic factors (e.g., c-selection), or whether the choice stems from different meanings introduced by each complementizer. We investigate these issues in a corpus study of Tanzanian Swahili, which uses two distinct complementizers, *kwamba* and *kuwa*, to introduce a finite indicative clause under a clause-embedding predicate (1). The complementizers are reported to be in free variation, with no interpretive distinction (Ashton, 1944); (Thompson & Schleicher, 2006) a.o.

(1) *Hamisi a-li-ni-ambia*   **kwamba/kuwa** *a-na-penda   kusoma*
     Hamisi 1SM-PAST-1SG.OM-tell COMP/COMP   1SM-PRES-like read.INF
     'Hamisi told me that he likes to read.'   (Mpiranya, 2015:220)

Using a regression-based analysis of Tanzanian Swahili data extracted from the Helsinki Corpus of Swahili 2.0 (≈25 million words), we show that the choice of complementizer in (Tanzanian) Swahili is affected by two factors in the matrix clause known to influence complementizer selection cross-linguistically. We conclude that the choice of *kwamba* or *kuwa* does not depends purely on syntactic selection. Instead, we suggest that it depends on who believes the embedded proposition is true.

Predicate class. The lexical semantics of the embedding predicate is well-known to influence the choice of complementizer (Kiparsky & Kiparsky, 1971); (Hooper & Thompson, 1973); (Noonan, 1979) a.o. This has been interpreted to indicate that complementizer choice is purely a function of syntactic selection. In our study, we coded embedding predicates into the predicate classes from Hooper & Thompson (1973). Comparison across classes in the corpus reveals a clear distinction (see Figure). We find that Doxastic Factives
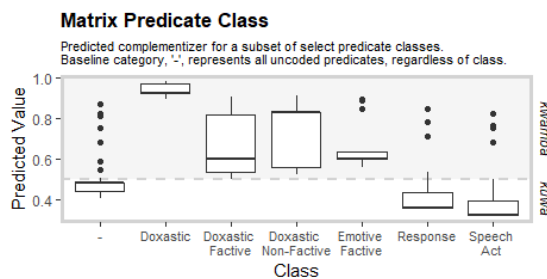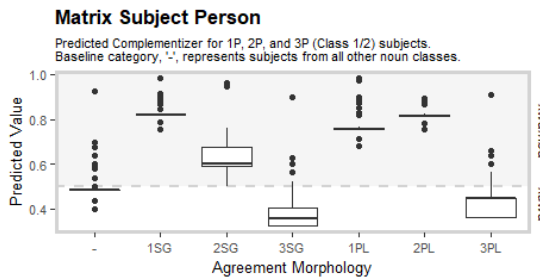

(-*jua* 'know'), Emotive Factives (-*hofia*, 'fear'), and (non-factive) Doxastics (-*kuta*, 'find') are strongly correlated with *kwamba*, while Speech Act Verbs (-*sema* 'say') and Response Predicates (-*kiri* 'admit') are strongly correlated with *kuwa*. Proportionally, 58% of Doxastic Factives (n=1255), 92% of Doxastics (n=74), and 52% of Emotive Factives (n=304) occur with *kwamba*, while, conversely, 70% of Speech Act Verbs (n=7270) and 65% of Response Predicates (n=1374) occur with *kuwa*. The data supports a dichotomy among embedding predicates. There are *attitude verbs*, which necessarily attribute to an argument (the subject) a belief about the embedded proposition. These verbs are correlated with—though do not require—*kwamba*. And then there are *reportative verbs*, which do not require anyone to have a belief about the embedded proposition. These verbs are correlated with—though do not require—*kuwa*. Note that this result is *prima facie* surprising, since *kwamba* is itself

historically a Speech Act verb meaning 'to say/tell,' though in standard Swahili it no longer has this function. *Kuwa* means (synchronically) 'to be.'

Matrix subject. A second factor which has been shown to affect complementizer choice is the person of the matrix subject (Givón & Kimenyi, 1974). In our study, we find that 1st person subject morphology (sg. or pl.) on the matrix verb correlates with *kwamba*, while third-person subject morphology correlates with *kuwa* (see Figure). Though included in the Figure below, 2nd person subjects were ultimately excluded from our analysis due to

Matrix Subject Person
Predicted Complementizer for 1P, 2P, and 3P (Class 1/2) subjects.
Baseline category, '-', represents subjects from all other noun classes.

insufficient sample size (n=379). Of all tokens involving a 1st person matrix subject (n=4366), 71% occur with *kwamba*, and 29% with *kuwa*, while for all tokens involving a third-person subject (n=13103), 30% occur with *kwamba*, and 70% with *kuwa*. The primacy of person becomes clear when we include both lexical class and person in the model. We find that person is *always* the strongest predictor of complementizer choice, potentially overriding any other influencing factors.

Discussion. In contrast to what is generally reported, our findings demonstrate that *kwamba* and *kuwa* in (Tanzanian) Swahili are not in free variation. We interpret the facts above to argue against a purely syntactic explanation for complementizer choice. Instead, we suggest that the C system in Swahili provides a way to express *relative belief* in the embedded proposition P. Specifically, *kwamba* is linked to a local attitudinal *anchor*, who commits to a belief in P. In the context of an attitude predicate (ATTPRED), the use of *kwamba* asserts that the local subject believes that P is true (2).

(2) [X ATTPRED [*kwamba*$_X$ P]]   (3) [X REPPRED [*kwamba*$_{Spkr}$ P]]
     → X believes P is true.   → Speaker believes P is true.

In the absence of any other attitude holder in the matrix clause (i.e., with purely reportative predicates REPPRED), *kwamba* is used to encode speaker belief in the truth of P (3). While *kwamba* is anchored to an attitude holder, we suggest that *kuwa* is entirely neutral. By using *kuwa* in an attitude report, the speaker avoids over-committing to the beliefs of the attitude holder. We thus explain the sensitivity to person (e.g. 1st person subjects correlate with *kwamba*, and 3rd person subjects *kuwa*) because the speaker is intrinsically more knowledgeable of their own beliefs about P than they are the beliefs of a 3rd person subject, and therefore more likely to use *kwamba* in a speaker-oriented attitude report.

Selected references. Ashton, E. O. (1944). *Swahili Grammar (Including Intonation).* | Diercks, M. (2013). Indirect agree in Lubukusu complementizer agreement. *NLLT* 31:2 | Givón, T. & A. Kimenyi (1974). Truth, belief and doubt in Kinyarwanda. SAL 5 | Hooper, J. B. & S. A. Thompson (1973). On the Applicability of Root Transformations. *LI* 4:4 | Kiparsky, P. & C. Kiparsky (1971). Fact. | Massamba, D. B. (1986). Reported speech in Swahili. | Noonan, M. (2007). Complementation. | Thompson, K. D. & A. Folárìn Schleicher (2006). *Swahili Learners' Reference Grammar, Second Edition.*