

The Gulf of Guinea creoles: a case-study of syntactic reconstruction

The scepticism toward syntactic reconstruction in the realm of the Comparative Method is related to inherently different nature of syntax compared to phonology and morphology (e.g. Lightfoot 2002). Especially over long time stretches, reanalysis and (areal) contact make syntactic comparisons more difficult. This talk focuses on the Gulf of Guinea languages (West Africa) and intends to show that creole languages can be quite successfully used for syntactic reconstruction.

The Gulf of Guinea creoles (GGCs) constitute a young language family of four mutually unintelligible languages that are spoken on three different islands, namely São Tomé, Príncipe and Annobón. It is normally assumed that these four creoles descend from a proto-language that came about in the early 16th century on the island of São Tomé as the result of contact between Portuguese and several African languages, mainly from the Edoid family and western Bantu (e.g. Hagemeyer 2011). The African imprint on these creoles is very significant in both the lexicon and the grammar, which shows from features such as ideophones, implosives, labiovelars, body part reflexives, and heavy verb serialization. The languages only started to be documented in the second half of the 19th century (Coelho 1880-1886; Schuchardt 1882, 1888, 1889; Negreiros 1895) but lengthy descriptions and other studies date back to the last quarter of the 20th century (e.g. Ferraz 1979; Hagemeyer 2007; Lorenzino 1998; Maurer 1995; 2009; Zamora 2010).

Criticism to syntactic reconstruction is generally overruled when there is identity between sister languages. There are several reasons why the contemporary GGCs show identity in many domains of their grammars. First, the branching of the proto-language happened at a shallow time-depth (at most 500 years) promoting identity (e.g. Pires & Thomason 2008). Second, after the catastrophic event whereby the proto-language came into being, the languages went through a period of relative isolation in a context of previously uninhabited islands. Therefore, it can be concluded that the new languages were less susceptible to contact. Third, the GGCs are strongly isolating languages (no inflectional morphology, no morphological Case, etc.) with a rather rigid syntax. It is therefore expected that many patterns have remained diachronically stable.

For the reasons above, syntactic identity in the GGCs often means that identical word order patterns can be constructed from cognate items, thus yielding structural cognacy.

However, there are also a number of cases where identity between the contemporary sister languages fails to apply across the board. Here I will focus on three different syntactic domains: (i) sentence negation patterns; (ii) preverbal TAM-systems; and (iii) word order patterns in the Noun Phrase. In a nutshell, it will be argued that a discontinuous negation pattern can be reconstructed to the proto-language by taking into account well-attested directional processes, such as Jespersen's Cycle (Jespersen 1917). To understand and reconstruct the core TMA-system, the burden can be placed on grammaticalization, with some degree of fusion taking place, and in the case of word order in the NP it will be shown that archaic patterns hold the key to the original structure.

To achieve these goals, I will use data from published sources, from annotated corpora of the four GGCs that are currently being built and from my own fieldwork. The final goal of this investigation is to unravel how differentiated syntactic patterns spread into the different GGCs. This is a highly complex issue which has to take into consideration some of the following questions: (i) what the relation is between the two creoles spoken in the centre (on the island of São Tomé) and the two peripheral creoles, spoken on Príncipe and Annobón; (ii) whether and how we can distinguish between genetically related features and features brought about by contact (in particular with the lexifier).