

Susanne Michaelis (Max Planck Institute for Evolutionary Anthropology, Leipzig)

Sampling in contact linguistics: What is a typical creole feature?

Creolists sometimes claim that there are typological generalizations across contact languages (e.g. Bickerton 1981, McWhorter 2001, Bakker et al. 2011). But they have generally overlooked that cross-creole generalizations require representative samples, especially when working quantitatively. Sampling for genealogical and areal control has been a much discussed topic within world-wide typology, but not yet in comparative creolistics.

In all available comparative creoles studies, European-based Atlantic creoles are strongly overrepresented, so that typical features of these languages are taken as typical creole features, e.g. serial verbs, or 3PL pronouns as nominal plural markers. But many of these creoles have the same genealogical/areal profile, i.e. European (lexifier) + Macro-Sudan (substrate(s)). I therefore propose a sampling method that controls for genealogical/areal relatedness of both the substrate(s) and the lexifier, which I call "bi-clan" control. A "clan" is a language family or linguistic area, and a "bi-clan" is a combination of a lexifier clan and a substrate clan. For example, Atlantic English (pidgin)creoles belong to the bi-clan English + Macro-Sudan, Oceanic English (pidgin)creoles belong to the bi-clan English + Oceanic.

Using APiCS data, I will show that there are implications of bi-clan sampling for the notion of so-called "typical creole features". For instance, directional serial verb constructions (with 'come' and 'go') are present in 52% of the APiCS languages, but only in 35% of the bi-clans, simply because of the oversampling of the bi-clans English/French/Portuguese + Macro-Sudan. On the other hand, features which have been classified as rare creole features (e.g. inclusive/exclusive distinction in pronouns, purely aspectual tense-aspect systems) may simply seem to be rare because of the underrepresentation of the non-Atlantic creoles. For instance, 13% of the APiCS languages have a purely aspectual tense-aspect system, whereas 75% show a mixed aspectual-temporal system. When represented in terms of bi-clans, the figures change: 20% have a purely aspectual system whereas only 46% show a mixed system. Therefore, bi-clans give a much more realistic picture of the structural properties of pidgins/creoles, and any notion of "typical creole feature" must be checked against a genealogically/areally representative sample of contact languages.