

Bayesian typology

Gerhard Jäger (Universität Tübingen)

In a landmark paper, Maslova (2000) argued that the synchronous frequencies of a typological variable do not reveal distributional universals. As there is no guarantee that the underlying dynamic process has reached equilibrium, observed frequencies may reflect properties of ancestor languages rather than functional tendencies. As a remedy, Maslova proposes to estimate the transition rates between types from diachronic data and to compute the equilibrium distribution analytically instead.

Probably due to the sparsity of diachronic typological evidence, this program has not been realized so far. Techniques from the *phylogenetic comparative method* (cf. Nunn, 2011) in computational biology, however, paired with the newly available electronic typological data sources and Bayesian inference, afford an alternative way to realized Maslova's goal.

Once a typological variable and a collection of languages has been fixed, the workflow is as follows:

1. Infer a (distribution of) phylogeny(ies) from lexical data.
2. Estimate the transition matrix between the values of the variable.
3. Calculate the equilibrium distribution of this Markov process.

In the talk, I will present the method as well as several case studies pertaining to word order and case marking typology.