Words and Bones: correlations between phenotypical and vocabulary-based linguistic distances between populations

Gerhard Jäger (Universität Tübingen, Tübingen)
— joint work with Katerina Harvati and Hugo Reynes-Centeno

In this talk I will present some preliminary results of an ongoing collaboration between linguistics and paleoanthropology.

Based on measurements of various features of skulls it is possible to quantify to what degree various human populations differ with respect to their facial features, the shape of their neurocranium etc. Likewise, the large scale vocabulary database collected by the Automated Similarity Judgment Program (ASJP) provides the basis for calculating a measure of the distance between different languages based on their vocabulary. As both phenotypical and linguistic features are inherited within populations, we expect phenotypical and linguistic distance measures to be correlated with each other. Among the cranial features considered, the neurocranium is known to correlate strongest with neutral genetic markers, while facial features are the subject to selective pressures. Therefore we expected to find a strong correlation between linguistic distances and the neurocranium and at best a weak correlation of linguistic distances with the face. To test this hypothesis, we considered data from ca. 130 populations across the globe for which both cranial and linguistic data are available. On short distances – i.e. within language families – the data confirm our initial hypothesis. However, across language family boundaries the pattern is reversed. There is a significant correlation between linguistic distances and facial features. The correlation between languages and the neurocranium is weaker, and it disappears if we control for facial similarity and geographic distance.