

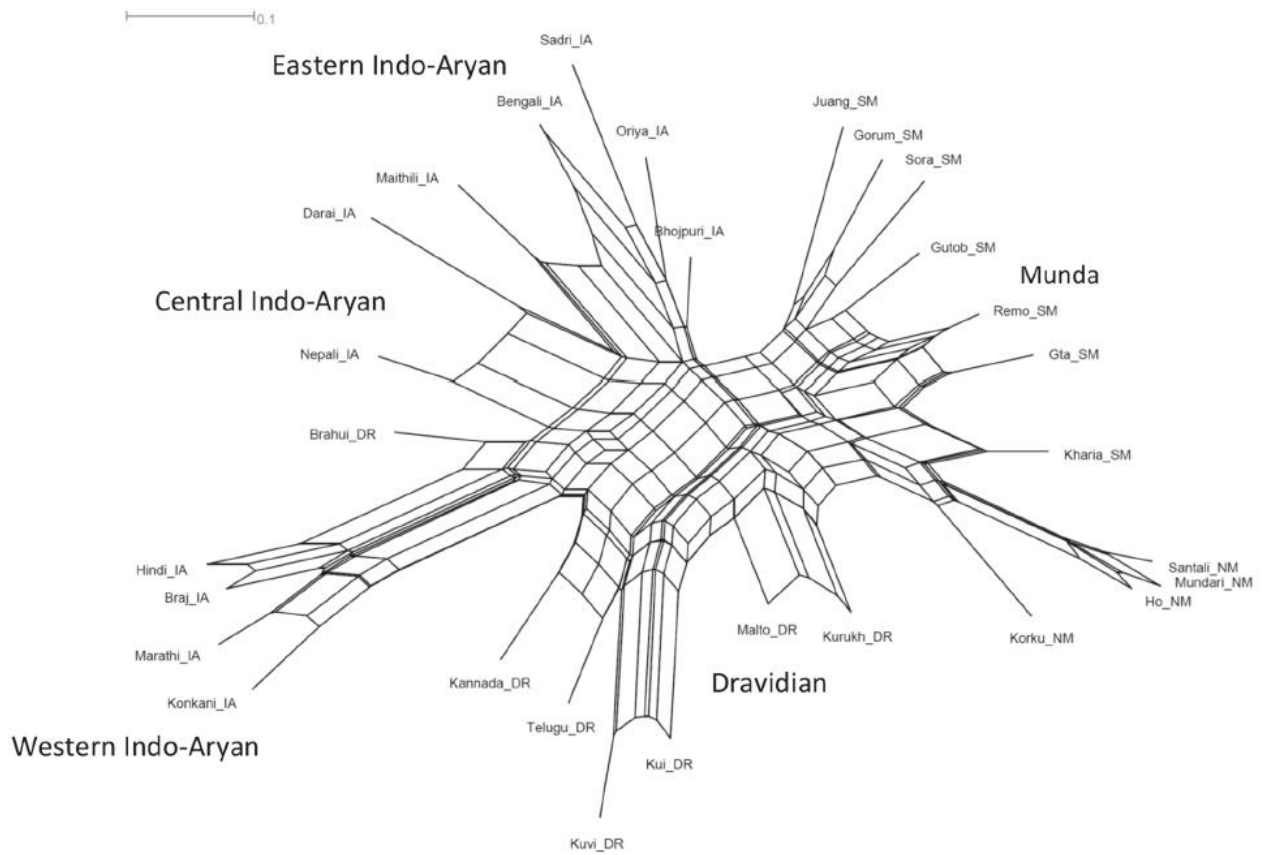
## **Language contact in prehistoric South Asia, with special reference to eastern South Asia**

In Peterson (2016) it is shown that there is a dramatic structural split within the Indo-Aryan branch of Indo-European, with very clearly definable eastern and western groups (see NeighborNet network, next page, from Peterson, 2017: 235). What is especially notable about this split is that it is much deeper than any other comparable split within the three main language (sub-)families of sub-Himalayan mainland South Asia, i.e., Indo-Aryan, Dravidian and Munda (Austro-Asiatic) or even between these three groups (cf. UPGMA tree, next page, from Peterson, 2017: 237).

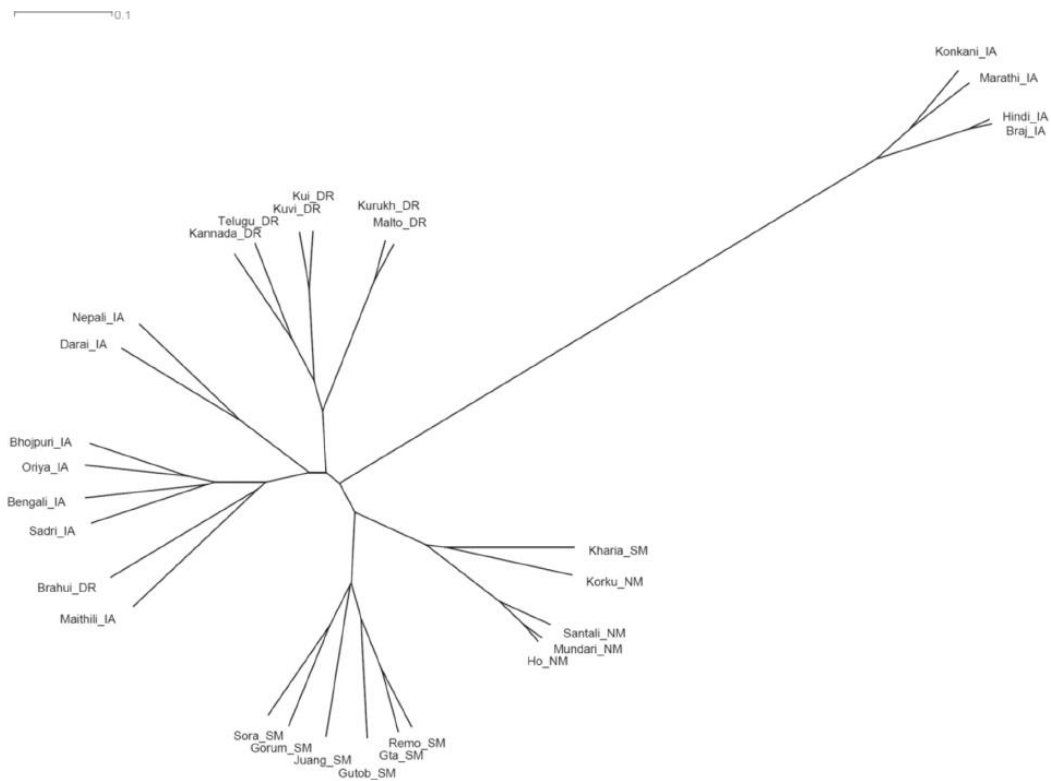
While western Indo-Aryan shows no strong structural similarities with any other language family of the subcontinent, the eastern branch is structurally quite close to the members of the Munda group of Austro-Asiatic languages in the eastern half of the subcontinent. Peterson suggests that this distribution is due to a prehistoric Austro-Asiatic substrate in the eastern half of northern South Asia, which resulted as speakers of Austro-Asiatic languages switched (presumably quite quickly) to the economically and politically dominant Indo-Aryan languages of the time, as speakers of these latter languages spread eastwards into the plains of northern India. However, with 29 languages and 28 features, Peterson's (2016) database is quite small and can only be viewed as a pilot study into this topic.

In our talk, we will improve upon the data contained in Peterson's study, summarizing work from an ongoing research project into the linguistic prehistory of the subcontinent, based on a much larger database with currently ca. 70 morphosyntactic features per language in a much larger number of languages from all families of sub-Himalayan mainland South Asia, including isolates. This will thus allow us to determine not only larger areal and genetic trends in the data for sub-Himalayan South Asia as a whole, suggesting earlier patterns of contact, but also to locate micro-areal trends and show how these smaller areas fit into the larger picture.

Peterson, John. 2017. Fitting the pieces together – Towards a linguistic prehistory of eastern-central South Asia (and beyond). *Journal of South Asian Languages and Linguistics*, 4(2): 211-257.



**Figure 5:** NeighborNet representation of the structural similarities of selected languages of South Asia (29 languages, 28 morphosyntactic features).



**Figure 6:** An unrooted UPGMA tree of the structural data for all 29 languages.