

## **Basic Word Order and Language Contact**

It is commonly suspected that SVO word order is associated with high-contact languages, after the observation that a number of high-contact languages display this word order. However, if high-contact was indeed the driving factor, we would expect to see SVO word order to systematically appear in pockets throughout the world's languages, since languages are in contact everywhere, and high-contact scenarios occur, and can be assumed to have occurred in past, in pockets everywhere on the planet. However, informal inspection of the occurrence of SVO word order instead reveals a high concentration of SVO along certain lineages and certain areas, not prima facie associated higher contact than areal concentration of other word orders. This calls for a systematic investigation.

To this end we will perform three tests regarding basic word order, as follows. Using a recently compiled database of basic word order for over 4600 languages (1) we perform the following tests:

\_ **Word Order Borrowing:** Using known family relationships and geographical coordinates, we can do perform maximum-parsimony analysis of the attested word order data infer 1) likely word orders of proto-languages at every stage, and 2) likely word order borrowings. This allows us to find out whether SVO word order is borrowed more often than other word orders.

\_ **Word Order Borrowing:** For most of the 4600 languages, the practically usable sociohistoric data are speaker number estimates. Using these and the analysis in the previous point we can investigate whether a smaller language is more likely to have borrowed from a larger neighbour (i.e., by direct influence) or vice versa (i.e., by substrate effects).

\_ **Contact-Favoured Word Order:** Suppose A is the parent of a language B which neighbours another language C. If there is a word order favoured by contact per se (i.e., without any of the involved languages actually displaying it, which is required to infer direct borrowing), it should be overrepresented regardless of the configuration of A and C.

<sup>1</sup>This database is exhaustive in the sense that there is no published data on the basic word order of the remaining languages.