

*The value of phonological innovations in cladistic language classification:
evidence from Formosa*

Laurent Sagart

Centre de Recherches Linguistiques sur l'Asie Orientale, CNRS, Paris

In Sagart (2004), I observed an implicational hierarchy in the main Austronesian etyma for the six numerals 5-10 in the AN languages of Taiwan, such that a reflex of *puluq 'ten' always implies the presence of a reflex of *Siwa 'nine', which always implies the presence of a reflex of *walu 'eight', which always implies a reflex of *enem "six", which always implies a reflex of *lima "five", which always implies a reflex of *pitu "seven. Thus:

puluq '10' >> Siwa '9' >> walu '8' >> enem '6' >> lima '5' >> pitu '7'

From that observation I extracted a new higher phylogeny for the Austronesian family. That phylogeny entails the view that the main AN numerals for 6-10 were innovated one after the other in early post-PAN times, each of them defining a clade nested inside the preceding one. The spatial pattern evinced by the sequence of innovations showed a coherent settlement pattern beginning in the NW of the island (where Taiwan is in sight from the coast of Fujian), and circling the island counterclockwise, with a first expansion southward along the west coast, a passage to the west coast around the southern tip of the island, and a northward progression on the east coast.

The linguistic evidence included etymologies for numerals 7-8-9 (a further proposal will be made here for '10'), as well as a list of basic-vocabulary and morphological characters whose distribution in Taiwan and among AN languages in general supports the model. In this paper, I will add to this body of evidence by presenting facts from two genetic systems –HLA-DRB1 and mtDNA–, as well as from archaeology.

Attention will be drawn to the unruly and problematic behaviour of the main phonological innovations that have been used to classify Formosan languages: for the most part they do not fit with any of the tree's nodes. Explanations for this situation will be considered. It will be proposed that they are best regarded as areal events having spread across language boundaries through bilinguals, whereas basic-vocabulary and morphological characters are relatively protected from such spreading by their known resistance to borrowing.

Reference

Sagart, L. (2004) The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43,2: 411-444.