

# The PAN Words for '7', '8', '9' and Austronesian Phylogeny

Laurent SAGART

CNRS, Paris

An examination of the numerals for '7', '8' and '9' in the Formosan languages suggests that the standard reconstructions \*pitu, \*walu and \*Siwa are reduced versions of earlier (=PAN) additive expressions meaning 5+2, 5+3 and 5+4, preserved as a paradigm in Pazeh, a west coast language: thus \*pitu '7' is reduced from PAN \*RaCep-i-tuSa 5+2 (> Pazeh xasebidusa). A set of six simple phonetic changes converting the hypothetical long PAN forms into the more recent short ones will be presented. The displacement of the long expressions by the three short forms occurred very early in AN history, when the AN horizon was limited to the west coast of Taiwan. These three innovations form a compatibility set (= are compatible with the same phylogenetic tree), and are thus helpful in determining the higher subgrouping of the AN family. Five more compatible lexical characters are added, resulting in a set, or 'clique', of eight compatible lexical characters (all basic items): the tree they define has these characteristics:

- the first languages to break off were Pazeh, Saisiat and Luilang; the AN homeland must be in NW Formosa where these languages are/were spoken;
- the early ANs settled Taiwan from the NW, circling the island's coastal plains in a counter-clockwise movement, resulting in gradual and unidirectional accumulation of lexical innovations;
- Malayo-Polynesian's next of kin in Taiwan is Ketagalan; MP is not a primary branch of the AN family.

Moreover, the innovated forms for '7', '8' and '9' are reflected in those Tai-Kadai languages which have not borrowed the Chinese numerals: this adds to the growing body of evidence that the Austronesian-related component in Tai-Kadai comes from a branch of Austronesian, a language closely resembling PMP in its lexicon, but more conservative in its phonology.