

# The Internal Relationships of Formosan Languages

Paul Li

Academia Sinica

Different subgrouping hypotheses of the Formosan languages have been proposed, as based on different types of linguistic evidence, including (1) three main subgroups: Atayalic, Tsouic and East Formosan by Dyen (1963), as based on lexicostatistic evidence, (2) two main subgroups: Rukai and the rest by Starosta (1995), as based on morphological evidence, and (3) nine main subgroups by Blust (1999): Atayalic, East Formosan, Puyuma, Paiwan, Rukai, Tsouic, Bunun, Western Plains, and Northwest Formosan, as based on phonological evidence. Which one of the above is the most acceptable? I shall discuss problems, supporting or counter-evidence for each of the above subgrouping hypotheses, and suggest a revised subgrouping hypothesis of my own.

The problem of lexicostatistic classification is obvious: Mutual influence among the Formosan languages is almost inevitable, and it is not easy to distinguish between early loanwords and inherited words. Hence the percentage of cognate sets shared by each pair of languages is not a very reliable criterion.

I have found further supporting evidence for Blust's subgroups of East Formosan (Li 2004) and Western Plains (Li 2001, 2003). Nevertheless, the main problem with Blust's subgrouping is that there are too many subgroups. It is extremely unlikely that Proto-Austronesian would split into ten subgroups (including Malayo-Polynesian) all at once at the earliest stage. Notice that each of his four main subgroups consists of only a single language: Puyuma, Paiwan, Rukai and Bunun. We may be skeptical, especially when we consider the fact that the Japanese anthropologists could not distinguish between Puyuma, Paiwan and Rukai in an early stage of their work on the Formosan aborigines.

If no phonological evidence can be found to establish a closer relationship between some of the subgroups, we had better look for other types of evidence, such as morpho-syntactic.

Starosta's binary classification seems to be feasible and looks fine at least as a working hypothesis: The first split is Rukai, the second split is Tsou, and so forth. However, scholars who are familiar with the Southern Tsouic languages will be skeptical about the strange relationship among Saaroa, Chamorro and Kanakanavu in that Chamorro is more closely related to Kanakanavu, as shown in his family tree. In fact, Saaroa and Kanakanavu are so closely related that mutual intelligibility is very high.