The Vietnamese etymological dictionary
And “new” Language families

Sikkhamarn, Krisanaa

Abstract

The Austroasiatic language family which includes Vietnamese has been the subject of comparative investigation since the last century. However, it is only now, at the end of the twentieth century, that linguists are able to undertake truly detailed comparative study of the family. From this point of view, Austroasiatis is a typical ‘new’ language family, in contrast to the ‘old’ families with long and successful comparative traditions: Indo-European, Semitic and many others. The label ‘new’ does not indicate that the family is young, and formed only by closely related languages, but rather provides us with information about the level at which comparative studies in the family’s linguistic prehistory have been conducted.

The following features are common to ‘new’ family:

- Lack of detailed synchronic descriptions for the majority of languages included in the family. There are usually good and reliable descriptions (phonological studies, grammars and dictionaries) for just a few of the major languages of the family, with the rest known only superficially. In the Austroasiatic family less than 15 languages are known well enough to be included in comparative study.

- No generally accepted genetis classification. In the Austroasiatis case, linguists are able to identify primary groups of closely related languages, like Vietmuong or Munda, but it remains unclear how these groups are related to each other.

- Reconstruction are available for some of the primary groups of the family. Only three reliable reconstruction of Austroasiatic groups are available to me: Vietmuong (Sokolovkaja MS), Monic (Diffloth 1984) and Katuic (Pejros 1996).

- There is no convincing reconstruction of the proto-languages of the family, and linguists usually have to deal with hypothetical structured which are not based on thorough comparative investigation.

- A lack of reconstruction means that it is difficult (and sometimes even impossible) to identify forms borrowed from one related languages into another.
comparative investigation into the language’s history—is based on an intuitive lexicon.

An etymological dictionary for a particular language—a compilation of morphological reconstructions—creates the etymological classification of the family, and enable us to interpret its evidence. This is basic upon which a linguist can conduct a successful investigation of a provisional comparative dictionary, and on which ability linguistics depends on.

It is clear, however, that everything in comparative linguistics depends on (a) reconstruction of lexical and grammatical morphemes of the proto-

(1) Recognition of the system of proto-phonemes and

(2) Reconstruction of the system of proto-phonological distribution of the proto-language

(3) Identification of systemic proto-phonological correspondences which are

(4) Establishment of a set of proto-phonological correspondences which connect all

languages of the family;

(5) Reconstruction of a provisional comparative dictionary, which includes

languages of the family;

Investigation must involve:

Applying this principle to the study of new families, the first stage of our

Comparison is thus central principle of modern comparative studies.

Comparisons between their systems,

each and every part of the languages, one by one, as sure to discover all possible relationships between the systems of all languages. This necessitates a knowledge of which connect the systems of all language. The necessary of this reconstruction are clear: the aim of a proper comparative study is to find formal patterns and regular patterns. Therefore, for this purpose of possible patterns of the chosen languages. The reasons for this

(1) comparison should be conducted between part of language (A) and B, and C. (A)

(2) without any attempt to find in its context part of language (B) it is not a valid procedure.

One of the basic principle of modern comparative linguistics is the aspiration
studies and on a deep knowledge for the historical phonology of the language and its family. Does this imply that attempt to compile an etymological dictionary for a language of a 'new' family is a priori premature? I do not think so and for several years now I have been completing an etymological dictionary of Vietnamese. The theoretical foundations of this project are discussed below.

As mentioned above, a 'new' family know only partially with detailed reliable descriptions being unavailable for most of its language. In dealing with such a family, then, we need a strategy which will meet the principle of completeness. One strategy is to include data from every languages or dialect mentioned in the literature, regardless of the possibility that this data is by no means complete. If in our example, a languages is know only by a list of 200 words, we could include it in our investigation, but we would not expect to be able to draw detailed conclusions about this language. If the number of such poorly recorded languages is significant, then a provisional comparative dictionary and phonological correspondences based this strategy will be complete, but not adequate for a reconstruction.

Another possibility is to concentrate form their comparison. In such a strategy the emphasis is on the complete interpretation of the data rather than on the completeness of the data itself. Both of these strategies are represented in the literature, but I think that the second one is more appropriate in the circumstances of 'new' language families such as Austroasiatic. My project is therefore base on a study of several languages chosen as primary sources, and an attempt is made to identify all possible comparisons between these languages and Vietnamese.

The staring point of the project is a list of common Vietnamese morphemes represented in the major modern Vietnamese dictionary. To it I have added some archaic morphemes which interesting etymologies. Recent borrowings, mostly from European languages, are not included. Altogether the list comprises about 5000 entries.

Vietnamese, together with many close related languages to the Vietnamese group of Austroasiatic. In principle, this group should be investigated before a comparison with other Austroasiatic languages begins. However, it is not possible to compile a full scale Vietmuong comparative dictionary because published data is not available for most Vietmuong languages (Barker 1993). With the exception of Vietnamese, only Muong dialect has been described in any detail (Materialy 1987), although there are short dictionaries of Ruc (Nguyen et al. 1988) and Thavung (Ferlus 1979). Nonetheless the history of the group has been the subject of intensive investigation, and numerous reconstruction have already been proposed (Barker 1963; Barker 1970; Ferlus 1975; Thompson 1976; etc.). In my opinion the best reconstruction is Sokolovskaja's (MS), which is base on interesting and important data including Nguyen Van Tai's dissertation, and unpublished field materials of joint Russian-Vietnamese linguistic expeditions. Sokolovskaja identifies more than 700 cognates showing good semantic and phonological correspondences between the selected languages. The Proto-Vietmuong reconstruction includes many consonantal clusters, and some disyllabic roots. There is no evidence of tonal oppositions, and the tones of modern languages are explained as having developed from two suffixes *-x and *-?. These suffixes, which follow the obligatory final consonant of the root and are uncertain grammatical meaning, have been maintained in Arem. Their existence in Proto Vietmuong was first suggested by Haudricourt (1954).
### Vietnamese

Figure 2 represents the development of modern Vietnamese from the Vietnamese system suggested by Sokolovskaya.

- N - any other initial consonant or O.
- C - finals - p, f - t or - k.

<table>
<thead>
<tr>
<th>a</th>
<th>a</th>
<th>a</th>
<th>m*</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>a</td>
<td>a</td>
<td>d*</td>
<td>p*</td>
</tr>
<tr>
<td>C*</td>
<td>C+N*</td>
<td>P+N*</td>
<td>N*</td>
<td></td>
</tr>
</tbody>
</table>

### Vietnamese

Figure 2

(Vietnamese initials)

(continued)