TYPOTLOGICAL COMPARISON OF THE SOUTHEAST ASIAN INSULAR AND PENINSULAR LANGUAGES:
GRAMMATICAL STRUCTURES

Natalia F. Alieva

0. Introduction

0.1. The languages of the region comprising the Indochinese peninsular and the insular world of the South Seas present a great number of structural types, from completely analytical up to highly synthetical, with many intermediate variations.

The structural synchronic comparison of this multitude of types will be the more needed and useful for pragmatic purposes the more widely the connections among the states and peoples of the region will develop. With this the consequences of the era of the colonization split in the region are to become less and less felt.

0.2. It cannot be denied that there were already made great and successful investigations in comparative studies of this region languages, especially in historical comparativistics. However the more pragmatic typological approaches are still to be found and worked out upon the abundant new linguistic material.

0.3. In the structural comparisons the attention of linguists was mostly directed to the phonetics and phonology, and primarily, quite naturally, to such things as tones and pitch/register characteristics, preglottalization, pre- and post-aspiration and other specific features of the languages of Indochina. As to the AN languages that are usually have no complicated phonetic phenomena of this kind, the investigation was directed to the phonological correspondences among the languages and to the allophones inside some of them, and, of course, to the problem of nasals, as prenasalization and nasalization of consonants at the morphemic boundaries present an important and characteristic feature. Specifics of pronunciation are usually described in corresponding handbooks and grammars, some of them with the first stage of instrumental investigation.

0.4. My studies in comparing Southeast Asian languages that are summarised in this paper were in many respects preceded and stimulated by the general project of typological comparisons of the languages of the East and Southeast Asia which was during some decades (beginning from 1960) conducted by the scholars of the Language department of the Institute of Oriental Studies, Russian Academy of Sciences, Moscow.

This work was started with the publishing of the series of short descriptions of various languages of Asia and Africa (many of them firstly described); a number of conferences were conducted with special attention to the methods, history and methodology in the researches of the linguistic typology: some books of articles were published typologically regarding definite themes (morpheme, morphology, pronouns, reduplication, compound words, syntax, word-order) simultaneously with the series of conferences and publications on general typology made in Leningrad/St. Petersburg (these books are mostly included in the Addition to Bibliography at the end of this paper and to [Alieva 1998]). The investigation of the SEA languages was a new branch in Russian / Soviet linguistics, and special attention was directed to it.
1. The morpheme level

1.0. Any description of the word-level structures must be naturally preceded by some remarks about morphemes.

A whole group of problems concerns the morpheme structures, with the important difference between two main types, one being the monosyllabic peninsular and the other being the disyllabic insular.

1.1. The utmost monosyllabism existing in Viet-Muong, Tai languages has two special interrelated features: (1) the prevailing two-, three-and four-phoneme models of one-syllable morpheme, with numerous positional restrictions, predict the relatively small number of segmental material units possible in a language, and (2) this shortage is balanced with the help of suprasegmental devices for diversification which allow to denote far more great quantity of semantic notions. Registers (pitch) and musical, or contour, tones serve as the phonologically supra segmental devices.

1.2. The disyllabic morpheme models, also with more or less numerous positional restrictions, present the main and general type of lexical and full pronoun morphemes in the most AN languages. One-syllable and three-syllable lexical morphemes exist either, in various proportions. The minor morphemes, both free (particles and prepositions, pronoun clitics) and bound (affixes, duplicators) are usually monosyllabic. In the AN languages the quantity of phonemes is not great, especially few are vowels, and definite position restrictions variously characterise the morphemic structures in different languages. The stress is in some languages purely phonetic, but in some other, not many, it is phonological, distinguishing lexical and grammatical meanings.

1.3. A number of the AN languages at the Northern periphery, on islands close to the peninsular – Moklen, Mokken, Urak Lawoi, as well as Acheh on North Sumatra and cognate to it Chamic group in Vietnam and Cambodia form a separate subtype. They have in more or less degree endured evolution from disyllabic to monosyllabic morpheme patterns which proceeded in combination with the shift of the phonetic stress from penultimate to ultimate syllables with the gradual disappearing of the penultimate. This phenomenon systemically correlates with the evolution of other important features: therefore it is possible to speak of a special typological subtype inside the AN family that is usually called "continental" or "mainland" subtype.. Also I need at once to say that the contemporary spoken Cham should not be included in this subtype, as it has already lost main AN features, developed into completely analytical system and actually belongs to the Indochinese type.

The conclusion in regard to the morpheme level must be that the main opposition between peninsular and insular languages is that of monosyllabic and disyllabic morpheme structures.

2. On the methods of comparison

2.1. In the consideration of the languages of this great region a linguist cannot compare simultaneously all the languages of the isles with all the languages on the peninsular. We need to have previously some genera-lization of typical structural features, which means to formulate definite conceptions of types, each being a system
of typologically relevant features, and precisely these types are to be compared among themselves. Summarizing the opinions existing up to now, I see it enough founded to speak about three main types: the Malayan (MT), the Philippine (FT), the Oceanic (OT): I tried to give arguments and to enumerate their specific features in my monograph [Alieva 1998: 170-206].

I am far from presuming that the named three types present in them the whole variousness of systems of the thousand AN languages. Possibly a few more types and undoubtedly not few subtypes are still to be formulated, without or with considering genealogical data.

The features of the IT are also presented in matrix form in [Alieva & Bui Kh. The 1999: 160-163], which was preferred by other studies.

The features of a type are sometimes exposed bellow as abstract linguistic notions, but sometimes they are inevitably exemplified with concrete language facts. And in such cases I will prefer to use data from the languages exemplary for a definite type: from Malay/Indonesian for the Malayan type, from Tagalog for the Philippine type, from Maori for the Oceanic type, though other languages are not excluded. For the Indocheinese type I shall try to give examples from several languages belonging to different genetic families: Vietnamese and Khmer for Austroasiatic, Thai (Standard) for Tai (Tai), but the language most known to me is Cham as it was studied upon self collected field materials which study resulted in a book [Alieva & Bui Khanh The 1999]. The languages of the Burman group in Indochina have many own properties, therefore they are not joined by me with IT.

2.2. The investigation of languages of the SEA region is now rapidly developing in several directions, and representatives from various linguistic schools take part in this work. Inevitably a typologist is faced to different methods and world-views (and in many of them to the lack of systemic approaches.

A type for us is the type of language systems. Inside a type categories of morphological, morpheme and word level regularly correlate with the syntactic structures.

One example of a regulation for such inter-level correlations can be given. In linguistics, firstly in the Russian turkology, a frequentual rule has been formulated stating the interdependence in a language system between the prevailing position of affixes in the word and the prevailing position of dependent words in the phrase.

By N. Baskakov, later by N. Gadjieva and B. Serebrennikov (see Bibliography) a rule of frequentual nature was proposed of direct correlation between morphological and syntactical levels on the basis of alleged logical semantical abstractness of the units at both levels. I have studied the problem further and wider, beyond of Altaic, and propose to formulate a rule - frequentual, not universal - about the interrelation between language levels in the following way:

1 - Suffixes and postpositional grammatical markers correlate with the left-side position of the dependent words in phrases (attributes to nouns and actant-nouns to verbs); the predicative verb is in the final clause position.
II - Prefixes and prepositional grammatical markers correlate with post-posed, right-sided attributes to nouns and actant-nouns to verbs; the verb is in the initial clause position.

This thesis looks as too ideal and systemic, however its reality is demonstrated by many languages: the formula I is relevant for many Altaic languages and to some degree to Sino-Tibetan, while the formula II is relevant for many AN, first of all for Malay, which stimulated me to this proceeding. At the same time, more or less free word-order in phrases is present in languages with productive left- and right-sided affixes, such as inflective Russian and agglutinative Tagalog. This is a negative argumentation for the same conclusion.

3. The word level.

3.0. The units that are conventionally called "words" can be examined concerning their morphemic structure and their semantic and functional properties.

3.1. The word-structure.

From the point of view of the word-structure the line of difference lies between analytical IT and OT, on one side, and FT and MT, more or less synthetical, on the other. The last two have rich, in FT abundant, inventories of agglutinative models, including prefixes, suffixes, infixes, confixes and combined affixes. Such models are normally absent in analytical languages.

3.1.1. Monomorpheme word.

Plain words equal to one morpheme, be it the IT monosyllabic syllabomorpheme, or disyllable in AN or Khmer, are found in every language, but with great difference in percentage.

The OT, as the IT, have analytical structure, but the character of analytism of these two types is principally different. The OT languages usually have no productive affixes, but the cognate to AN affixes morphemes serve as numerous free grammatical markers in the form of left- and right-sided particles. The IT languages, on the contrary, have non or few grammatical markers of particle, not denotive, nature, using in these functions various lexemes. For inst., my analysis by Greenberg's quantification method of the texts in Maori (OT) and Cham (IT) gave interesting results: the ratio of morphemes to words for Maori was 1.03, while for Cham it was 1.02, the ratio of particles to words for Maori was 0.44, for Cham it was 0.14. [Alieva 1992: 13-14].

3.1.2. Affixation

From the functional point of view, various affixed patterns are fulfilling sometimes derivational, sometimes grammatical functions, which is a usual thing. A specific feature of the AN agglutination is the syncretism, or combining both functions in one pattern.

Ex.: Malay: paku 'nail' (noun) > memaku'to nail' TRANS. ACT, dipaku 'be nailed' TRANS. PASS.

The properties of the morpheme/base with full lexical denotive meaning (not of pronoun, numeral, particle nature) play important role in the word-forming process which is sometimes connected with the phenomena of the "noun vs verb" classification. In the FT (ex. Tagalog) such classification seems to be absent and the majority of full morphemes are conventionally considered to be nouns, but in MT we may speak, basing