

THE ENIGMATIC ETHNOLECTS OF THE MLABRI (YELLOW-LEAF) TRIBE

Jürgen Rischel^(*)

1. Introduction

1.1. *The Mlabri language*

The language of the hunter-gatherer groups of Northern Thailand and Westernmost Laos has been dealt with in several publications since Bernatzik (1938). As is already apparent from the literature there are different varieties of Mlabri with intriguing relationships to each other. The reason why I take the issue up once again in this paper is that we now have contemporary data on three varieties of the language, as against only one accessible variety in the 1980es, and two in the 1990s. Since the three extant varieties fall beautifully in line with earlier records, the whole linguistic scenario now presents itself in a much clearer light. The present paper inevitably overlaps to some extent with Rischel¹ (2000) but has a rather different emphasis.

Although it is the overall lexicon of Mlabri, rather than the presence of language-internal lexical variation, that is most directly relevant to comparative Austroasiatic studies, it is certainly essential also from that perspective to ascertain that it does make sense to speak of one language and thus of one lexicon. This requires consideration of the dialect or sociolect scenario. At the same time this scenario may be of some general interest as it illustrates what may happen in the language of a nomadic tribe speaking a purely oral language.

1.2 *Research history*

By way of introduction I shall give a survey of the earlier records, more or less duplicating the survey in Rischel (2000). Afterwards these are lined up in relation to contemporary data. The main emphasis, however, will be on similarities and dissimilarities between the three extant varieties of Mlabri, the recent data being far more extensive for all three varieties than any earlier data.

The first round of documentation comprised three separate achievements¹. In Bernatzik's well-known book (1938, 1941) on the "Yumbri", their culture and language there is a word list which unfortunately uses a very unfelicitous notation and has posed severe difficulties of interpretation. (II) Kraisri Nimmanhaeminda (1963), who had encountered a group of "Mrabri",¹ produced a word list using both Thai-based and Roman transcriptions; contrary to the impression one gets from his own statements the former is much superior to the latter in terms of phonological adequacy (Rischel 1989a). (III) Michel Ferlus took down a short word list in IPA phonetic notation from tribespeople staying in a Tin (Lua) village in Sayaburi Province of Laos in 1964; the material was referred to in Ferlus (1974) but was never published².

The next round of documentation began in the early 1980s. (IV) Dr. Sören Egerod and I collaborated with Dr. Theraphan Luangthongkam doing fieldwork on the Phi Tong Luang language, whose proper name had now been identified as Mlabri; we published various papers

^(*) Royal Danish Academy of Sciences and Letters

¹ Dr. Ferlus later generously put the list at my disposal, and I exploited it when writing a monograph about Mlabri (1995), which makes references to his data throughout its Part II: Dictionary.

² The project in Laos is jointly supervised by the Director of the Lao National Institute of Research on Culture, Mr. Houmpanh Rattanavong, and myself. During field sessions in the forest I have been accompanied by the junior cultural anthropologist Mr. Khammanh Siphanhxay, and by Mr. Vieng Khamcanh of the cultural division in Sayabouri Province.

language, whose proper name had now been identified as Mlabri; we published various papers including a Mlabri-English word list (Egerod and Rischel 1987). (V) More recently I recorded and published a more extensive word list representing another variety of Mlabri (Rischel 1995, Part II: Dictionary). (VI) In 1999 and again in 2000 a further group of Mlabri people was encountered in Laos (Rischel 2000, in press) and a considerable amount of material was recorded¹. Supplementary fieldwork has been scheduled.

Back in the early sixties, when Kraisri had gathered his material, it became an issue whether "Mrabri" was the same language as "Yumbri". Kraisri himself assumed this to be the case and further documented the lexical relatedness of this language to Tin and Kmhmu for both of which he had collected fairly extensive word lists (Nimmanhaeminda 1963). Smalley (1963) tried to apply the lexicostatistical method to the data and found that "Mrabri" showed only a 35 per cent relatedness to "Yumbri", which was of the same order as its percentage of relationship to Kmhmu and Tin. Dr. Smalley also made the interesting experiment of attempting to restate some symbols in Bernatzik's awkward notation on the basis of typological knowledge rather than from a "Mlabri" perspective. Schuhmacher (1969) on the available evidence concluded that the two typologically similar groups could not be the same people.

In later work it has become increasingly apparent that the sources reflect one and only one "Khon Pa" ("Phi Tong Luang") language. Ferlus (1974) stated that the data he had collected in Laos in 1964 showed most lexical similarity to Bernatzik's list but found the notation of the latter to be so poor that it could essentially be used only to verify the existence in it of words known from other (later) sources.

In the late eighties Egerod and I, now having access to many more lexical items of the Mlabri language than those available to Kraisri and Smalley in 1963, made a direct attempt to interpret Bernatzik's notation as a distorted representation of this Mlabri. Taking some very confusing underdifferentiations in Bernatzik's notation in account, we could identify so many shared lexical items as to show beyond doubt that "Yumbri" and Mlabri must represent the same language (Rischel and Egerod 1987); this conclusion was subsequently endorsed by Smalley (1994). Still, there was much guesswork and, as I have found later, several errors in our pairwise identifications of lexical items. The subsequent data from other varieties of Mlabri prompted two rounds of revisions (Rischel 1989a, 2000). The result each time was that the former identification of "Yumbri" as Mlabri got an even more solid foundation,² while at the same time it became increasingly clear that some of the apparent oddities in Bernatzik's data represent genuine divergencies within Mlabri rather than errors on his part.

1.3 The present research situation

Today the pattern of inter-group variation within Mlabri is transparent in itself. There are three different varieties, each associated with one or more groups of people such that no group speaks more than one variety of Mlabri. One variety, which may be called A-Mlabri, is represented in

¹ The linguistic unity of the Mlabri is accompanied by conspicuous tokens of ethnic identity, such as their material culture with its near absence of indigenous tool technology, their traditional use of lean-to shelters, and their social pattern (also cf. note 7 below). For themselves the criterion of self-identification is life in the forest (Mla' Bri', *mluq briiq* = person forest i.e. 'forest people', 'Yumbri', *ym briiq* = stay forest i.e. 'we live in the forest'). There are, however, other tribes in mainland and insular Southeast Asia who exhibit a more or less similar, forest-bound survival culture.

² For typographical simplicity and on-line accessibility I here use a transcription of Mlabri which is entirely composed of ASCII-symbols (see Rischel 1982, 1995 for a phonologically equivalent IPA notation):

The consonants which occur both syllable-initially and syllable-finally exhibit four oral places of articulation: labial, dental, palatal, and velar. Aspirated voiceless stops are rendered as *ph th ch kh*; *ch* is often realized as a palatoalveolar or even alveolar sibilant. Plain voiceless and voiced stops are rendered as *p t c k* and *b d j g*, respectively. The corresponding nasals are rendered as *m n ny ng*. Mlabri has labiovelar and palatal glides *w y*, a lateral *l*, an apical trilled *r*, and two laryngeals: the approximant *h* and the glottal stop *q*. The glottal stop symbol is redundant prevocally and is then omitted in this transcription.

the data of Kraissri (1963) and Egerod and Rischel (1987); another, B-Mlabri (which I have referred to as "Minor Mlabri") is represented in the data of Ferlus (unpublished) and Rischel (1995); a third, C-Mlabri, is represented in the data of Bernatzik (1938) and Rischel (2000).

With respect to speakers there is a skewness in that we know for sure that the older and more recent data for A-Mlabri stem from the very same group of people, and similarly for B-Mlabri, whereas this is not the case for C-Mlabri. The group met by Bernatzik was never refound, and it has not been possible so far to establish any direct link between the people he talked to on the Thai side of the border between Nan and Sayaburi provinces in 1937 and the Mlabri group recently encountered on the Lao side of the border. Still, they clearly speak the same variety of the language (Rischel 2000). Thus, the pairings of older and more recent data into three branches of Mlabri are linguistically overwhelmingly clear and will not be at issue here.

Since the recent data collections are much richer and technically more reliable than the earlier ones, the dialectal and sociolectal variation within Mlabri can be studied conveniently with exclusive reference to recent data. All the older data sets are, however, highly significant especially since we are dealing with the strictly oral language of a nomading hunter-gatherer tribe.

They enable us to (i) see how far back in time the present degree of dialectal or sociolectal split can be safely projected, (ii) see what happens over a time span of four or six decades with coexisting varieties of one language.

The topic for the remainder of this paper is the scenario of constancy and variation which we can deduce from the now available data on the Mlabri language,¹ with a view to the relationship between linguistic variation and ethnic diversity. In addition I shall venture a few speculative suggestions as to how the scenario may have come into existence.

The Mlabri in Northern Thailand and Western Laos number less than 200 persons in total. The area in which the various groups migrate along mountain ridges is not immense, and it would be perfectly possible for them to make repeated, occasional contact if they so wished.

In addition to the above there are two sets of more complex consonants which occur only syllable-initially, viz. glottalized voiced stops: *qb qd* and aspirated (more or less voiceless) continuants: *hm hn hny hng hl*.

There are also two postaspirated and more or less voiceless segments which occur only syllable-finally: a fricative or approximant *yh* (which vacillates between more alveolar or more palatal articulation) and a lateral *lh*. If a compact transcription system is to be designed it is possible to lump initial *ch* with final *yh*, and initial *hl* with final *lh*, but no relevant insight is gained by doing so.

Several, more or less complex consonant clusters occur syllable-initially but the combinatorics need not be stated for the purposes of this paper.

As for vowels, there are four degrees of aperture (disregarding some skewnesses and complications which are immaterial to the present paper), *a* being the maximally open (low) vowel. The close to half-open vowels fall in three series, each with three degrees of aperture: front unrounded *i e ê*, back unrounded *ĩ ẽ ä*, and back rounded *u o ô*. Some pretonic syllables have a schwa-vowel which is here rendered as *ë*.

The most controversial feature of Mlabri syllable structure is vowel length since long and short vowels have merged more or less in modern usage, especially in the dialect here called A-Mlabri. Long vowels are rendered as double vowels; the notation must be taken with some reserve because of the elusiveness and variability of vowel length in Mlabri. Word and sentence prosody is here disregarded; stress falls predictably on the last syllable of a word or stress-group (with grammatically motivated exceptions which need not concern us here).

¹ This etymology of *rt.lat* was suggested by my late colleague Søren Egerod long before it became possible to visit the C-Mlabri.

The placement of syllable boundaries may be distinctive or at least ambiguous in the present transcription system, and accordingly, syllable boundary is indicated, viz. by a dot (e.g. *VC.V*; *V.CCV*). In vowelless syllables the last consonant is syllabic, e.g. *l* in *k(r)l.kiil* 'knee'.

One might perhaps a priori assume that a language spoken by such a tiny population within a limited area would either exhibit very little variation, namely if the groups interacted e.g. by intermarriage, or that it would on the contrary split into very different dialects if the groups stayed apart consistently. However, in the case of Mlabri it is not a clear-cut dichotomy of this kind. The complexity of the scenario stands out if we consider phonology and lexicon separately.

2. The phonological closeness of the three varieties of Mlabri

I have repeatedly made the observation is that there is extremely little segmental difference between A-, B- and C-Mlabri. If words are transcribed phonemically using the same kinds of conventions for all three varieties of Mlabri, they often look completely alike across the board. This is not just an illusion created by phonemic abstraction from phonetic detail; I have been communicating repeatedly with speakers of all three varieties and have felt all the time that if a word used in one variety is also present in another there is generally little or no difference in segmental phonetics.

Numerous words sound alike all across A-, B- and C-Mlabri; I shall mention a few shared words for illustration: *aar* 'in advance, first', *ba.tit* 'close together', *bèèr* 'two', *crëw* 'to call', *chë.mòny* 'star', *chm.bèp* 'mouth', *dë.kat* 'feel cold', *dëmòy* 'a single one', *èèw* 'child', *gany* 'sunshine', *gèèng* 'house, lean-to', *hling* 'to cough', *hng.keq* 'tick', *jäyh* 'it tastes good', *jëèng* 'foot', *kr.lap* 'split bamboo forceps', *k(r)l.kiil* 'knee', *lat* 'to lick', *mat* 'eye', *mëm* 'father', *mïy* 'fat, oil', *nony* 'complete', *pëy* 'to eat soft fruits', *poolh* 'barking-deer', *rèèlh* 'root (of tree)', *rwaay* 'tiger', *tèk* 'to hit', *throoc* 'water snail', *uuy* 'woman', *wëèng* 'chin'.

This cross-Mlabri equivalence sometimes applies even to (short) clauses: *oh a jak* I PERF go, i.e. 'I am off', *oh chi thapuul* I ache stomach, i.e. 'I have a stomach-ache', *mèh di leh* you IMP come, i.e. 'come here', *mèèq hot* rain fall, i.e. 'it is raining', *aac päär jak* bird fly go, i.e. 'the bird flies', *tm.ooq määj jak* cobra creep go, i.e. 'the cobra creeps' are phonologically alike and communicatively functional utterances in all three varieties of Mlabri.

Hopefully the examples above suffice to illustrate that the shared lexical material is to a large extent genuine Mon-Khmer and at the same time a language specific in its phonology. The various varieties of Mlabri clearly share a long history within Mon-Khmer, and they have not separated phonologically to any considerable extent.

There are, though, quite conspicuous differences in prosody, viz. in sentence intonation and in the degree of final lengthening (e.g. if a word such as *pm.poo* 'elephant' occurs utterance-finally), the latter phenomenon being extreme in A-Mlabri, less in C-Mlabri, and auditorily negligible in B-Mlabri. As a result A-Mlabri and B-Mlabri may sound like different languages if one listens mainly to the prosody without paying any attention to the single words, many of which are in fact the same.

The above remarks about phonological uniformity within Mlabri apply to the general sound pattern and to the majority of shared lexical items. There are, however, several instances of phonological fluctuation such that a word has one phonological shape in one variety of Mlabri and a slightly different phonological shape in another, although the overall sound pattern would allow for both pronunciations in each of the varieties of the language. The occurrence of such differences is unsurprising since Mlabri has been traded down through generations as a strictly oral language, but they may present a challenge to language comparison since it is essential to establish the etymologically primary forms of Common Mlabri, using internal evidence as far as it goes. This is sometimes possible. An example is the numeral 'six' which is *taal* in A-Mlabri (in homonymy with the A-Mlabri word for 'day') but *thaal* in B-Mlabri (I have not come across any C-Mlabri speaker who could count that high up); the former appears to be primary since the latter is explicable as due to influence from *thëèng* 'five'. Another example is the word for