

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

'tongue', which is *r(i)t.lat* in A- and B-Mlabri but *līt.lat* in C-Mlabri. The latter form is obviously the primary one, the word having originated as a reduplicating nominalization from the verb *lat* 'to lick'.¹

In some words there are more drastic aberrations in phonology, to the extent that it is controversial to speak of "the same word". In some cases different etyma or different word formation processes may be involved, but that it is often not transparent at all. The phrase meaning 'to go for a stroll or visit' is *jak gwaa* in both B- and C-Mlabri but *jak gwèèng* in A-Mlabri. The latter shows a conspicuous phonological affinity to the A-Mlabri phrase *jak galèèng* 'go where, i.e. 'where are you going?', which may be old Common Mlabri since Bernatzik (1938) quotes "tshakaleng" as part of a greeting phrase used by the "Yumbri". Present-day speakers of C-Mlabri, however, say *jak ga.nèng* with [-n-], as do B-Mlabri women (B-Mlabri men say *jak gi.nèng*). I suppose *ga.lèèng* and *ga.nèng* coexisted as synonyms or near-synonyms in old Mlabri, and one or the other was lost over time.

In the case of phonological fluctuations, non-trivial agreement or discrepancy between earlier and present-day data is of course a criterion for the alignment of these sets of data into different branches of Mlabri. For crucial evidence of this kind supporting the identification of C-Mlabri with Bernatzik's "Yumbri" see Rischel (2000); for the alignment of the other sets of earlier and more recent data cf. Rischel (1989a).

3. The lexical distance or closeness among the Mlabri varieties

In regard to lexical comparisons among the varieties of Mlabri, the picture is confounded by recent loanwords of diverse origins (e.g. from Kmhmu).

If we disregard recent loans, we find that an extensive Mlabri vocabulary is shared by speakers of all three varieties. It is, however, often the case that a word which is frequent in daily usage in one variety of Mlabri, turns up in another variety only in specific spheres of language use or as an obsolete or stigmatized word. As long as only A- and B-Mlabri were accessible the basic lexical unity of Mlabri was not so obvious since these two varieties differ in daily lexical usage to the extent of making mutual comprehension difficult, but it stands out very clearly now that the third variety: C-Mlabri is yielding a rich set of contemporary data.

3.1 The lexical position of C-Mlabri

The newest and thus most interesting finding being that it is Bernatzik's 1938 "Yumbri" that is continued as C-Mlabri in westernmost Laos, it makes sense to focus on the relationship of that language variety to A-Mlabri on the one hand, and to B-Mlabri on the other hand. This is in fact the only safe strategy at the present early stage of work on C-Mlabri. In principle, one might instead search for lexical idiosyncracies in C-Mlabri by making records of lexemes which are unmatched in the files for A- and B-Mlabri, and records of lexemes shared by A- and B-Mlabri which have failed to show up in C-Mlabri. It would, however, be premature to draw general conclusions on that basis because we still know too little about C-Mlabri lexicon outside the sphere of basic communication. As for daily language usage, my experience from extensive conversations contradicts whatever expectations one might have about a continuation of Bernatzik's "Yumbri" being particularly aberrant from A- and B-Mlabri. On the contrary, I have found C-Mlabri to be surprisingly easy to understand as it shares some vocabulary with A-Mlabri, other vocabulary with B-Mlabri, and very much vocabulary with both (disregarding Lao loans). C-Mlabri mediates communicatively between A- and B-Mlabri, as it were. For me this justifies the asymmetric strategy adopted here.

¹ The basic cultural unity of the Mlabri is confirmed above all by the existence of common formulaic language in rituals and of very distinctive elements of traditional mythology and cosmology shared by the A- and B-Mlabri, and at least to some extent also by the C-Mlabri.

In the exposition below I shall focus on instances from daily conversational language where C-Mlabri sides with either A- or B-Mlabri, but not with both, in order to see if there is evidence for a Stammbaum with consecutive bifurcations rather than a ternary branching. The criterion for inclusion of data will be that a word found in only two of the varieties can be matched with a functionally equivalent synonym in the third variety (so as to avoid errors due to holes in the data). The statement I can make so far is qualitative not quantitative. Only a few examples of each kind are given by way of illustration (asterisks mark recent loans from Northern Thai or Lao).

C-Mlabri sides with B-Mlabri against A-Mlabri on a number of frequently used synonyms such as A *nòon** vs. B/C *èm* 'sleep'; A *khün** vs. B/C *glēh* 'ascend', A *thalèew* vs. B/C *īm* 'bathe'; A *hnguh* vs. B/C *yěem* 'sit, stay'; A *mē.ün** vs. B/C *jriw* 'tomorrow'; A *brany* vs. B/C *chòòq* 'dog'; A *nqlq.nglèq* vs. B/C *ku.kòòq* 'neck'; A *taal* vs. B/C *ta.wen** 'day'; A *ciin** vs. B/C *thăc* 'meat'; A *chr.kèng* vs. B/C *qyoc* 'hen, chicken'.

On the other hand, C-Mlabri sides with A-Mlabri against B-Mlabri on some other perfectly usual words such as B *toc* vs. A/C *ek* 'accept, take'; B *bong* vs. A/C *ěeq* 'eat vegetable'; B *thuuc* vs. A/C *krăp* 'sting'; B *pa.yok* (< *yok**) vs. A/C *pa.duääw* 'lift', B *glaq* vs. A/C *tăny* 'speak'; B *qyen** vs. A/C *tè(è)k* 'cold (of objects)', B *jrăäk* vs. A/C *wěek* 'drink, water'; B *kam.pong* vs. A/C *glěeq* 'head'; B *thrèng* vs. A/C *căny* 'tooth/teeth'.

This intermediate position of C-Mlabri is apparent with expressions such as *īm wěek* 'bathe water, i.e. 'take a bath': A-Mlabri has *thalèew wěek*, in B-Mlabri one says *īm ruäng* (if one bathes in a stream).

3.2 The inadequacy of superficial lexical comparison

Lexically contrasting tables such as the above are suggestive of a certain genetic distance. But in the case of Mlabri, at least, they give a scewed picture since quite a few of the words turn eventually out to be known in all three varieties. In order to explicate this it is necessary to go into some detail; in return, the examples chosen illustrate the general trend:

Among the B/C-Mlabri words above *chòòq* and *èm* are known to elderly A-Mlabri speakers as obsolete words, and the former may be current as a pejorative term; *qyoc* is known in the meaning of 'wild fowl'. Among the words specific to A-Mlabri several are obvious loans and thus widely known as extraneous words.

As for "genuine" A-specific words, *nqlq.nglèq* clearly rang a bell when I mentioned it to an elderly B-Mlabri speaker but he explained to me that it is very difficult to say so I "had better say *ku.kòòq*" like they themselves prefer to do! As for the word *chr.kèng*, it is known throughout all varieties of Mlabri in the meaning of 'wing'; in the derived meaning of 'hen' it is probably an euphemistic word: a noa-word, which suggests that originally it came into use about wild fowls, for which the regular Mlabri-word was *qyoc*.

Turning to the A/C-Mlabri words, *ek* is known but seems less used in B-Mlabri; the B-word *toc*, in return, is used in A-/C-Mlabri in the meaning of 'fetch'. The A/C-Mlabri word *ěeq* I have encountered in an elaborated B-Mlabri expression with two parallel near-synonyms: *ěeq bong*; the B-word *bong*, on the other hand, is a frequent word also in A-/C-Mlabri but is here used only about the consumption of meat. As for *glěeq*, this word does in fact occur in several collocations in B-Mlabri although the neutral B-word for 'head' (at least in the usage of males) is the Kmhmu-word *kam.pong*. Finally, whereas *căny* is the standard word for 'tooth' in A-/C-Mlabri but seems totally absent in B-Mlabri, its (near-)synonym *thrèng* exists in A-Mlabri but is used only to refer to the lower (front) teeth (also the C-Mlabri seem to recognize the word).

3.3 Overall lexical relatedness

The conclusion so far is that from a comparative or diachronic perspective the lexical differences among the varieties of Mlabri are superficial. I therefore wish to qualify the statements I have made earlier (e.g. Rischel 1989c, 1995) about the lexical distance between varieties of Mlabri: it must be emphasized that it is a matter of different lexical usage in

everyday speech. The differences are striking but they shrink if a larger register of speaking-styles, and not only the active but also the passive linguistic competence of mature speakers, is taken into account. All across A-, B- and C-Mlabri the knowledgeable, i.e. elderly persons, especially such that master a repertoire of traditional narrative or ritual texts, may have much overlap in their total lexical competence^{1, 7}.

In the face of such a scenario, lexicostatistical methods can hardly be used to shed light on genetic distances between languages or dialects. If, however, one forgets about preconceived lists of allegedly "basic" words and instead applies statistical comparison to words that are frequent in running speech in everyday usage, there is a good chance of making meaningful predictions about something else, namely about the degree to which intercommunication between speakers (in this case speakers of A-, B- and C-Mlabri) without previous experience in such situations would be successful. Genetically close languages or language varieties may be less close to each other when it comes to communicative function; I made this point about Mlabri in Rischel (1993).

3.4 Parameters of lexical differentiation

The lexical trifurcation of Mlabri is complicated by the intersection of parameters of lexical use. I have already mentioned obsolescence versus currency as one such parameter; it is probably useful to distinguish three levels here: (i) totally current words, (ii) words which can be used but have been replaced in daily usage, (iii) obsolete and often also stigmatized words known only to certain elders. It takes some digging to get access to words of category (iii); some speakers get irritated if confronted with this old vintage because they feel that such words are not part of their conception of current proper Mlabri. Even words of category (ii) may escape a field worker's notice for a long time. For firewood' A-Mlabri has a standard term *hng.keeq*, but there is a synonym *uulh* (an etymon found widely with -s in Mon-Khmer) which occurs in Bernatzik's list. It was only after retrieving the latter as the standard term in B-Mlabri that I discovered that it is common Mlabri and accepted as a live word also in A-Mlabri, though it has yielded to *hng.keeq* in daily usage.

A more intriguing and difficult parameter is gender. Mlabri men and women have more or less different vocabularies, though the information on this parameter is so far insufficient for comparative study.

It has not so far been possible to retrieve much by way of mythology or cosmology from the C-Mlabri, either because they have lost many of their spiritual traditions or rather because they are not ready to reveal them to outsiders whom they have known only for a short time.

I have extremely little data for A-Mlabri since my field sessions have been entirely male dominated for reasons of social etiquette, but fairly extensive metalinguistic information for B-Mlabri (see Rischel 1995, Part II: Dictionary). Unfortunately that was obtained almost exclusively from a male speaker and does not seem to match actual usage completely. The most tangible information is from C-Mlabri, since there is clearly no social code restricting my verbal interaction with C-Mlabri women, but the information on gender-specific usage is meagre and also awaits processing. An amusing example is 'nose'. The Mlabri word for 'nose' is a well-known Mon-Khmer etymon *mòh*; a form containing this lexeme occurs also in Bernatzik's list, but in modern C-Mlabri only women can use the word. C-Mlabri men hear the word *mòh* from women but get extremely embarrassed if they are asked to pronounce it themselves. It seems that it evokes an obscene association, so the men insist on using a quite different word: *lip.luap*.

The obsolescence and gender parameters enter the comparative-diachronic lexical scenario in complex ways, as illustrated by the words for 'nose'. The gender specific use of some vocabulary

¹ There are some interesting exceptions. One is *burthor* or *burthol* 'hair on the head', which is reserved for use by females everywhere. There may also be a shared tendency for two synonyms for 'elephant': *pm.poo* and *chaang* to be distributed in such a way that the former is male language, the latter female language

may be a very recent phenomenon since it is often not shared by different varieties of Mlabri¹. When digging into the vocabulary of B-Mlabri (with which I am most familiar) I have often found that a word which I thought was exclusively A-Mlabri or C-Mlabri occurs in B-Mlabri as well but only in female usage. C-Mlabri is particularly interesting since it sometimes subsumes both A- and B-usages but makes a lexical gender difference: whereas 'blanket' is *pol* in A-Mlabri and *gīncay* in B-Mlabri, women say *pol* and men say *khn.cay* in C-Mlabri. It is also sometimes the case that (allegedly) male and female language in one of the language varieties differ in that one uses a Mon-Khmer word, the other a Tai loanword.

4. The diachronic perspective

4.1 *The ethnolectal scenario, its causes and time-depth*

All three varieties of Mlabri share a number of old loanwords preserving features of Pre-Tin phonology (cf. Rischel 1989b) or of old Tai phonology (e.g. *hlek* 'iron'). Even a word such as *khòt* 'spear', which according to its whole phonological makeup seems to be a fairly late loan from a dialect of Tin², is common Mlabri. Unlike very recent loans e.g. from Lao, earlier layers of loanwords thus reflect the common history of Mlabri.

The old loanwords are also interesting in showing that Mlabri phonology has changed only little over several centuries. Consistent with this, its most recent history is conspicuously static. If forms in Nimmanhaeminda (1963), Ferlus (1964) and Bernatzik (1938) differ from contemporary data, then this is a matter of transcription conventions rather than phonological change. In the case of C-Mlabri one can even observe live patterns of dialect-internal phonological variation going back to Bernatzik's time (Rischel 2000). The diachronic stability over the time span we have access to is thus very high.

The presence of three stable varieties of Mlabri seems to reflect a long-standing polarization among tribal subgroups. They have prejudices about each other which have been kept alive up to this day, so they are aware of but carefully stay clear of each other. One might call the subgroups clans but I think it is more appropriate to speak of ethnic subgroups, and I have chosen to speak of the three language varieties as ethnolects. It is impossible to say how far back in time the present scenario should be projected but the high stability of the differential data over forty to sixty years suggests that the trifurcation is much older than that. The ethnolects remain strangely interlocked considering the apparent lack of mutual contact. This observation may be relevant to the study of oral languages for which there is little or no time depth in the recorded data.

4.2 *Stammbaum relationships?*

Finally, there are a few speculations as to whether the three ethnolects reflect one ternary branching or two binary branchings with some time gap in between. The lexical evidence is so far inconclusive. Although B-Mlabri and C-Mlabri seem lexically closest to each other, this need not reflect the original state of affairs since A-Mlabri is clearly very innovative, with formations such as *brany* for 'dog' (versus *brang* 'horse') and *chr.kèng* 'wing' used in the meaning of 'hen, domesticated chicken'. Altogether, the lexical variation over the ethnolects is suggestive of a formerly richer Mlabri lexicon, with several synonyms or near-synonyms (often due to borrowing). According to this scenario, it was the case in each separate ethnolect that its lexicon shrank more or less by selective generalization of one or another of two synonyms; at the same

¹ In genuine Tin-words *kh-* comes from unaspirated **k*, which would surface if Mlabri had borrowed the word early; the final *-t* is a dialect-specific feature of modern Tin. Bernatzik (1938) supposed that the Mlabri acquired spears through fairly recent contact with the Hmong and does not mention the Tin connection.

² The variation can be seen already in Bernatzik (1938). His spellings suggest that it started as allomorphy, *ak* being a bound variant before words beginning with a velar. In contemporary contemporary C-Mlabri, however, there is individual variation: some say *at*, others say *ak*. In A-Mlabri I have heard *at* only from a few elderly informants; the form *ak* is entirely dominant in the speech of young people, and some even generalize the substitution of *k* for *t* in the definite article to another grammatical word specific to Mlabri: the first person possessive pronoun *ot* (derived from *oh* 'I'). These persons say *ok^hw* 'my child' for conservative *ot^hw*, for example. Altogether, the histories and variational patterns of these strange prenominal determiners deserve closer study.

time it was rescued or enriched by neologisms and by late borrowings from surrounding languages. The lexical substitution process is clearly related to a stigmatization of certain expressions, either so that a word became entirely undesirable in a certain ethnolect or so that it was confined to use by one of the genders within that ethnolect. This fits the data, as far as I can see, and it accounts for the strange lexical divergencies among the ethnolects, but it implies that lexicon is of little help in establishing a pedigree.

4.3. Phonology revisited

In the light of the lexical complexity, phonology becomes quite crucial in the comparison of ethnolects although there is little to work with as a result of the general phonological conservatism of Mlabri, coupled with the shallow time depth we have access to.

The phonological evidence is slightly in favour of a first split between B-Mlabri and the rest. For one thing, A-Mlabri and C-Mlabri share an interesting development in the Definite Article, which is continued as *at* in B-Mlabri but exhibits a conspicuous variation between *at* and *ak* both in A-Mlabri and in C-Mlabri all the way back to Bernatzik.¹ That might also reflect later contact between these two groups, however. Another thing is that B-Mlabri alone has a drastic sound-shift by which *aa* in a handful of words has changed into *ii*, e.g. *kliir* 'sky' for A-Mlabri and C-Mlabri *klaar*. If (as I presume) this came about by diphthongization **aa* > **ia* followed by the areally very well-attested monophthongization **ia* > *ii* it must have required some time for the sound change to be completed.

Unfortunately we have no useful data on the Mlabri before Bernatzik (1938). The existence of a tribe hiding in the mountain forests of Northern Thailand and referred to locally as Phi Tong Luang, 'Ghosts of the Yellow Banana Leaves', was first mentioned in print by Seidenfaden (1920, 1927) and Kerr (1924). Neither author provided any useful information on the language the tribal people spoke,² so we have no solid proof that it was Mlabri, but that is overwhelmingly likely. According to Seidenfaden their speech was said to be "piping", which may reflect the extremely long falling pitch contours later observed as characteristic of A-Mlabri (cf. Rischel 1982, Egerod and Rischel 1987). The prosody of A-Mlabri and C-Mlabri is, however, impressionistically fairly similar, whereas B-Mlabri has a markedly different speech rhythm and intonation (perhaps due to bilingualism stemming from close contact with certain highlander villages over time).

5. Conclusion

Mlabri, including so-called "Yumbri", is one language comprising at least three distinct varieties, which may be called ethnolects.

The Mlabri ethnolects of which we have any knowledge are all still spoken, and they have now been documented over time spans of many decades. Each ethnolect shows a remarkable stability over time, considering the purely oral status of the language. At the same time the ethnolects are closely related, to the point of being sociolectal varieties of one dialect. The lexical differences among them, though considerable, are for the most part rather shallow and reflect some still operative parameters of social distinctness.

Diachronically, the information is so meagre that the ethnolectal trifurcation of Mlabri cannot so far be given any definitive interpretation.

BIBLIOGRAPHY

¹ Intriguingly, the only precisely formulated piece of information about the language is the paradoxical contention that "the letter R does not occur" (Seidenfaden 1927:47). That certainly does not apply to Mlabri which has [r] both initially and finally in many words, even in a number of Tai loans which occur with [l] or [h] in neighbouring languages. It may be a matter of sharing a stereotype about northern speech with some Tin dialects.

² The difference in shape between *kóh* and *kôh* is due to tone sandhi.

1. Bernatzik, Hugo A. 1938 *Die Geister der gelben Blätter*. München. 2nd ed. Leipzig 1941.
2. Egerod, Sören 1982 An English-Mlabri basic vocabulary. *Annual Newsletter of the Scandinavian Institute of Asian Studies*. 16:14-20.
3. Egerod, Sören and Jörgen Rischel 1987 A Mlabri-English vocabulary. *Acta Orientalia*. 48:35-88.
4. Ferlus, Michel 1964 Field notes from Laos (personal communication).
5. Ferlus, Michel 1974 Les langues du groupe austroasiatique nord. *ASEMI*. V.1, see p. 39ff.
6. Kerr, A.F.G. 1924 The Ka Tawng Lûang. *JSS*. 18.2:142-144.
7. Nimmanhaeminda, Kraisri 1963 The Mrabri language. *JSS*. 51.2:179-184.
8. Rischel, Jörgen 1982 Fieldwork on the Mlabri language: a preliminary sketch of its phonetics. *Annual Report of the Institute of Phonetics, Univ. of Copenhagen*. 16:247-255.
9. Rischel, Jörgen 1989a Fifty years of research on the Mlabri language. A re-appraisal of old and recent fieldwork data. *Acta Orientalia*. 50:49-78.
10. Rischel, Jörgen 1989b Can the Khmuic component in Mlabri ('Phi Tong Luang') be identified as old T'in?. *Acta Orientalia*. 50:79-115.
11. Rischel, Jörgen 1989c What language do 'The Spirits of the Yellow Leaves' speak?: A case of conflicting lexical and phonological evidence. *Ann. Rep. of the Inst. of Phonetics, Univ. of Copenhagen*. 23:87-118.
12. Rischel, Jörgen 1992 Isolation, contact, and lexical variation in a tribal setting. in E.H. Jahr ed.: *Language contact. Theoretical and empirical studies = Trends in linguistics. Studies and monographs*. 60:149-177. Mouton de Gruyter.
13. Rischel, Jörgen 1993 Lexical variation in two 'Kammuic' languages. *Pan-Asiatic Linguistics, Proceedings of the Pan-Asiatic Conference on Languages and Linguistics*. III:1451-1462. Bangkok: Chulalongkorn University.
14. Rischel, Jörgen 1994 Connectives in Mlabri (Les connectifs en mlabri). *Résumés. 27ème Congrès International sur les Langues et la Linguistique Sino-Tibétaines*:51. Paris: Centre National de la Recherche Scientifique.
15. Rischel, Jörgen 1995 *Minor Mlabri, A Hunter-Gatherer Language of Northern Indochina*. Copenhagen: Museum Tusculanum.
16. Rischel, Jörgen 1997 Typology and reconstruction of numeral systems: The case of Austroasiatic. Jacek Fisiak ed.: *Linguistic Reconstruction and Typology = Trends in Linguistics. Studies and Monographs*. 96:273-312. Mouton de Gruyter.
17. Rischel, Jörgen 2000 The dialect of Bernatzik's (1938) 'Yumbri' refound? (to appear in *Mon-Khmer Studies*.)
18. Rischel, Jörgen ms. (August 2000) The language of the forest tribe in Phiang District.
19. Rischel, Jörgen and Sören Egerod 1987 'Yumbri' (Phi Tong Luang) and Mlabri. *Acta Orientalia*. 48:19-33.
20. Schuhmacher, W.W. 1969 Zur Klassifizierung der Phi Tong Luang Sprache. - Ein kommunikationstheoretischer Versuch. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*. 22.4:360-363.
21. Seidenfaden, E. 1920 Further notes about Chaubun, etc. *JSS*. 13.3:47-53.
22. Seidenfaden, E. 1927 The Khâ Tong Lu'ang. *JSS*. 20:41-48.
23. Smalley, William A. 1963 Notes on Kraisri's and Bernatzik's word lists. *JSS*. 51.2:189-201.

24. Smalley, William A. 1994 *Linguistic Diversity and National Unity. Language Ecology in Thailand*. University of Chicago Press.

KẾT NỐI MỆNH ĐỀ TRONG VĂN BIA RAMKAMHAENG VÀ TRONG TIẾNG THÁI HIỆN ĐẠI

Pranee Kullavanijaya và Stanley Starosta

Trong văn bia Ramkamhaeng, các mệnh đề trong phát ngôn được kết nối bằng những liên từ như *thâa/thîi*, các giới từ mở rộng/ extension prepositions (liên từ phụ thuộc) như *phîa* và *phr* ? và câu làm trạng từ như *as ctī* và *kôc/kóh*. Trong bài nghiên cứu này, các tác giả sẽ đưa ra một phân tích cú pháp các kết cấu có chứa những từ này và những từ liên quan về phương diện phái sinh hay đổi vị, các tác giả cũng sẽ thảo luận những thuộc tính ngữ nghĩa và đối chiếu hệ thống các phương tiện kết nối mệnh đề trên các văn bia theo cách thức mà các phương thức kết nối mệnh đề đã thay đổi hoặc được thay thế hoặc còn đang tranh chấp với các công cụ kết nối mệnh đề trong tiếng Thái hiện đại.