

WHAT LANGUAGE DO "THE SPIRITS OF THE YELLOW LEAVES"
SPEAK?: A CASE OF CONFLICTING LEXICAL
AND PHONOLOGICAL EVIDENCE*

Jørgen Rischel

This paper (which in part summarizes two papers to appear in Acta Orientalia but which presents separate information as well) deals with some issues raised by descriptive and comparative linguistic work in northern Thailand. The putative Austroasiatic languages "Yumbri" and "Mrabri" (more correctly: Mlabri) have been assigned to "Khmuic" within the Mon-Khmer languages, but the relationship between these two idioms has been a controversial issue. On the basis of recent fieldwork all existing data on "Yumbri" and "Mrabri" can be shown to reflect one and the same language Mlabri in spite of wide discrepancies in notation; these do not even reveal major phonological dialect differences whereas there are conspicuously different lexical usages. This has not so far been properly understood because of difficulties in the interpretation of earlier data which were all gathered by amateurs. - As for the tentative genetic classification of Mlabri as Khmuic, the lexical evidence used to substantiate this claim now turns out to be controversial: a large number of the Khmuic words in Mlabri are rather direct reflexes of an early stage of Tin, a language that has been assigned to the Khmuic branch of Mon-Khmer. Thus, it is either the case that Mlabri and Tin are sister-languages (forming a "Tinic" branch of Khmuic) or that Mlabri has early borrowings from Tin.

* The work on "α-Mlabri" and much of the work on Bernatzik's "Yumbri" was done in close collaboration between Professor Søren Egerod of The East Asian Department, U. of Copenhagen, and this author.

1. Introduction.

Mlabri (i.e. /mla:ʔ bri:ʔ/ 'people of the forest') is a small tribe of hunter-gatherers living in the northern part of Indochina. Of this tribe there are less than two hundred adults and children who live in the eastern provinces of northern Thailand; it is unknown to what extent and in what number there are Mlabri in neighbouring countries. The Mlabri in Thailand are now rapidly giving up their former nomadic life partly because there is not enough forest left to provide food for hunter-gatherers, which makes them dependent on villein service on the fields belonging to Miao (Hmong) peasants, and partly because the Thais have recently engaged in a program of acculturation (school training etc.).

The Mlabri speak a language (also called Mlabri) of which a brief phonetic sketch was given in an earlier issue of ARIPUC (Rischel 1982).

Some fifty years ago Hugo Bernatzik (1938) wrote a report about his encounter with a mysterious tribe called the "Spirits of the Yellow Leaves" in the mountains of Northern Thailand. He included two rather short word lists in the language of these hunter-gatherers, who according to Bernatzik called themselves "Yumbri". Later reports (Kraisri 1963) have dealt with a related language called "Mrabri", more correctly Mlabri (see, e.g., Rischel 1982), which is likewise spoken by tribal people referred to as the "Spirits of the Yellow Leaves" (= "Phi Tong LUang" in Thai).

In linguistic handbooks one finds the designations Yumbri, Mrabri, and Phi Tong Luang; the recent language map of Thailand (1977) just gives the (correct) cover term Mlabri. However, since Bernatzik's data differ considerably from those of later sources, it has been a matter of dispute whether these terms refer to different languages or whether they all refer to one and the same language.

Kraisri Nimmanhaeminda, who had worked with the Mlabri in 1961 and 1962 (Kraisri 1963), assumed that "the Yumbri and Mrabri are the same people" and found that "the Yumbri and Mrabri languages are close to Mon-Khmer languages and they should belong to this group". The obvious difficulty with linguistic comparisons involving the older Yumbri and Mrabri data is that neither Bernatzik nor Kraisri used a professional phonetic or phonemic notation.

Smalley (1963) made an impressive attempt at a phonological restatement of both Bernatzik's and Kraisri's data, assuming that these languages were

typologically close to well-known northern Mon-Khmer (Austroasiatic) languages. He further made a lexicostatistic comparison of the Yumbri and Mrabri data on the basis of Swadesh' list, and compared these with some Austroasiatic languages including Kammu (or Khmu?) and Tin (or T'in), languages assigned to the Kammuic or Khmuic branch (on lexicostatistic evidence, Thomas and Headley 1970). Smalley concluded that Yumbri and Mrabri must be two different languages and that "the Mrabri are linguistically just as closely related to the Khmu? and the Tin as to the Yumbri". He further made the remark that he very much doubted that Mrabri and Yumbri would be mutually intelligible (Smalley 1963, p. 191).

Thus, some twenty-five years ago the prevailing opinion was that there are two languages Yumbri and Mrabri spoken by groups of the same kind of people, and that these languages belong to Khmuic. However, Ferlus (1974, p.47-48) referring to data from Laos (see below) found that Yumbri and Mrabri are undoubtedly one language which has branched into dialects because of the way its speakers live. Rischel and Egerod (1987) published evidence for an identification of Yumbri and Mrabri as one language: Mlabri, a result which is strongly at variance with Smalley's analysis of the earlier data, but which was based on new lexical data stemming from rather extensive fieldwork.

Nobody so far has suggested an alternative to the classification of Yumbri/Mrabri as belonging to the Khmuic branch, although this classification is in fact much more controversial than our identification of Bernatzik's and Kraissri's data as specimens of one and the same language. There are obvious affiliations with Khmuic, but in principle these might be due to influence from a neighbouring language on Mlabri (Yumbri/Mrabri) at some period in time, and in fact there is very strong evidence in favour of an old layer of influence from Tin, as will be shown below (also see Rischel forthcoming b).

The description and classification of an almost unknown tribal language spoken by maybe some 200 persons may seem of limited phonetic or linguistic interest. But I think the difficulties researchers have had with the proper classification of Mlabri may be of some interest also to theoreticians. It is thus the purpose of the present paper to use the Mlabri language to illustrate the obstacles and pitfalls that linguists may encounter when interpreting old fieldwork data and when attempting to classify languages for which there are no very old records available.

2. "Yumbri", "Mrabri" and Mlabri.

2.1. The transcription problem.

Rischel and Egerod (1987) gives a survey of the research done on Mlabri up to that time. In addition to Bernatzik's "Yumbri" data from 1938 and Kraisri's "Mrabri" data from 1963 there is a small word list which Michel Ferlus took down in IPA notation in Xagnabouri province of Laos in 1964 from Mlabri speakers, who according to Ferlus referred to their language as kǎmlua? (lua? is a name associated with several groups speaking Mon-Khmer languages). Ferlus found their language to be rather close to Kraisri's Mrabri, and in cases of disagreement between the two sources the word was often found in Bernatzik's Yumbri. Quite recently Dr. Ferlus has kindly let me study his word list, of which some specimens are given below. Furthermore, Jesper Trier (1986) has recorded the oral ritual texts of the Mlabri and given some specimens in his own transcription, with English translations.

In order to illustrate the phonetic relationship between the data in these sources and our own Mlabri data a few lexical items which happen to be shared by most sources will be cited here exactly the way they are transcribed by the various authors:

| | Bernatzik 1938 | Kraisri 1963 | Ferlus 1964 | Trier 1986 | Egerod & Rischel |
|----------|-------------------|-----------------|----------------|---------------|---------------------|
| father | ěmūm | merm | mǎm | mǝm | mɤm |
| mother | ěmū | merh | mǎ? | mǝ | mɤ? |
| man | ěyǝgn | yǝn | - | yom | jo:n |
| eye | mǎt | mǎd | mǎt | (mad) | mat |
| tree | lǎm | lam | lǎm | lam | lam |
| wind | rmǝt | - | rmɤt | rui-mud | rmwt/rumwt |
| spear | kǝt | kod | khǝt | kot | khǝt |
| spirit | - | wok | wǝk | wok | wǝk |
| die/dead | bǝl | bǝl | bǝl | bul | bwl |

[Commentary to some individual words: (1) As for 'father', 'mother', 'man' Bernatzik has an initial vowel ě-. It occurs in his list with several nouns and may be a particle which we know as /ʔat/ or /ʔak/. (2) As for the word 'eye', Trier's form is translated 'to see'. (3) For 'wind' Kraisri has the Thai word lom. (4) As for 'spirit', Bernatzik has some quite different words occurring in phrases which are difficult to interpret.]

At first glance the correspondences given above may give the impression of considerable phonetic distance, such as between dialects or even between different but closely related languages. However, it is obvious that quite a few of the differences are simply artefacts of the different notational conventions the researchers have used. In fact, only two sets of data can be directly compared, viz. those of Ferlus and of Egerod & Rischel who use the IPA alphabet; it is noteworthy that these very sets are extremely alike though one was taken down in Laos twenty-five years ago, the other in Thailand during the last few years. It should be noted that Ferlus did not indicate vowel length but instead indicated brevity in some cases by an arc over the vowel (example: $\overset{\frown}{o}$); Bernatzik did the same but also indicated long vowels by a stroke over the vowel symbols (such as \bar{o}). Among the purely conventional differences I should also mention the use of [y] in some sources versus IPA [j] in others, and the (inconsistent) use in Bernatzik's data of underlining to indicate more open vowel quality (\underline{e} being intended as a more open vowel quality than e, for example). The symbol ŋ is used in some cases by Bernatzik to indicate a velar nasal, but he sometimes writes gn instead in words that have a velar nasal in Mlabri.

However, there is no simple way to assess the phonetic or phonemic meaning of the data that is given in non-IPA notation since with the exception of Bernatzik the authors give no separate phonetic information, and Bernatzik says little more than has been mentioned already about his own notation. One may get a little closer to an interpretation by confronting a general typological knowledge (of sound systems in northern Mon-Khmer languages and of the sound system of contemporary Mlabri par excellence) with our knowledge about the various untrained authors' backgrounds. - Bernatzik and Trier are immediately comparable on this point.

Before going into details of the various authors' transcriptions I shall present the inventory of Mlabri phonemes according to a very simple-minded phonemicization, cf. Rischel 1982 and also cf. Egerod and Rischel 1987, p.36ff. (In Egerod and Rischel 1987 vowel length was not indicated since we were not sure how to phonemicize with regard to quantity; the length mark given in this paper agree with our most recent revision of the data.)

As for the vowels, Mon-Khmer languages may exhibit three or even four degrees of aperture; Mlabri has a distinction of four in the back unrounded series /w ɤ ʌ a/. Moreover, there are three different sets: front unrounded, (mid or

back) unrounded, and back rounded, e.g. /i w u/. This together makes for a vowel system in Mlabri which is of the same complexity as the Danish one but with very different vowel qualities:

| | | |
|---|---|---|
| i | w | u |
| e | ɣ | o |
| ɛ | ʌ | ɔ |
| | a | |

The consonant system (the maximal system, viz. in initial position) may be set up as follows:

| | | | | |
|----|-----|------|----|---|
| ph | th | ch/ɕ | kh | |
| p | t | c | k | |
| b | d | ʃ | g | |
| ʔb | ʔd | | | |
| hm | hn | | hŋ | |
| m | n | ɲ | ŋ | |
| hw | hl | | | h |
| w | l r | j | | |
| ʔw | | ʔj | | ʔ |

Word finally only the following occur in contrast: /p t c k m n ɲ ŋ lh ɕ h w l r j ʔ/.

The digraphs /hm hn hŋ hw hl/ indicate essentially voiceless consonants (given as /m̥ n̥ ŋ̥ w̥ l̥/ in my 1982-paper); the digraph /lh/ indicates a (final) lateral with devoicing increasing toward the end of the segment (in 1982 I used the symbol l̥).

Ferlus' notation of 1964 is largely - and in fact surprisingly - consistent with this representation of Mlabri, although there are some differences in individual wordforms. As for purely technical differences in notation I shall mention that he uses ɛ or ɛ̣ for our /w/, and ɛ̣ for our initial /ch/ or /ɕ/ (the latter phoneme is in fact a sibilant more often than an affricate, and it is never a pure stop in Mlabri, so the representation /ɕ/ or /ɛ̣/ is preferable from the point of view of phonetic realism). As said above Ferlus does not indicate length but sometimes vowel brevity, cf.

| | Ferlus | Egerod and Rischel |
|----------|--------|--------------------|
| earth | bɛ̣ʔ | bɛʔ |
| mushroom | het | het |
| fish | kaʔ | ka:ʔ |

The fact that there is not complete bi-uniqueness between our transcriptions could be due either to difficulties in assessing length (which is extremely difficult for Mlabri) or to differences in the linguistic usages of informants, or both. Anyway,

the phonetic similarity between his kām̄luaʔ of Laos and our Mlabri should be apparent.

Let us look then at Kraisi's data from 1963. Kraisi used a notation which seems in part based on English semi-phonetic transcription. However, for the sake of readers with better command of Thai he added transcriptions using Thai letters. It turns out that these Thai transcriptions are generally more adequate and much more informative, which is not very surprising, since the sound system of Thai is typologically reasonably close to that of Mlabri. If we take the words listed above and transliterate the Thai letters to IPA symbols in accordance with Standard Thai pronunciation, the forms will look approximately as follows (for convenience our own forms are given again):

| | Kraisi 1963 | Egerod & Rischel |
|----------|----------------|---------------------|
| father | mɔ̄m | mɔ̄m |
| mother | mɔ̄ʔ | mɔ̄ʔ |
| man | jo:ŋ | jo:ŋ |
| eye | má:t | mat |
| tree | lām | lam |
| spear | khó:t | khɔt |
| spirit | wó:k | wɔk |
| die/dead | bw:l | bwl |

(the accent marks over some of the vowels in the transliterations of Kraisi's forms reflect his use of Thai tone marks to indicate that he has heard high or falling tone in the rendering of these words, which is probably a matter of citation intonation; as for the word 'spirit' it occurs only as the first part of phrases in his list).

It is impressive how well these words (though certainly not all words on his lists!) agree with our transcription and with that of Ferlus (see above). Kraisi was in fact transcribing the very dialect that we have documented in Egerod and Rischel (1987) and elsewhere; this is directly confirmed by the fact that we have recently had his informant Ai Pla as our informant and had our entire word list rechecked with him. Not much seems to have happened with the pronunciation over the time span of twenty-five years separating the two fieldwork sessions, although it is interesting that Kraisi uses the symbols for long vowels more often than is warranted by Ai Pla's pronunciation nowadays.

Then to Bernatzik and Trier. Being both Westerners without a professional background in the use of linguistic field methods and phonetic transcription they are in principle comparable enough to be considered together here.

Bernatzik was an Austrian (whose English was not very good, it is said), and Trier is a Dane. We would predict that both of them would either fail to notice a glottal stop consonant or not have any consistent way of indicating it. Mlabri has an initial contrast between a glottal stop and other stops, and a final contrast between glottal and non-glottal syllable termination. The prediction is borne out, cf. Bernatzik atí corresponding to (conservative) Mlabri ?at ti:ʔ '(the) hand', bæ = Mlabri bæʔ 'earth', Trier ø = Mlabri ?x(:)ʔ 'to eat', tar = Mlabri ta:ʔ 'grandfather, uncle', etc.

Mon-Khmer languages have an abundance of different manners of articulation of initial stops, some of them exhibiting contrastive voicing and aspiration and glottalization and even prenasalization; Mlabri has four manners: /ph p b ʔb/ and even vestiges of a fifth: /mb/. Austrian German and Danish have only a two-way contrast between /p t k/ (which are aspirated in Danish but not in Austrian) and /b d g/ (which are voiceless). One might thus expect under-differentiation and possibly also inconsistency in the notation of the initial stops by these authors since they use ordinary letters (in Bernatzik's case with several added diacritics which, however, do not ever serve the purpose of distinguishing between manners of articulation). This also is borne out by a comparison with our Mlabri data. Bernatzik distinguishes well between /p t c k/ and /b d ɣ g/, as one might perhaps expect, but rather less so between the aspirated and unaspirated stops, cf. kěyót = Mlabri khəj joc 'egg of wild fowl' or 'hen's egg' (depending on dialect) vs. kək = Mlabri ke:c or ke:t 'ear', kót = Mlabri khot 'spear' vs. kón = Mlabri ko:n 'to snore'. (The material is too small to show further details on this.) - Trier, on the other hand, turns out to vacillate when transcribing words that occur in our data with unaspirated tenues (/p/ versus /b/ etc., a contrast that is missing in Danish) cf. his pung = Mlabri pu:ŋ 'to blow' vs. bor = Mlabri pɔʔ 'to push' vs. bung = Mlabri boŋ 'to eat (meat)' or gaep = Mlabri kɛp 'stone' vs. gaeng = Mlabri gɛ:ŋ 'windscreen', etc. The aspirates are identified with his Danish aspirates, spelled p t k: thus he writes kai = Mlabri khəj 'egg', which is indeed predictable. As for glottalized stops these are rendered as plain stops, e.g. ding = Mlabri ʔdiŋ 'big' like diŋ = Mlabri diŋ 'elder sibling'.

Mlabri has voiceless nonsibilant continuants which must cause difficulties for untrained listeners. Bernatzik does not indicate voicelessness in such cases, cf. *kək lút* = Mlabri *ke:c hlu:t* 'deaf', *kí mǎ* = Mlabri *ki:ʔ hmɛʔ* 'new moon'. Trier has *kl* for the voiceless lateral in *kli* or *klæ* = Mlabri *hlek* 'iron'.

It is a further complication that Mon-Khmer languages and also Mlabri have four oral points of articulation in stops and nasals both initially and finally, viz. labial, dental, palatal, and velar. One may here expect some difficulties with the unfamiliar palatals (which invite a transcription as clusters with "j" or "y") and also with the velar nasal in initial position. By and large Bernatzik and Trier agree with our data as regards point of articulation in consonants, though with several discrepancies some of which at least must be downright errors. - Bernatzik mentions explicitly that it is often difficult to hear the stops in final position (as indeed it is in Mlabri, which has unreleased stops in this position as is typical of languages in the area).

The vowel system of Mlabri is of about the same complexity as that of Danish, as said above, but it is certainly more complex than that of Austrian German. Because of the special character of the English vowel system even a good command of English would be of little help here. Thus we may expect Bernatzik and Trier to have had considerable difficulties in matching the perceived vowels with letters in the Latin alphabet, and we may a priori expect their vowel notations to be more or less underdifferentiating and more or less inconsistent.

The notation of vowels does indeed exhibit great discrepancies between the various sources, and it is very difficult to decide what is due only to different conventions (such as Trier's use of the letter *r* in "ar", "or" to indicate open vowel qualities) and what reflects genuine phonetic differences among dialects.

If we assume that Bernatzik was transcribing Mlabri the following obtains:

The Roman letter symbols "i e u o a" generally have a straightforward phonetic interpretation if compared to the spelling conventions of languages such as German. There are serious shortcomings in Bernatzik's transcription, however. Thus "ü" in his Yumbri wordforms may mean short unstressed /u/ or /w/ (occasionally other vowels as well, though not often). There is in his transcription system

apparently no separate (simplex or complex) symbol for the highly frequent vowel phoneme /w/, nor for /ʌ/; the whole series /w ʌ/ of Mlabri is rendered by "u" and "o" with or without relevant diacritics.

As for Bernatzik's diacritic marks over and under vowel symbols the following information is found in the introduction to his word list (he used similar conventions in transcribing other languages, cf. Bernatzik 1947, p. 4):

- (a) ´ (the acute accent) is a stress mark ("Betonungszeichen")
- (b) ˘ ("hachek") indicates brevity
- (c) ¯ (horizontal stroke) over a letter indicates length
- (d) underlining under a vowel indicates open pronunciation, e.g. e like German e in "Märchen"
- (e) a dot under a vowel indicates close quality, e.g. o like in German "Sohn" (sic!, but cf. below)

Comment to (d)-(e): Bernatzik's plain and underlined vowel symbols "e e o o ɔ" tend to correspond one-to-one to the Mlabri vowels /e ε o ɔ ʌ/, but he seems not to be very consistent, to judge from the Mlabri correspondences. The pattern is also somewhat obscured by idiosyncratic spellings. The counterparts to Mlabri /o/ and /ʌ/ are often not distinguished, and the underlining for openness is often used in a way which runs counter to the vowel qualities in Mlabri, in particular, the symbol "o" is often used for Mlabri /o/ although it is meant to represent a more open vowel, i.e. /ɔ/, to judge from Bernatzik's own explanations. The following examples may illustrate the degree of overlap and mismatch we encounter in some words when comparing "Yumbri" and Mlabri:

| | Bernatzik | Egerod & Rischel |
|--------------|--------------------------------|------------------|
| penis | t ^o n | ?don |
| to see | d ^o gn | dʌŋ |
| to be wet | tš ^u k ^o | chukko? |
| to be scared | kr ^o u | krʌw |
| to cry | b ^e t | be:c/be:t |
| to beat | t ^e k | tɛk |

Trier has mostly i, y, u, ø, a (ar) corresponding to our /i w u ʌ a/; ae occurs for /e/ and /ε/, u, o and å are all used for /o/, and both o and å are used for Mlabri /ɔ/. The remaining central vowel /ʌ/ in Mlabri has a variety of reflexes in Trier's notation, cf. gor = Mlabri gʌh 'here', wål = Mlabri wʌl 'to return', gam = Mlabri gʌm 'don't' (Trier: 'not'). - All of this is indeed

understandable considering the skewness of the Mlabri and Danish vowel systems in relation to each other, and considering the awkward representation of the vowels of spoken Danish in Danish spelling.

Conclusion about Bernatzik's and Trier's data as putative representations of Mlabri:

If we stick to the obvious and indisputable cognates in Bernatzik's and Trier's data and in our own data, the more or less idiosyncratic notational habits of each author may account for most of the apparent discrepancies. As for Trier's data the possibility of tying it up with our own data, in spite of difficulties with the notation, is entirely in accordance with expectations, since it has turned out that there is a strong overlap in our choice of Mlabri informants (incidentally, a photograph in Trier 1986 shows Mr Ai Som, who also served as an informant for us). We are dealing with the same dialect.

What then about Bernatzik? To the extent that we can recognize his lexical items they look like a rather distorted transcription of Mlabri, perhaps of a dialect with somewhat deviating phonology compared to our Mlabri. I shall return to this issue later.

2.2. The lexical aspect.

As said above we share informants with both Kraisi (1963) and Trier (1986), which firmly establishes the identification of all three sources as representing the same language and even the same dialect. Moreover, Ferlus' data is so very similar to ours in phonetic form that judging only from the obvious cognates one would not hesitate a moment to say that this is the same language and perhaps even the same main dialect.

The transcription of obvious Mlabri cognates in Bernatzik's data can likewise be construed to represent the very same language. On this basis we set out in 1987 to identify as many items in his list as possible, and this seemed so successful that we concluded that his Yumbri is indeed Mlabri (Rischel and Egerod 1987). More recently (Rischel 1988) I have shown that part of the residue of unexplained forms in Bernatzik's list can likewise be interpreted as Mlabri, though the relationship between Bernatzik's data and our own is not quite as simple as we had assumed.

Since several of the words could be identified as plain Mlabri we tried hard to look for Mlabri equivalents to the remaining forms. In some cases we found equivalents that seemed plausible if we allowed for gross inaccuracies and inconsistencies in Bernatzik's transcription and in some cases also for distorted translations.

With the most recent additions to our Mlabri data (from speakers of different age and with different linguistic usages) it has turned out that some of the bad matches were in fact wrong guesses on our part. One embarrassing example, which is given here to warn other field workers, is the Mlabri word for 'tooth':

It has been very enigmatic to us why B would give 'tooth' as "ätréñ". In our Mlabri Vocabulary (1987) the corresponding entry is /(?ak) ɣʌ/. To make such forms fit we must assume either a set of multiple deviations in phonetic form across dialects (which would have no obvious parallels in other items and thus would seem less likely) or we must assume that B's rendering was unusually sloppy. We had, I think, silently worked on the latter assumption under the influence of the general scepticism about the reliability of Bernatzik's data.

This scepticism turned out to be quite unfounded as regards the entry for 'tooth', and as it turned out this is not a case of variation either. It is so that there is a word /thrɛ:ŋ/ meaning 'tooth'. To my great surprise Ai Pla told me that this word is current in his group in the meaning of 'lower teeth' (/ʔat 'thrɛ:ŋ/ with the particle mentioned above = B's "ätréñ"), whereas they would always use /ɣʌ/ as the general term for 'tooth'. The reason why we had not come across /thrɛ:ŋ/ long before may be that we had always pointed at our upper teeth when asking about the word for 'tooth'! - It should be kept in mind that we were working with members of a tribe which is notorious for its shyness and limited contact with modern civilization, and that we had had severe difficulties communicating with the Mlabri about their language. This was true also of a session in which we attempted to elicit forms from Bernatzik's word list. Firstly, there was the problem of explaining the meaning of the word we were searching for, and secondly, there was the difficulty of determining how Bernatzik's spelling was to be interpreted phonetically if we did not know the word in advance. The latter was a great obstacle (see later on the conjectures caused by the inclusion of -Mlabri").

In several cases we had not understood Bernatzik's forms because they were obsolete or rarely used in

the type of Mlabri we had been working with, and because of difficulties with Bernatzik's notation. A single example may suffice for illustration of the difficulties:

Bernatzik's word for 'fire': "tškántūy" did not a priori look too plausible and we guessed that there must be some misunderstanding here. It was only after the publication of Rischel and Egerod (1987) that I became aware of a Mlabri word /ʔulh/ 'fire(wood)' (= B's "-ūy") which we had not so far recorded because it was more or less obsolete in the usage of the Mlabri speakers we had consulted so far: they always used another word: /hŋke:ʔ/. Our Vocabulary already contained the word /çingap/ but we had not recognized that this was what Bernatzik was transcribing because the final part of his noun phrase made no sense to us (I have later established that Ferlus 1964 has /ʔulh/ as well; in fact the word is still current in the dialect of Mlabri referred to as β-Mlabri below). It was now clear that the whole entry "tškántūy" equals Mlabri: /çingap ʔulh/ or /çingap ʔulh/ and means 'smouldering charcoal in the fireplace' (that this is idiomatic, has been verified with informants).

The word for 'to blow' in B's list is "būñūy", but only the first part could be identified, viz. as /puŋ/ 'to blow'. However, with the advent of the form /ʔulh/ 'fire' it became likely that Bernatzik had heard something like /puŋ ʔulh/ (/ - ʔujh/?), a well-formed phrase meaning 'to blow on the fire'.

This identification of Bernatzik's word for fireplace as equivalent to Mlabri /çingap/+ /ʔulh/ implies that he has used the very strange spelling "-nt" for the palatal nasal, and "-y" for Mlabri /lh/, i.e. the symbol for palatal glide instead of the unfamiliar voiceless lateral; maybe "Yumbri" had a voiceless palatal (i.e. /ʔujh/), like Tin??

Anyway, it is true of several of the lexemes in "Yumbri" that fail to resemble Kraisi's "Mrabri", that these turn out to exist in current Mlabri, though often as archaic or quite obsolete words.

Until recently it could not be decided to what extent the differences in the published data on Mlabri reflect dialect differences or changes in the language taking place over the time span of some fifty years of Mlabri studies. More recently I myself happened to meet some Mlabri speakers whom we had not previously encountered and whose linguistic usage turned out to differ significantly from that of our previous informants. There were only very minor differences in phonology, except for the prosody; by and large wordforms shared by

the two dialects could be given the same segmental transcription except for marginal differences, the most conspicuous being that the newly encountered dialect has /w/ in some of the words in which the other has /a/, ex. /klw:r/ versus /kla:r/ 'sky'. In the following I shall refer to these two dialects as α -Mlabri (being the one previously familiar to us = the one described by Kraisi in 1963) and β -Mlabri (being the newly encountered one).

The strange thing about α -Mlabri and β -Mlabri is that there are very considerable differences in the lexicon, in fact to the extent that communication is rather impeded on the first encounter. There are numerous doublets of synonymous words, of which one is used in α -Mlabri, the other in β -Mlabri. The differences encompass all parts of the lexicon.

Nouns are well represented (though not much more than other word classes) among the doublets, e.g.

| | α -Mlabri | β -Mlabri |
|---------|------------------|-----------------|
| woman | lɲguh | mwlh |
| wife | mjɛ: | hmaj |
| water | wɛ:k | ɣrɛ:k |
| meat | ci:n (loanword) | thac |
| pig | cəbut | chi:ŋ |
| mouse | hnɛl | hwɛ:k |
| neck | ŋlwʔŋlɛʔ | kukɔʔ |
| skin | goguh | naŋ (loanword) |
| blanket | pol | kncaj |
| brain | glɛ:ʔ ʔɔ:n | dɔm |

(α literally: soft head)

There are numerous differences in verbs (and in particles) as well, e.g.

| | α -Mlabri | β -Mlabri |
|--------------|------------------|-----------------|
| to speak | tɔŋ | glɛʔ |
| to sit/stay | hɲuh | jɛ:m |
| to bathe | thale:w | ʔwm |
| to throw | dor | kwm |
| to sing | malam | ʔɛh grnap |
| to run | ra:p | mujthɔj |
| to run fast | jɛlɔj ra:p | mujthɔj jare:w |
| (to be) soft | ʔɔ:n (loanword) | biɛt |
| to know | mɛc | bɛ:n |
| to drink | wɛ:k | ɣrɛ:k |

(cf. 'water' above)

or doublets of near synonyms, of which one lexical item seems to cover the whole semantic range in either α - or β -Mlabri:

| | α -Mlabri | β -Mlabri |
|------------------|------------------|-----------------|
| to eat meat | boŋ | boŋ |
| to eat rice etc. | ʔx:ʔ | boŋ |
| to like | mak | mak |
| to love | mak | hlah |
| to take/catch | toc | toc |
| to fetch/bring | ʔek/ʔec | toc |
| to come/come out | leh | leh |
| to come/approach | leh | pruk |

sometimes one item is shared with (old) Thai:

to be good ʔdi: (or theh) theh

Still, most of the vocabulary seems to be shared, and with (extensive) lexical adjustments it is possible to communicate as if it is indeed one language. With most phrases I have elicited for β -Mlabri it would take some substitutions for them to be possible in α -Mlabri (cf. β -Mlabri /joc non ni ɣe:ŋ/ 'there are no chicken at [our] home', which would be /chrkɛŋ hla:k ni ɣe:ŋ/ in α -Mlabri), but it is easy to elicit (or make) phrases that are apparently equally understandable in α - and β -Mlabri.

A few may suffice here:

ʔot ɣe:ŋ prem 'my house is delapidated'
 ʔot ʔe:w be:c (α more often: be:t) 'my child cries'
 ʔoh pɣ:ʔ ju:k 'I have some rice'
 mlaʔbri:ʔ toc ɣo:ʔ ʔak cuak ʔe:ʔ 'the Mlabri
 takes the digging stick and digs for taro'
 mla:ʔ ʔak ʔjak loŋ bri:ʔ 'the man goes to shit
 in the foresæ
 ki:ʔ ʔathap pmpo: 'it is full moon'

Ferlus' small vocabulary of "kămluaʔ" from 1964 is on the whole very similar to β -Mlabri (distinctly more so than to α -Mlabri).

Let us return finally to Bernatzik. It turns out that β -Mlabri is on many (but not all) points closer to his vocabulary of fifty years ago than is the more well-known α -Mlabri. Some of the words that are current in "Yumbri" and in β -Mlabri but not in α -Mlabri are known to elderly speakers of α -Mlabri as more or less obsolete words. Occasionally, it is the other way round: there are some few words that are shared by "Yumbri" and α -Mlabri but seem not used or even unknown in β -Mlabri, ex. the word for 'fur': Bernatzik pōl, α -Mlabri /pol/, for which β -Mlabri uses a quite different word: /knɛɛj/, or the word for 'water' (or 'to drink'): Bernatzik wō, α -Mlabri /wɣ:k/ as against β -Mlabri /ɣrɔ:k/, Ferlus /ɣrɔk/, /ɣrək/. However, such cases are few in comparison with the cases in which it is β -Mlabri

(and Ferlus' kǎmlua?) that sides with Bernatzik. These correspondences are sometimes quite tricky because of the occurrence of "u" and "o" (with or without diacritics) for Mlabri unrounded vowels. Examples with Mlabri /w/ are:

| | Bernatzik | β-Mlabri |
|----------|-----------|----------|
| heavy | dyŭm | ɟum |
| to throw | kŭm | kum |
| tomorrow | gryú | ɟrw |

We have not retrieved these etyma in α-Mlabri.

In some cases it is so that β-Mlabri agrees better with Bernatzik on the semantic content of a word although the word exists also in α-Mlabri. When I got access to the β-variety of Mlabri it turned out that /thrɛ:ŋ/ is here the cover term for 'teeth' in general (α-Mlabri /ɟɛp/ being not used at all), which agrees with Bernatzik's translation, whereas the word means specifically 'lower teeth' in α-Mlabri, as said earlier.

2.3 Conclusion about "Yumbri" and "Mrabri"/Mlabri.

We have seen that the various sources: Bernatzik, Kraissri, Ferlus, Trier, Egerod and Rischel, and finally Rischel alone (for β-Mlabri) are mutually related by a combination of similarities. There is firstly external evidence such as overlapping use of informants (Kraissri * Trier * Egerod/Rischel for α-Mlabri). Then there is internal evidence, viz. (i) phonetic similarity (Kraissri' Thai-letter version * Egerod/Rischel for α-Mlabri; further Ferlus * Rischel for β-Mlabri) and (ii) lexical similarity (Kraissri * Egerod/Rischel for α-Mlabri; further Bernatzik * Ferlus * Rischel for β-Mlabri which in part coincides lexically with α-Mlabri as documented in Rischel and Egerod 1987). - The degree of lexical similarity between the data of Kraissri and Egerod/Rischel and the data in Trier's analysis of ritual texts is less transparent.

There is some evidence from the lexicon for a gross bipartition into two dialects or dialect groups, viz. α-Mlabri comprising the linguistic data of Kraissri, Trier, and Egerod & Rischel, and β-Mlabri comprising the linguistic data of Bernatzik, Ferlus, and Rischel. However, as said above, this is not a clear-cut bipartition (since Bernatzik sometimes agrees better with α-Mlabri). It should be taken into consideration that we are talking about very small subgroups of a migrating people living in a rather restricted area of northern Thailand and adjacent Laos, and that the various data represent

a time span of fifty years. The conspicuous lexical differences could well have developed in a rather short time. We may assume quite local usages, and these may have at least two different sources:

(1) recent loanwords replacing old Mlabri words: β -Mlabri has the Thai word /naŋ/ instead of α -Mlabri /goguh/ 'skin'); in α -Mlabri the Thai word /nɔ:n/ is replacing the archaic /ʔɛm/ = β -Mlabri /ʔem/ 'to sleep', etc.

(2) metaphorical expressions replacing lexemes, exx.: /joc/ means 'chicken' in β -Mlabri; in α -Mlabri this is now used only about wild fowl, the domesticated chickens (of neighbouring tribes) being called /ɔrkeŋ/, which also means 'wing'.

This leaves a considerable residue of lexical differences such as α -Mlabri /tʌŋ/ vs. β -Mlabri /glaʔ/ 'to speak' or α -Mlabri /wɛ:k/ vs. β -Mlabri /ɔrʌ:k/ 'water; to drink', or α -Mlabri /ŋlʔwŋlɛʔ/ vs. β -Mlabri /kukɔʔ/ 'neck' (/ŋlwʔŋlɛʔ/ is known but is considered rather awkward in β -Mlabri usage). It takes more research to determine how many of these discrepancies are due to loan from other languages. Some, at least, undoubtedly reflect former pairs of synonyms or near synonyms such that for each of these one member survives in α -Mlabri and the other in β -Mlabri (with or without modifications of their "original" semantic ranges).

The strange thing about α -Mlabri and β -Mlabri is that a very great part of the lexicon is shared, after all, and that they have similar syntax and phonology, in fact so similar that I have produced sentences which β -Mlabri speakers could accept and respond to (in β -Mlabri) by using my knowledge of α -Mlabri and attempting to avoid the words I knew to be current only in α -Mlabri. - In a segmental phonological transcription of the shared lexicon one hardly needs to make a consistent distinction between two dialects (let alone two languages): the majority of the entries would have the very same phonological shape. This is true of the types of Mlabri that are currently spoken; it remains a postulate, of course, that Bernatzik's strange notation should be construed to reflect largely the same type of pronunciation (with allowance for differences of detail such as exist also in modern Mlabri) rather than a more aberrant dialect. The major arguments for the former alternative are that Bernatzik's transcription can be shown to be underdifferentiating in a way which is explicable from his background, and that the inexplicable notational discrepancies between Bernatzik and the other sources are random and unsystematic so as to suggest that they are due to imperfect percep-

tion of the phonetic values in individual words, which certainly would not be surprising given the circumstances under which his fieldwork was made.

In any case, allowing for rapid lexical divergence as a consequence of the way the Mlabri are split up in small groups, it seems warranted to refer to all sources mentioned here as specimens of the same language and even the same main dialect (with subdialects such as α - and β -Mlabri and "Yumbri"). This clashes with Smalley's (1963) definition of Bernatzik's Yumbri and Kraissri's Mrabri as two different languages which he judged to be hardly mutually intelligible. There is a real crux here.

The first question is whether "Yumbri" is clearly closer to (Kraissri's "Mrabri" and) our Mlabri than to either Tin or Kammu, the two other languages that make up the Khmuic branch of Mon-Khmer. Let us look at a couple of lexemes for which all the languages involved have obvious cognates and see how phonetically similar they are. Later in this paper numerous Mlabri-Tin cognates will be cited to show how these languages are closely related with regard to part of their lexicon but still clearly distinct from each other in phonetic form. Mlabri is much closer to Tin than to Kammu; the following two examples may give a hint as to the degree of mutual similarity (Tin = Mal dialect from my own field notes, Kammu = Southern Kammu cited from Svantesson 1983);

| | "Yumbri" | Mlabri | Tin | Kammu |
|------|----------|---------|-------|-------|
| rain | mɛ̃ | mɛ:ʔ | mi:ʌʔ | kmaʔ |
| foot | ätʏŋ | ʔat ɣ:ŋ | cɣŋ | ʃɣaŋ |

(Mlabri /ʔat/ is a prenominal particle).

These examples are typical in showing that when either Tin or Kammu disagrees with Mlabri the Mlabri form (from the data of Egerod and Rischel) is the one that is closest to Bernatzik's data.

This closeness of "Yumbri" and Mlabri with regard to phonemes/letters has a counterpart in lexicon. With the data now available (which are far more extensive and more accurate than those available to Smalley in 1963) it is apparent that there is much greater similarity between Bernatzik's data and our Mlabri data than between either of these and Tin or Kammu (Lindell 1974, Svantesson 1983).

There are thus good reasons for the assumption (to which we have adhered all along) that "Yumbri" is a kind of Mlabri, and that Bernatzik's notation can be interpreted so as to be a (very imperfect)

rendering of a Mlabri dialect, whereas it is impossible for me to imagine a way of construing Bernatzik's forms to be either Kammu or Tin.

How, then, could it be that Smalley arrived at a quite different conclusion? In his paper he first emphasizes that ultimately, "questions of linguistic relationship must be solved by painstaking comparative analysis" in which one looks for regularities of correspondence, but then says that "in light of our limitations here we will have to do something more provisional, less fully diagnostic," viz. a comparison based on "basic vocabulary" (Smalley 1963, p.190). Out of the longer list of words compiled by Swadesh he found sixty-six in Kraisri's lists and fifty-eight in Bernatzik. These were compared with each other and with data from Kammu (Khmu?), from three dialects of Tin, and from the more distant Mon-Khmer languages Khamed, Mon, and Lawa, the data being all provided by Kraisri.

The result was that the Mrabri data shared 25 basic words (out of the possible 66) with Kammu (Khmu?), and 23, 22, and 21 words, respectively with the three dialects of Tin, but only 20 (out of 58 possible) with Yumbri. The total number of shared words in the lists were 35 for Mrabri and Kammu, 45 for Mrabri and Tin, and 41 for Mrabri and Yumbri. (The figures for Mlabri compared with the more distant languages were all considerably lower.) "According to these figures", says Smalley, "the Mlabri are linguistically just as closely related to the Khmu? and the Tin as to the Yumbri", even though he mentions the possibility of faults in Bernatzik's data.

A real scrutiny of this line of argument would require a comparison of Smalley's sets of shared words with a new set worked out on the basis of our present understanding of Yumbri and Mrabri, since the use of lexicostatistics is crucially dependent on the concept of "shared word", which in turn is crucially dependent on the philological analysis of each set of data. With data as idiosyncratic as that of Bernatzik we would hardly now arrive at the very same figures as Smalley did.

I shall, however, argue along a different line. In my view, the glaring discrepancy between Smalley's lexicostatistic findings and our rather successful identification of the majority of Yumbri words as some kind of Mlabri, is a genuine and important fact. It shows that there is something wrong with the use of lexicostatistics, especially when it is applied to so small sets of data gathered by non-professional field workers. Not only is there the

possibility of all kinds of errors but more importantly, there is reason not to expect such small sets of data to be representative enough for a comparison of the kind that Smalley made. The most important aspect of the Yumbri-Mrabri comparison, however, is that it has now turned out that these represent a rather extreme case of lexical split between dialects that seem closely related in other respects. Lexicostatistics, it seems, was not designed to cope with this kind of phenomenon, at least not if the intention was to arrive at linguistic classifications which were congruent with classifications based on the "painstaking" comparative method (i.e., looking for regularities in cognate words), also cf. Huffman 1976.

The question remains what one shall call such varieties as "Yumbri" and present-day α -Mlabri (under which I subsume also Kraisri's Mrabri) and β -Mlabri. Ethnically, the speakers are all a kind of mlaʔbri:ʔ 'forest-dwellers'; those who speak β -Mlabri call themselves mlaʔbri:ʔ just like the α -Mlabri do (Bernatzik's term "Yumbri" may have to do with the expression /jɤ:m bri:ʔ/ '[who] live in the forest' used by the β -Mlabri). Because of the lack of agreement between phonological and lexical evidence, however, the mutual linguistic classification of the present varieties of Mlabri and of "Yumbri" becomes a paradox.

3. The relationship between Mlabri and Tin.*

As said earlier, Mlabri is generally classified as Khmuic, but it rests on shaky evidence.

As for Khmuic in itself, the pairing together of Kammu and Tin as a separate branch seems to be generally accepted, but in fact is only now that extensive, reliable data on both the northern and southern dialects of Kammu are becoming available, and for Tin the first major source is from 1978. The comparative study of this branch of Mon-Khmer thus has not proceeded very far, and the inclusion of Mlabri in the study of Khmuic will not only serve the purpose of placing this language *per se* but may also contribute to the understanding of the linguistic development of Khmuic as a whole. Mlabri being clearly much closer to Tin than to Kammu in terms of phonological correspondences, I have looked at the genetic relationship between these two languages in some detail.

My (very limited) first-hand knowledge of Tin stems from a field trip to three settlements in

* This section just summarizes Rischel (forthc. b).

Northern Thailand in the beginning of 1988. As for published information on Tin the main source (and indeed the only relevant source) is the authoritative work of David M. Filbeck (e.g. 1976, 1978, 1987). In my fieldwork I first worked out my transcriptions independently of Filbeck's phonemicization of Mal and Prai because I wanted to establish my raw-data on the same kind of basis as our Mlabri data in order to make comparison as realistic as possible. I did not make definitive phonemic analyses of the Tin dialects during my brief fieldwork sessions but stuck to a semi-phonemic (broad phonetic) transcription, which I think is a reasonable starting point for genetic comparison (provided that the transcription is consistent and provided that it does not miss phonemic contrasts, an inherent danger in this kind of work). Afterwards, of course, I have consulted Filbeck's writings quite extensively both with regard to the tricky question of dialect differentiation within Tin (see below) and with respect to details of the phonemic analysis. By and large, however, I have stuck to my own semi-phonemic transcription since it does not seem to be inconsistent with Filbeck's phonemicizations of the various dialects.

This approach may seem a rather roundabout one, but it gave the advantage of having a first-hand impression of the phonetics of Tin, and I have found that Filbeck's published data and my own data supplemented each other in a felicitous way. Filbeck's historical study of Tin (1978), which is the important source, presents only a rather limited number of lexemes. His inventory, which of course was carefully selected out of the total lexicon for his internal comparison of the Tin dialects, does not by far contain all the Mlabri-Tin cognates which I needed, so my own field sessions have served a purpose in this sense as well. On the other hand, my fieldwork was limited to three villages (representing two very different dialects plus one transitional dialect). It did not include a village whose dialect Filbeck has found to be particularly conservative on crucial points (his "Mal A", which he finds to be quite close in phonology to Proto-Tin). In comparing Mlabri and Tin, data from this last-mentioned dialect must certainly be taken into consideration.

According to Filbeck (1987) Tin is in fact just a common denominator for two (clusters of) dialects, which are more properly designated as Mal and Prai, respectively. I was introduced to a Mal village (Ta Noi) in Amphoe Pua and a Prai village (Nam Phi) in Amphoe Thung Chang, both in Nan Province, i.e., the same part of Thailand where the Mlabri live,

and I worked with Tin informants in both places. By a happy coincidence I also got access to speakers of a third type of dialect: that of the village Chuun (or Cuul, in Thai pronunciation /cuun/), which Filbeck has classified as Prai with a heavy Mal superstratum. Accidentally, I had as one of my informants a very fluent speaker of Mal as well as the Chuun-dialect: a young lady who was married in Ta Noi but born in adjacent Chuun. This access to three different types of Tin was absolutely essential (though not sufficient, see above) for the comparison with Mlabri.

I attempted to collect roughly the same lexical material from the three dialects. Some of this data consisted of names of body parts and other words that are frequently used in daily life, but in addition I looked specifically for words shared with Mlabri and in particular words that would be interesting in a comparative perspective. My own point of departure when beginning to take down such wordforms was a list of Tin words which Dr. Theraphan L. Thongkum had made during a casual encounter with a Tin informant (who clearly presented something of a dialect mixture). Her list was of enormous help in getting me started and in giving hints as to the general relationship with Mlabri.

In the sections below my Tin data is used to the extent that it is relevant to show the beautiful regularity of the correspondences between Tin and Mlabri. Since much of the data is irrelevant to the comparison with Mlabri I do not reproduce my word lists as such in this paper. Those who attach much importance to the lexicostatistic aspect of linguistic comparison may perhaps find the picture as presented here rather skewed. I must emphasize that a great part of the lexicon is not shared by Mlabri and Tin. However, there is considerable lexical divergence even within Tin (Mal vs. Prai). As shown above the same is true within Mlabri, so it is no wonder that the relationship between Tin and Mlabri must be of a very complex nature.

The findings from my genetic comparison between Mlabri and Tin will be summarized rather briefly here; a much more detailed account is given in my forthcoming paper in Acta Orientalia (Rischel, forthcoming b).

3.1. Regular phonological correspondences.

In transcribing Tin I have used largely the same typographical conventions as for Mlabri. The vowels which Filbeck writes as *i, ə* (our 1987 *ɪ, ɘ*) are here rendered as *u, ɤ*. Then there are some interesting

finals: an unrounded mid/back offglide (in the diphthong [αɨ̯], rendered by Filbeck as ɨ̯ but here written ɨ̯ (it is similar to the conservative Danish ɥ); a palatal glide (Filbeck's y) which is here written j, and a palatal glide with devoicing at the end, which I write jh (Filbeck's yh).

My impressionistic length marks may not always correspond to length (written with double vowel letters) in Filbeck's notation. Filbeck (1978) states that length is phonemic in both Tin Mal and Prai; his data shows some dialectal variation, however. - This paper only deals with segmental correspondences; quantity is not really accounted for as the analysis is controversial. On the whole, word prosody is ignored in the comparisons.

Since the final part of the syllable in Mon-Khmer words is generally more stable than the initial and medial parts it seems natural to begin with finals, then to proceed to initials, and to take the vowels last.

(i) The final stops and nasals.

Mlabri /-p, -t, -c, -k/ correspond to Tin /-p, -t, -c, -k/: Mlabri /kɫɰp/ 'cover of container', /mat/ 'eye', /ʔec/ 'to take', /ɨak/ 'to go' = Tin /khɰp/ (Prai /ŋkhɰp/ with a reflex of a prefix?), /mat/, /ʔec/, /cak/.

The Mlabri nasals /-m, -n, -ŋ/ correspond to Tin /-m, -n, -ŋ/: Mlabri /lam/ 'tree', /po:n/ 'five', /peŋ/ 'to shoot', /boŋ/ 'to eat' = Tin /lam/, /phon/, /phɰŋ/, /pɔŋ/ (the correspondence Mlabri /-n/ - Tin /-n/ is weakly attested in my material but uncontroversial).

It should be added it is only in Tin Prai/Chun that there are palatal counterparts to Mlabri /-c, -ŋ/: Mal has /-t, -n/, which is a secondary development.

(ii) Voiced nonnasal continuants in final position.

In Mlabri there is a contrast between two glides and two "liquids": /-w, -j, -r, -l/ in final position. Tin has the same inventory in some subdialects (of Mal as well as of Prai), but according to Filbeck (1976) etymological /r/ has been replaced, or is in the process of being replaced, by a glide or by zero in most dialects of Mal as well as Prai.

The three dialects I have considered for the purpose of this small study, behave differently with respect to final */r/. Quite generally, the

Prai dialect of Nam Phi' has preserved /-r/ in the speech of informants above a certain age. In the Chuun dialect and in the Mal dialect of Ta Noi final */r/ has become a glide after vowels other than /u(:)/, viz. /-w/ in Chuun and unrounded mid /-ɨ/ in Ta Noi. An example is Mlabri /tar/ 'rope' vs. Nam Phi' /thar/ (or /thal/), Chuun /thaw/, and Ta Noi /thaɨ/. This pattern is solidly established, cf. also Mlabri /par/ 'to fly' = /phar/, /phaw/, /phaɨ/, respectively, in the three Tin dialects.

To the extent that words having final plain /l/ in Mlabri have cognates in Tin, these have either /l/ like Mlabri or zero. Mlabri /-l/ corresponds to Tin /-l/ in e.g. Mlabri /wal/ 'to return' = Prai /wal/.

As for final glides there is a straightforward correspondence between Mlabri /-j, -w/ and Tin /-j, -w/, cf. Mlabri /mɔ:j/ 'one', /rwa:j/ 'tiger', /(m)bra:w/ 'coconut' = Tin /mɔ:j/, /wa:j/, Tin Chuun /pja:w/ (with */-r-/ > /-j-/; the Tin Prai of Nam Phi' has /pha:w/ which reflects a change of prevocalic /r/ > /h/ in Northern Thai).

(iii) Voiceless nonnasal continuants in final position.

Tin has a very small inventory of final voiceless continuants. Mal (Ta Noi) has a distinction between /-h/ and a palatal glide which ends in voicelessness and aspiration, i.e. /-jh/, cf. /mah/ 'you' vs. /mphajh/ 'to flick something away', and this is true of the Chuun dialect as well, whereas pure Tin Prai has only /-h/, the words with Mal /jh/ having a stop (/t/ or /c/) in this dialect.

Mlabri has a a richer inventory comprising three different entities, viz. /-h/, /-ɕ/, and a more or less voiceless (but never strident) lateral /-lh/. If we take the richer inventory of Mlabri as our point of departure the following apparently regular correspondences emerge:

To Mlabri /-h/ corresponds Tin /-h/: Mlabri /msh/ 'you' = Tin /mah/.

To Mlabri /-ɕ/ corresponds Tin /-jh/: Mlabri /lɔ:ɕ/ 'to steal' = Tin Mal/Chuun /lɔ:jh/ (it may be, however, that this is a loanword in Mlabri). The reflex /-t/ in Prai /lɔ:t/ shows the secondary development continuant > stop.

To Mlabri /-lh/ corresponds the same Tin entity /-jh/: Mlabri /po:lh/ 'barking deer' = Mal/Chuun /phɔ(:)jh/. Again, Prai has got /-t/: /phɔ:t/.

It should be noted that Mlabri distinguishes four voiceless dentals and palatals in final position: two stops /-t, -c/ and two continuants /-lh, -ç/. These reduce to three in some dialects of Tin and to two in others: Chuun has /-jh/ for both /-ç/ and /-lh/ but /-t/ and /-c/ corresponding to Mlabri /-t/ and /-c/, respectively. The Mal of Ta Noi has reduced to two items: /-jh/ corresponding to both /-ç/ and /-lh/, and /-t/ corresponding to /-t/ and /-c/ in Mlabri. Finally the Tin Prai of Nam Phi' has reduced to two items in a different way: it has /-t/ corresponding to both /-ç/, /-lh/ and /-t/, but it has /-c/ corresponding to /-c/ in Mlabri. The only way to get order out of chaos is to take the pattern in Mlabri to be the one underlying Proto-Tin, and to assume that there was already in Proto-Tin a merger of the two entities I have represented in Mlabri as respectively /-lh/ and /-ç/, whereas the other mergers must be later developments (Chuun showing a strange pattern of interference between Mal and Prai, as one should expect from Filbeck's characterization of it).

(iv) Checked vs. open syllable.

The relationship between Mlabri and Tin on this point is not at all straightforward. To Mlabri /-ʔ/ corresponds Tin /-ʔ/ in some cases but open syllable with a long vowel in others, cf. on the one hand Mlabri /mɛ:ʔ/ 'rain' = Tin /miʔ/, Mlabri /boʔ/ 'breast' = Tin /poʔ/, on the other hand Mlabri /blu:ʔ/ 'thigh' = Tin /blu:/, Mlabri /ti:ʔ/ = Tin /thi:/. For a common proto-language we would have to reconstruct three different syllable terminations: one giving Mlabri /-ʔ/ or /-:ʔ/ and Tin /-ʔ/, another giving Mlabri /-:ʔ/ but Tin /-:/, and a third giving Mlabri and Tin /-:/.

I shall not go into this intricate matter here.

(v) Initial stops.

In Mlabri as well as in Tin, initial stops (and nasals) show a contrast between four points of articulation: labial, dental, palatal, and velar (plus predictable laryngeal /ʔ-/). Except for some complications with the palatal versus velar points of articulation there is a trivial one-to-one relationship between Mlabri and Tin:

Labial corresponds to labial: Mlabri /boʔ/ 'breast' = Tin /poʔ/, dental to dental: Mlabri /ti:ʔ/ 'hand' = Tin /thi:/, palatal to palatal: Mlabri /ɣ:ŋ/ 'foot' = Tin Mal /cɣŋ/, and velar to velar: Mlabri /ka:ʔ/ 'fish' = Tin /kha:/ (there seem to be only minor discrepancies with respect to place of articulation).

As to manner of articulation it is so that Mlabri has a four-way contrast between aspirated and unaspirated voiceless stops (the types /ph-, p-/), and plain and preglottalized voiced stops (the types /b-, ʔb-/).

Tin (Mal as well as Prai) has a well-established contrast between the two voiceless series: aspirated and unaspirated (the types /ph-, p-/). Both series also occur with prenasalization (the types /mph-, mp-/); the unaspirated type /mp/ has more or less voicing of the stop (even fully voiced [mb]).

Tin moreover has voiced initial stops occurring in loanwords from Thai, ex. /bɛ:k/ 'to carry on the shoulder' (in Mal and Chuun; Prai has /tɔ:r/). In such loanwords Mlabri has glottalized voiced stops, e.g. Mlabri /ʔbɛ:k/ 'to carry on the shoulder', which suggests old borrowing (there is plain /b-/ in Modern Thai /bɛk/ but */ʔm-/ > */ʔb-/ is reconstructed for Proto-Thai).

Apart from the clusters with initial nasal (which have a special status anyway since they often clearly correspond to presyllables in Mlabri) the system in Tin differs from that of Mlabri in lacking the contrast between preglottalized and plain voiced stops, whereas the pattern of voiceless stops is conspicuously similar in the two languages. This is deceptive, however. As will have been apparent already from the examples given earlier the series of aspirated and unaspirated voiceless stops in Mlabri and Tin do not line up etymologically; on the contrary there is a pervasive sound shift involved with the result that aspirated voiceless stops in Tin correspond to unaspirated voiceless stops in Mlabri, and unaspirated voiceless stops in Tin correspond to plain voiced stops in Mlabri:

| | | |
|-----------|---|------------|
| Mlabri: | | Tin: |
| p- (etc.) | = | ph- (etc.) |
| b- (etc.) | = | p- (etc.) |

The upper type of correspondence may be illustrated by Mlabri /pɛ:ʔ/ 'three' = Tin /phɛʔ/, the lower by Mlabri /boŋ/ 'to eat' = Tin /pɔŋ/. This is irrespective of point of articulation; it is similar with velars, for example: Mlabri /kwr/ 'thunder' = Tin /khwr/, Mlabri /gɛ:ŋ/ = Tin Mal /kiʌŋ/.

If we compare with other Mon-Khmer languages it is immediately clear that Mlabri is closer to Proto-Mon-Khmer, i.e., Tin has undergone a sound shift changing unaspirated to aspirated, and voiced to voiceless initial stops (a shift in this direction, especially with regard to the change voiced > voiceless, is found in many Mon-Khmer languages).

As for the prenasalized and more or less voiced (unaspirated) initial stops in Tin (Filbeck's /mp-/, etc.) these sometimes occur in roots shared by Mlabri, and then as plain voiced stops. However, there is generally the additional complication that the word contains a pre-syllable in the form found in Mlabri, cf. Mlabri /ɕ(w)mbɛp/ 'lip' = Tin Mal /mpɛ:p/, or Mlabri /rwɛŋga:p/ 'mouth' = Mal /ŋka:p/. A similar correspondence between pre-syllable in Mlabri and homorganic nasal in Tin is observed with the consonants that have become aspirates in Tin: Mlabri /thapu:l/ 'belly', /ɕŋkɛ:r/ 'nail' = Tin /mphul/ (or /phu(:)l/), /ŋkɛr/. Mlabri thus supports Filbeck's assumption that the complexes of nasal plus stop in Tin may have complex origins.

It is particularly interesting that Tin has a causative homorganic nasal: /mpɔ:l/ 'to kill' (vs. /pɔ:l/ 'to die'), /ŋkoh/ 'to hit (with the knuckles)'. Filbeck (1978, p. 29) suggests that "Mon-Khmer */p-/ 'causative prefix' became the nasal */m-/ in Tin" with later place assimilation. Here Mlabri gives direct proof of the origin: /pabul/ 'to kill' (vs. /bul/ 'to die'), /pagoh/ 'to cause to break'; this shows that there still was a /p/ at the time when Tin and Mlabri began to move apart with respect to the phonology of such words.

This still leaves the aspirated voiceless stops and /ɕ/ (= /ch/) in Mlabri unaccounted for. They do not fit into the overall pattern, and in fact there is a conspicuous absence of old cognates in Tin to the (not numerous) monosyllables with these initials in Mlabri (some are recent loans).

(vi) Other initial consonants.

As for nasals and oral continuants there is on the whole good agreement between Mlabri and Tin in cognate words, with the important exception of the glottal manner features. The correspondences being otherwise trivial, I shall here concentrate on that very point.

Mlabri has a distinction between voiceless and voiced nasals: /hm-/ vs. /m-/, etc., cf. /hmuk/ 'tattoo' vs. /muɔj/ 'fat', and likewise between voiceless and voiced oral continuants: /hl-/ vs. /l-/, /hw-/ vs. /w-/, etc. In words shared with Thai the voiceless initials correspond beautifully to traditional Thai spellings (e.g. /hmuk/ 'tattoo' and /hleɔk/ 'iron' = Central Thai /mùk/, /lèk/ still spelled with "hm-, hl-").

In the Tin dialects I have studied the voiceless nasals and oral continuants /hm-, hl-/ etc. all have voiced counterparts, so that there is a

merger here, cf. Mlabri /hɛʔ/ 'new' = Tin /mɛʔ/ vs. Mlabri /mɛ:ʔ/ 'rain' = Tin /miʌʔ/. However, according to Filbeck (1978) there still is a contrast between initials such as /hn-/ and /hl-/ (or /lh-/) and /n-/ , /l-/ in some Mal dialects. Anyway, Mlabri agrees with the scenario that must be posited for Proto-Tin.

Mlabri moreover has a distinction between (rare) preglottalized and plain voiced glides in initial position: /ʔw-, ʔj-/ vs. /w-, j-/. Tin has only a plain series. There is some (but in my material scanty) evidence of a merger involving a process preglottalized > nonglottalized, cf. Tin /jak/ (combined with /mpuat/ 'bee' to mean 'wax', literally 'dung of bees') = Mlabri /ʔjak/ 'dung, shit'.

(vii) Vowels.

Mlabri and Tin agree on a distinction between front unrounded, back unrounded, and back rounded vowels. Mlabri has four degrees of aperture, cf. the unrounded back series of contrasting vowels: /w ɣ ʌ a/, whereas Tin has only three steps which may be represented as /w ɣ a/ (I somewhat arbitrarily use the symbol "ʌ" for the second part of the diphthongs /iʌ uʌ/ in Tin; Filbeck writes /ia ua/, as he wants to minimize the number of phonemic symbols and thus has only one choice, viz. to assign the second part to the /a/ phoneme).

As for the correspondences within cognates it is so that Mlabri and Tin mostly agree on the vowel features front-back and rounded-unrounded. There is not the same degree of regularity with regard to degree of aperture: Mlabri and Tin sometimes agree on this point, but in other cases they differ though (as far as I can see) never by more than one step in aperture. Finally it is quite often the case that Tin has a diphthong where Mlabri has a nonhigh monophthong.

There is to some extent a many-to-many relationship between Mlabri and Tin vowels. Thus, Mlabri /ɛ/ = Tin /iʌ, ɛ, a/, whereas Tin /a/ = Mlabri /ɛ, a, ʌ/; Mlabri /w/ = Tin /w, ɣ/, and Mlabri /ʌ/ = Tin /ɣ, a/, whereas Tin /ɣ/ = Mlabri /w, ɣ, ʌ/.

These various types of relationships may be briefly illustrated by the following examples:

Front series: Mlabri /ti:ʔ/ 'hand', /leh/ 'to come', /hɛʔ/ 'new', /bɛ:k/ 'bear' = Tin /thi:/ ('lower arm plus hand'), /leh/, /mɛʔ/, /piʌk/;

Unrounded back series: Mlabri /kwr/ 'thunder', /bwɪ/ 'to die', /glɪ:ʔ/ 'head', /mɔc/ 'to see', /pɔr/ 'to fly', /ɟak/ 'is going to' = Tin /khur/, /pɔl/, /k(ɪ)wʔ/, /mɔc/ (or /mɔt/), /phar/ (Ta Noi /phaɪ/, Chuun /phaw/), /cak/;

Rounded back series: Mlabri /mu:k/ 'to smell', /boʔ/ 'breast', /boŋ/ 'to eat', /bɔh/ 'ashes', /lɔ:ɣ/ 'to steal' = Tin /muk/, /poʔ/, /pɔŋ/, /puh/, /lɔ(:)jh/ (Nam Phi' /lɔ:t/).

This data testifies to the close genetic connection between Mlabri and Tin as regards the shared vocabulary, but the mutual relationship is such that we cannot consider the vowels of one of these languages to be the ones that occurred (in the etyma in question) in the proto-stage of the other language. It is not surprising in a Mon-Khmer context to find that it is the vowels in particular that present a complex and opaque picture.

3.2. Conclusion concerning Mlabri and Tin.

We have seen that even on the basis of limited data it is possible to establish a rather well-defined network of phonological correspondences between Mlabri and Tin (there is in most cases much more evidence for the regularities I have dealt with in this paper than is apparent from the presentation). Now, what does that tell us?

The phonetic comparison with Tin does not in itself give an answer to the question whether Mlabri is simply an offspring from a Khmuic ancestral language, or whether it has a different origin. On the former assumption the evidence certainly suggests that Mlabri and Tin are very closely related and together form one branch (the other being Kammu). On the latter assumption the Khmuic appearance of numerous words in Mlabri means that a substantial part of its Mon-Khmer lexicon must be borrowings from Khmuic, and that they stem from a time after the separation of Kammu and Tin.

A likely source of the old layer of shared words which have been the object of this paper, is Pre-Tin, a stage which Filbeck (1978) posits as preceding Proto-Tin (the common ancestor of Mal and Prai). However, if Mlabri reflects Pre-Tin it has consequences for the way in which this language should be reconstructed. Thus, if Mlabri has proof value (with regard to the relevant words) this means that Pre-Tin must be construed to represent a stage where this shift voiced > voiceless had not yet taken place (unlike Proto-

Tin). Comparative evidence - including old loans from Thai into Mlabri/Tin - tells us that Tin has undergone a complex Lautverschiebung:

| | | |
|----|---|----|
| p | > | ph |
| b | > | p |
| ?b | > | b |

whereas Mlabri altogether remains on the stage prior to this complex change. My suggestion, then, is that "Pre-Tin" may well be an adequate label for the left column in the Lautverschiebung chart, whereas Proto-Tin (as well as Modern Tin) is represented by the right column. - When confronted with Filbeck's conception of Pre-Tin this only requires minor modifications of the scenario, the proto-stage common to Tin and to the relevant lexical stratum in Mlabri being on a few points closer to Proto-Khmuic than Filbeck's Pre-Tin.

As for the finals Mlabri can likewise be shown to exhibit conservative features when compared to Modern Tin (see above concerning Mlabri /-lh/ and /-s/ which have the same reflex in Tin). By and large, the consonantism in its entirety supports the notion of (a pre-stage of?) Pre-Tin as a common source. - The vowel developments are more tricky but still consistent with the idea if we assume a complex underlying vowel system which was somewhat different from both that of Modern Tin and that of Mlabri.

Altogether, Mlabri is extremely conservative in its phonology compared to most neighbouring languages, with its richness of manner distinctions in consonants and its corresponding lack of tonal contrasts.

Thus it seems that in many ways Mlabri holds the key to an ultimate understanding of the phonology of Pre-Tin and of early Khmuic in general. Mlabri is also of relevance to Thai studies because of its preservation of old sound values in loanwords. These old Thai words in Mlabri may in part stem from its association with Tin in a proto-stage.

It would be tempting to postulate that Mlabri is a kind of Tin, viz. a fossilized offspring of a proto-language immediately preceding Filbeck's Pre-Tin. Above, I have mentioned that Pre-Tin could be conceived so as to accommodate Mlabri. If Mlabri is indeed an offspring we should rather call its ancestor Proto-Tinic. "Tinic" would then be a sub-branch of Khmuic comprising Tin (Mal + Prai) and Mlabri. This is not far-fetched, but it leaves us rather at a loss as regards the differences between Tin and Mlabri that we find along

with the conspicuous lexical similarities.

One would perhaps expect these differences to be explicable in terms of influence from Miao (Hmong) since the Mlabri are now most closely associated with the Miao, but that association is less than a hundred years old, and Miao does not seem to have exerted much lexical influence on Mlabri. Mlabri is now much influenced (in phraseology and lexicon) by Northern Thai or Lao, but that does not account for its special features either (Mlabri has recent loanwords from Tin as well).

We must consider an alternative explanation of the words exhibiting regular old correspondences between (Proto-)Tin and Mlabri, viz. that these were not "originally" part of the core vocabulary of Mlabri but only represent an early superstratum from Tin.

At present too little is known about the extent to which Mlabri shares its lexicon with Tin, but off-hand the superstratum hypothesis seems quite attractive.

What kind of language Mlabri may have been prior to such an exposition to Khmuic - if that is what has happened - is so far unknown.

Acknowledgements: I am grateful to the Carlsberg Foundation for supporting our Mlabri project and to the National Research Council of Thailand and the Tribal Welfare and Development Centre in Nan for generous assistance.

References:

- Bernatzik, Hugo 1938/41: Die Geister der gelben Blätter (München 1938; Leipzig 1941)
- Bernatzik, Hugo A. 1947: Akha und Meau
- Egerod, Søren & Jørgen Rischel 1987: "A Mlabri-English vocabulary", Acta Orientalia 48, p. 35-88
- Ferlus, Michel 1964: [unpublished word list of a Mlabri dialect taken down in Laos]
- Ferlus, Michel 1974: "Les langues du groupe austroasiatique-nord", ASEMI V,1
- Filbeck, David L. 1976: "On */r/ in T'in", Austroasiatic Studies I, p. 265-283

- Filbeck, David L. 1978: T'in: A Historical Study (= Pacific Linguistics Series B. No. 49, Canberra: Australian National University)
- Filbeck, David L. 1987: "New ethnic names for the Tin of Nan Province", Journal of the Siam Society 75, p. 129-138
- Huffman, Franklin E. 1976: "The relevance of lexicostatistics to Mon-Khmer languages", Austroasiatic Studies I, p. 539-574
- Kraisri Nimmanhaeminda 1963: "The Mrabri language", Journal of the Siam Society LI,2, p. 179-184 with appendices
- Lindell, Kristina 1974: "A vocabulary of the Yuan dialect of the Kammu language", Acta Orientalia 36, p. 191-207
- Rischel, Jørgen 1982: "Fieldwork on the Mlabri language: a preliminary sketch of its phonetics", ARIPUC 16, p.247-255
- Rischel, Jørgen, forthcoming a: "Fifty years of research on the Mlabri language: a re-appraisal of old and recent fieldwork data", to appear in Acta Orientalia 50
- Rischel, Jørgen, forthcoming b: "Can the Khmuic component in Mlabri ("Phi Tong Luang") be identified as Old T'in?", to appear in Acta Orientalia 50
- Rischel, Jørgen & Søren Egerod 1987: "'Yumbri" (Phi Tong Luang) and Mlabri", Acta Orientalia 48, p. 19-33
- Smalley, William A. 1963: "Notes on Kraisri's and Bernatzik's word lists", Journal of the Siam Society 51,2, p. 189-201
- Svantesson, Jan-Olof 1983: Kammu Phonology and Morphology (Travaux de l'Institut de Linguistique de Lund XVIII)
- Thomas, David and Robert Headley: "More on Mon-Khmer subgroupings", Lingua 25, p. 398-418
- Thongkum, Theraphan L. 1983: [Unpublished word list made with a Tin informant in Nan]
- Trier, Jesper 1986: "The Mlabri people of northern Thailand: social organization and supernatural beliefs", Contributions to South-east Asian Ethnography 5, p.3-41