Mlabri and Mon-Khmer

Tracing the history of a
hunter-gatherer language

By Jørgen Rischel

DET KONGELIGE DANSKE VIDENSKABERNES SELSKAB
udgiver følgende publikationsrækker:

THE ROYAL DANISH ACADEMY OF SCIENCES AND LETTERS
issues the following series of publications:

	Authorized Abbreviations
Historisk-filosofiske Meddelelser, 8°	Hist.Fil.Medd.Dan.Vid.Selsk.
	(printed area 175 x 104 mm, 2700 units)
Historisk-filosofiske Skrifter, 4°	Hist.Filos.Skr.Dan.Vid.Selsk.
(History, Philosophy, Philology,	(printed area 2 columns,
Archaeology, Art History)	each 199 x 77 mm, 2100 units)
Matematisk-fysiske Meddelelser, 8°	Mat.Fys.Medd.Dan.Vid.Selsk.
(Mathematics, Physics,	(printed area 180 x 126 mm, 3360 units)
Chemistry, Astronomy, Geology)	
Biologiske Skrifter, 4°	Biol.Skr.Dan.Vid.Selsk.
(Botany, Zoology, Palaeontology,	(printed area 2 columns,
General Biology)	each 199 x 77 mm, 2100 units)
Oversigt, Annual Report, 8°	Overs.Dan.Vid.Selsk.

General guidelines

The Academy invites original papers that contribute significantly to research carried on in Denmark. Foreign contributions are accepted from temporary residents in Denmark, participants in a joint project involving Danish researchers, or those in discussion with Danish contributors.

Instructions to authors

Manuscripts from contributors who are not members of the Academy will be refereed by two members of the Academy. Authors of papers accepted for publication will receive galley proofs and page proofs; these should be returned promptly to the editor. Corrections other than of printer's errors will be charged to the author(s) insofar as the costs exceed 15% of the cost of typesetting.

Authors receive a total of 50 free copies. An order form, quoting a special price for additional copies, will accompany the page proofs. Authors are invited to provide addresses of up to 20 journals to which review copies could profitably be sent.

Manuscripts can be returned, but only upon request made before publication of the paper. Original photos and artwork are returned upon request.

Mlabri and Mon-Khmer

Tracing the history of a
hunter-gatherer language

By Jørgen Rischel

Abstract

Mlabri and Mon-Khmer. Tracing the History of a Hunter-Gatherer Language.

This monograph deals with a Southeast Asian language, Mlabri, which is spoken by hunter-gatherers. I have earlier pointed out that there is an old linguistic bond between them and a mountain villager population called the Tin. Both Mlabri and Tin are classified as belonging to the large Mon-Khmer language family. Quite recently, a team of biologists have argued that the speakers of Mlabri are of villager origin and have undergone cultural reversion to foraging less than a thousand years ago. That puts the relationship between the Mlabri and Tin languages into focus since the biologists have referred to my linguistic data and pointed to the Tin villager population as a possible source of origin of the hunter-gatherers.

Scrutinizing the lexical evidence for a Mlabri-Tin connection, I claim that it goes back to language stage I call Proto-Tinic, and I attempt to set up sound-laws accounting for the modern reflexes of the old shared vocabulary in Mlabri and Tin. It is pointed out, however, that it is only a minor part of the total lexicon of Mlabri that can be taken care of by the Tinic hypothesis. Another part reveals contacts with other languages in the northern mainland of Southeast Asia, contacts which in part antedate the Tinic connection, but there remains a large lexical residue of so far unknown provenance.

The grammatical structure of Mlabri is presented in a comparative format. It is shown that Mlabri grammar has no particular resemblance to Tin grammar. On the contrary, it exhibits a number of phenomena which are alien to the language branch, Khmuic, to which Tin belongs. Again, the evidence is suggestive of very early contacts with other languages (especially of the Palaungic branch), outside the area in which the hunter-gatherer group has been living for probably many centuries. It is possible that the Mlabri language is a remnant of a separate branch of Northern Mon-Khmer. There are also a few features that are strongly suggestive of influence from Sino-Tibetan languages, whereas the alternative assumption: that Early Mlabri did not belong to the Mon-Khmer family at all, lacks support.

Contents

Preface

This monograph is about a Southeast Asian language, Mlabri, spoken by hunter-gatherers. Its speakers, the Mla Bri, live in the vicinity of mountain villagers speaking languages that are jointly referred to as Tin (Lua'). It is well-known from other parts of the World that hunter-gatherer groups tend to live in a kind of remote symbiosis with sedentary, food-producing groups and may even at some point in their history replace their own original language with the language of a neighbouring group of higher status. Linguistic evidence suggests that there was some kind of symbiosis between the Mla Bri and the Tin many centuries ago, but one can glimpse a language: "Early Mlabri" behind that connection. The presence of the Mla Bri as a hunter-gatherer group is in itself enigmatic, and for researchers they occupy a special status among hunter-gatherer groups world-wide because of intriguing biological evidence, so their linguistic prehistory is of interest far beyond the mere classification of the language.

The Mla Bri of our time number less than 400 individuals in total, most in Thailand and a few in Laos. Those living in Thailand are now being assimilated to modern society. Their language obviously belongs to the endangered languages of the World.

Some years ago I started a project aiming specifically at clarifying the relationship between the Mla Bri and the Tin from a linguistic and a general cultural perspective. It was supposed to run for another couple of years, and it would be premature to draw definite conclusions about culture this soon. My health is, however, declining rapidly, so I had to discontinue my fieldwork and publish my documentation of the linguistic connection with Tin in the form it has now. I include some very provisional observations on the affinities between Mlabri and more distant languages of the Mon-Khmer family and beyond that family, as a further background for my tentative conclusions about Early Mlabri.

Unfortunately, I have had limited access to recent literature on Mon-Khmer during the period of chronical illness in which this monograph was being written. Without doubt, better possibilities of exchanging data with colleagues and visiting libraries, such as the renowned David Thomas Library in Bangkok, would have improved the quality of my excursions outside the few languages on which I have concentrated in my own work: Mlabri, Tin and Khmu. While acknowledging the short-

comings of this monograph I hope that my completion of it will never-theless be considered worth while.

I wish to express my gratitude to colleagues and friends working on one or another language of the Mla Bri area, or on comparative Mon-Khmer, all of whom have been helpful to me in a variety of ways: Gérard Diffloth, Michel Ferlus, Damrong Tayanin, David Filbeck, Arthur Holmer, David Jordan, Thongpeth Kingsada, Kristina Lindell, Allen Long, Eugene and Mary Long, Håkan Lundström, Ilya Peiros, Suwilai Premsrirat, Frank Proschan, Jan-Olof Svantesson, and Thera-phan Luang Thongkum. I owe much to my late colleague Søren Egerod, with whom Dr. Theraphan and I collaborated on Mlabri in the 1980es. While acknowledging all their help and generosity I must emphasize, however, that my comparative-historical work on Mlabri and Tin is my own responsibility.

Søren Egerod belonged to a remarkable generation of now deceased scholars, who took over from Henri Maspéro and Kurt Wulff and cre-ated the foundation for present-day work in Southeast Asian linguistics. They included, among others, Paul Benedict, William Gedney, André G. Haudricourt, Mary W. Haas, Eugénie R. Henderson, R. B. Jones, Fang-Kuei Li, Stan Starosta, and David Thomas, all of whom I have had the privilege to meet in person decades ago, or more recently. It was a stimulus for me to talk to and listen to each of them.

Over many years I received research permissions from the National Research Council of Thailand. More recently, I was a visiting scholar at two universities in Thailand: Naresuan and Mahidol, and that affiliation granted me permission to continue my fieldwork. The Institute for Research on Culture of the Lao PDR was instrumental in helping me to do fieldwork in Laos in 1999-2001.

My Southeast Asian studies over many years would have been im-possible without travel grants from the Carlsberg Foundation. I greatly appreciate the patience with which the Foundation has continued to support me up to this date. I also gratefully acknowledge the support I received several years ago from the Danish national institution Rådet for Ulandsforskning, and from the Swedish-Danish Einar Hansens Fond.

This monograph is dedicated to my wife, Anna-Grethe, without whose help the project would never have been completed.

Birkerød, 7 May 2007   Jørgen Rischel
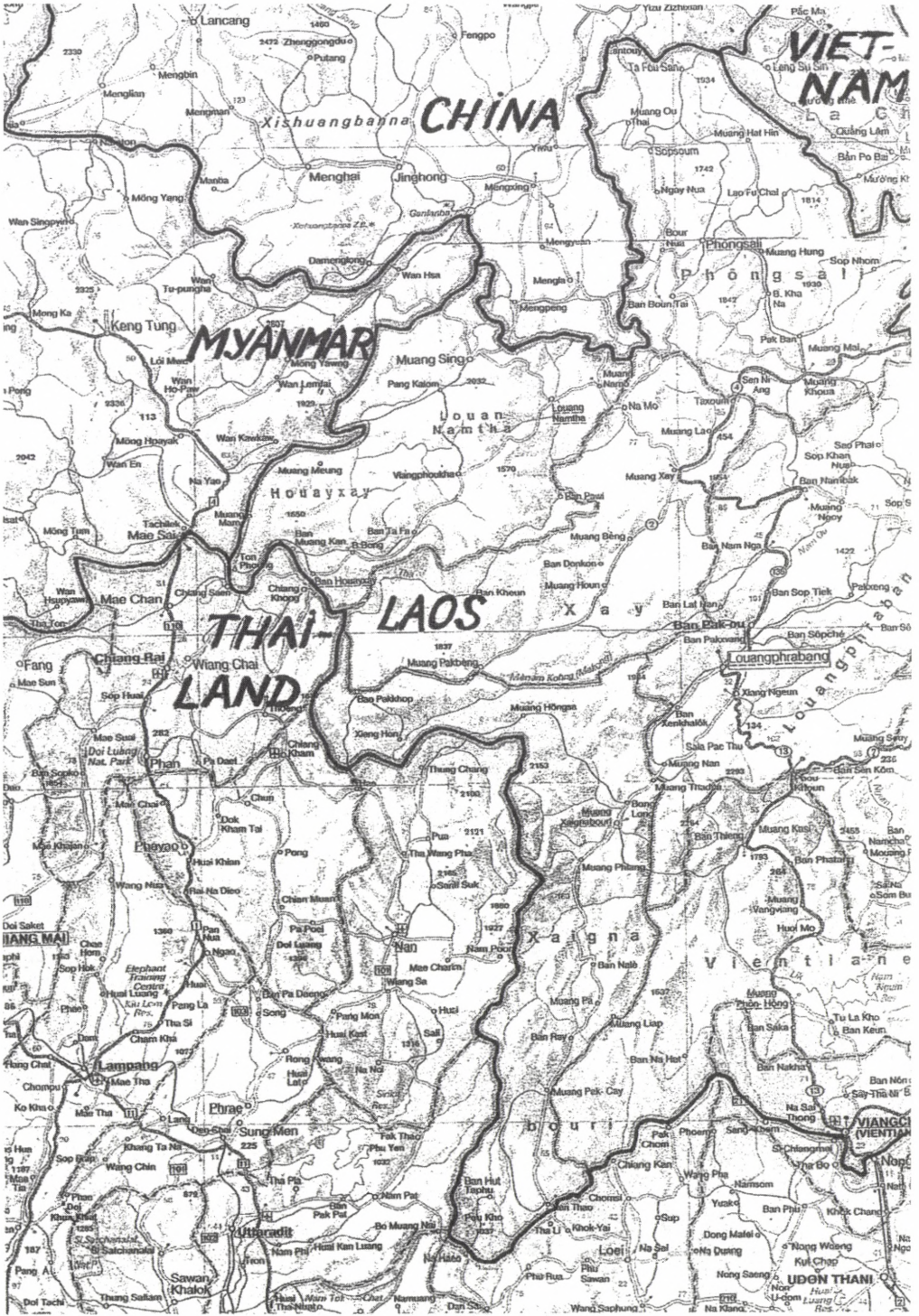
Map 1: The West-Khmuic Area.

Map 2: The Mlabri-Tin Area.

PART 1:

THE MLABRI ISSUE

Preamble

The origin of the tiny hunter-gatherer group Mla Bri in Southeast Asia is enigmatic. They live in an environment which is otherwise occupied by villagers (in the mountains and in the lowlands). The origin of their language, Mlabri, is no less enigmatic. In various papers over the years I have pointed out that a certain part of its old vocabulary is shared with a neighbouring cluster of languages, Tin Mal and Tin Prai, whereas the majority of the vocabulary has a different and mostly still unknown origin. Several features of Mlabri grammar, however, look more like relics of a language from a different branch of the Mon-Khmer language family than the Khmuic branch to which Tin indisputably belongs. A few grammatical phenomena in Mlabri even point to ties outside Mon-Khmer.

In a recent study published in *Public Library of Science, Biology 2005* a team of biologists have suggested, on the basis of mitochondrial and other genetic evidence, that the Mla Bri have a villager origin. Looking for possible ancestors they quote my published linguistic work as well as various pieces of personal communication as support for an old connection to the ethnic group known as Tin, mountain peasants settled in largely the same geographical area as the Mla Bri.

The paper by the biologists attracted considerable attention because it seems to be the first case in which a "cultural reversion" from a food-producing to a food-collecting lifestyle is attested by hard evidence. My linguistic research is not in conflict with that assumption but suggests that the existence of the Mla Bri may not have commenced with the founder event, which is supposed to be less than a thousand years old.

The Mlabri language is well-known by the linguistic community (though perhaps better under the name Phi Tong Luang) because of its strange status as the language of out-of-context hunter-gatherers. Over the last three or four decades virtually all linguists concerned with the old languages of Mainland Southeast Asia have taken for granted that Mlabri is a Khmuic language, because (i) it is spoken in a Khmuic environment and (ii) there are obvious lexical affinities to Khmuic. It has been my conviction for many years that the main reason for this classification as Khmuic is that Mlabri has a significant layer of words from an old stage of Tin, which I call "Tinic". Thus, one can only approach the genetic placement of the language if one has first "peeled off" this Tinic layer.

In this monograph, Part 1 presents the general setting in which the Mlabri language is spoken. It also presents some methodological considerations that are relevant to language comparison and to the Mlabri case in particular. Part 2 presents salient features of Mlabri phonology, morphology and syntax and considers to what extent the various structural properties of the language are compatible with the reconstruction of a common ancestor of Tin and Mlabri: Proto-Tinic. This part also presents an abstract repertory of vowels and consonants to be used in lexical Mlabri-Tin comparisons. Part 3 lists the interesting lexical cognates between Mlabri and Tin and considers in much detail whether it is possible to posit regular phonological correspondence rules between the two languages. Part 4, finally, looks at the history of the Mlabri language and its lexicon from a broader perspective. At the end it sketches a possible scenario for the early history of the Mla Bri people and its language.

The historical scenario

Way back in time, Mainland Southeast Asia was inhabited by a great variety of ethnic groups practising agriculture and – for the most part – speaking Austroasiatic languages. Almost a thousand years ago there was a massive migration of lowlanders from South China speaking Tai languages and practising wet-rice cultivation. In the northern lowlands the Austroasiatic-speaking groups were widely replaced by these Tai-speaking groups, either so that the former inhabitants were assimilated to the newcomers or so that they withdrew into mountaneous regions.

Among the Tai-speaking peoples, the Tai Yai (= Shan) and the Lue (Lü) stayed for the most part north of the Mekong river, whereas the Thais entered the lowlands to the east and south, spreading over what is now Laos and Thailand. Their language, "Ancient Thai", split into dialects such as Lao, Kammueang (= Northern Thai) and Siamese (= Central Thai), which acquired the status of separate languages by developing separate scripts.

Two of the Austroasiatic-speaking peoples were influential both culturally and politically: the Mons and the Khmers. The culture of the Tai-speaking newcomers was massively influenced by Mon and Khmer culture, this influence mediating the spread of Buddhism and of literacy among the Tai peoples. Especially in Thai proper (Siamese), there are numerous loanwords of Khmer origin. A third Austroasiatic group, the

Vietnamese, came to dominate the eastern side of Mainland Southeast Asia all the way from north to south; their language was strongly influenced by Tai and by Chinese so that Vietnamese now bears little resemblance to Austroasiatic languages in its phonological and grammatical characteristics.

The Austroasiatic languages spoken in Mainland Southeast Asia are generally referred to collectively as Mon-Khmer languages because of the two culturally dominant languages Mon and Khmer (the only ones of which there are old, written records). There are in total three main branches of Austroasiatic: (I) Mon-Khmer, (II) Nicobarese, spoken on the Nicobar Islands, and (III) the Munda languages spoken on the Indian subcontinent.

Mon-Khmer proper is a very large language family, which split into separate branches long ago. Most of the numerous Mon-Khmer languages are now spoken by fairly small populations in mountanous parts of Southeast Asia, and several of these "small languages", e.g. in northern Laos, are still known only from short word lists.

There are two autochthonous languages or language groups of Southeast Asia whose linguistic affiliation is unclarified. One is Andamanese, spoken by small groups that used to practise or still practise a hunter-gatherer lifestyle on the Andaman Islands north of the Nicobar Islands. They have probably been inhabiting those islands for tens of thousands of years, being a relic of an early immigration of modern humans to Southeast Asia.

The other group is the *Mla Bri* or Mlabri (in this monograph I render the ethnonym as Mla Bri, and their language as Mlabri, so that it is easy to distinguish between ethnic group and language). For centuries they have been living as hunter-gatherers in extensive mountain forests on both sides of the northernmost border between Thailand and Laos. In the 1980es the deforestation on the Thailand side had proceeded so much that the Mla Bri there, who number less than 400 people, had to gradually give up their former food-gathering lifestyle and switch to a sedentary life. In Laos there is still a group of slightly above twenty individuals who live in the forest and migrate withing a certain area there.

The origin of the Mla Bri has for several decades been an issue both archeologically, ethnographically and linguistically. It was speculated in the 1980es that they might be descendants of hunter-gatherers of the so-called Hoabinh culture predating agriculture. Another former suggestion was that they are descendants of prisoners who were released in the forest centuries ago and thus adopted a survival culture. As men-

tioned already in the Preamble above, a recent biological study of a Mla Bri group (which will be referred to further below) has shown that they have DNA-sequences in common with surrounding groups practising agriculture, to such an extent that it suggests a similar origin.

As to present or former hunter-gatherers, there are no other such groups close to the Mla Bri. The closest are the Ruc in the northwestern corner of Vietnam; they speak a language belonging to the same branch of Mon-Khmer as Vietnamese. Otherwise, there are such groups on the Malayan Peninsula; some of those speak Mon-Khmer languages of the so-called Aslian branch, others speak Austronesian (languages related to Malay).

I have not found it possible to characterize the Mla Bri as being unique in terms of their material culture. On the contrary, there are features of similarity in material culture among the various hunter-gatherer groups, such as the so-called wind-screen or lean-to or canopy, which many people associated with the Mla Bri. It is a shelter consisting of an oblique bamboo frame covered with leaves (not really a shelter against the wind but rather against rain and against the dew that collects in the cool nights high up in the mountains). A similar shelter is used in the forests on the Malayan Peninsula. Another tool that seems to be widespread is a small, round landing net for catching fish. I obtained a specimen from a Mla Bri in Laos; it looks just like a similar tool used by a distant ethnic group down south.

The use of primitive tools made out of bamboo and other plant material, such as a forge with air supply from bamboo tubes with pistons, which the Mla Bri use out in the forest, is not even unique to hunter-gatherers. Most are areally widespread tools shared by sedentary groups that practise agriculture but supplement their supply of food by also fouraging in the forest. If there is anything unique about the material culture of the Mla Bri in comparison with other ethnic groups of the area, it is rather the *scarcity* of tools and techniques, even when it comes to the most basic ones. The Mla Bri do, for example, not make pottery, and their main specialty is a highly developed skill in weaving tools out of split bamboo, rattan and various plant fibres.

The question, then, is: what about the affiliations of *the Mlabri language*? For probably thousands of years the area in which the Mla Bri live has been inhabited by sedentary groups speaking languages of the so-called Khmuic branch of (Northern) Mon-Khmer. The western part of the area that is traditionally Khmuic is found in Map 1. By far the largest group is Khmu or Kammu, who number around half a million

people living for the most part i northern Laos but also in adjacent regions of Thailand, China and Vietnam (other renderings of their ethnonym are Khmu' and Kmhmu).

Another fairly large group, called Tin or Lua', live in mountaneous regions of Nan Province on the Thai side of the Thailand-Laos border and in Sayaburi (Sainyabuli, Xaignabouri) Province on the Laos side, an area that overlaps with the area in which the Mla Bri groups live. The Tin area is situated in the lower central part of Map 1 and shown in more detail in Map 2.

Linguistically the Tin branch into Mal (in Thailand) and Prai (both in Thailand and Laos).

Yet another Khmuic language, Phong, is spoken in Phongsali (Phongxaly) Province in northernmost Laos (see Map 1, top). Further languages of the Khmuic branch include Theen and Khabit (perhaps more properly Bit, "Kha" being a pejorative Lao term for tribe or slave) in Laos, Ksingmul, spoken in rather different varieties in Laos and Vietnam, and Iduh (Tayhat) spoken by a tiny population in Vietnam.

Proschan 1996, who has worked on a linguistic survey of the area, gives useful information on these groups; his account differs from the above by classifying Khabit/Bit as Palaungic rather than Khmuic. Palaungic is another large branch of Northern Mon-Khmer; there are lexical affinities between the two branches, e.g. between the Palaungic language Rmeet (Lamet) and Khmu. For comparative work on Khmuic languages, Palaungic is useful particularly because Diffloth (1980) has made a reconstruction of the proto-language of one of its sub-branches, Waic, and furnished his presentation with a Proto-Waic Etymological Lexicon. Some Palaungic languages are spoken fairly closely to the Mlabri area, others further off to the northwest. Another large branch of Northern Mon-Khmer is Katuic, spoken to the southeast of the Mlabri area.

Because of the geographical proximity of the Mla Bri to Khmuic-speaking peoples the expectation would be that they speak a Khmuic language, and the general consensus has been that this is borne out: several words in Mlabri have close cognates in Khmuic languages, or they even occur in a specifically Khmuic form. One phonological feature that defines Khmuic (Gérard Diffloth, personal communication) is the loss of medial *h before a full vowel, i.e. in words such as Mon-Khmer *( )*mha:m* 'blood' (a form *maham* is reconstructed, for example, for Proto-North-Bahnaric, Smith 1972); Mlabri has *mɛ:m* and thus qualifies as Khmuic in this sense.

If the loss of medial *h never occurred outside Khmuic it is of course irrefutable that there is a Khmuic component in Mlabri lexicon, but I have repeatedly (e.g. Rischel 1995: 46) aired the opinion that the genetic placement of Mlabri is an unsettled issue. Its early classification as Khmuic was too hasty and based on too little evidence, mostly in the form of lexical "look-alikes" (words in two or more languages that look sufficiently alike to be suggestive of etymological cognacy).

I am not the only one or the first one to suggest that the genetic affiliations of Mlabri are not as simple as suggested by the current classification as Khmuic. In the Salzner Language Atlas (*Sprachenatlas des Indopazifischen Raumes*, Wiesbaden 1960) "Yumbri", i.e. C-Mlabri, was set up as an Austroasiatic group of its own; Pinnow (1963: 152) speculated that it might belong to the Palaung-Wa languages, i.e. what is now called Palaungic (in Part 4 I shall look at evidence for such a connection). To be true, the only source for knowledge about Mlabri until Kraisri Nimmanhaeminda's A-Mlabri word list of 1963 was Bernatzik's word list of "Yumbri", but in spite of its notational inadequacies that list is with few exceptions a lexically adequate specimen of Mlabri, as I have recently demonstrated (Rischel 2005, *Appendix*).

Quite early in my work on Mlabri I formed the opinion that this language has a complex and controversial relationship to its closest neighbours (a rather detailed survey of the most salient factors was given in Rischel 1995, Chapter 2). The Khmuic words in Mlabri fall into groups of words that are so different in their resemblance to Khmu or to Tin that it does not take much comparative research to see that some are recent loanwords from one of these languages whereas others have a deeper etymological history. Moreover, an astonishingly large proportion of the Mlabri vocabulary has not (yet) been identifiable as Khmuic at all.

Even more striking to me, after working for years on the language, was the gradual revelation that Mlabri syntax does not look Khmuic. In comparative work on Mon-Khmer almost all emphasis is on phonology and word-formation, which is understandable since many of the languages are known only from short word lists. When it comes to languages of controversial affiliation, that limitation of scope is detrimental to the proper appraisal of genetic or other ties among neighbouring languages. I shall return to Mlabri syntax in Part 4 and just mention that this language uses relational particles ("function words", "grammatical words") in ways which I do not recognize in Khmu and Tin.

Already twenty years ago (see Rischel 1989b) I realized that a significant part of the Khmuic words in Mlabri are shared with Tin in a form that resembles "Pre-Tin" as reconstructed by Filbeck (1978). Filbeck reconstructed Proto-Tin by careful comparison of the extant Tin languages and dialects and then tentatively posited Pre-Tin by internal reconstruction backwards from Proto-Tin. An increasing number of words shared by Mlabri and Tin in this sense have been identified over the years along with my work on Tin as well. The numerical size of this particular set of words is almost an order of magnitude larger than the words I have identified as recent loanwords from Khmu or from Tin. Part 2 and Part 3 of the present monograph are mainly concerned with the reconstruction of a "Tinic" (Proto-Tinic) level, which must date back several centuries and probably close to a millennium, as the level at which Mlabri and Tin share the many words in question.

The main problem with the reconstruction of Proto-Tinic is that the vowel correspondences in Mlabri and Tin cognates are intriguingly irregular, suggesting that either Mlabri or Tin was not traded down smoothly from a Tinic ancestral language, e.g. because of interference from non-Tinic. This monograph tries to present the evidence as fully as possible without being able to cope fully with the Tinic paradox.

If one strips the Mlabri lexicon of loanwords and Tinic words, the residue, which is very considerable, shows both Khmuic and non-Khmuic affiliations. Thus a further reason for pursuing the Tinic connection is to arrive at a hopefully clearer picture of the non-Tinic features of Mlabri lexicon and grammar.

Returning to the Tinic issue, it may seem immediately surprising that a hunter-gatherer language would share a fairly old layer of words and also much more recent loanwords with the languages of sedentary food-producers in the region where they live. One expects them to live in isolation over centuries, and indeed, the stereotype about the Mla Bri is a legendary seclusiveness. Across the world, however, it is anything but unusual to encounter hunter-gatherer languages which have some kind of relationship with languages spoken sedentary, food-producing groups in the geographical vicinity of the hunter-gatherers. A trivial result is the occurrence of loanwords in the hunter-gatherer language as a result of barter trade and other contact; we observe that also in the Mlabri case. A more dramatic result is partial or complete language loss: substitution of a more prestigious language for the original language of the hunter-gatherers (a well-known example is the Pygmys in Africa, who now speak Bantu; their original language can no longer be retrieved).

In Asia, however, there are several instances of what seems to be "cultural reversion" to a lower technological level, not linguistic adaptation to a peer group on a higher technological level. Although it is customary to expect a unidirectional progression from a food-collecting level towards more and more advanced levels of food production, one has to envisage the possibility that poor peasants, who already do fouraging in the forest, may find it easier to survive by switching completely to that approach, perhaps because there is too little arable land available for the marginalized ethnic groups. In India this seems to have been the case with a Munda-speaking hunter-gatherer group, whose language still contains vocabulary related to agriculture, and there seem to be instances in the Philippines. The language of groups who have very locally switched to a hunter-gatherer lifestyle is, of course, likely to be related to those of neighbouring peoples who remain on a higher technological level.

The problem with this hypothetical scenario of cultural reversion is that so far there has been no hard proof of "backward" transition of whole ethnic groups to a survival culture. This is where the Mla Bri come in.

There was no real break-through in this field until the recent DNA-investigation. It comprised a number of tribal minorities in North Thailand, including a representative sample of the Mla Bri population in the western part of Nan Province (unfortunately, it did not include either Khmu or Tin people). The research team included an international group of evolutionary biologists with affiliation to the Max-Planck Institute for Evolutionary Anthropology in Leipzig and headed by Dr. Mark Stoneking, a Thai archeologist Surin Pookajorn (who had previously launched the hypothesis about the Mlabri as a very old, autochthonous group, also cf. Pookajorn 1992) and some high-ranking Thai officials. That resulted in the above-mentioned publication in *Public Library of Science*, which is specifically about the Mla Bri (Oota et al. 2005).

The biologists studied both the mitochondrie-DNA, which is inherited, across generations, from the maternal side, and DNA that is associated with the Y-chromosome and thus inherited, in each generation, from the paternal side. The mitochondrie-DNA is known to change only very slowly over time, as a result of random mutations. It was so invariant across the Mla Bri population that the researchers concluded that they are the descendants of one or at most two women, and that the origin of the tribe probably does not date back longer than some 500 to

800 years. Otherwise, one would statistically expect more variation due to mutations over time than observed in the data. The genetic material associated with the male Y-chromosome also showed conspicuously little variation and suggested that the tribe descends from one to at most four male ancestors.

The Mla Bri were accordingly singled out as *the* ethnic group, among all populations studied in this fashion so far, that showed the least genetic variation. That does not necessarily mean that the tribe arose by a *founder event*, as suggested by the authors of the above-mentioned paper. A main reason for preferring that explanation is the sharing of genetic material with other groups of the area who are not hunter-gatherers, a finding which can indeed be construed to suggest that the Mla Bri were not hunter-gatherers a thousand years ago.

Were it not for the just-mentioned piece of evidence one might prefer another possible explanation of the genetic uniformity of the Mla Bri, namely a *bottleneck* situation. The scenario, then, would be a cata-clysmic event in which the tribe had been decimated and just a couple of individuals survived, after which the tribe gradually recovered as an ethnic group, i.e. underwent a gradual increase in the number of indi-viduals, generation by generation, over 500-800 years. The near-eradi-cation of a formerly more numerous tribe would be consistent with some origin myths found among the Mla Bri themselves. They suggest that in old days the Mla Bri were numerous and all-dominant in the area; one story tells that most of them were killed by witchcraft by the Northern Thais. Such myths might, however, refer to fairly recent events – if they have any historical foundation at all.

The authors of the biology paper (Oota et al. 2005) searched, of course, for a cultural and linguistic connection to some other ethnic group of the area. Such a connection is furnished by my demonstration of ancient vocabulary shared between the Mla Bri and the Tin. I was able to furnish the authors with a cultural piece of evidence as well: a Tin story according to which two children or youngsters, a boy and a girl were expelled from a Tin village long ago (according to one version because they were siblings and were assumed to have committed incest). They survived, the story goes, by floating down the river and fleeing into the deep forest. Thus they became the ancestors of the Mla Bri. The young age of the ancestors at the time they left the village is supposed to explain the "childishness" (according to one version) and the "slow way of speaking" of present-day Mla Bri.

That story is consistent with the biologists' findings, and it may be of

considerable age since I have heard the story both in a Mal-speaking village and in a Prai-speaking village. These two groups separated linguistically centuries ago, and for lexical reasons Mal and Prai are not mutually intelligible, so it does not seem likely that the story spread across the linguistic boundary recently. It is rather a vestige of a very old narrative tradition. On the other hand, I think one should be careful not to consider it a historical source. I have not heard anything like that from the Mla Bri, and among the Tin it is a purely oral tradition. Literary historical records are, to my knowledge, silent about the history of the Tin centuries ago.

Incidentally, most people of the Tin communities have never heard the story about the origin of the Mla Bri. I recorded it from a very old Prai-speaking lady and later came across a likewise very old Mal-speaking lady who happened to know it. It might just as well have vanished from the collective memory of the Tin.

In my view the story is interesting mainly by asserting that the Tin regard the Mla Bri as very distant relatives. The same holds true, in some sense, when it comes to the attitude of the Mla Bri toward the Tin, but it is more remarkable that the Tin would recognize as relatives an ethnic group that is lower on the social ladder than they themselves are. Like the presence of Tinic words in Mlabri, this mutual attitude might be a reminiscence of an old symbiosis between the Mla Bri and the Tin, a hunter-gatherer group and a food-producing group. It might, however, equally well reflect a direct genetic bond between the two groups.

When the biology paper about the Mla Bri was about to be published I strongly advised the authors to leave open the possibility of another lineage beside the Tin lineage (which in itself is still a hypothesis) because the Mlabri language bears witness of a mixed origin. They loyally explicated that in the paper.

All of this is the background for the present monograph, which attempts to probe into the linguistic Mlabri-Tin relationship and to reveal to what extent that relationship accounts for, or does *not* account for, the lexical and grammatical characteristics of Mlabri. At the end of Part 4 I shall make some suggestions as to the possibility that Mlabri has linguistic ties not only beyond Tinic but even beyond Khmuic.

Ethnic groups of the Mlabri–Tin area and their languages

Sedentary ethnic groups

The area in which one encounters Tin communities comprises Nan Province on the Thailand side and Sayaburi Province on the Laos side. The Mla Bri are found not only in these two provinces but also just southwest of Nan Province, i.e. in the adjacent, northeastern part of Phrae Province; their groupings will be surveyed at the end of this section. The whole area is seen on Map 2.

I have previously referred to the area as being geographically "three-dimensional" (Rischel 2003b). It has a considerable horizontal extension, but the distribution of ethnic groups across it is linked to differences in altitude. In this section I shall survey the ethnic groups that are sedentary in the sense that they live in rather permanent settlements.

Central parts of Nan and Phrae Provinces are taken up by lowlands, and the same is true of certain parts of Sayaburi Province. That is where one finds the Tai-speaking lowlanders: the Northern Thais, the Lue (who are more or less assimilated culturally to the Northern Thais), and the Lao. On the Thailand side, there are also several lowland villages with highlanders of different ethnicity who were forced to move away from the war zone between Thailand and Laos in the 1970es and were resettled at a safer distance, mostly fairly close to their former homes up in the mountains. Some of the large Tin settlements are of this kind; there is a large settlement of Prai-speaking people inside the township of Tung Chang northerly in Nan Province, and a large settlement of Mal-speaking people just outside the township of Pua, between Nan City and Tung Chang.

As for Tin villages on the Laos side, many villagers fled across the border into Thailand and stayed for years at a huge refugee camp at Nam Yaaw between Pua and Santisuk (which is east of Nan, cf. Map 2), from where they eventually immigrated to the United States. There is now a large Tin settlement with inhabitants speaking Laotian "Lua'" (Tin) at Santa Rosa in California.

Returning to the topography of the Mlabri-Tin area, there is a long mountain ridge going north-south along the Thailand-Laos border, divided up by deep valleys and with broad foothills on both sides. This

mountaneous area is where the Tin groups have had their villages on the inner hillslopes, often at high altitudes, for probably many centuries. They are essentially alone in occupying this particular area (except for very recent and very local areal development programs attracting other people). Both on the Thai and the Laos side they have probably been living mainly to the north, where the mountain peaks are tallest, from around the northeastern tip of Nan Province (close to Muang Hôngsa on Map 2) and down to an impressive massif east of Pua.

As mentioned already, the Tin communities fall into two main groups: speakers of Prai and speakers of Mal. Prai and Mal are from the perspective of mutual intelligibility different languages rather than just dialect clusters. The Prai-speakers normally refer to themselves as Lua', with an ambiguous ethnonym (the term "Tin" is in itself unfortunate since it is a Thai word meaning 'dialect').

As for Prai-speaking Lua' communities on the Thai side, they still have villages both in the northern and southern parts of that mountain region, with preservation of traditional long-houses in some of the remote villages east of Pua. As for the Tin in Laos, there are also villages with an ancient lifestyle. Cultural information on the mainstream culture of the Laotian Tin can be found in Elliott et al. (1992). They probably all speak Tin of Prai type. That is the type of Tin I recorded during an extreme brief session with a few Tin settlers from Laos at Santa Rosa, California, and it is confirmed by a couple of linguistic specimens in Elliott et al. (1992). Some claim that the Prai-speaking Tin spread from the Laotian side into Thailand not very long ago.

The stronghold of the Mal communities, on the other hand, is an amphitheatrical massif northeast of Pua, the Ban Sakad area, where several villages lie in a semi-circle along the slopes. They are now for the most part connected via a more or less passable road but they were less stationary in former time. One village, Ban Kwet with around 600 inhabitants, is remarkable by having the phonologically most conservative dialect of all, a dialect called "Mal A" by Filbeck (1978), which in his view is virtually identical with Proto-Tin in its overall phonology. The current village headman of Ban Kwet told me that during the last few generations their village was moved up and down along the steep hillslope (there is a constant risk of mudslides during the rainy season in this area) while they kept exploiting largely the same area without encroaching on the areas available to other villages.

The location of Ban Kwet may have been at a different place even before that, for its dialect is in stark contrast to the dialect(s) of the sur-

rounding villages, which are phonologically much more advanced. That is Filbeck's "Mal C". The third main dialect, "Mal B", was spoken to the south of there and is now found in a large lowland settlement. Finally, one village, Cun, at the southern end of the core Tin area, had a dialect which is transitional between Mal and Prai; its inhabitants have now been resettled in the same settlement as speakers of Mal B.

If we look now at the eastern outskirts of the mountain ridge, in Sayaburi Province, there is a great ethnic variety. The landscape changes gradually from the real lowland through a densely forested and more or less sloped area, into the real mountains. There are villages with different kinds of Tai (mainly a western dialect of Lao) spoken, and there are Hmong villages (cf. below), also at fairly low altitudes. Some of the Lao villages have been established as spearheads cutting into the forest; one of them, Ban Navên southwest of Muang Phiang, is situated way inside the forest. New settlements of this kind are no longer permitted; the huge forest still contains a wide variety of animals: elephants, tigers, etc., and it is protected by a government decree.

Switching to the other, western side of Nan Province, there is another mountain ridge (see Map 2) going north-south and then turning off towards the southwest into Uttaradit Province. It has extensions, with valleys in between, in the Doi Luang area of Phayao Province west of Nan Province and, with less altitude, in the district of Rong Kwang in the eastern part of Phrae Province (southwest of Nan Province).

The villagers of this whole mountain ridge are for the most part of Hmong (Meo) ethnicity but there are also several Mien (Yao) villages. Especially the Hmong villages are traditionally at very high altitudes. The Hmong and the Mien migrated into Laos and Thailand roughly a century ago from South China, and many still keep family ties and cultural ties with their kinsmen in China. These ethnic groups have a tradition of slash-and-burn farming and must migrate cyclically over a considerable expanse of land in order to be able to exploit the soil without impoverishing it. They are now no longer permitted to do that (though deforestation of the hillsides by burning still happens on a large scale), so they have had to settle in permanently situated mountain villages. That is why I count them among the sedentary highlanders.

In the central part of the western mountain ridge, just west of Nan City, there are also villagers of Khmu ethnicity, but quite few compared to the other highlander groups. The Khmu are, on the contrary, the dominant ethnic group in all of northern Laos (information on Khmu culture can be found in Tayanin and Vang 1992). They all speak one language

though with minor dialect differences; these differences tie up completely with geographical location and will be surveyed later.

The Mla Bri

The Mla Bri that have been observed by outsiders fall into three distinct subgroups, which for convenience I have labelled "A", "B", and "C".

The A-subgroup comprises almost all extant Mla Bri, more than 300 persons. They have been roaming the forests of the western mountain ridge (e.g. west of Nan City) as long as their own memory and outsiders' records go back. This group was visited by several expeditions over the years, including an international expedition headed by Kraisri Nimmanhaeminda 1962, who worked out the first good word list (published 1963) but unfortunately called the Mla Bri "Mrabri", and a Thai expedition headed by Surin Pookajorn. The report from the latter (published in English 1992) contains a sketch of the Mlabri language by Theraphan Luangthongkum. Other early documentations of their language are by Søren Egerod (Egerod 1982) and myself (Rischel 1987) and by the two of us jointly (Egerod and Rischel 1987). The early word lists suffer from a serious phonological inadequacy by ignoring vowel-length.

The A-subgroup of Mla Bri lived as hunter-gatherers, getting some food and clothes through barter trade with mountain villagers. In recent decades an increasing number have been working as underpaid farmhands for the Hmong, and during the last two decades most of them became settled close to Hmong villages. They are now being gradually integrated into the mainstream society, mainly through the introduction of local schools for them. Over the last few years, an extensive documentation of their transitional culture and their way of conversing about it has been made in connection with a forthcoming documentary. As for the Mlabri Text of Sakamoto (2005) it is stated in the Introduction that the material was recorded in Phrae, which defines the language as A-Mlabri (in its way of rendering the distinctive sounds of the language, however, it is not consistent with the information on A-Mlabri presented in this monograph).

The B-subgroup had for decades been moving forth and back across the Thailand–Laos border east of Pua, being in much contact with Tin villages (some of them spoke fluent Tin). After ending up in the Nam Yaaw refugee camp they gave up their former lifestyle of food-

gathering in the forest and associated with Hmong villages on the Thailand side. One elderly man, whom I met in 1988, insisted that he was born in the Bo Kluea District on the Thailand side. His wife and daughter were, however, encountered in a Tin-village in Sayaburi Province back in 1964 by Ferlus (Ferlus 1974); at that time they claimed that they were speakers of "Kamlua'", i.e. Tin, but Ferlus established that they were Mla Bri (I have since had the younger of the two women as an informant; she spoke both Mlabri and Tin fluently). I met a total of eleven members of the B-subgroup in different Hmong villages on the Thai side and worked intensively with most of them in the interval 1988-2002. The adults claimed that they had had a miserable time during warfare between Thailand and Laos and were now the only surviving members of the subgroup. By 2002 almost all adult speakers were gone; their children mostly spoke broken Mlabri because Hmong had been their vernacular.

There is one family that descends from the B-subgroup in the Tin-settlement in Santa Rosa, California. An elderly lady there (who has a daughter and a granddaughter as well) was fully aware that her mother was a Mla Bri who had married a Tin, but she did not remember anything else but her mother's name *ʔi ɓua* (a Tai-style name, as expected, with Mlabri pronunciation). Later, when doing fieldwork in the Bo Kluea area I got decisive confirmation that her mother was the elder sister of the above-mentioned elderly B-Mlabri man, and that she had indeed "run away with a Tin man", thus bringing shame over the family. Most unfortunately, her daughter remembered none of her maternal language. The incident is nevertheless interesting in confirming how few Mla Bri of this particular subgroup there were already several decades ago. In fact, I have only encountered members of the same extended family though in different villages.

I have documented the B-variety of the Mlabri language in a separate monograph (Rischel 1995).

The C-subgroup was visited by Hugo A. Bernatzik before the Second World War (Bernatzik 1938). His ethnographical account bears the stamp of its time but is of lasting value (Rischel 2005). Bernatzik also collected some word lists, unfortunately in a very imperfect notation (interpreted in Rischel 1989a; the word lists were republished with identification of most words in them as genuine Mlabri, in Rischel 2005). There is a group of some twenty Mla Bri in the forest in the western part of Laos, whom I visited three times over the years 1999–2001. I established that their variety of Mlabri is identical with the one

encountered by Bernatzik more than sixty years earlier at a location on the Thailand side, but in fact not very far away from where the surviving group is now (Rischel 2000).

A quasi-ethnographical account of the almost unknown C-group was given by Chazée (2001), who travelled in the area and must have obtained his demographic information by speaking to people in nearby Lao villages. After working with these Mla Bri while conversing in their own language (they understood my mixture of A- and B-Mlabri perfectly) I have reservations against Chazée's presentation, which seems ill-conceived in its information about individual persons, with misrepresentation of their names, for example.

There seem to have been additional subgroups of ethnic Mla Bri. Firstly, there was at least one group south of the Tin area in the long mountain range between Thailand and Laos. It is a large, uninhabited area visited at intervals by Thai soldiers on patrol, and I have been told repeatedly over the years that the soldiers occasionally observe evasive groups of people with outfit like the Mla Bri. That is the same area in which Hugo Bernatzik met a Mla Bri group speaking MlaC in 1937 and not far from the present location of the small MlaC-speaking group. A member of the MlaC-speaking group in Laos told me that he and two others once, many years ago, were in the same area and met Mlabri-speaking people whom they did not know but whom they could talk to in Mlabri. Because there were soldiers present they had to part company.

Moreover, I have met and photographed a young man in Laos who was captured as a small boy when he lagged behind his parents who tried to flee from soldiers, apparently in approximately the same area but further south. That was some time in the late 1970es, probably. He was raised by a Lao family and understandably, he did not remember any of the language of his childhood, but the only likely identification of his ethnicity is Mla Bri.

At the southern end of the same mountain range there have been Mla Bri in Uttaradit province on the Thailand side until recently. One elderly man, who is still alive and has been a consultant for my data on Mlabri, tells that he migrated from that area into the present location of the MlaA-speakers. He speaks MlaA but with discernible differences in lexical usage (often noticed and commented on by his kinsmen).

North of the area in which the MlaA-speakers have been roaming the forest in the last generations there was still another group. An elderly MlaA-speaker has told me a story which he had from his parents: once,

in the northern outskirts of the area they used to gather food in, they met a couple of Mla Bri who spoke in a different manner but still understandably. After exchanging a few words, however, both parties broke up out of fear and never met since. I believe in that story because my source cited a couple of words which his parents had heard and found amusing. One of them was *mul^h*, which I recorded from his mouth maybe fifty years later, with exactly the pronunciation the word has in MlaBC and with its correct meaning there: 'woman'. This word is unknown in MlaA, which uses a borrowing from Tai: *mǝɲiŋ*.

People in Phayao Province west of Nan Province used to tell that farmers felling trees on high slopes sometimes heard crying children or other sounds from the "spirits of the yellow leaves" up in the mountains; that is maybe half a century ago. I doubt it that those belonged to the same group as the present-day MlaA-speakers.

Finally, both Prai-speaking and Mal-speaking Tin villagers in the mountains between Thailand and Laos have told many stories about their former encounters with the Mla Bri (described with all the traditional characteristics of the Mla Bri, so there is not doubt about the identification). Some of these encounters took place rather far north of the areas in which the known Mla Bri have been living in recent time. Again, one cannot exclude the possibility that they formed a distinct group.

All of this information points unequivocally in one direction: the Mla Bri used to exist over a much larger area, probably with a larger number of distinct groups than the three groups: A, B and C which I have been able to distinguish on linguistic grounds.

Basic information on the Mlabri language

Mlabri is in various respects different from other languages classified as belonging to the Khmuic branch within Mon-Khmer. There is no reason to go into detail with that here because it will be taken up in Part 2 and Part 4. There is so far very little material to illustrate the similarities and differences among the three varieties of Mlabri. A syntactical study, with rich exemplification from all three of them, appeared in Mon-Khmer Studies (Rischel 2006). A dictionary comprising all three varieties of Mlabri, likewise by myself, is in preparation and almost complete. It will hopefully appear in late 2007 or early 2008.

As for the three varieties of Mlabri, they are extremely similar in

terms of phonology, indeed so similar that it makes sense to speak of them as ethnically defined sociolects within one dialect. The sound system is the same all over; but the pronunciation of individual words may vary in terms of vowel length or vowel quality, for example.

Only the sentence or phrase intonation is characteristically different. MlaA-speakers have a high-pitched intonation in vivid speech, with a long declination at the end which rides on a final lengthening that may amount to more than two seconds and gives this kind of Mlabri a unique character (an early account of the Mla Bri describes their speech as whining and unmanly). There is another mode of speaking if one wants to express surprise or relate something important; the pitch is then low, and the voice quality rather breathy. The B-Mlabri speakers have (or had – most are now long since dead) an entirely different intonation: a rather monotonous pitch contour on a rapid succession of words. High pitch at the end is used to signal commands. Finally, the intonation of the C-Mlabri speakers is characterized by a quickly rising-falling pitch on the last few syllables before the pause at the end of a chunk of speech. As in other languages of the area pitch does not in itself cue interrogative status of utterance. Interrogative constructions either contain a pronoun or adverb of interrogative meaning (these vary to some extent across Mlabri) or a final interrogative particle, which is *lɛh* all across Mlabri.

The real difference between the three varieties of Mlabri is in lexicon. Especially B-Mlabri differs from the other varieties and to such an extent that it would hamper communication if the groups were interacting rather than avoiding each other but also the C-Mlabri is considered very aberrant by the A-Mlabri. I have tried to let speakers of A-Mlabri listen to the other varieties. Some of them then hesitate to recognize what they hear as proper Mlabri and then immediately give up attempts to understand what they hear, even suggesting that it may be some kind of Khmu (!). When I first met a speaker of B-Mlabri I tried to make myself understood in my imperfect A-Mlabri, to which he answered back in Northern Thai. Only after some adjustment did I learn to avoid words that were confusing to the addressee when putting questions to B-Mlabri speakers about their language. With C-Mlabri speakers it was very different; on our very first encounter they readily accepted the mixture of A- and B-Mlabri I tried to use then, and they did not even pay attention to the way we were communicating (in this case there was no alternative channel of communication since they spoke preciously little Lao).

The three subgroups of Mla Bri being traditionally distanced from each other by strong mutual animosity and fear, the words they differ on are of course stigmatized. I presume that the strong lexical cleavage one can observe now must have accompanied an ethnic polarization which developed over a not very long period. On the other hand, it may date back a century or more. Looking at earlier sources such as Bernatzik (1938), Ferlus (1964) and Kraisri Nimmanhaeminda (1963) for C-, B-, and A-Mlabri respectively, I find them to be in full agreement with modern differences in usage among the three groups. That is true both with respect to pronunciation (although Bernatzik's notation has been difficult to interpret) and lexicon. It is often the case, for example, that one variety has an old Mlabri word where the other variety has a recent loanword. Thus there are two words for 'meat': a loanword *ci:n* and an old Mon-Khmer word $t^h\lambda c$ (t^h- < *s*-), and two words for 'skin': a loanword *naŋ* and a Mon-Khmer looking word: *goguh*. In both cases the loanword is from Tai.

Another conspicuous polarization, especially in B-Mlabri, is between male and female lexicon. I do not know how old that dichotomy is. In a minority of instances I find the same difference between male and female speech in A- and B-Mlabri, which suggests that it is of quite some age, but in other instances either the male or the female word in one variety of Mlabri is unacceptable (or is not understood at all) in another variety.

There is lexical variation even within one variety of Mlabri spoken by the same gender. As one would expect, that is primarily a matter of differences to do with age. Elderly speakers have a command of several words which are more or less unknown to the younger generations, which suggests that the lexicon is undergoing change. That is true of course when it comes to terminology relating to hunting and gathering but also outside that realm. Even within the same generation, however, speakers differ very much as to their active or passive vocabulary. Some know many words, others fewer words. Mlabri speakers who are aware of their linguistic usage may choose to discard a word that is commonly used by others, claiming that it is not real Mlabri but Khmu, Tin or Northern Thai.

In comparing Mlabri with Tin I started out limiting my scope to words that occur in at least two varieties of Mlabri so as to avoid "noise" from idiosyncratic neologisms. In my further search for cognates I abandoned that principle, however, because some obviously old Mon-Khmer words in Mlabri happen to be preserved in only one

variety. When quoting Mlabri I always state the provenance of the word-form in question, i.e. in which group or groups I have encountered that particular form. If there is no such tag on a word it means that it is found all over Mlabri (similarly with provenance tags on Tin words).

Ethnonyms. The name of the Mla Bri and of their language Mlabri (more elaborately: *çmbɛp mlabriː?* 'Lips of the Mlabri'), occurs in distorted form in international reference works up to this day, although it was given correctly as Mla Bri in a Thai–English source (Jerry W. Gainey and Theraphan L. Thongkum's Language Map of Thailand Handbook) as early as 1977. The most widely used names referring to this ethnic group and its language are Mrabri , Yumbri and particularly Phi Tong Luang, all of which are inappropriate terms.

The term Yumbri goes back to Bernatzik (1938). The Laotian National Census, as summarized by researchers at the Institute of Ethnology in 1985, used it as the name of the small ethnic group in Laos, which I found to be speaking the very same dialect as the people met by Bernatzik (Rischel 2000, 2005: Introduction). I have argued that is the Mlabri phrase *jɤːm briː?* 'live forest', characterizing the Mla Bri by their lifestyle. This was almost dramatically confirmed when I was accompanied by a group of officials to my first encounter with the Mla Bri in Laos. In everybody's presence the Mla Bri confirmed that they are *mla? briː?*, 'forest people', but that they *jɤːm briː?*, 'live in the forest'. Reportedly, this seemingly caused the ethnonym to be changed in favour of Mla Bri.

As for "Mrabri" there is in fact no such form as Mrabri ; it is a distortion of Mlabri (an unfortunate conjecture due to Kraisri Nimmanhaeminda). As for "Phi Tong Luang ", this is a pejorative Thai designation meaning 'Spirits of the Yellow Leaves'; it has been popular among the lowlanders and its use has reinforced their prejudices against the Mla Bri.

Besides Mla Bri, there are two Thai-based names which are acceptable to the tribespeople themselves: one is "Khon Pa", which is a direct translation of *mla? briː?*; the other, which is being used increasingly, is "Yellow-Leaf People" (playing on the traditional Thai designation but avoiding the offensive allusion to ghosts).

Other Mla Bri-like tribes and ethnonyms. There are various tribal people in Laos and Vietnam whose material culture resembles that of

the Mla Bri and who are often referred to as Yellow-Leaf Tribesmen (according to Proschan 1992: 15 "Toong Luong" can even refer to the Lahu, an ethnic group speaking a language outside Mon-Khmer).

Matisoff (1991: 215) in his linguistic survey of endangered languages in Laos makes a distinction between Mla Bri and Kha Tong Luang / Phi Tong Luang . The language of the former is classified as Khmuic, and its speakers are said to total 300 (exactly the same statement is made about Mla Bri in Thailand, so I suppose that the figure is meant to include speakers in both countries). The Kha Tong Luang or Phi Tong Luang , however, are a total of 200 speakers of a language of the Viet-Muong group, according to Matisoff.

To this I wish to add that in Vietnam as well there is evidence of traditional Yellow-Leaf People, e.g. the Ruc people speaking a language of the Muong subgroup.

Finally, the Thai term Kha Hok , i.e. Spear Tribe , is sometimes used by the Tin and even by some Mla Bri to refer to allegedly savage Mla Bri who carry spears and have tattoos; according to these attributes the A-Mla Bri of Thailand are also Kha Hok, although they have recently discontinued the use of spears. The Kha Hok are widely claimed to be dangerous people, but that is in contradiction to the notorious shyness and meekness of the known A-Mla Bri in Thailand. Altogether, it is a very confusing term, like Kha Tong Luang. Proschan (1992: 45-46) cites a number of sources for the use of Kha Hok to refer to a group of Khmu who wear only loincloths rather than pants but he notes (p. 43, footnote) that the term implies the perspective of an outsider. It is hardly an ethnonym used by any group.

It is difficult to define exactly what it means to be a Mla Bri unless one takes language as the criterion. Although it seems intriguing that the different groups of Mla Bri tend to exhibit linguistic polarization, it is significant that they nevertheless speak of their language as being one and the same: *çṃbɛp mlaʔbriːʔ* 'the mouth of the Mla Bri'.

Basic information on Tin

Prai and Mal differ so much lexically that they are not mutually understandable, although the total area in which Tin is spoken, is fairly small (in Thailand one finds villages next to each other with Prai and Mal speakers).

Prai and Mal differ considerably in their lexicon. Also within Prai,

and likewise within Mal, there are distinct dialect differences.

If we look only at sound patterns, however, the different kinds of Tin are fairly similar in overall type, but there are phonological differences between dialects or subdialects. In the present context it may suffice to mention four which relate to genealogical splits of different depths within Tin: (i) *s*- versus *c*-, (ii) -*j*^{h} versus -*t*, (iii) retention of old aspirated or voiceless sonorants such as ^{h}*m*- or change of these into fully voiced sonorants; (iv) the retention or change or loss of *r in various positions in the syllable (for a detailed treatment of Tin dialectology from a historical perspective, see Filbeck 1978, also cf. Rischel 1989b).

Scholars think of Tin as one entity, both for purposes of language classification and when searching for cognates across Austroasiatic, and that makes sense since Prai and Mal are grammatically similar and share some early sound-shifts that together distinguish Tin from other Khmuic languages. Still, for those doing first-hand research on Tin it is two different things to work on Prai or to work on Mal, and in fact Prai and Mal have been approached separately by different scholars or missionaries. For all the reasons I have mentioned there is no such thing as a Tin dictionary, not even in the making.

The picture that emerges from recent study of Tin is that there has been a strange splitting-up within Tin so that much of the most central vocabulary (for example the word for first person singular) differs between Prai and Mal.

A Thai banker, Kraisri Nimmanhaeminda, published some word lists of various dialects of Tin and of Khmu along with a specimen of Mlabri after a journey in the Nan district in 1962. Two Thai linguists, Suwilai Premsrirat and Choltira Satyawadhna, have published textbooks for practical use. The very appearance of such publications is significant by provoking a wider interest in this ethnic minority, but because of their scope they have not in themselves advanced the scholarly study of the language.

Linguistic work on Tin has been dominated by foreign scholars. In the eighties and nineties some graduate students at Thai universities studied the Tin and their language as part of their university requirements. To my knowledge, such work has not resulted in scholarly publications.The basic comparative work on Tin has been done by the American Dr. David Filbeck of the Christian Mission to the Orient and of the Phayap University in Chiengmai (Filbeck 1973), who is an authority on Mal. Another American missionary group including the Rev. David Jordan (working at Ban Phae' Klang) has done valuable but

to my knowledge unpublished work on Prai grammar and lexicon. I myself have made rather extensive fieldwork on Tin phonology comprising most dialects of Mal and Prai in Thailand.

The basic ethnographical studies of Tin life and culture were made by Filbeck (1973) and Dessaint (1981). A study (in Thai) by Ch. Satyawadhna (1987) was reviewed rather critically by Filbeck (1987). There has not otherwise been much study of Tin culture. This has something to do with their cultural invisibility: the Tin look and behave much like the poorest ethnic Northern Thai living in dry-field areas.

Filbeck points to the impressive, amphitheatrical mountain area just east of Pua and Chiang Klang District towns (i.e. northeast of Nan City) as very old Tin area. The Mal people may have been there for several centuries. They do not seem to have any lore about coming from Laos; on the contrary, they associate one of their myths with a specific mountain Phu Huat in that area (Filbeck, personal communication). It may be different with the Prai, who are (or at least used to be) more numerous in western Laos than in Thailand. Most sources suggest that they came into Thailand only in the late nineteenth century (see references in Smalley 1994, p. 230).

Over the years I have visited twelve Mal-speaking or Prai-speaking Tin villages. Some of these I visited very briefly, mainly as a screening to ascertain the type of dialect and to see if they preserved the tradition of expressing numbers between 5 and 10 with reference to fingers on both hands (7 = 3+4, etc.), since I take this to be a token of conservative Tin ethnicity, shared by Prai and Mal speakers alike though now mostly forgotten. In other villages I stayed for several days to collect vocabulary data, to record specimens of Tin speech, or to make observations about language use.

Between Ban Kwet and another exit to the highway (further north) there is a village Ban Norng with a section called Ban Norng Kaw (Old Ban Norng) in which the dialect of elderly people shows interesting features (e.g. initial nasal consonants followed by aspiration: m^h-, etc.). The dialect has -*l* from Proto-Tin *-*r*, according to the data I could retrieve in the short time I was there.

Within Lua' (Prai-type Tin) in Laos there are considerable dialect differences from one village to another, according to Elliott et al.:

The villages are separate from each other and do things in their own ways. Each grows its own rice and takes care of its own people.

In Laos, we don't meet people from other villages very often. That's why each village has its own dialect. Sometimes we can understand people from different villages easily, sometimes we can't. We usually marry within our own village... (loc. cit., p. 83).

Basic information on the Khmu language

Khmu is spoken in a large area comprising the whole, mountainous northern part of Laos (down to the region where the country narrows abruptly) plus adjacent parts of Vietnam, China, the Shan State of Myanmar and Thailand. Around 90 per cent of the speakers live in Laos; in Northern Laos they form the majority of the population. General information on the language can be found in Premsrirat' *Thesaurus* (Premsrirat 2002b: xxix–lxxv).

The Khmu dialects, going east-west, differ mainly in terms of one single phonological feature: the development of voiced initial stops.

The original state of affairs is found far away from the Mla Bri–Tin area. In the southeastern part of the Khmu-speaking area in Laos, the so-called U-dialect has preserved an old distinction between voiceless and voiced initial stops, e.g. *p*- vs. *b*-. The Khmu dialect represented in the word-list by Kingsada and Kosaka (1999), which is spoken in the northermost province of Laos, Phongxaly, is remarkable by also having preserved the old voicing distinction.

In the northwestern part of Laos, however, the voiced stops have been devoiced and a tonal difference has arisen: high tone after old voiceless stops, low tone after formerly voiced stops. This is the type of Khmu that is being intensively studied at the University of Lund in Sweden (e.g. Svantesson 1983; Svantesson, Tayanin and Lindell 1994). Even further to the west, e.g. in a Khmu village Huay Nam Puk west of Nan City in Thailand, one encounters an intermediate stage: the initial voiced stops have been devoiced but the syllable has acquired a breathy voice quality with a more or less clear tendency toward lower tone than after originally voiceless stops.

Finally, there is a dialect area to the northeast, also far away from the Mla Bri–Tin area, in which Khmu has undergone a consonant mutation similar to the one that led to extensive tone splits in Thai, old voiced stops becoming not only voiceless but also aspirated: $*b- > p^h$-. This kind of Khmu is tonal. The entry forms in Premsrirat 2002a are based on a dialect of that area, "Khmu Rook" (dialect "lb").

Lexically, the differences among the dialects of Khmu are modest considering the vast area over which Khmu is spoken.

The loanword issue

It is one of my main tenets that the classification of Mlabri as Khmuic has been based on lexical evidence of little known provenance. In principle it is wise to be suspicious of borrowing whenever a Mlabri word has exactly or almost the same phonology as the corresponding word in a language from which Mlabri would be likely to be influenced, especially if the majority of cognates between the two languages are less similar. If one fails to spot words that are ambiguous as to their provenance the estimate of lexical closeness between two languages is exaggerated, of course. When doing comparative phonological work it is even more crucial since one might jeopardize one's reconstruction of the historical scenario by basing sound-laws on words that are in fact loanwords from a peer language within the same branch of Mon-Khmer.

Taking Khmu for a start, there is nothing unexpected in having loanwords from Khmu at the northern Thailand-Laos border considering that it is traditionally the dominant language in northern Laos. There is a handful of words that look so similar in Khmu and in Mlabri that they might formally be recent loans in Mlabri. On the other hand, languages may independently preserve old shared vocabulary in the same phonological shape if both are conservative .

In this monograph I am going to elaborate a hypothesis about a sister relationship between Mlabri and Tin. Off hand, that is an alternative to postulating a specific, old association between Mlabri and Khmu, and there is a potential conflict between the two assumptions. Evidently, however, one commits a methodological error by jumping to conclusions and branding all words in Mlabri as recent loanwords from Khmu if they are more similar to Khmu than to Tin. That makes the Tinic hypothesis self-fulfilling. It is safer to take the nature of the relationship between Khmu and Mlabri up later, *after* the substantiation of the Tinic hypothesis. That will be done in Part 4.

Another way to approach the loanword issue is to consider what would be likely to be borrowed by a hunter-gatherer group which in recent time has preserved a fully functional language for use in its natural environment and has only used basic vocabulary of external lan-

guages for communication on short and intermittent encounters. If the contact between the two ethnic groups was related to the exchange of goods the most useful borrowings would be terms for necessary utensils. Table 1-1 contains a few such words which are likely to be recent loans from Khmu, no matter whether one assumes that the two ethnic groups had more in common in the remote past, or not.

Table 1-1: Likely loanwords from Khmu

Mla *ǥɔːʔ* (in MlaC also *ǥɔːʔ*) 'hoe'; Khmu *sɔʔ* 'axe' (apparently, two different words for 'hoe' and 'axe' were mixed up in Mlabri)
MlaB *dɛj* 'soft bag'; Khmu *daj* '(soft) bag'
MlaAB *tr̩lɔh* 'pot'; Khmu *tr̩lɔh* 'pot'

There are several other words on which Mlabri agrees perfectly with the form and meaning the etymon has in Khmu. If these words in Mlabri are also in systematic agreement with Tin I have tentatively taken them to be part of the Tinic layer of words listed in Part 3. If they agree specifically with Khmu it may be because they are old shared vocabulary antedating the Tinic era of Mlabri or because they are recent borrowings from a conservative Khmu dialect like the words in Table 1-1.

Table 1-2 lists some such words that are in principle ambiguous. It is limited to words in initial stop because such words could otherwise be diagnostic of borrowing from Khmu, viz. (i) if the stop is voiceless in Mlabri but voiced in conservative Khmu (the word must then be a recent borrowing from a devoicing, tonal Khmu dialect), or (ii) if they agree on an aspirated t^h (in genuine Mlabri words that is a continuation of *s-) or (iii) if the Mlabri and Khmu forms share a diphthong that is otherwise unusual in Mlabri.

So far my observations are negative: whenever I have noticed words that are diagnostic of borrowing it has turned out that Khmu has them from Lao, and thus the words may have belonged to the lingua franca spoken by highlanders when addressing people from other ethnic groups.

Table 1-2
Some other forms shared between Khmu and Mlabri

Mla *cok* 'poke into (a hole)'; Khmu *cok* 'poke; dig out'
Mla *kɔh* 'rub off'; Khmu *kɔh* 'id.'
MlaA *kum* 'grave'; Khmu *kuːm* 'bury'
Mla *duʔ* 'escape'; Khmu *duʔ* 'id.'

It will be noted that these may all be of high age in Mlabri and just agree with Khmu because their structure or semantics did not invite any change.

Next there is the question of borrowing from Tin in recent time. It is well-attested that the Mla Bri were in intermittent contact with the Tin up through the first half of the nineteenth century and even later. As mentioned already, I have had confirmation from several elderly Tin villagers who used to live in the mountains close to the Thailand–Laos border, that in their childhood there were occasional visits to the village by small Mla Bri groups for the purpose of offering wax, meat, and the like from the forest and getting rice and old clothes in return. The Tin have traditionally been both fairly isolated and very poor, and the social distance between the Mlabri and the Tin has been small.

Although the Mla Bri are reported to have used some Tai (Lao or Northern Thai) words when addressing the villagers, such contact, which may have existed for centuries, would be likely to result in Tin loanwords in Mlabri. Such copying of whole lexical items from Tin into Mlabri would, of course, be much more likely to happen on a larger scale if it happened that some Mla Bri stayed in close contact with a Tin village over a prolonged time, the resulting additions to Mlabri lexicon being subsequently adopted by other kinsmen. The loanwords would then reflect the form the word has nowadays in either Mal or Prai.

Unlike the ambiguity in the extent of recent loans from Khmu there is one absolutely decisive criterion for recent borrowing from Tin although it applies only to words beginning in oral stops. Stops *mutated* without exception in Tin (formerly voiceless stops becoming aspirated, e.g. $p > p^h$, and formerly voiced stops becoming voiceless, e.g. $b > p$) thus setting the shapes of words in Tin apart from those in neighbouring languages.

This means that there is no way Mlabri and Tin can agree on the manner of articulation of initial stops except if Mlabri borrowed words from Tin after the time when the stops mutated, or if Tin had borrowed

the words in question e.g. from Khmu. I have found no evidence of any kind that Tin was ever significantly influenced by Khmu, so I take it that the borrowing happened from Tin into Mlabri.

Table 1-3 is limited to such loans from Tin (a couple of these etyma have a doublet in Mlabri which goes way back in time and therefore has a non-mutated initial stop; such data will be taken up in Part 2). It will be noted that the words in Table 1-3 are semantically highly variegated, suggesting that the Mlabri at some point not too far back in time had so extensive encounters with the Tin that it led to bilingualism.

Table 1-3: Mlabri words that can be identified as post-mutational loanwords from Tin

Mla *tuk* 'carry through to the end'; TinPrai *tuk* 'the end of something'

Mla *ce:t* 'touch (food) with the tongue-tip to taste it'; TinPrai *cɛ:t* 'taste food'

MlaAC *keh* 'weave (using split bamboo)'; TinPrai *keh* 'split bamboo into narrow strips (usable e.g. for weaving)'

Mla *ko:c* 'bamboo rat'; Tin *kɔc* 'bamboo rat'

MlaB *koh* 'attack'; TinPrai *koh* 'hit'

MlaC *klu:ʔ* 'CLASSIFIER for pots'; Tin *klu:ʔ* 'head' (Mlabri also has the etymon in an older form and meaning: *glɤ:ʔ* 'head')

MlaAB *kluh* 'drill'; TinMal *kluh* 'pierce'

MlaA *tʰet* 'cut; slash'; TinPrai (causative derivation?) *ⁿtʰet* 'cut (one's hand, or the like)'

Mla *tʰɛh* 'good'; TinPrai *tʰɛh* in *lek tʰɛh* 'iron for ignition' (likely to be the same word, with the basic meaning 'genuine')

MlaAC *tʰo:ŋ* 'bag'; TinPrai *tʰoŋ* 'shoulder bag'

Mla *kʰe:t* 'worry ; be afraid', cf. TinPrai *kʰet* 'quiet'

MlaB *kʰoŋ* 'be impervious ; have good magic'; TinPrai *kʰoŋ* 'id.'

MlaBC *kʰot* 'rest with bent or crossed legs', a contamination of two forms borrowed from Tin, one from Prai, the other possibly from Mal, cf. TinMal *kʰujʰ*, TinPrai *kʰut* 'sit' and TinMal *kʰot*, TinPrai *kʰo:t* 'curl up (in resting position)'

Mla *kʰɔ:t* 'spear'; TinPrai *kʰɔt* 'spear'

Considering the number of loanwords in initial stops it is entirely unlikely that there are no loanwords with initial continuants. It is, however, a priori difficult to distinguish those from inherited words shared

with Tin. For comparative purposes that may not be too grave since the more recent loans beginning in continuants must be a fairly small minority of the total set of words of that structure. As will be shown in Part 3, the number of old cognates with initial stops is of an order of magnitude larger than the number of borrowings have identified and put into Table 1-3. Thus, the old cognates may be expected to greatly outweigh the inherited cognates also when it comes to words with initial continuants.

If undetected recent loanwords nevertheless disturb the regularity of phonological correspondences they do so by showing invariant forms between Mlabri and Tin in cases where one expects differences (in vowels or initial clusters) due to sound shifts. Again, a conspicuous phonological similarity in a few look-alikes should make one consider whether they might indeed be recent loans.

Leaving the question about contacts with the Khmu and the Tin aside now, a third issue is contact with the contemporary peers of the Mla Bri: the Hmong. Since the Mla Bri have for decades been temporary doing work in the dry fields for the Hmong in order to get some basic necessities, one might expect several loanwords from Hmong. The Mlabri language, however, seems practically devoid of such borrowing. I presume that that reflects a social attitude: the Mlabri working for the Hmong feel a huge distance in prestige, the Hmong being always patronizing and often even abusing the Mla Bri. Although several Mla Bri can speak Hmong many others cannot, and the tradition is for the Mla Bri and the Hmong to address each other in the lingua franca, which on the Thailand side is Northern Thai. Mla Bri men traditionally speak that language with a limited vocabulary and without tones but otherwise with good pronunciation, whereas the Hmong traditionally (and often still) speak with a strong Hmong substratum resulting in segmental distortion.

There are some very recent Tai loanwords in Mlabri that bear the stamp of phonological distortion, in contrast with the general tendency of Mlabri to preserve the segmental phonology of loanwords faithfully. Examples are A-Mlabri *lika:* 'wrist watch' (for expected *nalika:*; Tin-Prai has this loanword with preservation of all three syllables) and *tʰəɲuʔ* 'radio', both with Tai phonology except that a whole, initial syllable has been dropped. Since Mlabri has extremely few words of more than two syllables, this might be a simplification happening within Mlabri proper, but such simplification might also suggest indirect borrowing via a highlander language with deviant phonology, possibly in

some cases Hmong or Yao or some Tibeto-Burman language. That is more likely when it comes to changes in place of articulation of finals, which is very unlikely in Mlabri but may have happened via a language tending to drop or substitute such finals in loanwords. An example of that is B-Mlabri *hlɔːp* 'haunt a place', with Tai phonology except that the final should properly have been *-k* (the same etymon occurs in A-Mlabri with the correct final and with causative prefixation: *baʰlɔːk*, clearly borrowed much earlier and therefore via a different source).

The last example above leads over to the final and quantitatively most important loanword issue: borrowing (directly or indirectly) from Tai. All over the northernmost part of Laos and adjacent Thailand the highlander minority languages have borrowed extensively from Lao. This is very much true of Khmu, for example. In languages spoken in adjacent Thailand, Northern Thai similarly plays a significant role as the lending language. In modern time even Central Thai plays a role in the speech of adults because of radio transmissions in that language and in a longer perspective in the speech of children in particular because all children now attend school.

There are many Tai loanwords also in Mlabri and in Tin. In Tin they are probably for the most part rather recent, and they are mostly just "noise" in the context of Tinic comparisons. Tai loanwords in languages with very conservative phonologies are interesting, however, especially if they begin in stops or aspirated continuants. Tai languages all had consonant mutations several centuries ago, and thus the phonetic quality of the initial in loanwords may reveal at what chronological period a particular word was borrowed by the minority language under consideration.

In principle, Tai loanwords in Mlabri may thus tell us something about older and more recent contacts with the outside world. The problem with such dating is that Mlabri may have borrowed most of the Tai words via some other highlander languages. In that case, the quality of the initial may show how early or late the words entered the intermediate language but if that language had no consonant mutations (e.g. if it was a conservative dialect of Khmu) it might preserve the Tai word in a fossilized form for a long time before passing it on to Mlabri, which then continued preserving the pronunciation carefully.

With this *caveat* in mind it is still interesting to look at loanwords with ancient Tai consonant qualities in Mlabri. Table 1-4 gives some examples (in this table the greater or lesser distribution of the words over A-, B- and C-Mlabri is ignored).

Table 1-4
Examples of loanwords with ancient Tai phonology

bɛː 'raft'
bɛːŋ 'expensive'
bɔː senior name tag < 'father'
dop 'fold'
jɔːn 'spoon'
jum 'group'
gɤj 'ever'
gɔːk 'pigpen'
gɔj 'moderately'
gɤːŋ 'middle'
grok 'mortar'
gruɤː 'things; belongings'

^{h}*miʌŋ* 'fermented tea'
^{h}*muː* 'group'
^{h}*num* 'young man'
^{h}*nuːn* 'jack fruit'
^{h}*lek* 'iron'
-^{h}*luːŋ* in *p^{h}a^{h}luːŋ* 'yellow' < 'yellow cloth'

The examples in Table 1-4 all have ancient forms that antedate changes such as **b-* > Northern Thai *p-*, Lao and Central Thai *p^h-*, and changes such as **^{h}m-* > Northern Thai, Lao, and central Thai *m-*.

The conservative dialects of Khmu present a similar picture, but the words preserved are not all the same. The dialect in Phongxaly for which a word list is presented in Kingsada and Kosaka (1999b) has such forms as *bup* 'to meet with', *gon* 'people', *gem* 'salty', of which only the third occurs in Mlabri. Conversely, this dialect of Khmu does not have the old form of the word for 'expensive' but a contemporary Lao form *p^{h}ɛːŋ*. An illustrative list exemplifying Tai loanwords with voiced initial stops in a Southern Khmu or U dialect (*ls*) as against voiceless or voiceless aspirated stops in other dialects (*lk, lb*) is given by Premsrirat (2002a: li), the examples being words such as *gem ~ kem ~ k^{h}em* 'salty'. Before listing these loanwords she makes a very intriguing statement, however:

> Some of the Tai-Lao loan words in Khmu have been phonologically modified to fit the Khmu phonological system of the Khmu dialect they belong to. (...) A loan word with original voiceless stop initials of Tai-Lao words occurs as a voiced initial in *ls* dialect, but as a voiceless initial with voice quality in *lk* and *lb* dialetcs. The stop initial in *lk* dialect is unaspirated whereas in *lb* dialect [it] is aspirated.

I can only construe this to mean that the words already had the present-time Lao pronunciation with voiceless aspirated initials at the time of borrowing. If so, the implication must be that all these words entered Khmu, and consequently also neighbouring languages such as Mlabri, at a fairly recent time. It is, however, a controversial assumption that some dialects mutated initial consonants backwards so as to make them fit the sound system. I shall demonstrate that for one set of consonants in one dialect:

Taking the *ls* dialect (U dialect), it does not just have voiced stops, of course. There is a series of voiceless stops as well but also a series of aspirated voiceless stops according to the excellent vocabulary of Dr. Suwilai (Premsrirat 2002a). Examples of the latter are $t^h\varepsilon$: 'slave' (ibid.: 231), $t^h i$: 'stingy' (ibid.: 229), $k^h i$: 'similar' (ibid.: 91), and $k^h u$:*l* 'hair' (ibid.: 96; incidentally, this word is suggestive of cognacy with *kul-* in A-Mlabri *kulmu:j* 'hair' of which the second part occurs reduplicated in the meaning of 'hair' in B-Mlabri: *mujmuj*). The claim might be that Lao words with aspirates that reflect Ancient Thai voiced stops ("low" consonants) came into the dialect first and were adjusted to the sound system, and that it was only later that the dialect acquired words with aspirates of a different origin (such as Lao words with "high" consonants, which were aspirates already in Ancient Thai), this time tolerating them. That is entirely implausible, however. I know of no instance in which a language or dialect has selected loanwords from a certain source on the basis of their history in the lending language rather than their usefulness.

Digression: If voicing of voiceless initials occurs it is either a general phonetic rule (a sound-shift affecting *all* words of that structure) or a sporadic, irregular and idiosyncratic phenomenon. To take an example of the latter, one of my B-Mlabri speakers consistently pronounced the particle *kan* 'if' as *gan*; such an aberration may be a case of faulty internalization of the pronunciation (he had severely impaired hearing and may have had that since childhood) or it may have arisen as a sandhi-phenomenon.

To conclude: there would be no reason and indeed no way for the *ls* dialect of Khmu to transform the initial stops of loanwords such as the word for 'salty' in order to make them fit the dialect. The alternative assumption: that highlander languages such as conservative Khmu and Mlabri preserve the old Tai consonantism, is overwhelmingly more likely.

The next set of Mlabri words borrowed from Tai are those with initial glottalized stops or glides, as exemplified in Table 1-5. It is an open question what chronological period these forms stem from since glottalization may have vanished much more recently in the lending Tai languages than the mutations of voiced stops and of aspirated continuants.

Table 1-5
Examples of Tai loans with preservation of glottalization

ɓa(ː) 'junior name tag'
ɓaːŋ 'flying lemur'
ɓit 'squeeze'
ɓɔːk 'tell; order to'
ɗaj 'able to; possible'
-ɗam in *pʰaɗam* 'black' < 'black cloth'
ɗiː 'good'
ʔjaː 'medicine; tobacco'
ʔjaːŋ 'roast'
ʔjɯːn 'stand upright'

Next there are words with a voiceless initial stop which reflects an ancient voiced stop. To the extent that such forms occur they must have been initially borrowed from Northern Thai *after* the devoicing (*b- > *p*-, etc.) happened in that language since they would have had aspiration if borrowed from Lao. There is a strange scarcity or almost absence of such words in Mlabri suggesting that until recently the borrowing was from Lao not from Northern Thai.

Finally there are all the loanwords with initials that did not change in the relevant time span in Tai, and which therefore cannot be dated. Some have forms that betray a certain age, for example by the preservation of a liquid after a stop, as in *kleːt* 'scales' (if this was a recent loanword it would have the form *keːt*). Others have undergone sound

changes typical for Northern Thai and Lao but are still old, e.g. *hɔːt* 'arrive' with *h-* < **r-* (this loanword is also of age in Waic according to Diffloth's reconstruction; Diffloth 1980).

There is a final consideration in connection with words that have Tai connections but are of old provenance in Mon-Khmer languages: the nature of the connection is not always straightforward. One type of correspondence is exemplified by *krep* 'pick (berries)', the Mlabri counterpart of a Tai word with the segmental structure *kep*. Another type of correspondence is exemplified by *kwaːc* 'sweep (with a broom: a *kr̩waːc*)', the Mlabri counterpart of a Tai word with a final dental. A third type of correspondence is exemplified by the words for 'grandfather' and 'grandmother', which in many M-K languages are *taʔ, jaʔ*, or the like, and which have open-syllable correspondences in Tai; in Mlabri the forms are *taːʔ* 'grandfather; uncle; old man' and *jaːʔ* 'grandmother; aunt; old woman', and they even occur in compounds with sandhi: *taraːm* 'divorced man', *jaraːm* 'divorced woman'.

Words of these categories are relevant in historical-comparative work. Although the default expectation is that there are Tai words in Mon-Khmer languages in northern Mainland Southeast Asia some such words are on the contrary (or are likely to be) extremely old Mon-Khmer borrowings into Tai.

Linguistic comparison within and beyond Khmuic: methodological considerations

When dealing with a historical scenario involving a complex of more or less well-attested phonological correspondences there is a proper analytical approach: to use the traditional comparative method that was developed so successfully in nineteenth century Indo-European studies.

In the case of Indo-European several branches of that language family are attested from early time in more or less extensive inscriptions or whole literatures. A further advantage is that some of the ancient languages are extremely conservative and thus reflect important features of Proto-Indo-European rather directly.

With the Mon-Khmer language family it is very different. Although there are ancient inscriptions in Mon and Khmer, the time depth is much smaller, and it is not the case that those old attestations of Mon-Khmer have a particularly "archaic" character compared to extant Mon-Khmer languages. More importantly, research on this language family

does not go back very long in time, and several branches of it have been explored very little so far. Several Mon-Khmer languages are known only from short word lists; in recent years additional languages have become accessible that way. It is often an issue in itself just to decide how to group these little-known languages within branches of Mon-Khmer.

A few sub-families such as Waic, North-Bahnaric, Monic, and Katuic have been studied from a comparative perspective to such an extent that the researchers have been able to make more or less definitive reconstructions of the phonologies of their proto-stages (Proto-Waic etc.). There is no such reconstruction available of Proto-Khmuic, let alone of Proto-Northern-Mon-Khmer.

All of this means that when doing reconstructions one has to tentatively establish as detailed a branching as possible within some division of Mon-Khmer and then work backwards from the smallest clusters of extant languages as defined by the branching, successively uniting them by nodes until one arrives at a family tree (this approach is demonstrated very explicitly e.g. in Smith 1972).

The reconstructed proto-language defines a certain branch or sub-branch of Mon-Khmer, as a kind of common denominator. The minimum requirement is that it should be typologically plausible. The next problem is how it ties in with evidence from outside the branch under consideration. Unless one takes a broader view all through the reconstruction there will often be a certain amount of arbitrariness in it. It may well turn out that subsequent comparative work on a broader scale casts doubts on parts of the reconstruction. For example, one of the languages used in the reconstruction may have consonant or vowel qualities that were considered as due to secondary developments but which agree beautifully with languages outside the sub-branch. For that reason I sometimes permit myself to cite forms from languages that may have played a minor role in the reconstruction of the sub-branch they belong to.

In my comparison of Mlabri with neighbouring languages that problem has presented itself in connection with the Katuic branch of Mon-Khmer in particular. There is a recent reconstruction and even an etymological dictionary (Pejros 1996a) of Proto-Katuic as a whole, but unfortunately that was unavailable to me during the completion of this monograph. In lack of that I have mainly consulted the large dictionary of a West-Katuic language Kui or Suai (Sriwises 1978) and occasionally also cite forms from three East-Katuic languages, Brôu, Pacŏh,

and Katu. Data from the latter three languages were presented and used for comparative purposes in an early reconstruction of Proto-East-Katuic (Thomas 1976), which has been available to me. Although the controversial features of that reconstruction can now be considered as obsolete, they still tell a lesson. When confronting its reconstructed PEK forms with Mlabri or Tin or Khmu, or for that matter with M-K in general, I found that they sometimes gave a so much poorer fit than forms from one or another of the individual languages that it raised doubts about the reconstruction. I shall illustrate that with reference to two reconstructed entities, a consonant *k^h and a diphthong *oâ ([oʌ]).

The consonant *k^h occurs, for example, in a word *k^huun meaning 'magic power'. That is obviously a cognate of Mlabri *gun* (same meaning) but the initial consonant is intriguing, considering the general appearance of Katuic languages. At closer inspection, Katu has a form *gun* that corresponds exactly to that of Mlabri, and Pacŏh has *kun*. The comparison with Mlabri suggests that an originally voiced stop was devoiced, with or without development of aspiration (rather than the other way round, as suggested in Thomas 1976: 11, 19). In this case, the issue is hardly settled by looking at Kui for the form occurring there in the meaning of '(having) magic power' is strongly deviating, but it is significant that it has 2nd register (breathy voice): $k^h ɔ ŋ$.

As for the diphthong *oʌ, it occurs in the PEK reconstructions before stops and *-h, e.g. in *moʌt 'eye' and *ploʌh, a classifier for sheets. The word for 'eye' is, however, *mat* across most of M-K including Mlabri, and the classifier for sheets is *kəpla^h* in Mlabri, *p^{h}lah* in Tin. Again, if one looks at individual East-Katuic languages the mystery is resolved: Katu has *mat* for the former, and Pacŏh has *plah* for the latter, whereas Brôu provides the evidence for the diphthong (the rules are given in Thomas 1976: 43, 59 but without discussion of possible alternative solutions). Kui sides with Mlabri in having /a/ in *màt* 'eye'. Again, on comparative evidence one would like to reverse the process *oʌ > a into a rule generating the diphthong in Brôu (*a > oʌ in environments to be specified).

Finally, I wish to present a case where the evidence from Mlabri can throw interesting light on conflicting forms e.g. in Katuic. The word for 'yawn' is reconstructed by Thomas (1976) as *hangqaap, i.e. *haŋʔaːp, on the basis of Brôu *saʔap*, Pacŏh *ŋʔap* and Katu *ʔaʔaːp* or *haʔaːp* (my transliteration of sa-ap, ng-ap, a-aap, ha-aap). I cannot see what has motivated the reconstruction *haŋʔaːp, but a presyllabic velar nasal is

supported by West-Katuic: Kui has *ŋ²aːp*. Mlabri has *tʰaŋaːp*, which is a regular reflex of a form *saŋaːp*.

If one compares all the forms cited above there is a strong case for reconstructing a main syllable *²aːp*. As for the presyllable, however, one observes that, with one exception, all the languages under consideration exhibit a subset of the elements *s-a-ŋ*, in that order, Mlabri alone having the maximum structure (although Mlabri has contracted the word so that the hiatus after *ŋ* vanished). This is suggestive of some kind of prefixation, though it is etymologically obscure. The only language that falls phonologically outside the pattern is Katu[HK] with *ha²aːp*, a form that obviously motivated the initial consonant in the reconstruction *haŋ²aːp*. Considering what often happens in Northern Mon-Khmer languages one would rather speculate that this aspirate stems from a weakening *s > h*, which may have happened because of the weak position in a presyllable. If that solution is adopted one ends up with a PEK reconstruction *saŋ²aːp*, which is an extremely plausible reconstruction of the Pre-Mlabri form as well.

In presenting these case-stories I do not want to challenge the basic strategy in comparison of remote cognates. It is in principle always the deepest reconstructions that are most relevant for comparative work across different branches of Mon-Khmer, but a *caveat* is in place.

There is another complication with the family-tree approach: the successively accumulated insights may sometimes force the researcher to redefine the boundary between neighbouring branches or sub-branches, e.g. moving a language from one to the other. In the northernmost Mon-Khmer speaking area, in which Mlabri is located, there is reason to speculate how well-defined the branches are, in particular the separation into a Palaungic and a Khmuic branch. That is worth keeping in mind if one sets out to trace the remote linguistic history of a migrating ethnic group such as the Mla Bri.

The family-tree approach has been successful with some language families without old language records, but it is difficult if the languages within a family exhibit two kinds of discrepancies: (i) large differences in overall lexicon, (ii) a lot of apparent or real irregularities in the phonological reflexes of shared cognates. That is the situation with Mon-Khmer as a whole and, on a smaller scale, even with Khmuic. It helps to explain why some researchers (even in contemporary scholarly papers) content themselves with demonstrating the relatedness between more or less remote Mon-Khmer languages by listing handfuls of shared words without much further analysis.

There are words which are widespread in rather similar forms across Mon-Khmer languages. Mlabri has several such "typical" M-K words in easily recognizable form and mostly with the expected meaning; some are given for exemplification in Table 1-6.

Table 1-6
"Typical" (Northern) Mon-Khmer words in Mlabri

bri:ʔ 'forest'
diŋ 'elder brother'
mat 'eye'
ti:ʔ 'hand'
jɤ:ŋ 'leg'
ʔjak 'excrements'
rɛ:l^h 'root'
ʔu:l^h 'firewood'
ɕɔ:ʔ 'dog'
po:l^h 'barking-deer'
t^hawa:ʔ 'monkey' (**s- > t^h-*)
rɔ:j 'housefly'
klʔa:k 'crow'
pʌ:r 'fly, v'
takat 'feverish (or cold)'
ɲa:m 'season'
^hnʌm 'year'

Some such lexical items are interesting by their mere existence in a hunter-gatherer language, for instance the series of numerals from one to ten in Mlabri: *mɔ:j, bɛ:r, pɛʔ, pon, t^hɤ:ŋ, ta:l, gul, ti:ʔ, gaj^h, gal* (for more information of these numerals seen in a wider context, cf. Rischel 1997a). Still, widespread and phonologically fairly stable etyma may be of limited help in comparisons of Mlabri with neighbouring languages unless the demonstration is accompanied by an in-depth analysis of the phonological relationships between the languages in question, and by an overall consideration of the lexicon (which may reveal that is for the most part *not* shared at all). It is often words with non-trivial phonology (for example, three of the words above illustrate M-K **-s* > Mlabri *-l^h*) or with a geographical restricted occurrence that yield most information.

Not all parts of a Mon-Khmer word are equally amenable to cross-

language comparison. Firstly, affixation (pre- and infixation) causes a lot of variation in forms and secondly, different parts of the syllable seem to have different degrees of stability over time.

Within the field of Mon-Khmer studies there seems to be a consensus that syllable-final consonants tend to be fairly stable and therefore a good point of reference in etymological work (accordingly, some compilers of word-lists arrange the words according to their termination and only secondarily according to the onset). Syllable-onsets vary more, but it is largely possible to establish regular phonological correspondences by applying the comparative method. I have found that to be borne out by my comparisons between Mlabri and Tin.

It is different with vowels, however. It was pointed out long ago (Shorto 1976) that there are strange variations in vowels across Mon-Khmer which must date back a long time. Comparing Mlabri and Tin I have found that although the time depth here is shallow in comparison (less than a thousand years) there is a similar, intriguing variation in vowels within what I postulate as Khmuic.

As for Mlabri versus Tin some discrepancies in the cognates are suggestive of the existence of old doublets with different vowels in an "ablaut"-looking relationship to each other, e.g. *plɤːm ~ *plɔːm for a leech species. Such doublets without any transparent reason (at the chronological level under consideration) have been suggested in historical reconstructions within other branches of Mon-Khmer as well; they may cause forms in closely related languages to look different because one language happened to preserve one variant, whereas the other language preserved the other (or perhaps both). Then there are more whimsical differences which suggest a complex scenario of lexical transmission. Maybe some Tinic words were not assimilated smoothly into Mlabri; maybe Tin had a turbulent history shortly after Mlabri received a Tinic component; maybe some of the words that strike me as shared by Mlabri and Tin did not in fact enter Mlabri from Tinic.

In Part 3 I shall look into the Mlabri-Tin vowel issue with a view to the possibility of resolving some of the complexities in the vowels by applying the comparative method. It will turn out that the result is only moderately successful. It is like solving equations with too many variables.

A priori, it is tempting to assume that discrepancies in vowel quality between two related languages Mlabri and Tin can be resolved by using a third, related language as a reference, a *tertium comparationis*. In that case the geographically widespread and lexically extremely well-docu-

mented language Khmu would at first sight be the obvious choice. A minimalistic rule of thumb might run as follows: If a vowel quality is shared by Khmu and either Mlabri and Tin, that is likely to be the Proto-Khmuic vowel quality, whereas there has been a change in the branch of Tinic (be it Mlabri or Tin) that deviates in terms of that vowel.

That is a grossly simplistic assumption, however. In languages of the world vowel qualities often change dramatically over time, and they may even change forth and back over sufficient time within one language branch or even within one dialect. Two languages may also exhibit parallel changes. Thus, there are several more or less plausible scenarios that could bring about the situation sketched above, and it takes a broad comparative approach to decide what is the likely scenario.

To complicate things one quickly observes that in spite of its dominance within the Khmuic sphere, Khmu is not the ideal choice of a reference language. It has undergone changes in the vowels (see exemplification below), and its pattern of short vowels is less rich than those of Tin and Mlabri, which have a degree of aperture more than Khmu (Mlabri has ε, Λ, $\jmath$ as distinctive vowels; in Khmu, short ε and $\jmath$ occur only as allophones of long vowels before laryngeals, and short Λ does not occur at all, Svantesson 1983: 14-15). Thus one can expect in advance that comparisons of Mlabri/Tin with Khmu will pose problems in reconstruction.

Ksingmul across the border between Laos and Vietnam well to the east of Mlabri area is another fairly well documented Khmuic language. It deviates quite much from the western Khmuic languages in its lexicon, and as for phonology, cognate words show considerable discrepancies between Ksingmul and other Khmuic in the consonants. In Ksingmul one encounters a voiceless palatal stop in forms that have a voiceless affricate or sibilant in other Khmuic, e.g. Ksingmul *ceːj* 'lice' (C-Mlabri *ɕɛj*), *coː* 'dog' (Khmu *sɔʔ*), and a voiced sibilant corresponding to a voiced palatal affricate in Mlabri or Tin, e.g. Ksingmul *zuŋ* 'man' (Mlabri *ɟjoːŋ*), *ziəŋ* 'house' (Tin *kiʌŋ*, *ceŋ*).

The Khmu reflexes in the palatal region of the consonant system do not on all points correspond to those of Mlabri or Tin either (I do not go into the complicated history of palatals in Mon-Khmer; Diffloth 1977 and Ferlus 1978 are important contributions to the issue). Khmu is not close to Ksingmul, however.

All of this invites a digression on branchings within Khmuic. It is entirely premature to posit a definitive branching (almost half of the

Khmuic languages are poorly documented, and the trail-blazing work of Thomas and Headley 1970 did not solve these issues). Looking at the overall vocabularies and overall structure of some of the languages one can make a few tentative observations, however. Ksingmul and Tinic clearly belong to different branches. There is another branching separating Khmu from Tinic and from the little-known language Phong in northernmost Laos. Khabit (Bit) belongs to a fourth branch, to judge from its very deviating vowel reflexes. I do not have data on Theen and Iduh (Tayhat) permitting me to consider where they belong in a branching tree.

All of this suggests that Khmuic falls into several branches, although it is unclear so far how "deep" the various branchings are in relation to each other. Even with the sparse evidence on Phong, however, it seems clear to me that Phong is close to Tinic, both phonologically and morphologically, and that Tinic and Phong together probably constitute a "Phongic" branch of Khmuic. Whether that putative branch also contains one or more of the remaining and in part little-known Khmuic languages remains an open question.

Returning to the problem of using a *tertium comparationis*, it can be illustrated by looking at actually attested forms. Taking the word for 'nose', for example, we observe that Mlabri has *mɔh* and TinMal has *moh*, whereas Khmu and TinPrai have *muh*. That would seem to suggest that Mlabri and some Tin have drifted away from an old Khmuic form **muh*. Looking beyond Khmuic, one does find forms in *u* very widely although both *u* and *ɔ* occur in Aslian, the second in fact in several languages (Benjamin 1976: 112). If we then look at Ksingmul, it sides perfectly with Mlabri in having the form *mɔh*. Considering that Ksingmul is a quite distant representative of Khmuic, there is reason to that speculate Mlabri and Ksingmul may have preserved an old form. Thus a more attractive proposal is that the Mon-Khmer word for 'nose' had two distinct variants: **muh* and **mɔh* of which the former is far more widespread than the latter. I would not be happy to posit **muh* as *the* Khmuic form, unless one decides to define as Khmuic whatever agrees with Khmu.

To take another example, the second person singular pronoun in Mlabri is *mɛh*, an etymon that is well represented in Khmuic but with strange vowel offsets: TinPrai has *mah*, Phong has *mɔ*, and Khmu has *meː* (acc. to Suwilai Premsrirat in the more specific meaning of 'you, male'). If we go outside Khmuic, we have other instances of the second person singular pronoun. On one side, Katu has *meh*, and on the other

side, Lawa has *meʔ* (reportedly only in poetic vocabulary); Mitani (1979) reconstructs *mi* for Proto-Palaung. The Mlabri and Khmu forms seem to furnish the best point of departure for reconstruction of a Khmuic ancestral form, quite unlike the Tin and Phong forms, but the variation is conspicuous and makes rigid comparisons difficult.

For the comparison of Mlabri with Tin and with various other Khmuic languages it is worth keeping in mind that there are several indications, as in the cognate sets above, that Mlabri phonology is conservative. That may also have been true of Proto-Tinic compared to some other Khmuic. Looking, for example, at the Tinic word for 'feel pain': MlaBC *çoʔ*, TinPrai *soʔ*, one observes that Khmu exhibits a higher vowel: *suʔ*, whereas Phong has diphthongization and loss of the glottal final: *sǝw*. There is a good case for positing **soʔ*, as in Tinic, as a valid reconstruction of the old Khmuic vocalism.

A similar example is the word *preʔ*, which occurs in identical form in Tin, where is has the expected meaning 'hot (spicy)', and in Mlabri, where it means 'chili' (in the phrase *preʔ gem*). This Tinic form is in perfect accordance with Diffloth's (1980) reconstruction of Waic **preʔ* and its preservation e.g. in contemporary Lawa *pjeʔ*. Again, the Tinic form is a perfect candidate for old vocalism. Other Khmuic languages have drifted in mutually opposite directions: Khmu has *priʔ* (Khabit has a remotely similar form *wiʔ*) and Phong has *prɛj*.

The correspondences being of the same nature in *suʔ* ~ *soʔ* ~ *sǝw* and *priʔ* ~ *preʔ* ~ *prɛj* we are looking at a regularity. With more data there might a solid ground for positing sound-laws relating forms such as these across Khmuic.

As exemplified by the forms above, Khmu often has a high vowel, which may in some cases be due to a vowel raising particular to Khmu. This illustrates a point that is made repeatedly in this monograph: although Khmu is conservative in various respects one should beware of taking that language as the exemplar of Khmuic if it is a matter of vowel comparisons between branches of Mon-Khmer.

In conclusion I dare postulate that Khmuic shows a lack of unity when it comes to vowel qualities, although the correspondences are not just chaotic. There may be a unitary explanation, possibly in terms of a phonation type causing diphthongs which later monophthongized in some cases. Such "register" phenomena play a decisive role in phonological differentiation within Waic, as shown by Diffloth 1980, and also elsewhere in Mon-Khmer. In the case of Khmuic, however, I fail to see any pattern suggesting such an explanation that would work across the

board (hopefully, all this will appear in a very different light once reconstructions of both Proto-Palaungic as a whole and Proto-Khmuic as a whole have become a reality).

When we pool all the cognates shared by Mlabri and Tin, such vowel variation faces us from one end of the vowel space to the other. There is nothing like a one-to-one correspondence between vowels of more or less similar vowel quality, not even after one has weeded out words that are likely to be recent loanwords and concentrates on words that are likely to be old cognates. Some correspondences are attested by many forms, others by very few forms (exemplification of that is given in Part 3), and that means that one can to some extent establish a set a major correspondences but the "minor" correspondences must of course also be accounted for somehow, in some cases as due to borrowing.

In Mlabri and Tin, a vowel in one of the languages often corresponds etymologically to two or more vowels in the other (cf. Table 2-6 in Part 2), and this complication works both ways. This is in contrast with the scenario depicted by Diffloth for Waic:

> The regularity of sound changes within Waic is actually quite impressive, as soon as spurious cognates are mercilessly abandoned (...) This regularity, and the apparent rarity of intra-Waic borrowings, are probably due to certain social values shared by speakers of Waic languages, and to the geographical dispersal connected to such values (...)

(Diffloth 1980: 94).

One should not take Diffloth's statement to mean that there is uniformity across Waic e.g. when it comes to vowel qualities. On the contrary. What one observes when looking at individual etyma and their reflexes in the various Waic languages, is that there are enormous differences precisely in vowel qualities, much larger differences than between Mlabri and Tin. In his comparative study, however, Diffloth brings order out of an apparent chaos, mainly by positing a number of language-specific (or even dialect-specific) diphthongizations all having a common origin: breathy phonation in words with original voiced initial (so-called "register"). The fascinating thing about the widely different appearances of shared words in the Waic languages or dialects is that the differences among them then turn out to exhibit *regularity*. Accordingly, they can be all derived from a well-defined set of proto-forms.

As will hopefully be convincingly demonstrated in Part 3 of this

monograph, it is possible to bring some order out of chaos in the case of Mlabri and Tin vowels as well, not by finding an underlying principle such as register but just by distinguishing (perhaps somewhat heavy-handedly) between regularities and more or less explicable aberrations from the regularities.

PART 2:

MLABRI AND TINIC:
PHONOLOGY
AND MORPHOLOGY

Evidence for Tinic

As mentioned repeatedly in Part 1 it is my working hypothesis that the moderately Khmuic appearance of Mlabri is primarily due to an old connection with Tin dating back to a chronological period which I call Proto-Tinic. The present Part 2 as well as Part 3 below are devoted to a detailed comparison of Mlabri and Tin in that framework, whereas a broader view of the possible sources of Mlabri lexicon and grammar will be taken in Part 4.

It will be immediately apparent that the Tinic hypothesis does not by itself give a satisfactorily explanation of the overall appearance of the Mlabri language. It is, however, the only likely linguistic correlate to the biologists' hypothesis about a founding event involving village people.

The Tinic connection is also quite significant from a purely linguistic perspective because of the amount of shared lexicon (I list a couple of hundred cognates in Part 3, and there are undoubtedly many more which I did not or could not detect with the data readily available to me), and especially because many of these show by their structure that they did not enter Mlabri recently from Tin. Thus the exploitation of this connection is an important first step in any attempt to unravel the early history of the Mlabri language.

By way of introduction to the pursuit of the language history of the alleged Tinic branch it may be called for to give some specific evidence for the existence of the alleged bond between Mlabri and Tin at the Proto-Tinic level, as against other sub-branches such as Khmu.

One type of evidence is shared phonological innovations. One such feature is the change in quality of *aː, cf. Mlabri *beːr*, Tin *piʌr* 'two', as against Khmu *baːr*. Another is the change of a final sibilant *-s into an aspirated, more or less voiceless sonorant, cf. Mlabri *poːlʰ*, Tin *pʰɔːjʰ* 'barking deer' as against Khmu *puəs*. On the whole, however, Mlabri phonology is very conservative, so if indeed there was such a thing as a Tinic level supplying Mlabri with lexicon that must have had a rather conservative phonology as well.

Another criterion is numerical evidence. There are large differences in lexicon among the Khmuic languages, and Mlabri does not share all that many etyma with any of the other languages. The number of etyma shared by Mlabri and Tin, however, is several times higher than the

number of etyma shared by Mlabri and other well-attested languages such as Khmu and Ksingmul. That is what I found by looking at general vocabulary with well-defined and tangible meanings ("content words" rather than "function words"). I have spotted some two hundred such words as possible candidates for being of Tinic origin (they are listed in Part 3 with their Mlabri and Tin forms).

All of this does not imply that Mlabri and Tin together formed a branch of Khmuic. In terms of language structure Mlabri and Tin are *not* particularly similar. The lexical comparison only suggests that Mlabri has a significant part of its vocabulary, maybe some two hundred words, from an early stage of the Tinic branch. The real nature of the association between Mlabri and Tin is a difficult issue the discussion of which will be postponed to Part 4.

Proto-Tinic and Pre-Tinic

The term *Proto-Tinic* is used by me to refer to the chronological stage that immediately preceded the cleavage between (Tinic) Mlabri and Tin proper. My assumption is that at that point, Tinic was already rather different from (Proto-)Khmu. To account for that accumulated divergence it is necessary to posit an even earlier, hypothetical pre-stage of Tinic: *Early Tinic* or *Pre-Tinic*, which was close to but already distinct from Khmu. From a strict Stammbaum-perspective Pre-Tinic would be the stage just after Tinic split off from other Western Khmuic.

Because of structural similarities with at least one other Western Khmuic language, Phong, I cannot exclude the possibility that Pre-Tinic should be renamed *Pre-West-Khmuic*, thus excluding Khmu from the specifically western Khmuic languages. The information on Phong being so scanty, however, I narrow the perspective to Tinic proper (Mlabri-Tin) and retain the term *Early Tinic* or *Pre-Tinic*. In what follows no attempt is made to reconstruct what exactly Pre-Tinic was like. The frame of reference which I try to establish for the purpose of relating Mlabri to Tin, is the later stage: *Proto-Tinic*, as defined chronologically at the beginning of this paragraph.

The present part of the monograph, Part 2, gives an overview of the structural features of Proto-Tinic to the extent that these features present themselves already by a cursory comparison of Tin (Mal and Prai) on the one hand and the Tinic component in Mlabri on the other. There are a host of difficult problems that would need to be solved in order to

establish a real reconstruction of Proto-Tinic. Most of these have to do with multiple correspondences between the vowels of Mlabri and Tin. I shall address these issues in Part 3 below.

The canonical structures of Tinic words

Mlabri and Tin have an altogether different appearance when it comes to word structure types, and both are different from Khmu, for example. I suspect that neither the pattern in Modern Mlabri nor the pattern in Modern Tin can be projected back to Proto-Tinic in its entirety. Still, it is illuminating to demonstrate the profound difference in canonical structure between the modern languages.

Mlabri. Mlabri has numerous monosyllabic words but also a proliferation of sesquisyllabic words, i.e. words in which a minor syllable (a reduced presyllable) precedes the major syllable that carries the word stress. A few words have two minor syllables before the major syllable, e.g. *tr̩ləgr̩k*, the name of a mythical hero. There are also several words with full vowels in two consecutive syllables; I refer to these (occasionally) as disyllabic words. In this monograph the term "presyllable" is used as if synonymous with "minor syllable". The distinction between sesquisyllables and disyllables is problematic; many words vary over both structure types (as for my conventions in transcribing presyllables, also see Appendix II).

The distinction between monosyllabicity and sesquisyllabicity, on the other hand, is very clear, consistent and in part non-predictable, i.e., it is phonemic in some instances in Mlabri. There are structural restrictions, however. I shall present the pattern in very compact form, largely without illustrating examples (some of the structures that are crucial in a Tinic context will be taken up later with some exemplification)

Wordforms whose segmental template exhibits a dual onset, i.e. *CC-*, are *predictably monosyllabic* if the second consonant is a sonorant, provided that there is no intervening vowel between the two consonants. Examples: *trup* 'cover, *v*', *mlaːʔ* 'person; man', *rwɤt* 'pour water over something'. Mlabri has a contrast between glide+vowel and diphthong after an initial consonant: *rwɤt* above contrasts with *ruɤt* 'pond that dries up intermittenly'.

In contrast, there are numerous *sesquisyllabic* words beginning in *CᵊC-* but in these, the parasitic vowel between the two consonants is

absolutely obligatory, and there is very often variant forms with a full
vowel, e.g. *tˀguk* = *toguk* 'frog species'.

It is different with words in *CC-* if there is a sonority drop from the
first to the second consonant, e.g. if the first consonant is a trill and the
second a nasal or stop, or if the first consonant is a nasal and the second
a stop. The first consonant is then always syllabic, i.e. the wordform is
predictably sesquisyllabic (instead of syllabicity on the first consonant
there may occur an epenthetic vowel in between the consonants: *CˀC-*).
Examples: *ṛmut, rˀmut* 'wind', MlaA *hŋkeːʔ* 'firewood'.

As for wordforms whose segmental template exhibits a triple onset,
i.e. *CCC-*, most are *predictably sesquisyllabic*, namely with a syllabic
segment (often realized phonetically as vowel plus non-syllabic sono-
rant) in second position. This is true if the second consonant is a nasal,
and also if the second consonant is a liquid (*l, r*) and the third consonant
a true consonant (i.e. *not* a glide *w* or *j*). Examples: *çŋgaɲ* (MlaA also
çuŋgaɲ, MlaB also *çeŋgaɲ*) 'glowing charcoal', MlaAB *bḷwet* (MlaA
also *bulwet*, MlaB also *belwet*) 'swing forth and back'. The only *CCC-*
sequence that occurs in *contrastively monosyllabic versus sesquisyl-
labic wordforms* is *stop + liquid + glide*, e.g. MlaAC *krwɛc* 'scratch, *v*',
MlaA *brwac* 'swallow, *v*' versus *kṛwɛk* 'bear cat', *gṛwɛːc* 'finger; toe;
claw', MlaB *jṛwɤːl* 'boil in a pot'.

Most sesquisyllabic wordforms whose segmental template exhibits a
triple onset, tend to be realized phonetically with a vowel that breaks up
the *CCC-*sequence and thus takes over syllabicity from the medial con-
sonant segment. This epenthetic vowel may vary in quality (though
there are certain preferred vowel qualities depending on the particular
consonants in the *CCC-*sequence) and it normally occurs before the
consonant segment that would otherwise be syllabic (in fairly rare
instances, however, it can occur after that segment).

Sesquisyllabicity gives Mlabri a very distinctive Mon-Khmer ap-
pearance, but there is a preponderance of forms that are morpho-
logically unanalysable, and that gives Mlabri a distinct appearance
compared to Khmu and many other mainstream M-K languages.

As for bisyllabic words, there are reduplicatives with two full sylla-
bles in succession, but otherwise it is difficult to draw a well-defined
boundary between disyllabic and sesquisyllabic wordforms in Mlabri.
Wordforms which can be represented phonologically as containing a
syllabic consonant are often realized with a fully articulated vowel
before the consonant in question if the speaker is asked to pronounce
the word in isolation. Some words deviate even more from prototypi-

cally sesquisyllabic words that they are pronounced with a succession of two stressed syllables, and a higher tone on the first than the second syllable, in more distinct rendering. An example is *kṇdeːp* 'centipede species', in distinct pronunciation ['kin'deːp, 'kuun'deːp]. I suppose that this is a vestige of some structural feature in the non-Tinic layer of Mlabri. Obviously, some such wordforms may be old compounds although I have had little success so far in identifying their parts (the particular word for 'centipede' is discussed in Part 4 in the context of similar forms from other M-K but without any solution). With the present lack of insight into their history I can just mentioning the existence of a particular prosodic pattern carried by some Mlabri words.

Tin. Tin is altogether different. The vast majority of lexical items are monosyllabic; most exceptions being borrowings from Tai.

Within the M-K vocabulary the only deviation from monosyllabicity is constituted by words beginning in *si-* (*ɕi-*) followed by a stop consonant. There are several such words but only in Prai. In some cases they are matched by Mal words which have *h-* followed by a stop consonant in the most conservative dialect (MalA). I shall consider the origin of word-initial *si-* in Prai in Part 3 below. The situation facing us in Tin, with its strong preponderance of monosyllabicity is, according to my observations, also found in some other, little-studied Khmuic languages. It is clearly secondary and due to pervasive structural changes some time in the past in that cluster of languages.

There are good cognates between Mlabri and Tin which illustrate sesquisyllabicity in Mlabri as against monosyllabicity in Tin. Therefore, the reconstruction of Proto-Tinic can be based on the structural types found in the Tinic layer of Mlabri vocabulary. What happened to sesquisyllabic words in Tin belongs to the history of that language *after* the separation between Mlabri and Tin.

Phonology: syllable initials

Most parts of the system of initial consonants are shared by Mlabri and conservative Tin (TinMalA), with the reservation that there is an offset in voicing or aspiration between Mlabri and Tin within the system of stop consonants. Comparative evidence immediately discloses that this offset is due to developments within Tin proper ("consonant mutation"), so one can largely reconstruct the Proto-Tinic consonant system on the

basis of Mlabri to the extent that the Tinic vocabulary in Mlabri gives sufficient data.

There were four oral points of articulation in Proto-Tinic initial consonants: labial, dental, palatal and velar, plus laryngeal articulation.

One can further posit the following ten manners of articulation with certainty: voiceless stop, voiced stop, voiced nasal, pre-aspirated nasal, voiceless continuant (sibilants and *h), voiced lateral and trill (below, I use the antiquated label "liquids" as a cover term), pre-aspirated lateral and trill, voiced glide, pre-aspirated glide, and finally glottalized glide.

The overall appearance of the system differs from that of Mlabri in the absence of aspirated and glottalized stops and the presence of glottalized nasals. The chart of Proto-Tinic single initials is given in Table 2-1.

Table 2-1: Proto-Tinic Initials

*p-	*t-	*c-	*k-	
*b-	*d-	*ɟ-	*g-	
*m-	*n-	*ɲ-	*ŋ-	
*^{h}m-	*^{h}n-	*^{h}ɲ-	*^{h}ŋ-	
	*S_1- *S_2-			*h-
	*l- *r-			
	*^{h}l- *^{h}r-			
*w-		*j-		
*^{h}w-				
*^{ʔ}w-		*^{ʔ}j-		

In addition to these initials one may posit a glottal onset which occurs automatically if the syllable would begin in a vowel otherwise. (In the transcriptions below – both for Mlabri, Tin, and proto-stages of these – I render this pre-vocalic entity as raised *ʔ* so as to make it easy to distinguish it graphically from a glottal syllable termination *ʔ* in sesqui- or bisyllabic words.)

Aspirated stops and sibilants. Tin has a series of aspirated stops, but those are reflexes of the voiceless stops series above and thus do not enter the reconstructed chart for Proto-Tinic. There is likewise a series of aspirated stops in Mlabri though with a different history than the aspirates in Tin. The aspirated stops of Mlabri can be arranged as con-

stituting a four-way series, like the stop series posited for Proto-Tinic above, see Table 2-2.

Table 2-2: Aspirated stops in Mlabri

$$p^h\text{-} \qquad t^h\text{-} \qquad c^h\text{-} \qquad k^h\text{-}$$

Two of these Mlabri initials, p^h- and k^h-, seem to occur exclusively in fairly recent loanwords, most of which stem from Tai but some from Tin. They are thus irrelevant to the Tinic issue. The same is true of several words in t^h- and c^h-. There is, however, a considerable residue of words in these two initials which must be of age in the language, and they require comment below.

In the Tinic vocabulary Mlabri t^h- is the regular correspondence of a Tin sibilant, which is realized variously with alveolar or palatalized alveolar articulation and with more or less affrication, but which is generally represented by s- (e.g. Mlabri $thʌc$, Tin sec 'meat; pulp'). Comparison with M-K languages of the area and beyond it shows beyond dispute that it is an old sibilant, which I rendered in the chart above with the abstract symbol $*S_1$-.

The Mlabri consonant listed as c^h- above warrants this representation only in the sense that it fills a structural slot in the aspirated stop series. It is realized as a palatoalveolar (sometimes alveolar) sibilant or affricate: [$^t\!ç$-, $ç$-], and in the present study I render it as $ç$-. This consonant occurs in many words of M-K origin, and I represent it above with the abstract symbol S_2- (this sibilant goes back to $*c$-, cf. Ferlus 1978: 8). It is often difficult or impossible to decide whether words in S_2- belong to the Tinic layer in Mlabri. The reflex in Tin is the same as for S_1- above, except that some speakers of the most conservative Tin dialect: TinMalA, use an alveolopalatal affricate [$^t\!ç$-] in a handful of these words and insist on its being a different sound than the reflex of S_1- (I could hear the difference myself, and it was explained to me with reference to the letters of the Thai writing system: assuming that one would write Tin with Thai letters one must use different consonant letters for the two sets of words, they said). Thus there appear to be vestiges of a distinction between S_1- and S_2- even in Tin.

Were there any glottalized stops or nasals in Proto-Tinic? There are two glottalized voiced stops in Mlabri. They could be represented as $^?b$-

and *ʔd-* but since they are slightly implosive I allow myself to use the more convenient notation *ɓ-, ɗ-*. They are the regular reflex (except in very recent loanwords) of the so-called "mid" consonants *b-, d-* in Tai, which have in fact been reconstructed as formerly glottalized. A great many of the Mlabri words in these initials have such a Tai source.

The residue of words in glottalized initials are mostly etymologically obscure to me, one way or another. Examples are: *ɓoh* 'of such-and-such size', (MlaBC) *dʌlʰ* 'amount; number'. There are, however, words that are clearly of Mon-Khmer origin and they point to clusters of voiceless stop + voiced source as the source of the glottalized initials. One is *ɓor* 'pangolin (an animal resembling an armadillo)', another *diŋ* 'big'. The former word is *kmbuər* in Khmu, *ᵐbor* in TinPrai, which suggests that one should reconstruct two forms related by infixation: **kbor*, continued as Mlabri *ɓor*, and **kmbor*, continued as the Khmu and Tin forms above. The latter word is reconstructed for Proto-Waic by Diffloth (1980) in the form **kdiŋ*, which would regularly give Mlabri *diŋ* by the assumed general change

$$*p/*t/*k + *b, *d > ɓ, ɗ$$

in Early Mlabri. There is a further word which has Tin connections: Mlabri *diːŋ* 'gaur, or a gaur-like animal species'; TinMal has *tiːŋ*, and TinPrai has *tiŋ*. Diffloth (1984) reconstructs *kndiːŋ* for Monic; the word has also been borrowed into Tai as a bisyllabic word of the segmental structure **kradiŋ* (Central Thai *kratʰiŋ*). This combined evidence again suggests that an initial complex containing two stops shrank into a glottalized stop in Mlabri. For this word I would assume an old form without infixation: **kdiːŋ* underlying the Mlabri form, whereas Tin might well have truncated a more complex presyllable completely like in so many other cases.

Although the explanation of Mlabri glottalized stops as reflexes of clusters of two stops may eventually turn out to explain all occurrences in monosyllables (if some time an exhaustive etymological M-K dictionary becomes available), there is the added complication that there are examples with these stops in morphologically non-transparent forms in which they occur after a full syllable or presyllable. If these originally derive from compounds or forms with prefixes the same explanation may apply, of course, but if the first syllable is open there is the theoretical possibility that that syllable used to end in a glottal stop *-ʔ, which has fused with a following *b- or *d- into *ɓ, ɗ*. There are,

however, two forms below that have a glottal stop followed by a glottalized consonant, which speaks against such fusion.

I shall just list the forms without further analysis (in presyllables with consistently prominent though sometimes variable vowels these vowels are included since they are obviously relevant if one attempts to retrieve the first part of the word morphologically):

MlaB *cȓɓat* 'trousers'
MlaC *cȓɓat* 'tuber species'
MlaA *kmɓɤːt* (obsolete) 'aeroplane (probably likened to an insect)'
MlaC *gmɓɔːŋ* 'lump of hardened resin'
MlaA *cȓɓot* 'prop used to close a container'
MlaA (obsolete) *cȓɓot* 'mode of singing'
MlaAC *tʰmɓok* (also: *tʰumɓok*) 'leaf prop or other thing used as a lid; to prop, v'
MlaA *kaɓoːk* 'hoe'
MlaAB *paɓom* 'cover'
MlaAB *calɓuːt* 'spoiled and smelly (food)'

MlaA *cədaː* 'tree species'
MlaA *kaʔdaːt* 'a personal name'
MlaA *cakɗar* 'squirrel species'
MlaA *kmdaːr* 'unripe'
MlaA *kandɛj* (some say *kʰandɛj*) 'elephant's tusk'
MlaABC *kindeːp* (also *kundeːp*) 'centipede species'
MlaB *kṇɗup* 'orifice of snail shell'
MlaB *kḷɗol* 'bird species'
MlaAC *caɗup* (some say *caɗup*) 'back (of body or thing)'
MlaA *bakʰʌɗuːp* 'insert (into somthing close-fitting)'
MlaABC *kʰuʔɗuːn* 'cover term for earthworms'

My conclusion is that the old set of Mlabri words in glottalized stops are not of Tinic origin but represent a very old Mon-Khmer layer.

Tin has glottalized nasals initially, but these are only found in the most conservative dialect TinMalA, elsewhere one encounters voiced stops. I do not know the etymologies of such words as *ʔmut* 'shirt', *ʔniː* 'gall' and *ʔnɔːm* 'hide (oneself), v'. In some cases, however, one can spot a Tai origin.

The source of *ʔm* in TinMalA *ʔmaː* 'be angry' is Tai *ʔb- (the corresponding word in Tai means 'crazy'), and the source of *ʔn- in TinMalA

ʾnuʾ 'a lot; much' is Tai *ʾd-* (assuming a semantic shift from 'fierce' to 'much'). There are also instances of glottalized velar nasal; the examples I have are loanwords from Tai, and the source of the glottalized nasal in these cases is a glide *ʾj-/j-* (the reflex *ɲ-* occurs in Lao): *ʾnaː* 'tobacco', *ʾɲuːt* 'elastic'.

These various Thai correspondences suggest that the glottalized nasals in Tin may have arisen in Tin proper and may not directly reflect the phonology of Proto-Tinic.

I have just two examples in which the Mlabri glottalized stop corresponds to a glottalized nasal in Tin. Both are of Tai origin. One is Mlabri *ʾbɔːŋ* 'hollow iron handle of a knife into which a wooden handle can be inserted'; it may be the same Tai etymon (found in Lao) that survives as TinMalA *ʾmɔːŋ* 'spear', probably originally denoting the iron part of the spear. The other example is MlaAB *buːt* in *ɕalɓuːt* 'for food to be spoiled (stale; invaded by maggots)', the second part of which corresponds to TinMalA *ʾmuːt* 'for food to be spoiled' (TinPrai has a plain stop: *buːt*). It is not a problem for that comparison that it is the second syllable of a bisyllabic Mlabri word that it involved, for its first part is a recognizable morpheme of its own (cf. MlaA *ɕalpreːl* 'have a bitter taste', *ɕalmeːn* 'stink'). What is really interesting is that this borrowing from Tai has been so fully integrated into Mlabri that it underwent compounding and is now only preserved in the compound.

The outcome of these comparisons is that there is no compelling evidence for reconstructing a series of glottalized voiced stops or glottalized nasals in Proto-Tinic. This is a significant finding considering that there are several Mlabri words in *ɓ-* and *ɗ-* and a couple of TinMalA words in *ʾm-* and *ʾn-* that do not seem to have Tai cognates.

Clusters. There were at least six types of Tinic onsets consisting of two consonants: (i) stop plus liquid, (ii) stop plus glide (in Modern Tin even with prenasalized stop, e.g. *ⁿjw-*), (iii) sibilant plus liquid, (iv) sibilant plus glide (with uncertainty about the repertory, see below), (v) liquid plus glide (*rw-* and possibly also *lw-*; the latter exists in Mlabri), (vi) nasal plus glide (both Mlabri and Tin have words in *ŋw-*, for example). It is uncertain whether combinations of (vii) nasal plus liquid (which occur in Mlabri) existed in Proto-Tinic or have a different source.

As for initial consonant combinations consisting of sibilant + liquid the only safely attested cluster is *sr-* though *sl-* must also have existed at the Tinic level. Comparison of the Mlabri word *ʰluːt* in *keːt ʰluːt* 'hard of hearing; deaf' (not found in Tin) with Khmu *sluut* 'deaf' (pro-

nounced with an epenthetic vowel between the onset consonants) is suggestive of a sound change *sl- > hl- but unfortunately, I have no etymon shared by Mlabri and Tin which can be reconstructed with *sl-. Thus it is uncertain when or where that sound change happened.

As for initial consonant combinations consisting of sibilant + glide there is plenty of evidence for *S_1w- and *S_2w- in Mlabri but an absence or at least scarcity of good comparative data linking Mlabri and Tin in this respect.

There is a significant absence, both in Mlabri and Tin, of a cluster type which occurs in several M-K languages and also (with some etyma) in Khmu, viz. *voiceless obstruent plus nasal*. In Mlabri and Tin all such clusters have been weakened to aspirated nasals, with complete loss of information about the place of articulation of the initial obstruent. In Tin, the aspiration on nasals is preserved in MalA; all other offsprings of Proto-Tin now have plain nasals.

In my comparative Tinic data set there are three etyma of this kind with a historical stop before nasal. One example is the word for 'new': Mla *^{h}me?*, TinPrai *me?*, Khmu *hmme?*; Diffloth (1984: 155) reconstructs *t[ ]mii? for Monic. Another is the word for 'year': Mla *^{h}nʌm*, TinMalB *na:m*, Khmu *nim* (not the regular word for 'year', though; I have noticed this form in certain phrases listed in Premsrirat's Thesaurus); Diffloth (1984: 134) reconstructs *cnaam for Monic. A third, more problematic example is the word for 'rat', MlaA has *^{h}nel* (with an unexpected and unexplained final lateral), TinPrai *sine:* (the pseudo-prefix *si-* signals the loss of phonological material); Khmu has *kne?*, and Diffloth (1984: 61) reconstructs *kni:? for Monic.

As for sibilant + nasal there is one example: Mla *^{h}ma:l* 'personal spirit', TinMalA *^{h}ma:l*, TinPrai *sima:l* 'soul; essence of life', Khmu *hmma:l*, *ma:l* 'personal spirit'; the old sibilant reveals itself before an infix, cf. MlaB *çrma:l* 'soul', an early loan from some Khmuic.

There clearly happened a sound change at some point: *M-K voiceless obstruent + nasal > h + nasal*. This sound change was not limited to Tinic but also affected Khmu, for example. The examples above illustrate different outcomes for Khmu: (i) preservation of the initial stop, (ii) weakening of the stop to *h*, (iii) complete loss of the initial segment. That looks chaotic, but maybe (i) vs. (ii)-(iii) has something to do with sesquisyllabic and monosyllabic realizations; (iii) may be a further development of (ii) which was arrested if material (in this case *-m-*) intervened between the reflex of the obstruent and the nasal.

In Mlabri one might a priori expect initial clusters consisting of

voiceless obstruent + nasal to occur in the non-Tinic stratum. There is, however, not a single lexical item of this structure in Mlabri. That strongly suggests that the sound change is very old and also affected Mlabri already before Tinic time.

Triple consonant sequences. In Tinic there were word-initial three-consonant sequences consisting of a stop plus a liquid (*-r- or *-l-) plus *-w-; it is unclear to what extent they were realized as sesquisyllabic. As stated above (in the subsection on "Canonical structures"), Mlabri has a contrast between monosyllabicity: *Crw-* and sesquisyllabicity: *Crw-* (realized with syllabic -r̩- or with a brief vowel before -r-). Sesquisyllabic structures of this kind are, however, much more frequent in Mlabri than the corresponding monosyllabic structures. If the medial consonant is *-l- the word-form always comes out as sesquisyllabic, e.g. MlaAB *k̩lwʌŋ* (also *kulwʌŋ*) 'ball, egg'. TinMalA, in contrast, has a homosyllabic three-consonant cluster e.g. in *kʰlwaŋ* 'seed in fruit', which is the same etymon as the just-mentioned Mlabri word. Similarly, TinPrai has three-consonant clusters as in *krwaʔ* 'burp'.

Modern Tin does not exhibit sesquisyllabic wordforms at all, and since that word type is almost defining for M-K the obvious explanation is that Tin got rid of sesquisyllabicity at some point. That leaves the question open whether Proto-Tinic had any syllabicity contrast in *CCC*-sequences since my few examples with homosyllabic *CCC*- cannot be proved to be words of Tinic provenance.

Commentary to the scenario assumed for Tin: Since the general tendency in Tin is to simplify word structures it may seem controversial to assume that phonetic *CCC*-clusters *arose* on the way to Modern Tin, but in fact that happened with monosyllabic loanwords from Tai as well. Tin has an old loanword for 'cultivated banana' (the borrowing predates consonant mutation, and the lateral after stop is preserved), which I took down in TinMalA as *kʰlwaj* (also pronounced *kʰloj* by some); it must originally have been borrowed in the form *kluaj*.

Phonology: oral syllable-finals

Judging from the combined evidence of Mlabri and Tin (and of other Khmuic as well) the system of non-laryngeal finals in Proto-Tinic was of mainstream M-K type, with four points of articulation: labial, dental, palatal, and velar, and with five manners of articulation: stop, nasal,

sibilant, liquid, and glide. The main complication is the category of
sibilants; they will be commented on below.

The overall inventory of final comprised 14 entities, as shown in
Table 2-3.

Table 2-3: Proto-Tinic oral finals

*-p	*-t	*-c	*-k
*-m	*-n	*-ɲ	*-ŋ
	*-S_1, *-S_2		
	*-l, *-r		
*-w		*-j	

The finals shown in Table 2-3 all have distinct reflexes in Mlabri. The
same is true of (some) Tin except that -S_1 and -S_2 have coalesced in Tin.
Comparison with Khmu or with other, more distant M-K languages
shows a good fit with respect to finals, although there is some dental ~
palatal fluctuation across M-K and also in Tinic, both in the stop series
and the nasal series.

The most interesting finals are the ones that I render "abstractly" as
*-S_1 and *-S_2 (it should be understood that I make no claims about any
etymological connection between initial and final sibilants; the symbols
I use just happen to be the same). Mlabri has -l^h for the first (e.g.
po:l^h 'barking deer', a word that exhibits -s widely in M-K languages),
whereas Mlabri has -j^h, a more or less forcefully aspirated glide which
often approaches [ç], for the second (gaj^h 'nine'). The latter is an infre-
quent final, occurring mainly in some words with no obvious MK ety-
mology. In TinMal the reflex of both of the old sibilants (to the extent
that the second is attested at all) is -j^h. In TinPrai their common reflex is
-t, with a three-way merger between *-S_1, *-S_2 and the old dental stop
*-t. A remarkable thing about these reflexes is that there is no trace of
sibilants in any Tinic language or dialect (some M-K languages exhibit
a weakening of sibilant to [h] but there is no trace of such a develop-
ment either). I therefore take it that the change away from sibilants hap-
pened already at the Proto-Tinic level, before Mlabri parted company
with Tin. This makes it possible to posit the early, syllable-final loss of
sibilant articulation as a definiens of Tinic.

Vowel-length, final laryngeals,
and open syllables

Vowel-length. In Proto-Tinic, vowel-length was contrastive in front of a wide variety of finals. The agreement between Mlabri and Tin on this point is, however, rather poor.

In reconstructions I shall represent the non-alternating vowels as short versus long, and the instances of alternating length (except in cases where they *are* rule-governed) by abstract symbols. This the total set of representations of stable or alternating length in my reconstructions of Proto-Tinic is:

V = short vowel in both Mlabri and Tin
V: = long vowel in both Mlabri and Tin
$V(:)_1$ = long vowel in Mlabri, short vowel in Tin
$V(:)_2$ = short vowel in Mlabri, long vowel in Tin

The trickiness of length in Mlabri. When speaking about long vowels in Mlabri one must make the qualification that shortening occurs almost regularly when a word is non-final in a phrase (the accentuation of phonological phrases being on the last word, as is areally expected). Thus, although 'man; person' is *mla:ʔ* and 'ear' is *ke:t* the unmarked way of saying 'Mlabri' ("man forest") and 'deaf' ("ear deaf") is (*mlaʔ ˈbri:ʔ* or) *mlaˈbri:ʔ, ket ˈʰlu:t*. Conversely, vowels are lengthened in utterance-final or even clause-final position; especially in MlaA this leads to extremely overlong syllables spoken on a high-falling pitch, a feature that has often been noticed as characteristic of Mlabri in contrast to other languages of the area.

Needless to say, it is not possible to determine how old these patterns of phrasal vowel shortening and final lengthening in Mlabri are, but they might well have caused some words to switch, within Mlabri proper, from lexicalization with a long vowel to lexicalization with a short vowel, and vice versa. That may be one of the reasons why there is no much misfit in vowel-length between Mlabri and Tin. There are even some words that differ in phonological vowel-length over different varieties of Mlabri although the main tendency is for the vowel of a particular word to be either phonologically long or phonologically short across all of Mlabri.

The length contrast is well preserved in MlaB and MlaC and with

some speakers of MlaA, whereas other speakers of the last-mentioned variety have more or less merged phonologically long and short vowels so that the actual length of a vowel is mainly determined by the prosodic position of the word.

As for open syllables in Mlabri and Tin, stressed syllables always have a long vowel, whereas there are several unstressed function words with a short open syllable. Since there is next to no overlap in lexicon when it comes to unstressed function words, I cannot give cognate pairs of this kind.

Final laryngeals. Except for unstressed function words, short syllables in both Mlabri and Tin must be checked, either by an oral consonant or by a final laryngeal. For Proto-Tinic one can tentatively posit the same repertory of laryngeals as in Mlabri and Tin, cf. Table 2-4.

Table 2-4: Proto-Tinic final laryngeals

-h   *-ʔ*

It is, however, a problem in the reconstruction of Proto-Tinic by comparison with Mlabri and Tin that these two languages often disagree on the final laryngeal. This adds to the difficulties in historical comparisons caused by the frequent lack of agreement on vowel length.

If Mlabri and Tin were Tai languages, the conflicting distributions of vowel-length and syllable-final laryngeals might raise suspicions about a proto-stage with prosodic contrasts in phonation type. That makes no sense in a Mon-Khmer context; instead, one might expect secondary and language-specific, so-called register differences having to do with voicing in the syllable-initials and affecting the vowel quality (the question of register is taken up later). Another possibility would be differences in nasalization (that feature plays a role in Bahnaric) that once affected vowel quality and then vanished. It would, however, take a large repertory of phonation types to take care of the mismatches between Mlabri and Tin since there are a total of 23 correspondence types involving long versus short vowel and *-h*, *-ʔ*, or open syllable, if one wants the coverage of the data to be exhaustive. There is not, to my knowledge, any shred of evidence in favour of such constructs in the history of languages such as Mlabri and Tin.

Mlabri (particularly MlaA) has a forceful breathiness often used as

an affective marker on words denoting scary or funny animals (such as the crocodile, *kroːc̣*, and the duck, *kaːp*), as well as verbs denoting colossal size (*baj tʰaːl* 'big'). Clearly, that has no historical connections with the mismatches in Tinic rhymes.

In lack of evidence I altogether discard ancient syllable prosodies as an explicate with respect to syllable structure in Mlabri and Tin.

For the time being I shall just define an abstract symbol for each of the offsets in laryngeal syllable-termination and beg the question what Proto-Tinic really looked like in these respects. The total set of mismatches includes Mlabri *-h* versus open syllable in Tin, but as will be shown later this is in fact a morphological difference. Therefore, the only genuine mismatches have to do with final glottal stop versus open syllable. There are mismatches both ways, and these are symbolized as follows in my reconstructions of Proto-Tinic:

$$H_1 = \text{Mlabri: } \textit{ʔ}, \text{Tin: } \textit{open syllable}$$
$$H_2 = \text{Mlabri: } \textit{open syllable}, \text{Tin: } \textit{ʔ}$$

As I shall show in Part 3, the only way to approach a more insightful analysis is to go beyond Mlabri and Tin, i.e. to look at rhymes in other Khmuic languages and in languages beyond Khmuic.

The pattern of vowel qualities

Vowel qualities can be handled in terms of the traditional two "dimensions": a complex dimension involving place of articulation along the tongue body and degree of lip-rounding and a dimension generally referred to as tongue-height although it is really assessed auditorily. The vowel qualities combine with a binary quantity system of long and short.

(*Digression:* I have found no need for additional dimensions such as tense-lax in specifying the vowels of these two languages, except that in Mlabri there occurs a weak, optional schwa-vowel in pre-syllables which I rendered as raised "ᵊ". In transcriptions of Tin I use the vowel symbol "ə" to refer to a vowel which is similar to Mlabri "ɤ" but sounds less retracted and may be central rather than back. Finally, it deserves mention that in my *Minor Mlabri*, 1995, I represent the vowels preceding nasals in some words as high and lax but I would now restate these vowels as belonging to the next degree of openness.)

Place of articulation and lip-rounding. Both Mlabri and Tin have a three-way distinction between front-unrounded, back/central-un-rounded, and back-rounded vowels.

Degrees of openness (aperture) or vowel height. Basically, one can distinguish three or four degrees of vowel height. Tin has three, Mlabri has four in that Mlabri *ʌ* and *a* represent two distinctive degrees of height corresponding to Tin *a*. In addition to, or perhaps as a result of, the resulting difference in overall arrangement (topology) of the two systems there is a certain phonetic offset in height, the Mlabri vowels at each level of "auditory height" being higher than the Tin vowels at the structurally corresponding level (I shall return to that in Part 3 below).

Vowel phoneme inventories. Disregarding length, there are nine distinctive vowels in Tin: /*i e ɛ ɯ ə a u o ɔ*/ and ten in Mlabri: /*i e ɛ ɯ ɤ ʌ a u o ɔ*/ (diphthongs are left out of consideration in this exposition since they do not form a crucial part of the comparative data used to test the Tinic hypothesis). – A note on phonetic notation (see further Appendix II): the unrounded central/back vowels are rendered as *ɯ ɤ* in my transcriptions of Mlabri but as *ɯ ə* in my transcriptions of Tin. When talking about reconstructed Proto-Tin I use the symbols **i *ə*. The differences between these three sets of symbols are entirely insignificant for the correspondences and reconstructions.

Tentative vowel chart for Proto-Tinic. Considering the considerable overall similarity between the vowel systems of Tin and Mlabri it is a natural choice of working hypothesis to posit the common core of the two systems as a first approximation to the Proto-Tinic vowel system. At the lower end of the system, however, the hypothetical common denominator must necessarily look different depending on whether the extra vowel in Mlabri is supposed to be a feature specific to Mlabri (whatever its explanation) or supposed to reflect the existence of a Proto-Tinic vowel phoneme which is no longer present in Tin.

At this stage, before the relevant etymological data have been presented and discussed, the options must of course be kept open. It is necessary to recognize two different but equally possible configurations of the vowel system that is tentatively reconstructed for the Proto-Tinic stage, namely one corresponding to that of Modern Tin and another corresponding to that of Modern Mlabri. These alternative systems are shown in Table 2-5.

Table 2-5: Proto-Tinic vowels

Alternative I			**Alternative II**		
*i	*$ɨ$	*u	*i	*$ɨ$	*u
*e	*$ə$	*o	*e	*$ə$	*o
*$ɛ$	*a	*$ɔ$	*$ɛ$	*$ʌ$	*$ɔ$
				*a	

The problem with positing a reconstructed vowel system at this point is that it is so far typological rather than etymological. The moment one looks at cognates shared by Mlabri and Tin it becomes apparent that the vowel correspondences are not one-to-one; there is a variety of pairwise correspondences. Still, all the sets of etymologically corresponding vowels in Mlabri and Tin share an important feature: the vowels rarely differ by more than one degree of openness (such as $e \sim \varepsilon$ or $u \sim o$) and mostly agree on place of articulation and rounding. Front-back and unrounded-rounded correspondences (such as $e \sim ə$ or $ʌ \sim ɔ$) do not occur with the high vowels but only in the lower part of the vowel space where such fluctuation is to be expected on general phonetic grounds (poor auditory discrimination in the lower part of the vowel space).

I therefore conclude that there is an underlying number of (overlapping) regions in the vowel space each defining a vowel correspondence between Mlabri and Tin but that there is a priori no simple way to relate the vowels of Mlabri and Tin to each other in terms of sound-laws, at least not exhaustively. Accordingly, I prefer at this point to provide each of these regions or correspondence sets with a label to be used in reconstructed forms. The possibility of generalizing over vowel correspondences will be taken up in Part 3.

Correspondence sets. In Table 2-6 I list – without further analysis at the moment – the total set of vowel correspondences between Mlabri and Tin, as symbolized in my reconstructions of Proto-Tinic. There are a total of 29 correspondences according to vowel quality alone, on which are superimposed four possibilities of vowel-length correspondence (Mla $V \sim$ Tin V, Mla $V: \sim$ Tin V, Mla $V \sim$ Tin $V:$, Mla $V: \sim$ Tin $V:$). Note that there is no correspondence Mla $ʌ \sim$ Tin $ʌ$ (which occurs only in the diphthong $iʌ$).

Table 2-6: Mlabri-Tin vowel correspondence sets

a = Mla a, Tin a	i = Mla $ɯ$, Tin $ɯ$	$ɔ$ = Mla $ɔ$, Tin $ɔ$
$ɛ$ = Mla $ɛ$, Tin $ɛ$	$Ä_1$ = Mla $ʌ$, Tin a	o = Mla o, Tin o
e = Mla e, Tin e	$Ä_2$ = Mla $ʌ$, Tin $ə$	u = Mla u, Tin u
i = Mla i, Tin i	$Ä_3$ = Mla a, Tin $ə$	O_1 = Mla $ɔ$, Tin o/u
E_1 = Mla $ɛ$, Tin e	$Ë_1$ = Mla $ɤ$, Tin o	O_2 = Mla o, Tin $ɔ$
E_2 = Mla e, Tin $ɛ$	$Ë_2$ = Mla $ɤ$, Tin $ɔ$	O_3 = Mla $ɔ$, Tin a
E_3 = Mla $ɛ$, Tin a	$Ë_3$ = Mla $ɤ$, Tin e	U_1 = Mla o, Tin u
E_4 = Mla a, Tin $ɛ$	$Ë_4$ = Mla $ɤ$, Tin a	U_2 = Mla u, Tin o
I = Mla e, Tin i	$Ï_1$ = Mla $ɯ$, Tin $ə$	U_3 = Mla u, Tin $ɔ$
$ɤ$ = Mla $ɤ$, Tin $ə$	$Ï_2$ = Mla $ɤ$, Tin $ɯ$	

The index numbers used on the abstract symbols above are entirely arbitrary; they not even proceed incrementally along the relevant phonetic parameter but just reflect the order in which I happened to identify and label each correspondence set in the first place.

Diphthongs

When talking about diphthongs in languages such as Mlabri and Tin one can posit both falling diphthongs (i.e. with falling sonority) and rising diphthongs (with rising sonority). In this section I shall deal with these diphthongs in much more detail than other features of Mlabri-Tin phonology because that is necessary in order to introduce the topic *register*. Otherwise, however, diphthongs play a very marginal role in the data sets above, and therefore, questions such as the origin of diphthongs or the relationship between Mlabri monophthongs and Tin diphthongs are hardly addressed in Parts 3 and 4.

Falling diphthongs. The falling diphthongs glide from a non-high vowel quality towards a high vowel quality, e.g. [ai, aɯ, au]. It is often an analytical question whether such phonetic diphthongs should be interpreted as unit vowels of diphthongal realization, as "true" diphthongs, i.e. sequences of two vowel phonemes within one syllable, or finally as VC-sequences ending in a glide (the relevant glides in a SEA setting being [i̯, ɯ̯, u̯], in more convenient notation: *j*, *ɣ*, *w*). In handling Mlabri and Tin the third solution is the most straightforward, since it explains why final consonants (following after the diphthongs) are

structurally impossible in words with such diphthongs. Thus the falling diphthongs can be approached diachronically in the same way as other *VC*-sequences.

Rising diphthongs: introductory remarks. The rising diphthongs glide from a high vowel quality [i , ɯ, u] towards a non-high vowel quality (typically [a] but also [ɛ, e]). It is a problem especially with Modern Mlabri that there is sometimes vacillation between initial sequences of the types *consonant* + *high vowel* + *vowel* (i.e. *CVV*-) and *consonant* + *glide* + *vowel* (i.e. *CCV*-). For example, [miɤː, mjɤː] can both be heard in MlaA as realizations of a meaning 'wife', and [kuɤj, kwɤj] can both be heard as realizations of a word meaning 'cultivated banana'. These are both borrowings from Tai, however. Words of other provenance are normally (always?) pronounced consistently one or the other way, *either* with an initial cluster *or* with a diphthong, e.g. MlaA *rwɤt* 'pour water over something' versus *ruʌt* 'an indentation in the ground which is sometimes muddy'.

Both Mlabri and Tin exhibit variation over diphthong versus monophthong, but in Tin that is a matter of dialect-specific variation; in Mlabri it may occur only in Tai loans and reflects developments in the lending languages.

There is an important phonetic characteristic associated with initial *rw*- in Mlabri: the onset has largely the same pitch from the very beginning and it is continued into the following vowel. I consider this a specific characteristic of Mlabri, and it contributes to making it difficult for the field worker to discriminate auditorily between sequences such as *rwV*- and *ruV*-. To the Mlabri, on the contrary, the characteristic pitch is essential for the recognition of the onset as a Mlabri syllable-onset (I have tried a couple of times to pronounce *rwaːj* 'tiger' – a M-K word shared with Khmu – with non-characteristic pitch, namely with a slight pitch rise from the trill to the glide + vowel; the reaction of the Mlabri listener was that I was saying the word not in Mlabri but in Khmu).

Rising diphthongs from a Tinic perspective. Modern Tin has a set of rising diphthongs some of which Filbeck (1978) projects back to Proto-Tin, and Modern Mlabri has a much richer set, which also exhibits a contrast between short and long diphthongs (e.g. *ia* versus *iaː*, *uɤ* versus *uʌː*). Several of these diphthongs in Mlabri or in Tin are found in loanwords of Tai origin but some occur (in Mlabri or Tin or both) in words of other provenance. Most of the diphthongs found in Mlabri or Tin

words of non-Tai provenance cannot, however, be reconstructed as continuations of diphthongs at the Tinic level because the words in question fail to occur in the cognate pairs I have identified as being of Tinic provenance.

The only safely attested diphthong at the Tinic level is *ia*. There is a single word pair in my data in which this Tinic diphthong is seen to be a reflex of a Khmuic diphthong *iə*, namely the word for 'root': Mlabri *rɛːlʰ*, TinMalA *ᵑgriʌjʰ*. At first glance the Mlabri and Tin forms may not seem to fit well, but the finals exhibit a regular correspondence between Mlabri and TinMal (their common origin being *-S₁), and the complexity of the initial in Tin is suggestive of morphological complexity, maybe compounding with the etymon in question as the second part. One has to go outside Tinic in order to see that this etymon had a diphthong already at the Khmuic level. The Khmu form is *riəs*, and if going beyond Khmuic one also encounters forms with a monophthong, e.g. Reconstructed Monic *ris* (Diffloth 1984: 86, 295).

Mlabri *ɛː* is sometimes matched by Tin *ɛː* instead; in such cases the reconstruction of Tinic *ɛː* seems straightforward. Examples are: Mla *dimrɛːŋ*, TinPrai *ᵐbrɛːŋ* 'bamboo flooring' (with a modification of the beginning of the word that is typical of Tin) and Mla *çm̩bɛp*, TinMal *ᵐbɛːp* 'lip region' (with a vowel-length difference, which complicates the retrieval of a Tinic proto-form).

What remains now is a considerable number of cognate pairs in which Mlabri *ɛː* is matched by Tin *iʌ*, like in the word 'root' above but with a different Khmuic origin. An example is Mla *bɛːr*, Tin *piar* 'two', the Khmu form *baːr* (and similar forms beyond Khmuic) suggesting that one should reconstruct a long vowel *aː* for Proto-Khmuic.

The evidence thus suggests that Khmuic *iə* and *aː* coalesced in Early Tinic into a long or diphthongal entity continued as Mla *ɛ* ~ Tin *iʌ*. One might contemplate reconstructing a long monophthong at the Tinic level, were it not for the existence of Tinic *ɛː* continued as such both in Mlabri and Tin (see above). The most plausible options, then, is to reconstruct a diphthong in the *ɛə-ɛʌ-ɛa* range (with subsequent monophthongization in Mlabri and with subsequent differentiation in Tin) or a diphthong in the *iə-iʌ-ia* range (again with subsequent monophthongization in Mlabri). I have chosen to reconstruct this entity as *ia* at the Tinic level because that agrees well with what one observes in some other languages of the Khmuic branch, such as Ksingmul and Phong.

Ksingmul has *iə* corresponding to Mla *ɛː* ~ Tin *iʌ*, e.g. KsM *ziəŋ*

'house' ~ Mla *geːŋ* ~ Tin *kiʌŋ*, or KsM *ʔəmiə* 'rain' ~ Mla *mɛːʔ* ~ Tin *miʌʔ*. Phong has a monophthong *i* that is readily explained as a truncated reflex of *ia or *iə (< *aː), although one cannot exclude that it is a more direct reflex of Proto-Khmuic *aː (with raising and shortening only: *aː > i). The word for 'blood', Khmuic *maːm, for example, occurs as Mla *mɛːm* ~ Tin *miʌm* ~ KsM *miəm* ~ Phong *mim*.

The conclusion, then, is that one can posit a seemingly regular sound-change *aː > *ia (or possibly > *iə) from Proto-Khmuic to Proto-Tinic. Since the diphthongization is an innovation that is shared by several Khmuic languages but not Khmu, one might consider pushing it even further back and let it define a major sub-branch of Khmuic that did *not* include Khmu. The diphthongization of *aː into a rising diphthong is prototypical in the presence of *register* in M-K; Diffloth (1980: 42) states that Wa has *ea*, *eə* or *ɛ* from *a after a proto-voiced initial and points to a similar process *aː > *ea*, *iə* in Khmer (ibid: 37). I shall return to the register issue below.

Other examples of this early diphthongization in my comparative data-set are:

Mla *bɛːk* 'bear, n'; Tin *piʌk* 'id.'
Mla *bɛːr* 'two'; Tin *piʌr* 'id.'
Mla *tʰrɛːŋ* 'tooth'; TinMalA *tʰiʌŋ* 'id.' (Khmuic *sr-)
MlaBC *ɟiʔɛːŋ*, MlaA *ɟəʔɛːŋ* 'bone'; TinMalA *ʔiʌŋ*, Tin Prai *siʔiʌŋ* (etc.) 'id.'

The fairly small number of words with Mla *ɛː* ~ Tin *iʌ* in my data-set might convey the impression that *ɛː* is rare in Mlabri, and that *iʌ* is rare in Tin. That is certainly not the case; there are numerous such words in Mlabri and Tin considered separately, it is just that I have found rather few cognate pairs.

So far, so good but the evidence is conflicting since there are also several instances in which Mlabri and Tin agree on *aː* as the continuation of the vowel *aː found in Proto-Khmuic, thus strongly suggesting Tinic *aː as another regular reflex. Examples are:

MlaAB *ɟraː* 'skinny'; TinMlaA *kraː* , TinPrai *siraː* 'id.'
MlaB *gaːp* 'hold something in the mouth'; TinPrai *kaːp* 'id.'
MlaB *gwaː* 'search for; hunt for'; TinPrai *kwaː* 'id.'

MlaB *taːɲ* 'weave'; TinMalA *tʰaːɲ*, TinPrai *tʰaɲ* 'id.' (vowel-length is
    confirmed by Khmu *taːɲ* 'weave; plait')
Mla *kaːʔ* 'fish'; Tin *kʰaː* 'id.'
Mla *kwaːj* 'edible tuber'; TinMal *kʰwaːj* 'id.'
MlaC *tʰəmaːl* 'spirit of deceased person'; TinMalA *ʰma(ː)l* 'id.'; TinPrai
    simaːl 'personal spirit'
Mla *raːp* 'chase; run after'; TinMalA *raːp* 'id.'
Mla *rwaːj* 'tiger'; TinMalA *rwaj* 'id.' (Khmu also has *rwaːj* 'tiger')
Mla *laːʔ* 'shoulder'; TinMalA *pak laː* 'shoulder'; TinPrai *laː* 'arm'
Mla *rŋgaːp* 'mouth'; TinMal *ŋgaːp* 'id.'

It is immediately apparent that there are a variety of onsets in both sets.
Before focussing on the role of the onset it may therefore be worth
while to explore whether the presence or absence of diphthongization
could have something to do with the syllable termination. There is a
preponderance of *final nasals* in the first set, which makes one wonder
whether they might possibly have promoted a change in vowel quality
(> diphthongization) via nasalization of the preceding vowel. In general
phonetic terms, perceptible coarticulatory prenasalization of an open
vowel [a] is a well-known phenomenon. The data supporting the
assumption that final nasals played a role, however, amount to just three
words. There are just as many words in the set that do *not* end in a nasal.
I shall therefore leave this issue unsolved at the moment.

Register in Proto-Tinic

The most fruitful hypothesis about the Tinic split of **aː* into *aː* ~ *ia* is
that diphthongization is a reflex of so-called register. This entails the
hypothesis that the *phonation type of voiced initials* at some point in
time affected the vowel and eventually caused it to change quality. This
is a widespread phenomenon in M-K, but there are language-specific
differences as to which onset types caused diphthongization of an [a]-
vowel. Diffloth (1980: 39) makes precise that voiced onsets have dif-
ferent histories and thus count as different from a register perspective:
their effect on vowel quality may depend on whether they were origi-
nally voiced or whether they were reflexes of preglottalized or preaspi-
rated sonorants (nasals, liquids).

    Returning to Mlabri, the examples with diphthongization in the first

set have something relevant in common: there is a voiced stop or a sonorant or a glottal catch before the vowel at the Pre-Tinic level, which seems to support the hypothesis. Unfortunately, the other set, with unchanged *aː, contains not only voiceless stops but also voiced stops and sonorants as onsets. Clearly, it was far from always the case that *aː diphthongized after onsets with a phonation type that might affect vowel quality.

This forces me to return to the possible contribution of final nasals to incipient diphthongization. Within the small set of Tinic data at my disposal it seems that Khmuic *aː diphthongized most regularly if the vowel was affected by initial voicing and final nasality at the same time. That looks like a spurious generalization since there are three words which have undergone diphthongization *without* ending in a nasal. At closer inspection, however, these all have something special about them though it is hard to present a convincing case for register causing dipthongization in these three words in particular. I shall take them one by one:

The word for 'bear' is *bɛːk* in MlaAC whereas MlaB has two forms: *biuk* and *bɛːk*. A transition (*aː) > *ia > iu* does not occur in any other Mlabri word I know of, and that suggests that there may have been something aberrant about the early vocalism of this Tinic animal name. Phong, however, has *biʔ* with the expected reflex of diphthongization of *aː (*-k > -ʔ* is expected) and thus contradicts the assumption that there was something special about the word. There is a different word for 'bear' in Khmu and in Ksingmul; the word list available for Khabit has no such entry.

As for the word for 'two' above there is nothing strange about it from a Khmuic perspective, but at the level of comparative Mon-Khmer its onset has a complex history. The *b-* in Khmuic is the reflex of an old prefix causing glottalization of the initial voiced stop in some other M-K. Maybe it was continued in Tinic as a strongly voiced onset or cluster (e.g. a prenasalized voiced stop which was later simplified into a plain voiced stop)? The evidence from Phong is confusing; the list by Thê´ (2000: 223, 226) gives two alternative forms *bi* and *par*. Supposing that the latter is a borrowing from a Khmu dialect with devoicing, the remaining form *bi* has the expected reflex of diphthongization of *aː.

Finally, the word for 'rain' can be reconstructed at the Khmuic level with a presyllable beginning in a velar stop, which is still preserved in the Khmu form *kmaʔ*; there are also vestiges of it in Ksingmul *ʔəmiə* (possibly a notational error for *ʔəmiəʔ*). I have no explanation of why it

has vanished in Tinic. Phong has *kmiʔ* with the same onset as Khmu and with the expected reflex of **aː* due to register. There is something genuinely exceptional about the behaviour of the onset in Mlabri, however: a voiceless stop + nasal should have undergone an otherwise entirely regular sound-change: **km-* > *ʰm-* and not surface as plain *m-* in Mlabri. Maybe Pre-Tinic had a presyllable before the nasal so that voicing was somehow reinforced and conditioned the effect of register although the material had shrunk to **m-* already in Proto-Tinic.

These explanations are speculative and unconvincing. One could envisage other scenarios for these words which, on the contrary, would prevent diphthongization.

It would take a comparative investigation involving Khmuic languages on a broad scale to see whether there is some general, old feature that distinguishes the words that fail to undergo expected diphthongization from those that do. Maybe the explanation is simply that it was only an incipient sound change which did not make it across the lexicon. Early borrowing e.g. from Khmu or even from a non-Khmuic Northern M-K language into Pre-Tinic on the one side, versus retention of old indigenous vocabulary on the other side, might also be an explanation.

So far, all I wish to conclude is that **aː* > **ia* is likely to be *an irregular reflex of a register difference* that was operative in Early Tinic and some related languages and then vanished.

Concluding remarks: the conservative character of Mlabri phonology

Although the irregularities in phonological correspondences between Mlabri and Tin will only be presented in full (with attempts at solutions) in Part 3, the information given so far is hopefully indicative of the distinctly conservative character of virtually all aspects of Mlabri phonology. This appears very clearly if one compares Mlabri with reconstructed stages of other branches of Mon-Khmer. Proto-Waic is a case in point; I shall just quote a single form with discussion of its Mlabri parallel.

PW **ʔriʔ* 'deep' is discussed by Diffloth (1980: 77), who reconstructs an underlying more complex form in Pre-Waic and assumes that it has split up into two words with ensuing reanalysis and loss of the first part: Pre-Waic **ɟaruʔ* > *ɟəʔ ruʔ* > PW *ʔriʔ*. Mlabri has *ɟruʔ* 'deep

(said of water)' in complete accordance with reconstructed early forms of this etymon.

Morphology in Mlabri and Tin

In a M-K context, morphology means word-formation in terms of affix-ation, reduplication, and compounding. There are only three Khmuic languages that are amply documented in print: Khmu, Ksingmul and Mlabri. These, then, are the languages on which attempts to reconstruct Proto-Khmuic morphology must be based in the first place. Ksingmul, at least as documented by Solntsev et al. (1990), is clearly influenced by Vietic and very different from the more westerly Khmuic languages under consideration in this monograph. Thus we are largely left with Khmu and Mlabri in attempting to picture what Proto-Khmuic morphol-ogy was like.

As for the remaining languages classified as Khmuic the available data (at least the data available to me) give little evidence for patterns of word formation; most of those languages are only sparsely documented in terms of short word lists. Looking into Tin, studying dialects of both Mal and Prai, I have found a conspicuous paucity of words of transpar-ent morphological complexity, the vast majority of words being unana-lysable monosyllables. One cannot conclude from that, however, that Proto-Tinic had little or no productive morphology; the paucity of morphology in modern Mal and Prai might be due to a change of lan-guage type on the way from Proto-Tinic to proto-Tin. On the other hand, one must keep the possibility open that even features of Mlabri morphology that are clearly of M-K origin may have a non-Tinic source.

In the context of this study the immediate challenge is to decide to what extent Mlabri morphology can be projected back to Proto-Tinic. There will be a strong case for internal reconstruction of Tinic morphol-ogy if there are structural peculiarities in Tin words that can be explained in accordance with Mlabri morphology even if a considera-tion of Tin in itself does not lead very far.

Khmu is, as always in a Khmuic context, a tempting frame of refer-ence though at the same time a possible pitfall. Khmu morphology is richer than Mlabri morphology but the latter is more streamlined. One cannot a priori exclude the possibility that Mlabri borrowed some of its morphology from Khmu along with vocabulary but that seems rather

unlikely. Such borrowing of word-formation patterns would be inconceivable without extensive lexical borrowing but there is no evidence for very extensive lexical borrowing from Khmu at an early time. On the contrary, the overall impression, both when looking at Mlabri and when looking at Tin, is that these languages differ very much in lexicon.

It is a more fruitful hypothesis that Pre-Tinic morphology was basically similar to Khmu morphology but had begone to get simplified at the time of Proto-Tinic, i.e. before Mlabri and Tin split company. That trend, then, was continued in Tin proper but halted in Mlabri. The few remarks below are based on this hypothesis.

There is an abundance of sesquisyllabic words in Mlabri, most of which have presyllables that do not seem to match anything in Khmu, for example. Most of these words are not only of opaque morphological constituency but so far also of unknown etymology, so it is a tempting assumption that they stem from the non-Tinic layer in Mlabri. That, in turn, would suggest that the non-Tinic layer has a more remote Mon-Khmer origin since sesquisyllabicity is above all a characteristic of M-K.

It is beyond doubt, however, that some of the opaque words in Mlabri are Tinic and that they have various morphological origins. That suggests that Pre-Tinic morphology was rich in such word-forming devices as affixation, reduplication and compounding, which flourish to varying degrees in modern M-K languages. If so, the morphology we have in Mlabri is just a relic of Pre-Tinic morphology, even if there is much more left than in Tin. Productive word-formation was already simplified at the Proto-Tinic level.

Compounding

Compounding creating bisyllabic and trisyllabic words occurs in Mlabri but not necessarily as a carry-over from Tinic.

The distinction between phrases and compounds is not well-defined in Mlabri except if compounding leads to word-forms similar to bisyllabic or even sesquisyllabic simplex words, and then the identification of such a form as a compound is a matter of etymological rather than descriptive analysis. Such phonological compounds may arise from juxtaposed nouns, as in conservative MlaA *kuurmut, kr̩mut* 'tempest' from *kuur* 'thunder' + *r̩mut* 'wind', or from a juxtaposition of serial verbs, as in MlaA *huurleh, hr̩leh* 'pull out; drag out' from *hɤr* 'pull;

drag' + *leh* 'come'. By painstaking analysis one might succeed in iden-
tifying quite a few strange prefixes in Mlabri as being due to this histor-
ical process.

There is one area of Mlabri lexicon in which it is particularly inter-
esting and rewarding to pursue morphological analysis of complex
wordforms: **kinship terminology**. Going into that here would exceed
the bonds of the present monograph, however (some aspects of kinship
terminology are mentioned in Part 4 below because they point beyond
Tinic and even beyond Khmuic).

Both in Mlabri and Tin, a noun phrase of the structure N+N will typ-
ically be interpreted as head plus attribute but in Tin the second con-
stituent can function as determiner with the possibility of independent
referentiality: TinMalA *kiʌŋ ʔaw* 'father's house' (lit.: 'house' +
'father'), *kʰɛp məj* 'mother's shoes' (lit.: 'shoe' + 'mother'). This is not
the case in Mlabri; a phrase such as *ɛːw braɲ* , lit. 'child' + 'dog', will
be interpreted as meaning 'whelp', whereas Mlabri uses a construction
quite alien to Khmuic: N + + N to encode the meaning of specific pos-
sessor + possessum: *braɲ di ʔɛːw* 'the dog's young' (for referentiality in
Mlabri see further Rischel 2006). Therefore, a Mlabri expression such
as *mlaːʔ briːʔ*, lit. 'man; person' + 'forest' will be interpreted as head +
non-referential attribute, and especially if such an expression is used
frequently it will be pronounced as a phonological whole, with short-
ening of the first vowel. It can even be contracted into one word by loss
of the final glottal stop: *mlabriːʔ*.

The type of noun phrase exemplified by *mlaːʔ briːʔ* is another source
of phonological compounds in Mlabri, but in this case their composi-
tion is entirely transparent, and in their phonological realization there is
a continuous transition from phrase to compound.

It is not my impression that compounding occurs regularly in Tin, at
least not to the same degree. Tin seems to be characterized by simplex
words constituting the vast majority of its M-K lexicon.

Reduplication is fully alive in Mlabri; for a detailed exposition of redu-
plication types with examples, see Rischel 1995: 92-98, for its semantic
function also see the section on Expressives below. Here I shall just
mention that there is a range of types from (i) words consisting of two
syllables of which one is a full echo of the other, e.g. MlaB *bohboh* 'be
boiling' (this type is found mostly with verbs) over (ii) bisyllabic words
of which the first syllable has a simplified onset and a vowel in "ablaut"

relationship to that of the next syllable, e.g. *kupkwap* 'toad species', to (iii) sesquisyllabic words with a very reduced presyllable followed by a full syllable, e.g. *ḳlkiːl* 'knee'. There are also quasi-reduplicatives in which only the onset of the first syllable echoes that of the second, e.g. in animal's names: MlaA *bṃbwaːj* 'spider', *pṃpoː* 'elephant'.

Tin proper, on the contrary, has no preservation of any of these phonological structures in words of M-K origin, so if there was an abundance of reduplicatives in Proto-Tinic they have all shrunk to monosyllables or been dropped from the lexicon eventually.

Affixation

There is both prefixation and infixation in Mlabri, but only relics of these types of affixation in Tin. With the exception of a few complex lexical entries not matched by Tin, the two types of affixation in Mlabri are confined to one morphological category each: *prefixation forming derived verbs from verbs*, and *infixation forming nominalizations from verbs*. In both cases, however, the morphological category in question is of very high functional importance in the language.

Causative formation. Mlabri has just two transparent prefixes: *ba-* and *pa-*; in return those prefixes are fully productive. They form causatives from verbs, e.g. *hot* 'fall' – *bahot* 'let drop', *bul* 'die' – *pabul* 'kill; slaughter'. The two prefixes are clearly variants of each other. This is allomorphy by dissimilation, the voiced onset occurring if the stem it occurs on has a voiceless onset, and the voiceless onset occuring if the stem has a voiced onset. It is an interesting question how this dissimilatory pattern arose; other wordforms in Mlabri show no such behaviour (reduplicatives, on the contrary, show agreement in voicing between the onset of the presyllable and that of the main syllable).

Tin has a small number of verbs pairs differing in that one is derived from the other and encodes the grammatical meaning of causative. The causative formations are only relics; there does not seem to be any such productive formation.

The only causative formation in Tin that occurs in sufficient number to form a pattern, is marked by prenasalization. The following verb pairs have semantically close counterparts in Mlabri (and the stems correspond exactly, in accordance with sound-laws in Tin):

Tin $^mb\partial l$ 'kill; slaughter' ~ Mla *pabul* 'id.'

TinPrai $^nt^hec$ 'tear something' ~ Mla *batac* 'id.'

TinMalA $^\eta guh$ 'ignite the firewood' ~ MlaB *paguh* 'make something light up; illuminate'

If Tin is considered in isolation one would suggest that Tin has a causative prefix consisting of a nasal consonant which assimilates completely to a following stop. Since, however, a round of massive phonological simplifications of onsets happened on the way from Pre-Tin to Proto-Tin, prenasalized stops in Tin may have a variety of sources, and the nasal component might instead be a reflex of a stop or a presyllable, as in Mlabri. In short, if there were no further examples of causative formation in Tin, the Tin and Mlabri forms listed above might well go back to a shared set of forms in Proto-Tinic.

There is, however, at least one crucial example of prefixation of a stop in a causative verb in TinPrai:

lɔːj 'float' – *klɔːj* 'let drift'

(the latter form occurs also in TinMalB but apparently with some semantic shift toward a medial rather than causative meaning). The word form meaning 'float' in Prai occurs identically in Tai but it must be of some age in Tin in order to have acquired a prefixation that is nowadays exceptional.

The causative formation by means of prefixal *k-* is very significant because it is undoubtedly an old Tinic formation. The Khmuic language that seems closest to Tinic, Phong, has preserved various consonantal prefixes and among them a causative prefix *k-*, e.g.

b∂l 'be dead; die' – *kb∂l* 'kill'

t∂p 'bury' – *kt∂p* 'cover with a lid or a stopper'

tuŋ 'be afraid; fear' – *ktuŋ* 'threaten'

(note that the second example is causative derivation from a verb with inherent causative meaning, something resembling the so-called double-transitives i Eskimo).

Khmu has a variety of causative prefixes (Svantesson 1983: 103–111). The ones that exhibit some generality are *pn-* and *p-*. Prefixal *p-* (e.g. in *haːn* 'die' – *phaːn* 'kill') is of rather restricted occurrence and differs more from Mlabri *pa-/ba-* in its phonotactic behaviour than does

pn-. Ksingmul has causative formation with *pə-* (*siən* 'die' – *pəsiən* 'kill'), or *pəl-* (*ˀim* 'be breastfed' – *pəlˀim* 'nurse the baby').

It is significant that Khmu and Ksingmul side with Mlabri in having causative prefixes with an initial labial stop, in contrast with Phong and Tin where it is a velar stop *k-* that performs the function of prefixal causative formation (with at least one example in Tin).

There are two word-forms within my limited Tin data in which the causative is marked by *infixation* after the first of two initial consonants. Their Mlabri cognates are formed like other causatives, i.e. with the prefix *pa-/ba-*.

One of the Tin verbs with causative infixation is marked by aspiration of an initial stop, presumably as a reflex of an *h*-infix:

TinMalA *pʰlah* 'release; let escape' ~ MlaA *pablah* 'id.' (the verb base is **blah* which could not possibly yield *pʰlah* in Tin unless something was affixed)

The other verb with causative infixation has *-r-* after the initial stop (the stem from which the causative verb below is derived, occurred with another causative formation in a TinMalA form cited above):

Tin Prai *kruh* 'ignite' ~ MlaB *paguh* 'make something light up; illuminate' – cf. TinMalA *kuh* 'be ablaze' ~ Mla *guh* 'id.'

Relics of morphology in Tin must necessarily go back to Tinic unless the derived form as such is a loanword. I have found no basis for explaining any of the causative verbs in Tin as having been borrowed in causative form.

I conclude that Tinic had several ways of forming causatives, although their preservation in Modern Tin is so rudimentary. The causative affixes attested in Tin fail entirely to occur in Mlabri; the phonotactic behaviour of *-h-* and *-r-* is, moreover, alien to Mlabri affixation.

In this case, then, it is Mlabri that shows aberration from Tinic morphology by having *pa-* ~ *ba-* as the only way of forming causative verbs. This affix pair may not even have occurred in Proto-Tinic. It is a possibility that causative **p-* occurred in Tinic along with **k-* but there is no way to disclose that by looking at the Tin forms with prenasalization; they would come out the same no matter whether the causative prefix used to have a labial or a velar stop.

The most striking observation is, in fact, not that Mlabri has a labial in the causative prefix but that causative *k-* as well as causative infixation *is totally absent in Mlabri*. That is significant for an overall assessment of the extent to which Mlabri morphology is at all a continuation of Tinic morphology.

From a comparative perspective, causative formation by means of an initial labial is of high age in Mon-Khmer. Outside Khmuic, but still inside Northern Mon-Khmer, it occurs for example in Palaung as evidenced by reconstructed Proto-Palaung **jam* 'die' – **pjam* 'kill' (Mitani 1979). This is an "archaic" feature shared by Mlabri, Khmu and Ksingmul.

Nominalization. Mlabri has a set of nominalizing affixes of prototypical M-K phonology: *-r-*, *-rn-*, and *-mn-*, which occur in complementary distribution depending on the phonotactics of the derived verb (the distribution in all varieties of Mlabri is as stated in my *Minor Mlabri*, 1995:87-89).

The inventory of infix variants in Mlabri is reminiscent of Khmu but Khmu does not have the pattern of complementary distribution found in Mlabri. In Mlabri such infixation is used to form instrument nouns, e.g. *klap* 'hold something (e.g. food) by squeezing it between two prongs' – *kr̩lap* 'bamboo biceps', but also to form abstracts, e.g. MlaAC *tʌɲ* 'speak' – *tm̩nʌɲ* 'conversation', MlaB *glaʔ* 'speak' – *gr̩laʔ* 'conversation'.

Clearly, Mlabri has merged different infixes and now uses them as allomorphs of a generalized nominalizing infix. Jacobs (1963: 69) in comparing Old Mon and Khmer found a partially shared repertory of nominalizing infixes, which among others contains *-r-*, *-rn-*, and *-mn-* like in Mlabri, but with partially different sets of infixes for different purposes: nominalizing in general, denoting utensils, and denoting agents. It is, however, not so that each of the three affixes above has the same specific function(s) across Mon and Khmer.

Thus there may have been quite early tendencies in Mon-Khmer to let their nominalizing functions overlap, although some kind of regularity was re-established in individual languages.

There are also instances of other infixes in Mlabri (cf. below) but the deverbal, nominalizing infix with its three allomorphs is the only one that pervades the lexicon.

In Tin I have so far found *no* traces at all of deverbal infixation. Instrument nouns are not formed from verbs in this language; they are

etymologically unrelated to the semantically related verbs, as in TinPrai *wiː* 'comb, n' versus *saːc* 'comb, v' (*wiː* is a Tai loanword), unless one used a phrase to encode the meaning of the instrument noun, e.g. tinPrai *kwaːt* 'sweep' – *ɲuː kwaːt* 'broom' (explained as a fibre bundle used for sweeping).

It is unthinkable that Tin, being a full-fledged M-K language, did not have nominalization by infixation some time in its past, but that part of the morphology has been lost. It may or may not have been similar to the nominalization pattern in Mlabri.

Again, one can conclude that although Mlabri possesses old M-K morphology there is no evidence that it got it from Tinic.

Other affixation. The word triple MlaB *kreːl* 'short (said e.g. of a log)', MlaAB *k̩lreːl* (A:) 'short (said e.g. of a log)' – (B:) 'brief (of duration)' – MlaAB *krareːl* 'a couple of items, e.g. logs, of short length' shows relics of two additional affixes *-l-* and *-ra-*. The word pair MlaAB *klap* 'hold by squeezing' – MlaB *kn̩lap*, a component of expressions referring to a door that closes along one side, show a relic of an additional affix *-n-*. There are undoubtedly relics of other infixes as well in fossilized forms in Mlabri considering the very great number of words of the structure CC^{syll}CV(ː)(C) that exist in the language.

It should be noted that these additional infixes (also cf. Rischel 1995: 84-85) are in agreement with an important generalization to the effect that transparent infixation in Mlabri always causes sesquisyllabicity. I therefore consider it likely that they belong to a genuine but non-Tinic part of Mlabri morphology. Unfortunately, the data seem both too limited and too heterogeneous for an analysis of the function of *-l̩-* and *-n̩-* in word formation, at least at on the basis of Mlabri alone. The *-ra-* infix will be considered briefly in the section on Expressives below.

Expressives

There is a widespread use in Mon-Khmer languages of iconic word types: words the phonology of which signals specific connotations. Roman Jakobson's term "expressives" was introduced for such M-K words by Diffloth, who also outlined a general semantic theory about them (1976, 1979, referring back to earlier work). With respect to syntactic function, I think one should emphasize that expressives do not constitute a word-class like nouns and verbs; they may comprise words

of different functions but of course also include words that directly refer to a sound or the swiftness of a movement and thus may be called a kind of interjections. There are very many words of the last-mentioned kind in Khmu, for example.

It is difficult to delimit the class of expressives in a language such as Mlabri, but numerous words in its lexicon satisfy Diffloth's criteria either by having a particular phonological structure or by direct sound symbolism. I shall not go much into the various types of expressives in Mlabri, the main point here being to point to their existence. In Tin, I have not come across many lexical items that I would call expressives except of course for interjections (in a wide sense of the term). I have to admit that I have not been searching specifically for expressive vocabulary either in Mlabri or in Tin.

Reduplicatives. Reduplicated forms with a full echo syllable before the main syllable are typical expressives. As mentioned above in the subsection on reduplication, verbs with a repetitive structure (as in *bohboh*) occur in Mlabri; they may denote repeated or iterative action or have an inchoative meaning. There are also sound-imitating words of this kind, e.g. *tuktuk*, the MlaC word for the sound of lightning. Other words with a full first syllable have what I called vowel "ablaut" above; some of these are nouns denoting animals with a characteristic sound, behaviour or appearance (e.g. *kupkwap*, a toad species, *tuktek*, the MlaA word for a small frog species), others may denote similarity in shape with the referent of the basic noun, e.g. *kitke:t* or *kutke:t* 'ears on pots' from *ke:t* 'ear'. Many reduplications have a strongly reduced vowel in the first syllable; some of these are verbs denoting repeated action, e.g. *kwr̥tkrwɛc* (incidentally with a rare type of metathesis), a MlaA word for scratching repeatedly or intensively, though the simple form *krwɛc* used all over Mlabri denotes basically the same kind of scratching.

Some instances of reduplication (of various kinds) are found repeatedly within a certain semantic field without it being always obvious what the reduplication symbolizes. There are, for example, some such terms for mobile parts of the body, such as *goguh* 'skin', *kr̥lki:l* 'knee', *kukɔʔ* 'neck' = MlaA *ŋlʔŋlɛʔ*, MlaC *lutlat* 'tongue' (from *lat* 'lick', clearly the organ that licks repeatedly).

Affixation. Diffloth also includes some affixation within the realm of Expressives. An example is an infix *-ra-* cited from Semai by Diffloth (1979: 53), who also points to similarities in other languages. Diffloth

assigns it the meaning "simultaneous plural". As mentioned in the sub-section "Other affixes" above, infixal -*ra*- also occurs in Mlabri with a similar meaning. There is, however, only one instance of it which I would call prototypically expressive.

The few examples I have found are mostly from B-Mlabri but one occurs generally in Mlabri: *krareːl* 'a couple of items, e.g. logs, of short length' from *kreːl* 'short', said e.g. of a log' (the simple stem *kreːl* may now occur only in MlaB). The other examples are MlaB *braliːŋ* 'how fresh green!' (said e.g. of a young rice field) from *bliːŋ* 'green'; MlaB *braruul^h* 'for a group to leave' from *bruul^h* 'leave'; MlaB *kralip* , a component of certain expressions meaning 'close a door by queezing' from *klip* 'squeeze something between sharp edges' (retrieved without infix-ation only in MlaAC but as nominalization all across Mlabri: *kṛli(ː)p* 'tool with biting edges').

As for the meaning of -*ra*-, most examples above can be construed to imply simultaneity and that may be true even of the first example since the term *krareːl* is typically used when picking logs. As for plurality, it was very explicitly explained to me that the use of *bruul^h* or *braruul^h* depended on whether one or several persons were leaving. Although plurality may be less evident for the examples *braliːŋ* and *kralip* they are consistent with that generalization as well. The former word was explained to me as being appropriate about a newly grown rice field with many plants (not about a single item), and the latter word is used in expressions that refer to the movement of a door with hinges that *klip*, i.e. close in a more or less biceps-like manner.

Particular phonology. Both the complex initial and the final palatal contribute to giving a Mlabri word such as *krwɛc* 'scratch' a particular flavour. Both in Mlabri and to a lesser extent in Tin one can identify several such words with a conspicuous phonology that can be construed as expressive. As for Mlabri in particular, the final -*j^h* (which has ety-mological ties to a sibilant) may be thought of as expressive since it occurs in words about strange things, often words that have defied my attempts at finding parallels in other languages, e.g. the animal names *ɕwɛj^h* (described as a dark, rather large animal that climbs trees, so far unidentified), *gɛj^h* 'freshwater crab', MlaA *kroːj^h* 'crocodile'; MlaBC *təkɔj^h*, *tukɔj^h* 'giant gecko' (the last-mentioned word looks at first glance like a parallel of the reptile word occurring in Jahai as *takɔj* 'sail-fin lizard'; unless this is a chance similarity, the final has perhaps been transformed into the more expressive sound *j^h* in Mlabri).

I am very hesitant to identify expressives on the basis of syllable phonology, however, because there are likewise numerous words with awkward phonology that have very "ordinary" meanings. The MlaC word $\underline{t}l^htal^h$ 'sparks (from the firewood)' looks expressive not just by its reduplication but by the final consonant as well, but the expressiveness of the latter may be spurious, cf. MlaC $\underline{t}l^hn\Lambda l^h$ 'oblique rafters in a lean-to' for which it is hard to see any connotative meaning (except that it may originally be a metaphor).

Then there are words that are conspicuous by beginning or ending in the same consonant; that structure has a funny character in many languages but hardly in the languages under consideration here. Repeated occurrence of *w* (as in *wouw!*) may have that status, though; Diffloth (1979: 55) explicitly says that it belongs to the phonological patterns that occur in expressives whereas they are "not absolutely excluded from the prosaic vocabulary but very rare". That structure occurs infrequently in Mlabri, and I have not been able to identify it as specifically expressive. The word $\underline{t}l^hwe\textrm{:}w$, for example, is simply an epithet to the word *ple:?* 'fruit' in the term for 'longan fruit': *ple?* $\underline{t}l^hwe\textrm{:}w$. It may have an expressive origin, however.

From the perspective of Mlabri-Tin historical comparison the really interesting issue is generalized iconic use of certain means of expression. In the sections on Mlabri particles below I shall point to the exploitation of certain structures (monosyllables of *(C)V*-structure, final *-h* versus *-?*) which give particles and pronouns a distinct appearance.

In rounding off this section, I wish to emphasize that Mlabri lexicon invites a careful analysis from the perspective of use of expressive phonological devices. Apart from being interesting in its own right, the identification of as many words as possible with expressive phonology may be of help in the search for the roots of the Mlabri language. That is primarily because such words are likely to belong to Mlabri proper, being either remnants of the stratum of Early Mlabri or later neologisms, so that one would be wasting effort in searching for cognates of the expressive forms in Tin or other Khmuic languages (whereas the base forms from which some of them are derived, are obvious candidates for cognacy outside Mlabri).

Conclusion: how much morphology was there in Proto-Tinic?

One cannot get very far in reconstructing salient characteristics of Proto-Tinic morphology. Mlabri and Tin seem to have next to nothing in common by way of morphology, and the only sweeping generalization one can make is that Modern Tin has only rudiments of morphology left.

Clearly, some word-formation morphology must have existed in Pre-Tinic. The misfits between Mlabri and Tin suggest the following scenario: Tinic morphology was partially lost on the way from Pre-Tinic to Proto-Tinic and never restored in Tin. Mlabri, however, consolidated itself with additional linguistic input from another source than Proto-Tinic, restoring the most necessary morphology in a very streamlined form.

As for affixes, prefixes and infixes were reduced to allomorphs of one affix each. Just two morphological categories survived with some generality of use: causative formation by prefixation and nominalization by infixation. In both cases, the form with affixation always differs from the base word by being sesquisyllabic. That regularity, which makes affixation more transparent, does not seem to exist in neighbouring languages; it is a peculiarity of Mlabri which hardly reflects the structure of Proto-Tinic.

There is an abundance of M-K type reduplication in Mlabri, but the above scenario suggests that that feature of the language may not be of Tinic origin either.

Non-trivial elements of Mlabri syntax

Mlabri syntax is of mainstream Southeast Asian type in that it has serial verb constructions, a negative imperative, and classifier constructions. Some ways of expressing oneself have affinities to Tai, as one might expect. An example is the use of the verb that basically means 'give' as a function word, e.g. *bɔːk maʔ* 'tell give' i.e. 'give an order to do so-and-so' (incidentally, I have encountered the etymon *maːʔ* outside Mlabri only in Ksingmul: *mah* 'give'; the gloss has widely different counterparts in different M-K languages).

The lexical material used in Mlabri to implement various syntactic constructions or to encode syntactic relations does not look much like

what one would expect in a Khmuic environment, and even some of the construction types seem alien to Khmuic. (A more compehensive survey of some of the phenomena below is given in Rischel 1995: 133–198 with specific emphasis on B-Mlabri; most of what is said there applies equally to all three varieties of Mlabri.)

Classifiers. In Mlabri, classifiers are used with quantifiers and with certain size nouns (cf. below); probably much less often with demonstratives and thus with specific reference (I have not come across more complex constructions of Tai type such as *ʔʔlm̩baːʔ sɔːŋ toː gʌh* 'cow two body those' i.e. 'those two cows'). Some of the words that can function as classifiers are also ordinary nouns; as such they occur in noun constructions with specific reference, e.g. *mlaːʔ gʌh* 'that man'.

Mlabri classifiers occur mostly in numerical expressions, namely (i) with a small selection of preposed numerals: *mɔːj* 'one', (*bɛːr* or in quantifying mostly) *sɔːŋ* 'two', *pon* 'four' (understood as 'several' in quantification and often translated by the Northern Thai or Lao word for 'five'), (ii) with a circumscribed numeral quantifier: *sɔːŋ* (+ CLASSIFIER +) *ʰloːj* 'three', (iii) with postposed *dəmɔj* 'a single one'.

I have retrieved a total of 26 classifiers in Mlabri. Most of them are ordinary nouns that are also used as classifiers; some are recent loanwords from Tai (Lao or Northern Thai). The number of indigenous words that are solely used as classifiers in contemporary Mlabri, is quite small.

The classifiers are mostly different across the three varieties of Mlabri, except for five terms:

mlaːʔ, a noun meaning 'man; person' and at the same classifier for persons, particular men

gɛːŋ, the standard term for dwellings in MlaAB and at the same time classifier for houses in a village (clusters of traditional shelters of lean-to construction are rather counted as households, i.e. in terms of men: *mlaːʔ*)

joːŋ, a noun with a wide meaning referring to males, particularly senior males, and at the same time used as classifier for male persons

kəplah, the classifier for objects with flat sides and more specifically for sheet-like objects such as blankets (that classifier was listed in Part 3 as a Tinic word)

ʔuːj, a noun meaning 'woman; mother', also used as a classifier for women above infancy

Other classifiers for objects of a certain shape or function are

cɯn, MlaCB classifier for cutting implements

kɛːn, a noun referring to the interior of something but also used as a classifier for containers

klɯʔ, MlaC classifier for pots (a loan from Tin, where the word means 'head')

kwʌŋ, a noun referring to balls and the like and at the same time MlaBC classifier for round objects

kwɛk, a noun meaning 'axe' and at the same time MlaB classifier for time periods

lm̩bʌr, a noun meaning 'leaf' and at the same time MlaB classifier for fruits and crops and for cups (semantic transfer from the Tai classifier system)

pak, a noun meaning 'sheath' and at the same time MlaB classifier for crops of oblong shape

pɯːn, a MlaB noun meaning 'mat' (Northern Thai loan) and at the same time classifier for clothes

pleːʔ, a noun meaning 'fruit; nut' and at the same time MlaB quantifier for small round things such as pills

raːŋ, a noun meaning 'long (in comparison with the width)' and at the same time MlaA classifier for iron implements such as hoes and knives

rm̩bah, MlaAB classifier for paired items such as the two parts of an angled roof

toːn, MlaB classifier for logs and branches (a Tai loan)

tɔːn, MlaBC classifier for section of something oblong or stick-like objects (in MlaC used e.g. about candles)

As for all-purpose classifiers, there are the following two (I have not come across an all-purpose classifier in MlaB):

klɔʔ, used in MlaA (of unknown provenance; also used about husband and wife: *sɔːŋ klɔʔ*)

ʔan, used in MlaC (a Lao loan)

There is only one classifier used specifically about animals

toː, a noun meaning 'body' (a Tai loan), mostly used as MlaB classifier for animals (and disrespectfully for persons, e.g. about young girls)

Some classifiers for persons were listed above because they occur all over Mlabri; there are two other classifiers for persons in general:

k^hon, used in MlaC (a Lao loan)
$labo:?$, used in MlaB (and marginally in MlaC, of unknown provenance)

plus a few classifiers for persons according to age or gender:

$cron$, MlaA classifier for small children
$lame:t$, MlaB classifier for male persons

$lnguh$, a noun meaning 'woman; young girl' and at the same time MlaC classifier for young women
mul^h, a noun meaning 'woman' and at the same time MlaC classifier for women

There is practically no overlapping between the list of classifiers above and the classifiers used in Tin, an exception being Mlabri $kaplah$ – Tin p^hlah. Tin Mal and Prai have a larger number of classifiers than Mlabri and with more semantic differentiation across the kinds of items one would be likely to quantify in a village society. The Mlabri quantifiers clearly focus on persons and on some very basic implements and utensils. Those are the items that are essential to count in a subsistence culture such as their traditional hunter-gatherer lifestyle.

That peculiarity of the use of classifiers in Mlabri is in fact consistent with an ancient Mon-Khmer affiliation since the developed use of quantifiers probably entered Mon-Khmer from Tai at a comparatively recent time.

Negative particles/auxiliaries. The form of the negative imperative auxiliary is $g\Lambda m$. That etymon is of M-K provenance; it is also found in Khmer (kom) but does not suggest a Khmuic affiliation. Other negatives, in the general sense of 'not', are strangely different all across the gamut of Khmuic languages considered in this monograph, e.g. Khmu $?am$, Tin $?aj$, Phong ci. Mlabri has a variety of negatives: **met**, **ki**, cak which are used in different types of constructions and used somewhat differently across MlaA, MlaB, and MlaC; Mlabri also has a Lao borrowing $kobo$, which in some usages is mixed up with ki ($> kibo$

and even *kibi*), as a default negative particle. The word *met* or *mit* occurs as the first part of of negative phrases, e.g. MlaA *mit gɤj* 'never' (with a loanword in ancient Tai form as the second part); these prases are often strongly contracted, e.g. MlaA **met gur > muugur* or **met lɔːj gur > mətləgur* 'not yet'. (For more information relating specifically to MlaB, see Rischel 1995: 107-110.)

Answering with opposite polarities. Mlabri has a category of words used as single-word phrases in answering in the affirmative or in the negative. The affirmative word is *koʔ* 'yes'; the negative is *bah* 'no'. Certain words occurring in other constructions as verbs can also be used in these functions, e.g. the Lao loan *mɛːn* 'yes, that's correct' and the MlaA word for 'be missing': *ʰlaːk* 'no, that's not correct'.

Interrogative status of utterance. Mlabri encodes sentences are interrogative either by the use of an interrogative pronoun or adverb (*tu* is such an interrogative word) or by the presence of a sentence-final interrogative particle *lɛh* which prosodically falls outside the sentence proper, having a lower tone than the preceding stretch. This particle does not have M-K counterparts known to me, although interrogative particles as such are certainly not alien to Khmuic. Tin has an interrogative particle *cuk* in a similar function. By its form Mlabri *lɛh* is vaguely suggestive of direct or indirect influence from a Burmic language although that similarity might well be due to chance (the Mlabri particle is used in questions that do not contain interrogative pronouns or adverbs; in Burmese, on the contrary, *lɛ* is used in questions that are already thus marked; otherwise *la* is used).

Aspect. It is, to my knowledge, not a very widespread feature in Mon-Khmer languages to encode tense or aspect. If there is at all a grammatical expression of tense or aspect it is most likely to encode continuous action and the like.

Mlabri has a perfective particle *ʔa* which occurs extremely frequently in sentence constructions. It can encode the meaning of completed action in relation to the moment of discourse, e.g. *mɛh ʔa ʔek lɛh* 'have you fetched it?', but it often encodes the speaker's conviction and assertion (guarantee) that something is going to happen, e.g. *maʔuːn ʔoh ʔa leh* 'I come tomorrow!' In questions it can focus on the speaker's request to know whether something is going to happen, and thus the first sentence above is strictly speaking ambiguous; it could also mean:

'do you accept it?', and the positive answer in both cases is: *ʔoh ʔa ʔek* '(i) yes, I have!, (ii) yes, I do!' (the examples above are formally in MlaA).

When used in a perfective meaning in relation to the present the particle *ʔa* is felt by Mlabri speakers to be equivalent to the Northern Thai or Lao use of expressions that literally convey the meaning of 'already' but that makes no sense in its use about asserted future. It seems to me that this duality of usage is much more similar to a Burmese use of a verb meaning 'finish' as a perfective or assertive particle *bi* although the form of the syntactic constructions are quite different (the particle being attached after the verb in Burmese).

Mlabri also has an aspectual particle *ʔɛl* '(not) yet' which is little used except in combination with negative particles or other preverbal particles; it is obsolete in A-Mlabri.

The encoding of such aspectual meanings as those of *ʔa* is in itself not alien to Northern Mon-Khmer. The Palaungic language Pale is analysed by Janzen (1976a) as having complex aspectual-modal particles consisting of consonant plus vowel with distinct meanings. I cite two of his analyses:

h + ɛ = hɛ, denoting a definite realized action or state. *e.g.*, ʔou hɛ hau
   'I did go, I went' (Janzen 1976a: 664).
d + i = di, denoting a definite action which will take place in the future.
   e.g. ʔou di hau 'I will definitely go' (ibid.: 665).

What is special about Mlabri compared to Pale (and probably most other M-K) is that (i) it uses one grammatical particle to denote a range of meanings covered by two different particles in Pale, (ii) the form of the particle in Mlabri is *ʔa* not *ʔi* (*ʔi* is taken up for a different purpose, see below). Interestingly, there are two additional aspectual particles which can be considered as expansions of *ʔa* if one applies the same kind of analysis to Mlabri as that of Janzen for Pale. One is *ta* the other *ma*.

The particle *ta* has a variety of modal and/or aspectual meanings across the three varieties of Mlabri: resultative, characterizational, perfective or desiderative; to some A-Mlabri speakers it implies the notion of inadvertent (and undesirable) action. The particle *ma* may encode either an immanent or a realized action, depending on the context.

All the above-mentioned particles combine with other particles in preverbal position: *ʔa di, ʔi ʔa, ʔi di, ʔi ʔɛl, ʔi ta, ta ma*, encoding a variety of meanings which it would lead too far to go into detail with here.

Relational particles. Mlabri has a small but functionally very important set of relational particles. Some have the function of connecting clauses, others are prepositions.

Clause-connecting particles. The particles *ʔi* and *di* connect clauses (also single-verb clauses) in a causal or temporal relationship; the former can often be translated 'so that' or 'until', the latter 'in order to' (for another use of *di*, see on Possessive constructions below). As mentioned under Aspect above, they combine with other particles, e.g. *ʔi di*, *ʔi ʔa*, *ʔa di* to encode other aspectual-modal meanings.

I do not know the particles *ʔi, di* from other Khmuic languages. It is difficult to establish safe etymological connections to more distant languages because many particles are of *(C)V*-structure, and considering the limited phonological possibilities fortuitous similarities across languages are bound to occur. Still, it is suggestive that similar particles with aspectual-modal meanings occur in other Mon-Khmer. For example, a language as geographically distant as Sedang has a set of purposive particles (Smith 2000: 506) of which one is identical to Mlabri *ʔi* whereas all the rest are different. It is much more thought-provoking to look at Palaungic. In the aspect system of Pale according to Janzen (1976a) which was referred to above, the vowel of "d + i = di, denoting a definite action which will take place in the future" is analysed as an element that by itself "denotes unrealized (future) action" (ibid.: 665).

Prepositions. There is a small set of prepositions such as *ni* or *ɲi* 'in', *luŋ* or *lɔŋ* 'into' (the former variant only in MlaAB, the word may be the same etymon as Khmu *luəŋ* 'side'), MlaB *tu* 'from'. Tin and Khmu also have rather small sets of prepositions; all three sets are lexically different.

There is no reason to go much into the syntax and semantics of relational particles here (for more information, though relating specifically to MlaB, see Rischel 1995: 154-194). What is essential to state, is that connecting particles are used very much in Mlabri discourse, also in cases where some other surrounding languages would perhaps use serial verbs without any connecting words between. Similarly, prepositions occur in some instances where speakers of other languages of the area might choose to encode the relation between a verb and its complement without a preposition, for example MlaB *hot tu lam* 'fall (*hot*) from the tree (*lam*)'.

Possessive constructions. The expected possessive construction in a Northern Mon-Khmer language is POSSESSUM + POSSESSOR, and this is indeed what one finds in Tin. In TinMalA 'Mother's house', expressed by means of the words *məj* 'mother' and *kiʌŋ* 'house', runs as follows: *kiʌŋ məj*, with a construction that is used quite generally when two nouns stand in what is traditionally called a head-modifier construction (I phrase this with some reservation because there is much evidence in favour of rather considering a modifier with specific reference to be the head).

In Mlabri a construction consisting of two nouns in sequence will be understood as denoting a single concept, the second noun qualifying the meaning of the first, e.g. *mlaʔ briːʔ* 'person forest', i.e. 'forest people', MlaA *ʔɛːw braɲ* 'child dog', i.e. 'puppy'. It cannot be used if the second noun has a specific reference, as in the Tin example above. Instead, Mlabri uses constructions with a connecting particle *di*, i.e. of the type POSSESSOR + *di* + POSSESSUM. Accordingly, 'Mother's house' comes out as *mɤʔ di gɛːŋ*, and 'the dog's whelp(s)' comes out as *braɲ di ʔɛːw*. This is not, to my knowledge, similar to anything in Mon-Khmer but is most conspicuously similar to Southern Chinese, even to the point of identity in form of the particle *di*.

There is another way of encoding possession or 'belonging-to' in Mlabri, namely by means of a word *bʌr*, which is normally attached proclitically to the immediately following word, which it governs. Unlike *di* which connects the two referents in a possessive construction, *bʌr* (*br̩-*) occurs as a component of verbless predicatives, e.g. *br̩mɛh* 'it is yours', *br̩briːʔ* 'it is a thing of the forest; it belongs to the forest', *br̩ʔɔːŋ gʌh* 'it exists in this place (in this environment: *ʔɔːŋ*)'. Again, I know of no parallels in either Tin or Khmu. In its configuration (but not etymologically), the construction *bʌr* + *N* is clearly reminiscent of Tai, however, and the Mlabri language has actually adopted a loanword from Tai in largely the same function, e.g. *kʰɔːŋ briːʔ* 'it belongs to the forest'.

Concluding general statement about particles in Mlabri. Seen from a Khmuic perspective, Mlabri is conspicuously rich in particles, in particular particles of *(C)V*-structure used to convey aspecual or modal meanings or to establish relations between grammatical constituents. Tin also has relational particles but hardly as elaborated a system as in Mlabri, and it is my impression from Tin narratives that typical Tin discourse uses them much less than Mlabri. The prepositions and conjunctions in Khmu that are listed in Suwilai Premsrirat's *Thesaurus* (Prem-

srirat 2002b: 424-425) bear no resemblance to the Mlabri inventory dealt with above.

Looking specifically at the encoding of grammatical functions by means of words of *(C)V*-structure it is significant that there is an unmistakable similarity with Palaungic but *not* with the Khmuic languages closest to Mlabri.

Size-words and preposed modifiers. There are a couple of size-nouns or verbs which seem specific to Mlabri: *ɓoh* 'size' and MlaAC *baːj*, MlaB *blaːj* 'of such-and-such size'.

The word *ɓoh* is conspicuous by its phonology (most words with glottalized stops are of Tai origin if they have known etymologies); it can stand alone or in the phrase *ɓə ɓoh* as a term meaning 'it is this big', the size being then shown gesturally. It can, however, also occur with following *mat* and with or without a preceding modal adverb *lɔːj* to form the phrase *(lɔːj) ɓoh mat* 'small' (I do not know whether the word *mat* in this context is the noun *mat* 'eye', so that the phrase literally used to mean: 'of the size of an eye only').

As for the other term, *baːj* or *blaːj*, the MlaB form contains a fossilized *-l-*infix of unknown meaning. This must be a very old formation if it originated in Mlabri since all transparent infixes in Mlabri cause the derived word to become sesquisyllabic.

This word cannot occur as a sentence constituent on its own. It occurs in two types of constructions, however.

(i) *b(l)aːj* occurs with a variety of postposed words, the whole complex denoting very big physical size, typically the big size of scary animals. At least in MlaA, these expressions are generally said with an affective, rasping voice quality. Examples are: MlaAC *baːj gʌw* 'big' (in MlaC typically said of the moon), MlaA *baːj kuldul* 'big', *baːj raːŋ* 'long', *baːj tr̩nap* 'big', *baːj tʰaːl* 'long'; Mla B *blaːj gʌw* 'big', *blaːj kwaːn* 'full-grown', *blaːj wec* 'long'; MlaC *baːj gɯp* 'huge', *baːj kel* 'big' (perhaps female language), *baːj kwaːn* 'tall', *baːj waːŋ* 'having grown to a big size'.

A strange thing about these expressions is that only a few of the epithets are otherwise attested in Mlabri lexicon, namely MlaBC *kwaːn* 'long', MlaA *raːŋ* 'long' (also a classifier), MlaBC *wec* 'far away'. The remaining ones need belong to an early layer of Mlabri vocabulary; they are likely to be more recent innovations.

The typologically most interesting thing about *b(l)aːj*, however, is that it also occurs prenominally, as a modifier denoting very big size.

Examples with classifiers are: MlaA *baːj kəplah* 'big sheet', *baːj kɛːn* 'big container'. The occurrence of the size-word with ordinary nouns seems to be restricted to certain semantic spheres. I have recorded it with terms for body-parts, serious afflictions and impressive natural phenomena: MlaA *baːj glɤːʔ* 'big head', *baːj kum* 'a big wound', *baːj ṭrlah* 'a big scar'; MlaB *blaːj çəmɔɲ* 'big star', *blaːj gṛnuh* 'big flames'; MlaC *baːj cəboh* 'big mountain', and several other similar expressions across all three varieties of Mlabri. It is even possible to combine such expressions with a postposed adjectival verb denoting big size, such as *diŋ* 'big', e.g. MlaA *baːj wɤːk diŋ* 'a very large water stream'. The word *b(l)aːj* can also occur with the definite article preposed, e.g. MlaC *ʔak baːj taʔ diŋ* 'the Sun', literally 'the great grandfather big'.

Semantically and also syntactically, preposed *b(l)aːj* is the counterpart of the word *ʔɛːw*, a noun means 'child; offspring', which can be used – with no rigid semantic restrictions – in the same prenominal position as *b(l)aːj* to denote small size, e.g. MlaB *ʔɛːw çndɛh* 'small cup' (*çndɛh* is 'cup'). MlaA further has a word *ʔej*, which occurs in complex predicates denoting something small, e.g. *pjaʔ ʔej* 'small-sized thing' (*pjaːʔ* has an indefinite meaning often translatable as 'something') but also occurs in prenominal position in some fixed expressions, e.g. *ʔej taːk* 'small children'. Interestingly, constructions headed by *ʔej* have a second part which does not seem to occur outside this context, just as is the case with *baːj ṭrnap*, etc. headed by its antonym.

I take it that the syntax of these size-words is an archaic feature in Mlabri (cf. the discussion of different orders of the *determinans* and *determinandum* in Austroasiatic in Pinnow 1963: 143).

Pronouns, the definite article, and demonstratives. Tin has a mainstream type of pronominal system, with singular, dual, and plural forms of the first and second person pronouns and mostly with the expected etyma in these roles (e.g. 1.p.sg. *ʔəɲ* or *ʔeɲ*). There is nothing controversial in projecting that back to Tinic time.

Mlabri is deficient in terms of its pronominal system compared to typical Khmuic languages. It has no proper third person pronoun (if necessary, one uses respectful terms or periphrastic expressions instead), and it has no lexical distinction between dual and plural 1st and 2nd person pronouns. The latter distinction is made by adding a collective term to dual pronouns in order to encode the notion of plurality, as in *bah* 'you two' vs. A-Mlabri *bah tʰɤːŋ*, B-Mlabri *bah ɟum* 'you several'.

Table 4-7: Personal pronouns,
1st and 2nd person singular and dual

Mla *ʔoh* 'I' (cf. Khmu *ʔoʔ*)
Mla *ʔah* 'we (two)' (cf. Khmu *ʔaʔ*, Ksingmul *ʔaː*)
Mla *mɛh* 'you (thou)' (cf. Khmu *meː*, Tin *mah*, Ksingmul *mih*)
Mla *bah* 'you (two)' (mainstream Khmuic, cf. Ksingmul *baː*)

(the dual pronouns do not specify inclusiveness or exclusiveness).

The first person pronouns in Mlabri resemble those of Khmu (also cf. the discussion of laryngeal finals under Tinic phonology above). If they are indeed borrowings one has to explain why they end in *-h* not *-ʔ*, like the second person dual pronoun. That might be due to levelling (pressure from the system) and it might have been mediated by a sandhi phenomenon arising within Mlabri proper.

In any case, it is noteworthy that there is a tendency toward generalization of *-h* as a component not only of personal pronouns but also of certain demonstratives (deictic pronouns and adverbs). This applies, for example, to *gʌh*, *nɛh*, both translatable as 'this' or 'here', MlaB *ɲʌh* 'there', *rih* 'here; in this direction', MlaB *tih* 'in there; out there (on the other side of a partition blocking the sight)', *ɕugwʌh* (or *ɕuguʌh*)'down on the other side (of a hill)', MlaA *ləgʌh* 'up there', and MlaA *ləguh* 'up over there, far away' (some decades ago there also existed an intermediate term for intermediate distance: MlaA *ləgɤh* 'up over there', which now seems quite obsolete). This final *-h* enters a pattern with final *-ʔ*, as in *ɲʌʔ* 'that one; there', *ɕugwʌʔ* (or *ɕuguʌ(ː)ʔ*) 'down below (on this side)', MlaAB *ləgʌʔ* 'up there (e.g. on a hill)'. (Note: the main syllables that are variably pronounced *gwʌh* and *guʌh*, *gwʌʔ* and *guʌ(ː)ʔ* exhibit *-w-*-infixation, the base word being *gʌh*, *gʌʔ*; I cannot pinpoint what contribution to meaning the infix makes. Prefixal *lə-*; and *ɕu-* are reflexes of proclitic use of the words *lah* 'high up', *ɕuh* 'low down'; the former also occurs in *latɤŋ* 'up in the sky'.)

The duality of options: *h* versus *ʔ*, as in *gʌh* ~ *-gʌʔ*, *ɲʌh* ~ *ɲʌʔ*, also occurs in *bɳʌh* ~ *bɳʌʔ*, both conveying the meaning of 'long ago' in various fixed expressions (the first has been recorded only in MlaA, the second in MlaAB; the prefix *bɳ-* may possibly be a proclitic reflex of the old Tai loanword *ban* which in MlaB is also used in *ban ʰnʌm rɔːj* ʰ*nʌm* 'many, many years (ago)'). These expressions have temporal deixis (long ago in relation to the discourse time) and in that sense enter the same field as the personal pronouns and demonstratives.

There is final proof that *-h* and *-ʔ* are indeed *suffixes* on these pronouns and adverbs. A couple of the stems occur in bare form, and then
of course with a long vowel. In the elaborated MlaA expression *tak
bɲnʌʔ tak nɛː* 'very, very long ago', one encounters the pronominal
stem *nɛː* that also underlies *nɛh* 'here'. Similarly, there is an interrogative pronoun MlaAC *pʌː* 'which', obviously the stem that underlies *pʌh*
and *pʌʔ*.

Suffixation of *-h* and *-ʔ* undoubtedly developed within Mlabri
grammar; it is of course not a relic of an ancient Austroasiatic state of
affairs. This feature contributes to giving Mlabri a distinct typological
profile. So far I have not established a clear correlation of meaning with
the different terminations *-h*, *-ʔ* and zero in the various pronominal and
adverbial forms except that the bare stem is apparently associated with
nonspecific reference.

Considering the combination of three peculiarities in the pronominal
system of Mlabri: dual rather than plural pronouns, *-h* in contrast with
-ʔ, and the development of corresponding possessives in *-t* (see below),
it likely that these all reflect an Early Mlabri pronominal system which
does not look Khmuic at all. They certainly have nothing to do with the
Tinic component in Mlabri and must antedate it.

The cognacy between Mlabri *ʔoh* and Khmu *ʔoʔ* then becomes very
significant. The form *ʔoh* is probably a modification of **ʔoʔ* but that
need not be a borrowing from Khmu; it may not be of Khmuic origin at
all. Khmu does *not* have the typical Khmuic pronoun; that is a
widespread M-K etymon which occurs in Tin in the form *ʔɤɲ* 'I' but it is
unknown in Mlabri. The Mlabri and Khmu pronoun, on the other hand,
has cognates in Palaungic. I cited Mitani's (1979) reconstruction of
Proto-Palaungic **ʔɔ* 'I' above; for Proto-Waic Diffloth (1980) reconstructs **ʔiʔ* but some of the Waic languages do have a rounded back
vowel in this pronoun, and Diffloth (1980: 91) actually posits a rounded
vowel for a pre-stage: Waic **i* < Pre-Waic **u*, so we can establish that
in Northern Mon-Khmer this etymon had a rounded back vowel. The
presence of this pronoun in Khmu is, I think, a good example of affinity
between Khmu and Palaungic, an affinity which may have contributed
to distancing Khmu from the small and more typically Khmuic languages. Mlabri sides with Khmu but maybe independently.

In terms of grammatical system, the most interesting thing about
Mlabri personal pronouns is that those for 1st and 2nd person singular
have strictly corresponding possessive pronouns. A similar phonolog-

ical relationship is found between the 1st person dual pronoun and the definite article in Mlabri. The pattern is shown in Table 4-6.

Table 4-8: Personal pronouns with corresponding possessive pronouns

ʔoh 'I' – *ʔot* 'my'
mɛh 'thou' – *mɛt* 'thy'
ʔah 'we (two)' – *ʔat* 'DEFINITE'

The definite article in Mlabri defines the referent of a noun as salient and given or expected in the context of the discourse or the narrative (see Rischel 1995: 152-154; a very detailed exposition comprising all varieties of Mlabri is given in Rischel 2006). It has the form *ʔat* in MlaB and the usage of some MlaA and MlaC speakers (in MlaA chiefly senior persons); other MlaAC speakers use a form *ʔak*, which can be explained as due to generalization of an allomorph that arose before words beginning in a velar stop after MlaB branched off (for discussion, see Rischel 2006).

The most likely explanation of the two possessive forms is that they are contractions of personal pronouns plus the connective particle *di*, which in this context denotes possession: *ʔoh di ʔɛːw* > *ʔot ʔɛːw* 'my child(ren)'.

The definite article, which is unique in its geographical environment and probably developed within Mlabri, invites a similar analysis, with the 1st person dual *ʔah* denoting a couple forming a household in which certain items, for example the baby or the children, *at ʔɛːw*, are salient parts of the setting. That might explain how that form could change meaning from 'our (of the two of us)' to 'DEFINITE' ('our baby' = 'the baby'), thus creating a grammatical category that is unexpected in the environment of Northern Mon-Khmer languages.

Since the possessives and presumably also the definite article are formed with involvement of the personal pronouns, the presence of the latter in Mlabri must antedate their development. I would assume that the contracted forms are old, which speaks against Mlabri having borrowed its 1st person pronouns from Khmu (apart from the problems with assuming such borrowing from a language which has otherwise influenced Mlabri only moderately). As said already, it seems more

likely that the whole pattern of 1st and 2nd person pronouns with their possessive counterparts is an inheritance from Early Mlabri, suggesting that Khmu and Mlabri had early and probably mutually independent contacts with Palaungic.

PART 3:

A SCRUTINY OF THE LEXICAL EVIDENCE FOR PROTO-TINIC

Introductory remarks

In this part of the monograph the point of departure is the assumption that much of the lexical similarity between Mlabri and Tin are due to a partially shared Tinic level.

The first sections outline what happened to Mlabri and to Tin *after* that hypothetical level, as a set of prerequisites to detailed comparison of contemporary forms in Mlabri and in Tin Mal or Tin Prai.

The remaining sections consist of a presentation and discussion of a rather extensive set of cognates that qualifies as Tinic according to the hypothesis.

Phonological and morphological changes in Mlabri after Tinic time

Since both the phonology and the morphology of Mlabri are very conservative there are few phenoma to be dealt with here.

Changes in consonants

Very little happened to the consonants of Mlabri. Most importantly, one of the initial sibilants, the one I have labelled "S₁", changed into an aspirated dental stop t^h-, as in the term for 'meat': $t^h\Lambda c$, whereas the other sibilant, one I have labelled "S₂" was continued as a variable sound with a sibilant component: the palatoalveolar or alveolopalatal sibilant or affricate ç- [ç-, ᵗç-], as in çoʔuːm 'smell'.

Changes in vowels

There are few transparent vowel changes on the way from Proto-Tinic to (Proto-)Mlabri. As said repeatedly in this monograph, the great majority of vowel correspondences between Mlabri and Tin belong to a criss-cross pattern of pairs of etymologically connected vowels (or diphthongs) the historical basis of which can hardly be reconstructed without more progress being done in comparative Khmuic. Still, a few

observations can be made with special reference to Mlabri. It will turn out that they center more or less around the vowel area [ɛ]-[a].

Mlabri has developed a long vowel ɛː in several words which can be plausibly reconstructed with Tinic *ia < Khmuic *aː (e.g. Mlabri gɛːŋ 'house; lean-to' ~ TinMal kiʌŋ). This is one of the strong criteria for Tinic provenance. The picture is, however, complicated by multiple correspondences. Besides stemming from monophthongization of Proto-Tinic *ia, Mlabri ɛː has at least two other origins as seen from its correspondences in Tin:

ɛː ~ ɛː
Mla tɛːk ~ Tin tʰɛːk
Mla mlɛːt ~ Tin ᵐblɛt
Mla mrɛːŋ ~ Tin ᵐbrɛːŋ

ɛː ~ a(ː)
Mla gwɛːt ~ Tin kwat
Mla krɛːc ~ Tin kʰraːc

(for additional information on the cited wordforms see the tables of cognate sets later in this part of the monograph). The question is what vowel one should reconstruct at the Tinic level for each of the two sets above. That is a rather hopeless challenge if one has no wider framework in which to place the reconstructions.

In addition, however, there are one or two words (with other types of finals than the two sets above) in which Mlabri ɛː corresponds to Tin iʌ but without going back to Khmuic *aː, namely:

ɛː ~ iʌ
Mla rɛːlʰ ~ TinMalA ᵑgriʌjʰ (occurring across M-K in forms such as ris, riəs) 'root'
Mla ʔjɛːʔ ~ TinMalA ʔiʌʔ , TinPrai jɛʔ 'far away'

This correspondence set ɛː ~ iʌ is significant when reconstructing Tinic. It is my assumption that at least the word for 'root' had a diphthong *ia already at the Pre-Tinic level (it could be referred to as *ia₁ to distinguish it from the much more widely distributed, secondary diphthong *ia₂ < Khmuic *aː). The two entities coalesced in Early Tinic and therefore shared the further development into a long vowel ɛː in Mlabri.

The simplicity of that scenario is a main reason why I preferred, in Part II, to posit *ia (i.e. *ia₂) rather than *ɛː as the Proto-Tinic step in the development of Khmuic *aː to Mlabri ɛː and to Tin iʌ in words such as Mlabri *geːŋ* 'house; lean-to' ~ TinMal *kiʌŋ*.

There are probably more hidden Tinic etyma belonging to these four correspondence sets for Mlabri ɛː, words which I have recorded in Mlabri but which happen not to have survived in Tin, or which fail for other reasons to show up in my limited Tin data. In any case, as a result of all these correspondences Mlabri has a disproportionate number of words in ɛː compared to aː. From the perspective of general phonetic typology of vowel systems it would otherwise be a reasonable assumption that the most open vowel quality [a] dominated over [ɛ] across the vocabulary but that is not what we observe with the long vowels in Mlabri. Looking at monosyllabic words beginning in a single voiceless stop – a word type not prone to strange phonological developments – I found practically equally large sets of words with ɛː and aː.

That observation supports the assumption that many of the occurrences of ɛː in Mlabri are not originally open front vowels but come from other types of vowels or diphthongs. This means that one should perhaps look at Tin rather than Mlabri in attempting to reconstruct this particular part of the Proto-Tinic vowel system.

Then there is an intriguing pattern of vowel alternation in Mlabri. As mentioned in Part II, the word for 'bear' testifies to a sporadic transition Tinic (*aː) > *ia > iu beside *ia > ɛː since it is *beːk* in MlaAC whereas MlaB has two forms: *biuk* and *beːk* (the latter only as female language; Tin has *piʌk*, whereas the word is non-occurring in Khmu, which has another etymon *huəl*). There may have been an affective voice quality accompanying the word for 'bear'; other words for large or scary or repulsive animals are typically said with affective rasping voice quality in contemporary MlaA.

There is a parallel in that a few words have long *uː* in MlaB corresponding to long *aː* in MlaA, whereas MlaC has now one, now the other option:

'sky': MlaAC *klaːr*, Mla B *kluːr*
'wild banana species': MlaA *ʔjaːk*, Mla BC *ʔjuːk*
'mucous; for the nose to be running': MlaAC *maːr*, MlaB *muːr*
'personal spirit': MlaAC *ʰmaːl*, cf. MlaB *ʰmuːl* 'perform ceremony for personal spirit', MlaB *çɾmaːl* ~ *çɾmuːl* 'soul'

The vowel split might suggest a diphthong rather than *aː* at the Proto-Tinic level in these words as well. That is contradicted, however, by the preservation of monophthongal *aː* in the word for 'personal spirit' in Tin; as for the three other words the source of the vowel alternation cannot be ascertained since they have not been retrieved in Tin. In any case, it is hard to make any generalization that could account for the exceptional development in the four words. There is no reason to posit an affective voice quality in words for the sky and for a harmless plant species. So far, the behaviour of the words above must be left as an unsolved enigma in Mlabri sound history.

Mlabri morphology

Since virtually all of the morphology found in Mlabri, though clearly of M-Tinic type, may be of non-Tinic origin there is nothing useful that can be stated here about the change from Proto-Tin to Mlabri in terms of word-formation.

Phonological and morphological changes in Tin

Most of the changes I postulate below for Tin happened early, maybe shortly after Mlabri branched off. They do not readily reveal them-selves except by rather detailed comparison with other Khmuic lan-guages. When Filbeck (1978) reconstructed Proto-Tin by comparison between modern Mal and Prai dialects, and even when he recon-structed a more hypothetical pre-stage: Pre-Tin by internal reconstruc-tion, he did this on a strictly Tinic basis without involving the only well-attested Khmuic language, Khmu, for systematic comparison. Consequently, Filbeck had no chance of disclosing how much Tin must have changed from its earliest stage up to the latest common denomi-nator for Mal and Prai (the exception was consonant mutation, which is so conspicuous a feature of Tin but which Filbeck erroneously dated back to the time when Tin branched off from Khmu and other related languages).

Thus it should be understood that the presentation below is an attempt to identify elements of the early history of the Tin language, from Proto-Tinic to Proto-Tin, more or less from scratch, as a necessary component of the identification of Tinic features in Mlabri. It is done on

the basis of comparison with Mlabri, and to a much lesser extent by comparison with other Khmuic languages.

The history of Tin after Proto-Tin began to split into Mal and Prai, on the other hand, has been reconstructed in so much detail and with so excellent documentation by Filbeck (1978) that it would be redundant to recapitulate his results except to the extent that it is strictly necessary for the coherence of the present study. Altogether, this monograph goes into the late developments from Proto-Tin up to Modern Mal and Prai only to the extent that it is essential for one of two reasons: (i) if my comparative approach yield results that are at variance with his (and are of significance for the word comparisons with Mlabri), (ii) if the word comparisons I present are not sufficiently transparent unless some additional information is given about offsets between Mal and Tin, or between conservative and advanced Mal. I take it that the existence of minor discrepancies between Mal and Tin, or even within subdialects of one language, is trivial in itself and does not necessarily require comment.

Changes in consonants

Initials: the consonant mutation. By "consonant mutation" in Mon-Khmer languages one understands a set of sound-shifts which devoice and/or aspirate initial consonants (much along the lines of the consonant mutations leading from Indo-European to Proto-Germanic). The concept was introduced into Mon-Khmer studies by André Haudricourt (1965). This phenomenon is found in Tin.

Mlabri exhibits no consonant mutation. The old distinction between voiceless and voiced stops is also preserved in three other Khmuic languages that I have data from: Khabit, Ksingmul and Phong. As for Khmu, conservative dialects preserve old voiced and voiceless stops intact in syllable-initial position whereas other dialects have a scenario of tonogenesis accompanying devoicing of old voiced stops. That is probably a fairly recent phenomenon and has clearly nothing to do with consonant mutation in Tin, which was a much older and typologically quite different change. Since, however, the presence of devoicing phenomena slightly complicates the etymological comparison among the Khmuic languages I shall briefly review the devoicing scenario in Khmu.

Across the dialects of Khmu we find (i) the old M-K state of affairs:

voiced versus voiceless initial stops, (ii) an intermediate stage with breathy phonation of vowels after old voiced stops, (iii) complete coalescence of voiced and voiceless stops along with the development of tone.

The above picture covers the general, historical and geographical scenario of Khmu. One of the dialects in Laos ("Khmu Rook", cf. Premsrirat 2002: xxxi & xlv), however, took a different course and both aspirated and devoiced old voiced stops probably due to influence from Lao, in which that very process happened several centuries ago.

The existence of different reflexes of the old voicing distinction in Khmu complicates the comparative work on Mlabri. That is because there are loanwords from Khmu in Mlabri, and it is not known in advance which dialect or dialects Mlabri borrowed from. It should be noted that Mlabri always borrows words from tone languages *without* borrowing the tone.

Returning to consonant mutation proper, it seems to be unique to Tin among the Khmuic languages. The pattern of changes is as follows:

I.   *Old voiceless stops became aspirated*
IIa. *Plain voiced stops were devoiced in absolutely initial position*
IIb. *Plain voiced stops remained voiced elsewhere*
III. *Glottalized voiced stops became (prenasalized) voiced stops*

The resulting pattern of stop consonants, their phonemic status and phonetic realizations are discussed in Huffmann (1976a: 582-583). I here disregard the old glottalized series, since the words in question are mostly loanwords from Tai.

Mutation of the type we find in Tin can also be observed in distant branches of Mon-Khmer. That might seem to invite a scenario in which consonant mutation spread across a large area and also swept across Tin. Strangely, the areas exhibiting this type of mutation are geographically unconnected, and as said already, Tin is quite unique within Khmuic in having this phenomenon. Thus it is a more appealing assumption that we are faced with a spontaneous complex of sound changes within Tin proper.

It is, in my view, an open question how early or late the consonant mutations in Tin took place. They may have been separated by shorter or longer time spans, namely so that voiceless consonants acquired aspiration early, then the plain voiced consonants were devoiced, and finally – perhaps quite recently – the glottalized consonants lost their glottalization.

Consonant mutation is shared fully by all modern Tin dialects, and Filbeck (1978) took it to be defining, as it were, for even the earliest reconstructed stage of Tin, which he called Pre-Tin. I see no grounds for dating it so early. In theory, consonant mutation in Tin might even have taken place after the split into Mal and Prai, provided that the speech communities remained in so much contact that a sound change could spread in a wave-like fashion.

(The reason for placing it so early in this survey is not that it is an early sound-shift but that it affects a great many word-forms. It is convenient to have the workings of consonant mutation under control when looking at other sound changes happening in Tin and when spotting etymological correspondences with Mlabri.)

In the Mlabri context the possibility of dating the sound shifts is interesting because Mlabri has borrowed from Tin long after the Tinic period, but further study is needed to clarify the extent of that. Mutation of old voiceless stops in Tin must necessarily postdate the first strong presence of Tai lowlanders in the twelfth or thirteenth century AD since loanwords from Tai could enter Tin early enough to undergo mutation.

Early loanwords from Tai in Tin are diagnostic but do not yield a simple answer when it comes to absolute datings (cf. discussion in Rischel 1989b: 110-111 with reference to Filbeck's reconstruction of the scenario). The most important loanword providing pre-mutational evidence is the function word k^hap 'and' in TinMal; it would take considerable language contact and even bilingualism to borrow such a grammatical operator from Tai. Still, it happened early enough for mutation to apply to the form. That speaks for a rather late dating of mutation in Tin.

Another pre-mutational borrowing is the word for 'cultivated banana', which I have recorded in the TinMal dialect of Ban Kwet with mutation of its initial consonant: k^hloj or $k^hlu\Lambda j$ (< Tai *kl-). Both the preservation of the lateral and the mutation speak for an early date of borrowing into Tin but again, there is a *terminus post quem* since it postdates the significant presence of Tai banana cultivators. Independently and much later, the word for 'cultivated banana' entered Mlabri as well, viz. as $ku\gamma j$ or $kw\gamma j$. In this case, the loss of the lateral after k is diagnostic of fairly recent borrowing since the loss did not happen in Mlabri but in the lending language (in contrast, Mlabri faithfully preserves clusters in old loanwords from Tai, such as *kle:t* 'scales (on a fish)', which would have been *ke:t* if borrowed recently from Northern Thai, for example).

There is ample evidence for the regularity of the consonant mutation in Tin, the main changes being $*b > p$, $*g > k$, and $*p > p^h$, $*k > k^h$.

We can illustrate what happened in the old voiced series by taking the word for 'house', which is $ga\mathrm{:}\eta$ in conservative Khmu and $g\varepsilon\mathrm{:}\eta$ in Mlabri but $kia\eta$ in Tin Mal and thus shows mutation of $*g$ to k in Tin. As for the old voiceless series one can mention the word for 'thunder', which is kir in Khmu and kur in Mlabri but k^hur or $k^h\gamma\mathrm{:}r$ in Tin, i.e. with mutation of $*k$ to k^h.

Finals. The two final sibilants "$*\text{-}S_1$" and "$*\text{-}S_2$", which are still pre-served as $-l^h$ versus $-l^h$ in Mlabri) coalesced into one final in Proto-Tin. It cannot be determined what was its quality then but it is continued as $-j^h$ in Mal, whereas it coalesced further with $*\text{-}t$ (continued as a dental stop) in Prai.

Initial and final $*r$. Initial clusters with $*\text{-}r\text{-}$ undergo various changes both in Mal and Prai. These are at least in part areal phenomena to do with a widespread tendency in northern Southeast Asian languages to somehow "eliminate" the [r]-sound. This is done by substituting a lateral or a palatal glide for the trill or by simply dropping it. Since the TinMal data are partly from MalA and partly from MalB it is essential to state here that initial $Cr\text{-} > Cj\text{-}$ in MalB but is preserved in MalA.

Final $*\text{-}r$ is preserved in some Tin but undergoes change or is lost in other varieties of Tin. The variation in Mal is impressive: for p^har 'fly, v' [p^{h}ar] (conservative pronunciation) one hears [p^{h}ay], [p^{h}aw] or [p^{h}aj] depending on the dialect. In Prai, the trill is often simply dropped.

Filbeck (1978) attempts to handle some of the changes involving r in terms of a Stammbaum for the Tin dialects and subdialects but concedes that the developments do not readily lend themselves to such an approach. There are obviously areal tendencies involved.

Specific initial cluster simplifications

There are a number of such simplifications, some of which happened already before Mal and Prai split apart whereas others are later and in part quite dialect-specific. Only a very superficial presentation of some of the most salient changes can be given here.

I. Tinic *S₁r- > Proto-Tin *s-

A safe etymology attesting this sound change is the word for 'tooth' that occurs across Khmuic. Mlabri has $t^h r\varepsilon:\eta$ with t^h- as the regular continuation of *s₁-. The loss of the trill in Tin is attested both in Mal and Prai: MalA ¢iʌŋ, Prai sǝŋ or ¢ǝŋ. In Khmu, on the contrary, the first consonant has been weakened to *h* or lost: *hra:ŋ, ra:ŋ*. The original cluster is preserved in Khabit: *sruǝŋ*, thus further testifying to the validity of the reconstruction *sra:ŋ* for Khmuic.

For the sake of completeness it may be mentioned that Mlabri also has a different word for 'tooth', namely cʌɲ (I have not ascertained the origin of that word). The distribution in MlabriAC is such that cʌɲ is the default term for 'tooth', whereas $t^h r\varepsilon:\eta$ refers specifically to the lower teeth; in MlabriB the latter is the default term for 'tooth' (MlabriAB has a further term *gul* or *gu:l* meaning 'molar'). Across Mlabri, the term for 'tooth' (be it cʌɲ or $t^h r\varepsilon:\eta$) is also used in the generalized meaning of 'cutting edge' or just 'sharp edge'.

II. Tinic *ɟr-/*dr- > kr-, *ɟl- > kl-

Especially the clusters with initial palatals are (in a universal sense) articulatory complex and likely to undergo change. Although they were preserved perfectly in Mlabri they vanished from the phonotactics of Tin.

The changes above are essentially specific to Mal. Prai simplified such awkward clusters by substituting a joker presyllable *si-* for the first consonant (see example (i) below) or by dropping the second consonant (see example (iii) below). I know of only one example that is suggestive of a shift palatal > velar before a liquid in Prai and that looks like a spurious cognate: Mla *crɤw* 'call out', TinPrai *ŋgrɔ:* (for one thing, one would expect an aspirate in Tin; also the vocalism is strange). The etymon in Mlabri is, however, interesting for another reason: its cognate in Phong is *tru*, suggesting that Phong had an independent change *cr- > tr-*, as against *cr- > kr-* in Mal, and providing additional evidence for the historical primacy of the palatal articulation in Mlabri. Since the change in Tin happened after consonant mutation the path of change was *ɟr-/*dr- > *cr-/tr- > kr-* and *ɟl- > *cl- > kl-*. The cluster *kr-* stemming from *ɟr-* and *dr-* has more recently changed into *kj-* in some dialects.

That there must have been a set of changes affecting lingual stops plus *r* is seen from the complete absence in Modern Tin of such clusters as *tr-/*tj- and *cr-/*cj- versus the relative abundance of words in

kr- or *kj-*, depending on the dialect. Comparison with Mlabri shows what was the source of some of these words in *kr-* or *kj-*.

As for *ɟr- > *gr-* there are three words which illustrate the change. In all these cases, Mlabri has preserved the old state of affairs. (i) MlaAB *ɟraː* 'skinny; thin' versus TinMlaA *kraː* 'id.' (TinPrai has *siraː*); (ii) MlaC *ɟrɤh* 'to crow' versus TinMalB *kjɔh* 'to crow; cluck'; (iii) Mla *ɟruʔ* 'deep' versus TinMalA *kruʔ*, TinMalB *kjuʔ* (TinPrai has *cuʔ*).

As for *dr- > *gr-* there is the word for 'termite', Mla *druːɲ* vs. TinMal *kjoːɲ*. That is just one good example, but in return, it is certain that the Mlabri form represents an ancient M-K stage since Diffloth (1984:73) reconstructs exactly the same form *druːɲ* for such a distant branch of M-K as Monic. The agreement with Mlabri is even more remarkable because the word is not preserved with this phonology in any of the attested Monic languages or dialects that Diffloth cites as basis for his reconstruction. (A variation *dental ~ velar* similar to that between Mlabri and Tin occurs in the Monic forms cited by Diffloth, e.g. Literary Mon <drun, grun>. It is a specific, secondary development found in Modern Mon, according to Diffloth 1984:305: "*tr-* and *dr-* initials are often written <kr-> and <gr-> respectively and pronounced /kr-/ in the spoken language".)

The exemplification above only concerns clusters with *-r-* as the second element of the onset. There is slender evidence for a similar change of clusters with *-l-* as the second element. In Part 1, Table 1-2 I listed the TinMal word *kluh* 'drill, v', which has been borrowed into Mlabri. This etymon also exists as an inherited word in Mlabri but only as an instrumental derivative *ɟrluh* or *crluh* with infixal *-r-*, meaning 'wooden pin; twig'. The first variant is a pre-mutational form; it could be of Tinic origin but may also be older in Mlabri. The second variant looks as if it was borrowed from post-mutational Tin (if so, it supports the second of the two paths of change suggested above), but it may just be an idiosyncratic variant of the second unless the distinction between the two variants is spurious. In any case, the combined evidence points to a reconstruction *ɟluh* of the verb underlying the recorded form *kluh* and thus testifies to the sound change *ɟl- > kl-* in Tin.

Changes in vowels

The vowel *a* was raised to *ə* before palatals in Proto-Tin. This vowel is preserved in Mal, e.g. in the word for meat: *t'əc* (TinMalA) but was

fronted and raised further to *e* in Prai: *sec*. The same happened after a palatal, at least in some instances, e.g. in the word for 'house': **gaːŋ >*
*giaŋ > *kiaŋ > *kiʌŋ > *cʌŋ > *cəŋ > Prai *ceŋ;* Mal preserves the intermediate form *kiʌŋ*.

If Proto-Tinic ever had the vowel **ʌ* it has changed into *a* or *ə* in Tin proper. If so, it is unclear what conditioned one or the other reflex, and at what time the changes happened. Phonetically, the same sound [ʌ] developed secondarily in the diphthong **ia > iʌ* (which most authors writing on Tin render as *ia*). In the conservative MalA dialect of Ban Kwet I have heard some speakers use [iɑ] consistently in a few words although they use [iʌ] in most words; I take that this is a local idiosyncracy without wider historical implications.

The vowel offsets between Mlabri and Tin in cognate pairs suggest that some other vowel changes happened in Proto-Tin time (cf. discussion of some of the vowel correspondences in later sections), particularly instances of vowel lowering in the high (narrow) part of the vowel space. Since, however, the conditioning is unclear in most cases a definitive repertory cannot be given here.

Word formation

A number of simplifications happen, some of which are evidenced by the Mlabri-Tin cognate pairs presented below. The most important of these simplifications can be mentioned in very summaric form:

(i) The canonical word structure in Tin is simplified (a.o. so as to minimize sesquisyllabicity). It happens chiefly in two ways:

(i,a) There is a strong tendency towards strict monosyllabicity brought about by loss of presyllables.

(i,b) Consonant sequences arising by infixation are mostly simplified.

(ii) Tin ceases to have productive affixal morphology, and all word-formation processes disintegrate, the residue being a small number of verb pairs suggesting a fossilized causative formation by prefixation.

(iii) Tin develops quasi-prefixation as a repair measure accompanying the shrinking of pretonic material or simplification of awkward initial

clusters (often stemming from contraction of sesquisyllabic forms). There are two different situations:

(iii,a) Before an obstruent there occurs a feature of prenasalization no matter what kind of shrinking happened. Shrinking with such reflexes is not only observed by comparison with Mlabri; etyma shared with Khmu show the same picture, cf. Khmu ^{h}*nta:k* 'tongue' ~ Tin ^{n}*t^{h}a:k* 'id.', Khmu *kdah* 'forehead' ~ Tin ^{n}*dah* 'id.' (neither of these words exists in Mlabri).

The Khmu-Tin cognate pair for 'tongue' shows that the mutation of voiceless stops into aspirated stops in Tin is unbound: it occurs in clusters that eventually shrink, just as it occurs in absolutely initial position. This is also true if there is a vestige of causative prefixation, cf. ^{n}*t^{h}ec* 'break', causative of the same stem word as Mlabri *tac*. Initial voiced stops, however, are voiced when prenasalized even though they mutate into voiceless stops elsewhere. This creates differences in voicing such as that between the underived verb *pɤl* 'die' (*p-* < *b-*) and its derivative ^{m}*bɤl* 'kill' with causative prefixation. In the most advanced Mal dialect, MalC, prenasalization is entirely lost but the unaspirated prenasalized stops remain voiced (Filbeck 1978: 44; Filbeck nevertheless argues that the voiced stops with prenasalization in more conservative dialects are allophones of /p/ etc. and therefore writes them as voiceless in broad notation).

(iii,b) Before a sonorant the repair phenomenon is a full quasi-prefix *si-* in Tin, whereas preaspiration: ^{h}*C-* occurs in some instances in Mal. These developments are truly enigmatic. One is tempted to speculate that these different pseudo-prefixes are contracted reflexes of one repair component, possibly of the form *$s(V)N$-*.

(iv) In Tin, prenasalization becomes a marker on some Tai loanwords beginning in an obstruent (stop or sibilant). As with the reaction to shrinking mentioned above, stops keep their old state as voiced or voiceless when prenasalized, whereas they exhibit aspiration caused by mutation. Prenasalizing of initial voiced obstruents looks like a measure to protect voicing in loanwords, but that does not make any sense with voiceless obstruents. Rather, the prenasalization has degenerated into a diacritic marker which just indicates that there is something special about the status of the word.

As said above, most of these changes are such that are not, or only partially, captured by Filbeck's internal reconstruction (1978) because they appear clearly only if Tin is approached from a comparative perspective.

A reference material of Mlabri-Tin cognates with Proto-Tinic reconstructions

An exhaustive search for Mlabri and Tin cognates is – at least for me – an impossible task. I have a limited coverage of Tin lexicon (Mal as well as Prai) with focus on domestic vocabulary rather than more specialized areas such as names of animal and plant species in the deep forest, or vocabulary to do with traditional beliefs. As for Mlabri I have a somewhat better coverage of those specialized areas but that is of little help in comparisons. The identification of species in the forest is more often than not uncertain, and the religious beliefs of the Mla Bri differ from those found in traditional Tin culture.

Even during my most recent fieldwork on Mlabri I kept retrieving hitherto unknown vocabulary, e.g. words relating to phenomena of nature and to the way you behave in the forest in order to avoid bad luck. Thus the lexicon is open-ended, and it makes no sense to talk about the "total vocabulary". In any case, there is a limited number of shared semantic fields – related to life – within which it has been possible for me to search for Mlabri-Tin cognates. That biases the allegedly Tinic vocabulary in favour of a basic, domestic vocabulary.

Another bias, on the Tin side, is caused by insufficient search for relevant lexical data in Mal dialects. Over the years I put most emphasis on Prai because there was evidence for some borrowing from Prai into Mlabri, but for the Tinic hypothesis Mal is just as relevant, of course. If it had been possible for me to pursue my fieldwork a little bit further I would have devoted most of it to the MalA dialect spoken in Ban Kwet, which is the most conservative form of Tin, at least phonologically. Although several forms from that dialect are cited below, my coverage of the lexicon is still dismal.

Finally, it must be mentioned that my data for the C-variety of Mlabri are meagre compared to my data for the A- and B-varieties. It was planned to supplement the data but that is now beyond my capacities.

My limited reference material for substantiating the Tinic hypothesis,

with reconstructed Proto-Tinic forms and Mlabri and Tin reflexes, is listed in Table 3-1 through Table 3-5.

Table 3-1. Words reconstructed as monosyllabic
with initial oral stop

*p-

*-*pam* (in Tinic always with *-*m-p-*) 'butterfly'; MlaC *pr̩mpam* (male lg., female lg.: *pimpim*) 'butterfly', *r̩mpʌm* 'id.' (possibly different species unless one word is the cover term); TinMalA $^m p^h am$ in $^m p^h oj^h$ $^m p^h am$ 'id.'

*pɛʔ 'three'; Mla *pɛʔ* 'three'; Tin $p^h ɛʔ$ 'id.'

**pe(ː)ɲ* 'shoot'; Mla *peːɲ* 'shoot'; TinMalA $p^h əɲ$, TinPrai $p^h əɲ$ or $p^h eɲ$ 'id.' (Khmu has *piɲ*)

*piːH₁ 'hulled rice (grains)'; Mla *piːʔ* 'hulled rice'; TinPrai $p^h iː$ 'id.'

*pÄ₁r 'fly, N'; Mla *pʌr* 'fly'; TinMalA $p^h ar$, TinPrai $p^h aːl$ 'id.'

*pɤl ~ *pU₁l* 'crossbow'; MlaBC *pɤl*, MlaA *pol* 'crossbow'; TinMalA $p^h il$ (some apparently say $p^h əl$), TinPrai $p^h ul$ 'id.'

*poh 'split (wood etc.)'; Mla *poh* 'split'; Tin (Mal and Prai) $p^h oh$ 'id.'

*pO₁ːS₁ 'barking-deer'; Mla *poːl^h* 'id.'; TinMalA & TinCuun $p^h ɔːj^h$, Tin-Prai $p^h ɔːt$ 'id.' (Khmu has *puəs*)

**po(ː)ɪn* 'four'; Mla *poːn* 'four'; Tin $p^h on$ 'id.' (Khmu has *puən*, *puːn* depending on dialect)

*puŋ 'blow (produce an air stream with the lips); Mla *puŋ* 'id.'; Tin-MalA $p^h uŋ$ 'id.'

**pre(ː)₂m* 'old; former; long ago'; Mla *prem* 'id.'; TinMalB $p^h jeːm$ 'id.'

*preʔ 'hot (spicy)'; MlaAB *preʔ* in *preʔ ɡem* 'chili'; Tin (Mal and Prai) $p^h eʔ$ 'hot (spicy)'

*plah 'classifier for sheet-like objects'; Mlabri has the etymon with an unidentified prefix: *kəplah*, (MlaA also:) *kaplah*, *kɔplah*, (MlaC:) *koplah* 'classifier for sheet-like objects'; TinMalC and TinPrai $p^h lah$ 'classifier for big pieces of cloth' (apparently not used in Tin-MalA, which has the classifier *mɔʔ* instead); the monosyllabic etymon occurs identically in Waic

*plE₂ːʔ 'fruit; nut; grain'; Mla *pleːʔ* 'fruit; nut; grain' (also cf. *kɔːk); Tin $p^h lɛʔ$ (in Prai varying with $p^h ɛʔ$) 'nut-shaped thing; fruit etc.' ($p^h lɛʔ$ *lam* = $p^h ɛʔ$ *lam* 'fruit'); Khmu has *pleʔ* 'grain'

**plË₂(ː)₁m* 'landleech'; MlaC *plɤːm* 'landleech'; Tin (Mal and Prai) $p^h lɔm$ '(small) leech species'

*t-

*tac 'be torn; get a cut or fissure'; Mla tac 'get loosened; fall of; get a
   fissure', (MlaB:) 'cut oneself'; also with causative affixation: Mla
   batac 'tear something'; TinPrai ⁿtʰec 'tear something; be cut'

*taːɲ 'weave'; MlaB taːɲ 'weave'; TinMalA tʰaːɲ, TinPrai tʰaɲ 'id.';
   NB: Khmu taːɲ 'weave; plait'

*tar 'string'; MlaA tar 'string'; TinMal tʰar, TinPrai tʰaːl 'id.' (Khmu
   has tar in the meaning of 'head strap for carrying')

*tÄɜc 'be torn'; Mla tac 'be torn'; Tin in the causative derivative Tin-
   MalA ⁿtʰɔc 'cut', TinPrai ⁿtʰec 'tear' = Mla batac; cf. TinPrai tʰaːc
   'pick off with fingernail' suggesting that Mlabri tac might reflect a
   merger of two etyma *tac and *taːc

*teːk 'snail (species)'; MlaB teːk 'snail species'; Tin tʰeːk in TinMalA
   kʰlwaʔ tʰeːk 'inedible snail species' (kʰlwaʔ 'edible snail species'),
   TinPrai ko tʰeːk 'inedible snail species'

*tiːH₁ 'hand'; Mla tiːʔ 'hand'; Tin (all across Tin) tʰiː 'id.' (Khmu tiʔ
   'hand'; Khabit tiː 'hand', cf. etymon for 'eight'!)

*tiːʔ 'eight'; Mla tiːʔ 'eight'; Tin tʰiʔ 'id.' (cf. etymon for 'hand'!)

*tɤl 'cultivate; plant'; Mla tɤl 'cultivate; plant'; Tin (Mal and Prai) tʰəl
   'id.' (also cf. Khabit til 'cultivate'; Khmu tir 'sow (paddy grains)')

*tÏp 'bury'; MlaBC tuup 'cover, v' (with semantic interference from
   TinPrai təp 'cover; block'); Tin (Mal and Prai) tʰəp 'bury' (also cf.
   Khmu tiəp 'wrap up')

*tO₂c 'grasp'; Mla toc 'grasp; (B: 'take'); TinPrai tʰɔc (in tʰɔc ʔət cor-
   responding to Mla toc ʔek) 'id.'

*tO₂h 'pull out'; MlaB toh 'pull out (with a jerk, e.g. a hair)'; TinMalA
   tʰɔh 'pull out' (the Mlabri form looks like a borrowing from Khmu
   toh 'pull out' but considering its semantics it seems plausible that it
   goes back to Tinic which in that case sided with Khmu)

*tu(ː)ɪn 'rodent species'; Mla tuːn 'mole'; TinPrai tʰun 'big rodent
   species'; cf. Khmu tuːn tʰraːɲ (affinity to Tai but obviously old in
   Tinic)

*k-

*kaːH₁ 'fish'; Mla kaːʔ 'fish'; Tin kʰaː 'id.'

*kɛc ~ kE₃c 'metal ring on shaft'; MlaA kɛc, MlaB kac 'id.', TinPrai
   with infixation: kʰrɛc 'id.'

*keh 'wild goat'; MlaA kɛh (said with rasping voice) 'wild goat-like
   animal species'; TinMal kʰɛh 'wild goat' (cf. Khmu keh; the word
   has a Tai connection)

*keh 'receive a guest' (??); MlaB keh 'greet; receive cordially'; TinPrai
  kheh khaw 'turn one's back on a visitor' (the apparent reversal of
  meaning is not unique in a Mlabri context)

*ke:ŋ 'suspend (a rope)'; Mla ke:ŋ 'carry in a strap across forehead';
  TinPrai kheŋ 'stretch or hold up a rope'

*ke:t 'deaf' (?); Mla ke:t 'ear'; TinMal khe:t 'deaf'(TinMalA also hmɔj
  khe:t'deaf' where hmɔj means 'ear'; TinPrai has klup ntho:r 'deaf',
  where ntho:r likewise means 'ear'); the change of meaning in Mlabri
  may stem from a combination of two words originally both meaning
  'deaf': ke:t and hlu:t

*kir 'thunder'; Mla kuur 'thunder; there is thunder'; most Tin khuur or
  khə:r (TinMalC khaw) 'thunder', TinMalA also with trace of a prefix:
  ŋkhər 'thunder; there is thunder'; Khmu has kir 'thunder' just like
  Mlabri)

*kɔ:k in *plE2(:)₁ʔ kɔ:k (see the former entry); Mla pleʔ kɔ:k 'the thyroid
  cartilage; Adam's apple'; TinMalA phlɛʔ khɔ:k 'id.'; NB: this is a Tai
  word but must have been borrowed very early into Tinic so as to
  differ between Mlabri and Tin because of the mutation in Tin; if it
  were a recent loan one might expect something like mak kɔ:k (Tin-
  MalA children also say phrɛʔ ŋɔ:k as a probably newer formation
  involving ŋɔ:k 'neck; throat')

*koh 'cut apart; chop up'; MlaB koh 'cut apart'; Tin (Mal and Prai)
  khoh 'chop up'; NB: Khmu has koh 'chop'

*kO₂:S₁ 'wooden spear'; MlaC ko:lh (idiolectally ko:jh) 'wooden shaft of
  spear'; TinMal khɔjh, TinPrai khɔt 'spear' (the latter borrowed into
  Mlabri as khɔ:t; NB: the etymon seems lost in TinMalA, which has
  another interesting etymon in the meaning of spear: ʔmɔ:ŋ)

*kot 'encircle'; MlaAB kot 'encircle; bend around; hug', also as com-
  ponent of a reduplicated form: MlaA kotkwat 'hug and carry in one's
  arms' ; TinPrai kho:t 'curl up in resting position' (borrowed as khot in
  Mlabri, also cf. Tai)

*krE₃:c 'bite; cut'; MlaB kre:c 'cut with biting implement; bite', cf.
  TinPrai khra:c 'scratch with claws'

*krE₂:ʔ 'hit the mark'; Mla kre:ʔ 'hit the mark'; a kre:ʔ 'that is correct';
  TinPrai khrɛʔ 'correct'

*krɔ: 'ask for'; MlaA 'hope for', BC 'ask for'; TinPrai khrɔ: 'ask for'
  (connection to Tai but obviously old in M-K; Khmu has krɔ:)

*kru:t 'scrape'; Mla in the reduplicated form krutkrwe:c 'scratch'; Tin-
  Prai khru:t 'scrape'

*klE₄(:)₂p 'squeeze something between two pins'(?); Mla klap 'hold

something squeezed between two pins' (mostly as *?ek kɽlap klap* 'hold with a bamboo forceps'), cf. Tin Mal = TinPrai *kʰlɛːp* '(hold with a) forceps'

*klɽp 'small (woven) box; lid of a box'; Mla *klɽp* 'small (woven) box'; TinMal *kʰəp* = TinPrai *ⁿkʰəp* 'small (woven) box'; TinMalA (with trace of affixation) *ⁿkʰəp* = TinPrai *kʰləp* 'lid of a box'

*klɔk in *ple? klɔk* 'forest fruit species'; Mla *ple? klɔk* 'forest fruit species'; TinPrai *pʰɛ? kʰɔk* 'id.'

*klɔt 'knot'; MlaB *klɔt* 'knot; tie a knot'; TinPrai *silɔt* 'knot'

*kwaːj 'edible tuber species'; Mla *kwaːj* 'edible tuber species'; TinMal *kʰwaːj*, TinPrai *kʰoːj* or *kʰweː* 'id.'

*kwÄɪŋ 'something round'; Mla *kwʌŋ* 'something round, e.g. seeds of a fruit; CLASSIFIER for round things', MlaAB *k̩lwʌŋ* 'ball; egg; body-part of round shape'; TinPrai *kʰwaŋ* 'seed (in fruit)'

*b-

*bah 'for light to appear after darkness'; Mla *?a bah* 'it is early dawn'; Tin *pah sa?* 'it is early dawn', TinPrai *pah* 'appear', across Tin also as causative formation *ᵐbah mat* 'open one's eyes', cf. Khmu *bah* 'bright; morning'

*bi(ː)ɪ? 'full'; Mla *biː?* 'full; satisfied'; Tin (Mal and Prai) *pi?* 'full' (also in K: *bi?*)

*biak 'bear, n'; Mla *bɛːk* 'bear'; Tin *piʌk* 'bear'

*biar 'two'; Mla *bɛːr* 'two'; Tin *piʌr* (etc., in TinPrai also heard as *piʌ?*) 'two'

*bÄɪt 'dip water; scoop up water'; Mla *bʌt* 'id.'; TinPrai *pat* 'id.'

*bË2ːn 'be able to; get'; MlaB *bɽːn* 'be able to; it is permissible; get'; TinMal *pɔːn* 'get; be able to', *?aj pɔːn* 'cannot' (cf. Khabit *ban* 'know')

*bÏɪl 'die; dead'; Mla *buɪl* 'id.'; Tin *pəl* 'id.'; Tin (Mal and Prai) with causative prefix: *ᵐbəl* 'kill; slaughter', cf. Mla *pabul* 'id.' (e.g. MlaB *pabul ɕiːŋ* = TinMalA *mbəl siŋ* 'slaughter a pig')

*bO3? 'carry (a child) on the back'; MlaA *bɔ?* 'carry (a child) on the back'; TinPrai *pa?* 'carry on the back'

*bɔːŋ 'swell up'; MlaB *-bɔːŋ* in the reduplication *bungbɔːŋ* 'swell up; be swollen'; TinPrai *pɔːŋ* 'swell up; be swollen'

*bO1ː 'whisper'; MlaB *bɔ-* (presumably < *bɔː*) in *bəbɛːp* 'whisper'; TinMalA *poː* 'whisper'

*bO2ŋ 'eat'; Mla *boŋ* 'eat' (AC: 'eat meat'); Tin *pɔŋ* 'eat'

*bo(ː)ₗm 'jar'; MlaB *boːm* 'bottle'; TinPrai *pom* in *pom tʰeː* 'pottery jar'

*boh 'boil'; with reduplication in MlaB *bohboh* 'be boiling'; TinMalA ᵐ*boh* 'be boiling'; with causative affixation in TinPrai ᵐ*boh* 'boil something'

*boːʔ 'breast; breast feed'; Mla *boːʔ* 'breast; breast feed'; Tin *poʔ* 'breast; breast feed' (cf. Khabit and Khmu *buʔ*)

*brə(ː)₂Sₗ '(for the wind to) blow sweepingly (?); MlaC *brɤlh* 'blow (of wind); TinPrai *prəːt* 'sweep'

*brɔːŋ 'deceased person'; MlaC *brɔːŋ* 'long since dead'; TinPrai *prɔːŋ* 'spirit of a deceased person; ghost' (TinMalA has *ɕɔːj* for 'ghost')

*broʔ 'wound, n'; Mla *broʔ* 'wound'; TinPrai *proʔ* 'skin lesion or infection'

*broːc 'spit liquid out of the mouth'; MlaAB *broːc* 'id.'; TinMalA *proːc* 'id.', in TinPrai with traces of a prefix:ᵐ*broːc* 'spit liquor (as part of a ceremony)'

*blah 'escape'; MlaAB *blah* 'run off in different directions'; TinMal *plah* 'escape; disappear', TinPrai *plah* in *plah prah* 'escape'; also with causative affix: Mla *pablah* '(A) release; (B) brush or push away'; TinMalA (with aspiration as a trace of affixation) *pʰlah* 'release; let somebody escape'

*blÏ̈ɪt 'extinguished (fire)'; Mla *bluɪt* 'id.'; TinPrai ᵐ*bət* 'extinguished'

*blU₂ːŋ '(bamboo) shoot'; MlaA *bluːŋ* 'bamboo shoot'; TinPrai *poːŋ* 'sprout shoot'

*bluːHₗ 'thigh'; Mla *bluːʔ* 'thigh; upper part of leg'; TinMal(A) *pluː* 'leg' (NB cf. K: *bluʔ* 'thigh')

*d-

*duːc 'tired'; Mla *duːc* 'tired'; TinPrai with traces of prefixation: ⁿ*duːc* 'drunk' (possibly a causative with euphemistic meaning: 'made tired')

*dU₁k 'poor'; MlaAB *dok* 'poor'; TinPrai *tuk* 'poor'

*drU₂ːɲ/*drU₂ːn 'termite'; Mla *druːɲ*, TinMla *kjoːn* 'id.'

*ɟ-

*ɟak 'go'; Mla *ɟak* 'go'; Tin *cak* 'go'

*ɟen 'be skilled'; Mla *ɟen* 'be skilled'; TinPrai with causative affixation ⁿ*ɟen* 'teach' (apparently not known in TinMalA)

*ɟIŋ 'sew', Mla *ɟeŋ* sew'; TinPrai *ciŋ* 'sew'

*ɟʌS₂ 'taste good'; Mla ɟʌjʰ 'taste good'; TinPrai cɛt 'taste good', cf. Mlabri ce:t!

*ɟə(:)ŋ 'foot', Mla ɟɤ:ŋ 'foot'; TinMal cəŋ, TinPrai ceŋ 'foot' (cf. Khmu ɟiəŋ 'foot')

*ɟĬn 'light; bright'; Mla ɟun 'candle'; TinMal cən 'bright'

*ɟɔʔ 'small of size (?)'; MlaB ɟɔʔ 'a few; a small quantity'; Tin-MalA cɔʔ 'short (of physical size)'

*ɟo:k 'suck'; MlaAB ɟo:k 'suck'; TinMalA & TinPrai co:k 'suck (through a straw)', cf. TinMalA cok pʰɔj cuʌk 'collect saliva and spit'

*ɟu:r 'descend'; Mla ɟu:r 'descend'; Tin MalA cu:, other Tin cu:r, cu:l or cu: 'descend'; TinPrai with trace of prefixation (medial meaning?): ⁿɟul ⁿsep 'come downhill'; Khmu has ɟu:r 'descend' just like Mlabri

*ɟra: 'skinny'; MlaAB ɟra: 'id.'; TinMlaA kra:, TinPrai sira: 'id.' (cf. Khabit dʐa: 'thin')

*ɟrË₂h 'for a fowl to make its characteristic sound'; MlaC ɟrɤh 'id., e.g. to crow).'; TinMalB kjɔh 'crow; cluck'

*ɟruʔ 'deep'; Mla ɟruʔ 'deep'; TinMalA kruʔ, TinMalB kjuʔ, TinMalC ku:ʔ, TinPrai & TinCuun cuʔ 'deep'; NB: the occurrence in Mlabri may predate Tinic but the Tin forms are straightforward continuations of *ɟruʔ (Khmu also has ɟruʔ but is hardly a loanword); the same etymon occurs with a simpler onset in Waic: Proto-Waic *riʔ 'deep'

*ɟjʌl 'evening'; Mla ɟjʌl 'evening'; TinMalA kiʌl (TinPrai cɛl) 'evening; dark' (cf. Khabit n(ə)ŋdʑəl 'evening')

*g-

*ga:p 'hold (something edible) in the mouth'; MlaB ga:p 'id.'; TinPrai ka:p 'id.'

*ge:t 'chop'; MlaBC ge:t 'cut; chop'; TinPrai ke:t 'split pieces off a log'

*gaS₂ 'nine'; Mla gajʰ 'nine'; TinPrai (conservative lg. in Laos) ⁿgat 'nine' (NB: the etymon exists in Khmu but strangely deviant: with an old voiceless initial and with a variety of finals; the closest recorded form – by this author – is Ban Huay Puk Khmu kas)

*gIh 'scratch'; Mla in the reduplicated form gehgeh 'scratch'; Tin (Mal and Prai) kih 'scratch'

*giaŋ 'house'; Mla ge:ŋ 'house; lean-to'; TinMal kiʌŋ, TinCuunh ci:ʌŋ, TinPrai (Nam Phi') cɤ:ŋ, (DJ:) cəŋ 'house'

*gih 'be ablaze'; Mla guh 'id.'; TinMalA kuh 'id.'; also with causative affixation: MlaB paguh 'make something light up; illuminate'; Tin-

MalA *ᵑguh* 'ignite the firewood'; Tin Prai (with an unusual infix) *kruh* 'ignite'

*gÏr 'shaft; handle'; Mla *gur* 'shaft'; TinPrai *kəl* 'handle, n'

*gÏₗm 'suck on something (that is kept in the mouth)'; Mla *gum* 'id.'; TinPrai *kəm* 'id.'

*goh 'break (and fall down)'; MlaAB *goh* 'id.'; also with causative affixation: MlaA *pagoh* 'cause something to break; chop'; Tin (Mal and Prai) *ᵑgoh* 'strike something; knock on something'

*gul 'seven'; Mla *gul* 'seven'; TinPrai (conservative lg. in Laos) *ᵑgul* 'seven'; Khmu has *guːl* 'seven' in ceremonial language

**grË₄(ː)₂ŋ* 'sifter'; MlaB *grɤŋ* 'sifter'; TinPrai with trace of causative prefixation: *ᵑgraːŋ* 'sift'

*groːʔ 'make a sound'; Mla *groːʔ* 'make a sound; one hears a sound'; TinMalA *kroʔ* 'for an animal or bird to make its characteristic sound'; TinPrai *kroʔ* 'cry out; make a sound' (TinMalA also has another similar word *krɔh* '(for a rooster to) crow')

*glË₂h 'ascend': MlaBC *glɤh* 'ascend'; old-fashioned TinPrai *klɔh* 'ascend' (David Jordan, pers. comm.)

**glÏ₂(ː)ₗ* 'head': Mla *glɤːʔ* 'head'; Tin *kluʔ* (also: *kuʔ*) 'id.'; Mlabri clearly has the older form as evidenced by Khmu *glɤʔ* but the meaning of the word has apparently changed into 'hair' in Khmu (one would have expected something like *nsok for 'hair', as in Tin and more widely)

*gwaː 'search for; hunt for'; MlaB *gwaː* 'id.'; TinPrai *kwaː* 'id.'

**gwE₃(ː)ₗt* 'open a hole'; MlaB *gwɛːt* 'drill to hollow out', TinMalA *kwat* 'poke into with a pin'; TinPrai *kwat* 'open a hole with a knife'

Table 3-2. Words reconstructed as monosyllabic with other initials

*m-

*mat 'eye'; Mla *mat* 'eye'; Tin *mat* 'eye'

*miaʔ 'rain'; Mla *mɛːʔ* 'rain', *mɛʔ hot* 'it rains'; Tin *miʌʔ* 'rain', *leh miʌʔ* 'it rains' (note that an old velar prefix, which is evidenced for this word by Khmu, was absent already at the Tinic level)

*miam 'blood'; Mla *mɛːm* 'blood'; Tin *miʌm* 'blood', *leh miʌm* 'it bleeds'

*mÄₗc 'experience; see'; Mla *mʌc* 'see'; TinMalA *məc*, TinMalB *mət* 'see'; TinPrai *mec* 'know'

**mÄₗ(ː)ₗr* 'snake'; Mla *mʌːr* 'creep, v'; TinMalA *mar* 'snake'; Tin-

MalC *maw*; TinPrai *maːl* 'snake', cf. Khmu *mar* 'snake' (Mlabri has *mʌːʔ* as the word for 'python')

mË₃(ː)₂ʔ 'mother'; Mla *mɤʔ* 'mother', TinMalA *məj*, TinPrai *meː* 'mother'

mO₁ːj 'one'; Mla *mɔːj* 'one'; Tin (including TinMalC, whereas Tin-MalA has generalized a sandhi-form *maː*) *mɔːj* 'one' (Khmu sides with Tin)

mO₁h 'nose'; Mla *mɔh* 'nose'; TinMal *moh* 'nose' (TinPrai *muh*, cf. Khmu *muh* 'nose')

mu(ː)₁k 'smell (perceive the smell or fragrance of an object), V '; Mla *muːk* 'smell (with the nose)'; Tin *muk* 'smell (with the nose)'

mrɛːŋ 'bamboo flooring'; Mla *diŋ mrɛːŋ*, *dimrɛːŋ* 'bamboo flooring' (*diːŋ* denotes a bamboo species); TinPrai *ᵐbrɛːŋ* 'bamboo flooring'

mlɛ(ː)₁t 'press'; Mla *mlɛːt* 'pinch; strangulate'; TinPrai *ᵐblɛt* 'press (a button)'

*ɲ-

ɲɔːk 'net'; MlaA *ɲɔːk* 'plaited bag (like a net)'; Tin (Mal and Prai) *ɲɔːk* 'trap (for fishing)'

*ŋ-

ŋaʔ 'itch; it itches'; MlaB *ŋaʔ* 'id.'; TinMalA *ŋaʔ* 'id.' (the word occurs identically in Khmu but undoubtedly goes back to Tinic)

*ʰm-

ʰmɛʔ 'new'; Mla *ʰmɛʔ* 'new'; TinMalA *ʰmɛʔ*, other Tin *mɛʔ* 'new', cf. Khmu *ʰmmɛʔ* 'new'

ʰmÄ₂n 'girl (?)'; MlaC *ʰmʌn* in *ɛːw ʰmʌn* 'little sister'; TinMalA *ʰmɔn* in *kʰwan ʰmɔn* 'daughter'

ʰmuːH₂ 'areca nut'; Mla *ʰmuː* 'areca nut'; TinMal *ʰmuʔ* 'areca nut'

*ʰn-

ʰnÄ₁(ː)₂m 'year'; Mla *ʰnʌm* 'year'; TinMalB (Ta Noi) *naːm* 'year'

*S₁-

S₁iaw 'honey'; Mla *tʰɛːw* 'honey-like substance'; TinMlaA *siʌw* 'sweet of taste'; TinPrai *siw* 'honey-like substance' (*siw ᵐbuat* 'bee's honey')

S₁Ä₂c 'meat; flesh'; Mla *tʰʌc* 'meat'; TinMalA *çəc* 'meat'; TinPrai *sec* 'meat'

*S₁Ë₁:*p* 'insert (one's arm)'; MlaA *tʰɤːp* 'insert one's arm'; TinPrai *soːp*
sih 'insert one's arm'

*S₁Ë₂:*l*/S₂Ë₂:*l* 'vomit'; MlaBC *tʰɤːl* 'vomit'; TinMalA *çɔl* 'vomit'; Tin-
Prai *sɔːl* 'vomit'; cf. Khmu *həːl, hiəl* 'vomit'

*S₁Ë₂:*ŋ* 'five'; Mla *tʰɤːŋ* 'five'; TinPrai *çɔːŋ* 'five'; cf. Khmu *pʰiəŋ* 'five
(in ritual language)' (NB in this numeral, Khmu clearly continues a
bisyllabic form *pVsVːŋ > *pVhVːŋ, whereas Tinic has only the
main syllable)

*S₁*uːl* 'food'; MlaBC *tʰuːl* 'meat and/or vegetables eaten with staple
food' (also as causative derivation *batʰuːl* 'have (meat/vegetables)
with the staple food'); TinMalA *çul*, other TinMal *sul* 'food'

*S₁*maːl* 'personal spirit'; MlaC *tʰəmaːl* 'spirit of deceased person'; Tin-
MalA *ʰma(ː)l*, TinMalB *ʰmaːl*, TinPrai *simaːl* 'personal spirit

*S₁*riaŋ* 'tooth'; Mla *tʰrɛːŋ* 'tooth'; TinMalA *çiʌŋ*, TinPrai *səŋ / çəŋ*
'tooth'

*S₂-

*S₂*aS₁* 'itching (on the skin)', MlaA -*salʰ* in *br̥salʰ* 'itching'; TinMalA
çajʰ 'itching'

*S₂*uʔ* 'smelly'; Mla in *çoʔuːm, çuʔuːm* '(A) have a bad smell; (BC) be
fragrant'; TinMalA *çuʔ* 'have a bad smell'

*S²*uʌk* 'salt'; Mla *çuʌk* 'salt'; TinPrai *çuʌk* 'id.' (not a Khmu word;
Khmu has different etyma)

*r-

*ra:*p* 'run after'; Mla *raːp* 'chase; run after'; TinMalA *raːp* 'run after';
TinPrai *raːp* 'come up to'

*rE₂*h* 'tear across; rip'; MlaBC *reh* 'id.'; TinPrai *rɛh* 'id.'

*riaS₁ 'root'; Mla *rɛːlʰ* 'root'; TinMalA *ᵑgriʌjʰ* (*ʔuʌŋ*) 'root (of tree)'
(NB: this is Khmuic *ia not *aː, cf. Khmu *riəs*)

*rɨh 'clean up'; MlaB *ruhruh* 'clean the field; remove weeds' (redu-
plication); TinMalA *ruh* 'throw rubbish out'; TinPrai *ruh* 'remove
weeds'

*rɔ(ː)₂*c* 'rinse the interior of something (e.g. to clean the bowels of a
slaughtered animal)'; MlaA *rɔc* 'id.'; TinMalA *rwac* (sic), TinPrai
rɔːc 'id.'

*rwaːj 'tiger'; Mla *rwaːj* 'tiger'; TinMalA *rwaj*; other Tin *waːj* 'tiger',
cf. Khmu *rwaːj* 'tiger'

*l-

*lak 'gourd'; Mla lak 'gourd; tuber'; TinMalA lak in pʰlɛʔ lak 'gourd'; TinPrai lak in ᵐbah lak 'gourd'

*lam 'tree'; Mla lam 'tree'; Tin lam 'tree'

*laːH₁ 'upper arm'(?); Mla laːʔ 'shoulder'; TinMalA laː in pak laː = palaː, pəlaː 'shoulder'; TinPrai laː 'arm'

*lo(ː)ɩn 'personal spirit'; Mla loːn 'personal spirit'; TinPrai silon 'spirit; soul' (apparently not known in TinMalA, in which only the word ʰmaːl is used)

*loʔ 'search for; hunt for'; Mla loh 'search for; hunt for'; TinPrai loʔ 'hunt for' (the Mlabri form may have developed out of a sandhi form without final glottal termination, but its form is suggestive of a connection to Phong loh 'go out')

*luh 'be abused'; MlaB luh 'be abused' (?, cf. MlaAB paluh 'scold; abuse'); TinPrai luh 'offended; angry' (NB if this were from Tai one would expect ʰluː in Mlabri)

*ʰl-

*ʰli(ː)ɩŋ 'forget'; Mlabri with infixation: ʰrliːŋ 'forget'; TinMalA ʰliŋ in ʰliŋ com 'forget'

*w-

*wÄ₁l 'return'; Mla wʌl 'return home'; Tin wal 'return' (lost in MalA, which has toʔ for 'return')

*wË₂ːk 'water'; Mla wɤːk 'water'; Tin ʔɔːk 'drinking water'

*wË₂(ː)ɩŋ / *wË₄(ː)ɩŋ 'chin'; Mla wɤːŋ 'chin'; TinMalA wɔŋ, TinPrai waŋ 'chin'

*ʔj-

*ʔjak 'shit; waste'; Mla ʔjak 'shit; waste'; TinMalA ʔiʌk 'excrement'

*ʔjɛːʔ 'far away'; Mla ʔjɛːʔ 'far away'; TinMalA ʔiʌʔ (sic!), TinPrai jɛʔ saʔ 'do.'

*ʔV-

*ʔec 'take'; Mla ʔec, ʔek; 'take; accept'; TinMal ʔet, TinPrai ʔət 'take'

*ʔĬ₁m 'bathe'; MlaBC ʔum 'bathe'; Tin (MalA and Prai) ʔəm 'bathe'(cf. Phong ʔim 'bathe')

*ʔU₂ːS₁ 'fireplace; firewood; Mla ʔuːlʰ 'id.'; TinMal ʔoːjʰ 'id.'

Table 3-3. Words reconstructed as sesquisyllabic
with a syllabic consonant in the presyllable

(3-3,a) The presyllable has a syllabic nasal followed by a stop

*^{h}ṇto:r 'orifice'; MlaC ^{h}ṇto:r, MlaB ^{h}ṇtor 'orifice; duct'; TinPrai ^{n}t^{h}o:r 'ear'

*S₂ŋkɛr 'nail (on finger/toe)': Mla ȼŋkɛr 'nail'; TinMalA ^{v}k^{h}ear, k^{h}ɛr, TinPrai ^{v}k^{h}ɛl 'nail (on finger/toe)'

*lṃbah 'vegetable species or generic term for vegetables'; MlaB lṃbah 'cabbage'; MlaA in lṃbah p^{h}aku:t 'inedible plant species resembling the p^{h}aku:t vegetable'; TinPrai ^{m}bah in ^{m}bah khro? 'vegetable species', ^{m}bah t^{h}u:r 'mushroom species', etc. (apparently not known in TinMalA)

*S₂ṃbɛ(:)₂p 'lip region'; Mla ȼṃbɛp 'lip region'; TinMal ^{m}bɛ:p 'id.'

*^{h}ṇdË₂l 'heel'; MlaB ^{h}ndɤl 'heel', TinMalA ^{n}dɔl, TinPrai ^{v}ɟɔl 'heel'

*rŋga:p 'mouth'; Mla rŋga:p 'mouth'; TinMal ^{v}ga:p 'mouth'

(3-3,b) The presyllable has a syllabic nasal followed by a sonorant or by ? plus vowel

*kṃrɛh 'old'; MlaC ȼak kṃrɛh (in AB ȼak kṃru:j^h) 'old person'; Tin (Mal and Prai) ^{m}brɛh 'be old'

*gṇrɛ:H₂ 'pestle'; Mla gṇrɛ: 'curry' (with semantic shift); Tin (Mal and Prai) ^{v}grɛ? 'pestle'

*dŋrÄ₁ŋ/ɲṇrÄ₁ŋ 'house pole'; MlaB dŋrʌŋ, MlaA ɲṇrʌŋ 'house pole'; Tin (Mal and Prai) ^{v}graŋ '(house) pole'; the same etymon occurs with a deviating onset in Waic: Proto-Waic *?rɔŋ 'house pole' (Diffloth 1980)

(3-3,c) The presyllable has a syllabic nonnasal sonorant followed by a glide

*kḷwÄ₁ŋ 'round thing' (derivative of *kwÄ₁ŋ 'round thing' as in MlaBC kwʌŋ 'id.'); MlaAB kḷwʌŋ, kulwʌŋ 'ball-shaped object' with consistent sesquisyllabicity; TinMalA k^{h}lwaŋ 'seeds in fruit' (with a true consonant cluster [k^{h}lw-])

Table 3-4. Words reconstructed as sesqui- or bisyllabic
with an initial open syllable

*takat 'be feverish'; Mla takat 'be feverish; feel cold'; TinPrai k^{h}at 'fever' (Khmu has kat 'cold')

*ciˀic 'relaxed' (?); MlaB cuˀuc cɯmɯc 'merriful'; TinPrai siˀuc
   'lazy'
*jəˀiaŋ; MlaBC ɟiˀɛːŋ, MlaA jəˀɛːŋ 'bone'; TinMalA ˀiʌŋ, Tin Pr siˀiʌŋ,
   sijəŋ, jəŋ 'bone'; Khmu has cˀaŋ, and Diffloth reconstructs Proto-
   Waic *sˀaŋ suggesting that the voicing of the initial in Mlabri may be
   a secondary Tinic phenomenon)
*Sɪapaɲ 'small insect species'; MlaB tʰapaɲ 'small bug that bites at
   night'; TinMalA ᵐpʰɤɲ 'housefly'
*Sɪapuːl 'stomach'; Mla tʰapuːl 'id.'; TinMalA and some TinPrai ᵐpʰul
   'stomach', other Mal pʰuːl (Prai also pʰul)
*S₂oˀu(ː)ɪm 'smell, v'; Mla ɕoˀuːm 'smell'; TinPrai sijum 'smell good'
   (cf. Phong sĕˀum 'become stale')
*rəˀək 'upper part of chest'; MlaAB r̩ˀɤk (A mostly rəˀɤk) 'id.'; Tin-
   MalA ˀək 'id.'

Table 3-5. Words reconstructed as reduplicatives

*kukɔʔ 'neck'; MlaB kukɔʔ 'neck'; TinMalA ᵑkʰɔʔ (TinPrai ᵑkʰɔː)
   'neck' (NB: connection to Tai spurious since one would then expect
   *g- not *k-)
*bohboh 'be boiling' (cf. *boh in Table 3-1); MlaB bohboh 'be
   boiling'; TinMalA ᵐboh 'be boiling'

This ends the data set to be used in Mlabri-Tin comparisons for the pur-
pose of establishing sound-laws and other regularities. It should be
remembered that for words with initial sibilants or continuants it is not
possible to draw a rigid boundary between words that go back to Tinic
time and words that were recently borrowed into Mlabri.

Correspondences in rhymes
between Proto-Tinic and Mlabri

In cognates that seem to qualify as being of Tinic origin there is, on the
whole, an excellent agreement between the final oral consonants of
Mlabri and Tin, and likewise between the initial consonants, keeping in
mind that voiced and voiceless stops in Mlabri correspond to voiceless
and aspirated voiceless stops in Tin. The points of trouble in the compar-
isons are: (i) syllable termination, (ii) vowel quantity, (iii) vowel qual-

ity. The remainder of Part 3 consists of a long-winded analysis of the comparative data sets with respect to these three issues.

Syllable termination types:
-*h* , -*ʔ* or open syllable

The situation from a Tinic perspective. In several instances Mlabri and Tin agree on the syllable termination. Proto-Tinic must, for example, have had *-*h* in

Mla *bah*, Tin *pah* '(for light to begin to) appear'
Mla *mɔh*, TinMal *moh* 'nose'

but *-*ʔ* in

MlaB *jɔʔ* 'a few', TinMalA *cɔʔ* 'short (of physical size)'
Mla *jruʔ*, TinMalA *kruʔ* 'deep'

One of the conspicuous features of Tinic syllable structure in a M-K context is the occurrence of long open syllables. Some examples from the comparative data set will be given below. Thus I have to take issue with Diffloth's statement that open syllables do not exist in Northern M-K branches such as Khmuic except in borrowings. They do indeed exist in the lesser known Khmuic languages although they are more exceptional in Khmu, which typically has final -*ʔ* in the words in question (an open syllable in Khmu is exemplified by *krɔː* 'ask for', obviously not a loanword from Lao since it should then have the form **kʰɔː-* not *krɔː-*). In some cases both Mlabri and Tin exhibit an open syllable, e.g.

Mla *gwaː*, Tin *kwaː* 'search for; hunt for'
Mla *jraː*, TinMal *kraː*, TinPrai *siraː* 'skinny; thin'
Mla *krɔː*, TinPrai *kʰrɔː* 'ask for'

but often, it is so that Tin has an open syllable where Mlabri has a final glottal stop (it should be noted that in Mlabri, the form of the word for 'hand' is identical with that of the word for 'eight' above although Tin differs in its reflexes: *tʰiː* versus *tʰiʔ*):

Mla *bi:ʔ*, TinPrai *pʰi:* 'hulled rice'
Mla *blu:ʔ* 'thigh', TinMal *plu:* 'leg'
Mla *ka:ʔ*, TinMal *kʰa:* 'fish'
Mla *ti:ʔ*, Tin *tʰi:* 'hand'
Mla *la:ʔ*, TinMalA *pak la:, pala:* 'shoulder'

Finally, it seems quite exceptional for Mlabri to have an open syllable versus a laryngeally checked syllable in Tin. I have found this only with words that have -*ʔ* not -*h* in Tin, and only with words which one might suspect of being loanwords in Mlabri such as

Mla *ʰmu:*, TinMal *ʰmuʔ* 'areca nut'
Mla *gɲrɛ:* 'curry', Tin *ᵛgrɛʔ* 'pestle'

The last-mentioned example is furthermore strange in that its meaning in Mlabri has switched from the expected meaning of 'pestle (used when pounding spices for a curry)' to the meaning of 'curry' or perhaps rather 'soup'. Still, I tend to consider this word as belonging to the Tinic vocabulary because the Mlabri form is in other respects quite conservative. If it were borrowed from Tin after Tinic time (which might account for the semantic shift) it must have been borrowed very early indeed since the present-day form in Tin shows a restructuring of the initial part of the word that is typical of Early Tin, and which must predate Proto-Tin.

To sum up: there are four regular reflexes between Mlabri and Tin:

Mlabri open syllable ~ Tin open syllable
Mlabri -*ʔ* ~ Tin open syllable
Mlabri -*ʔ* ~ Tin -*ʔ*
Mlabri -*h* ~ Tin -*h*

The question is whether the correspondence Mlabri -*ʔ* ~ Tin open syllable is due to loss of the final glottal stop in Tin or introduction of it in Mlabri. There is no simple way to solve that issue in terms of merger in one of the languages or split in the other, for the phonological contexts in which the first three correspondence types above occur are not mutually exclusive. It is necessary, therefore, to go beyond Tinic.

If one looks at other Khmuic, Phong has a diphthong in 'hand': *təj*, and an open syllable in 'fish': *ka:*, Ksingmul likewise: *taj, ka:*, but Khmu has the glottal stop in both: *tiʔ, kaʔ*. For 'shoulder' Khmu has

bla?, likewise with the glottal stop (that form may have some historical connection with the Tin phrase *pak la:*, which in fact exists in a contracted variant *pəla:*). Off-hand it looks as if the Khmuic language divide themselves into two blocks with respect to this particular distinction, but the blocks are strange from a geographical perspective.

The syllable-termination crux is a first warning that the comparison between Mlabri and Tin does not lead to stream-lined phonological correspondences in all cases.

Since an open syllable is a more marked structure in this area than a stopped syllable one would expect a development *open syllable > stopped syllable* rather than the opposite. This suggests that the glottal stop in Mlabri and Khmu is a shared innovation. As I argue in various parts of this monograph, however, Mlabri and Khmu do not together form a sub-branch of Khmuic; if there is affinity between them it is either a very old feature or it is due to recent borrowing from Khmu into Mlabri.

In this case my assumption is that the words for 'hand', 'shoulder' and 'fish' in Mlabri are unlikely to be recent loans. On the contrary, it seems likely that they are very old in the language and not even part of the Tinic layer. I would take the situation in Tin, Phong and Ksingmul as indicative of the old state of affairs in Khmuic. The most likely explanation of the stopped forms in Mlabri and Khmu, then, is that these two languages have an old history that goes beyond Khmuic (in my view one has to accept that Khmu is not a full-fledged Khmuic language, paradoxical as that may sound).

Looking beyond Khmuic. As for laryngeal finals versus open syllable there is an interesting scenario in Palaung-Riang. Shorto (1952: 50) posits a loss of Proto-Northern-Mon-Khmer *-h* in Riang-Lang resulting in open syllables, whereas the final is continued as Palaung *-h*. Palaung, however, has an open syllable in several forms in which Riang-Lang has -*?* (like Khmu does in some of the shared forms). There are three pronominal forms in Shorto's paper (ibid.: 59) that illustrate this, and for which Mlabri has cognates in -*h* (I ignore whether the Mlabri and Khmu forms below are basically dual or plural forms, and I ignore differences in inclusiveness in Palaung):

1. p. sg. Pal *?ɔ:* ~ R-L *?ó?* ~ Khmu *?o?* ~ Mla *?oh*
1. p. pl. Pal *?ɛ:* ~ R-L *?é?* ~ Khmu *?a?* ~ Mla *?ah*
2. p. pl. Pal *pɛ:* ~ R-L *pé?* ~ Khmu *bɔ:* ~ Mla *bah*

About this -ʔ in Riang-Lang, Shorto keeps an open mind: "Whether Riang-Lang -ʔ should be regarded as an isolated survival of the proto-NMK pattern, or as a neologism, is less certain" (1952: 50).

Mitani (1979: 142) gives several examples of Palaung-Riang cognates which exhibit the difference between open syllable in Palaung and stopped syllable in Riang (but does not comment on the origin of the difference because the paper is concerned with vowels only). Several of these have Mlabri cognates. I list Mitani's Proto-Palaung forms with their Mlabri cognates, first words in which Mlabri has -ʔ, like Riang:

PP *ti 'hand' ~ Mla tiːʔ
PP *bri 'jungle' ~ Mla briːʔ
PP *ple 'fruit' ~ Mla pleːʔ
PP *he 'firewood' (NB: Riang has kʰ-) ~ MlaAB ʰŋkeːʔ
PP *-mɛ 'new' ~ Mla ʰmɛʔ
PP *su 'sick' ~ MlaBC ɕoʔ
PP *sɔ 'dog' ~ Mla ɕɔːʔ
PP *ma 'mother' ~ Mla mɤʔ
PP *ka 'fish' ~ Mla kaːʔ
PP *hva 'monkey' ~ MlaAB tʰawaːʔ

Most of these etyma also occur in Diffloth's lexicon of Proto-Waic (1980), and without exception they have a final glottal stop just as in Riang: PW *teʔ 'hand', *breʔ 'forest, etc.', *pliʔ 'fruit', *kʰiʔ '(fire) wood', *sɔʔ 'dog, *maʔ 'mother', *hwaʔ 'monkey'. This pretty much settles the issue: the final -ʔ goes back to Proto-Palaungic.

From Mitani's examples of words with open syllable there remain three words in which Mlabri has -h versus -ʔ in Riang; two of these also occur on Diffloth's Proto-Waic list and then with -ʔ like Riang. I give the Proto-Palaung, Proto-Waic and Mlabri forms:

PP *ʔɔ 'I' ~ PW ʔiʔ (< rounded vowel in Pre-Waic) ~ Mla ʔoh
PP *mi 'you (sg.)' ~ PW meʔ ~ Mla mɛh
PP *pɛ 'you (pl.)' ~ Mla bah

Importantly, these three forms are from the system of personal pronouns, like the forms cited from Shorto above, thus explaining why Mlabri has -h: it is suffixal.

Vowel quantity

It is useful to distinguish between the quantitative pattern in syllables with laryngeal termination, i.e. *-h* or *-ʔ*, and the pattern in syllables terminated by an oral final.

Vowel quantity in syllables with laryngeal termination. The distribution of long versus short vowels is fairly straightforward.

Before *-h*, vowels are always short both in Mlabri and Tin, so words reconstructed with Tinic *-*h* can at the same time be reconstructed with a short vowel. There is a minor problem with the data in that in the case of half-low or low vowels it is often difficult to hear the difference between open syllables with a long vowel (which can end in a weakly audible release of air) and syllables containing a short vowel plus *-h*. Open main syllables are rare across the lexicon compared to syllables ending in *-h*.

As for syllables with final *-*ʔ* in the Tinic reconstructions (i.e. also in Mlabri), some Mlabri words have a consistently long vowel, others a consistently short vowel; there are even minimal or near-minimal pairs such as *ʰmɛʔ* 'new' (TinMalA likewise: *ʰmɛʔ*) versus *mɛːʔ* 'rain' (Tin: *miʌʔ*). In this instance, comparative evidence shows that the first word had a short vowel, the second word a long vowel in Khmuic. Indeed, one would not expect a vowel-length contrast to arise spontaneously before a glottal stop.

As for words with Tinic *-*ʔ* in which Mlabri has a long vowel, the regular Tin reflex is a short vowel:

Mla *tiːʔ*, Tin *tʰiʔ* 'eight'
Mla *boːʔ* , Tin *poʔ* 'breast; breast feed'
Mla *glɤːʔ*, Tin *kluʔ* 'head'

(in the last example Khmu has a short vowel: *gləʔ*; in Khmu the word means 'hair').

The obvious conclusion is that the length contrast before *-ʔ* belongs to the Proto-Tinic level, and that vowel shortening before *-ʔ* has occurred in Tin. The word for 'rain' shows that the shortening rule must be qualified: in Tin, diphthongs were retained in this position (this is true no matter whether it was an old diphthong or a long vowel that had already diphthongized at the Tinic level).

Vowel quantity in syllables with an oral final. In several instances Mlabri and Tin agree on a long vowel, e.g.

MlaBC *geːt* 'cut; chop'; TinPrai *keːt* 'split pieces off a log'
Mla *mɔːj*, Tin *moːj* 'one'

and in several instances Mlabri and Tin agree on a short vowel, e.g.

MlaAB *dok*, TinPrai *tuk* 'poor'
Mla *buɪl*, Tin *pəl* 'die'

On the other hand, there are instances in which Mlabri has a long vowel but Tin a short vowel, e.g.

Mla *muːk*, Tin *muk* 'smell (with the nose)'
MlaBC *tʰuːl* 'meat and/or vegetables eaten with staple food' Tin *sul* 'food'

and a few instances in which Tin (or some Tin) has a long vowel but Mlabri a short vowel, e.g.

Mla *çɱbɛp*, TinMal *ᵐbɛːp* 'lip region'
MlaA *rɔc* 'rinse the interior of something (e.g. bowels)', TinMalA *rwac*, TinPrai *rɔːc* 'id.'

Finally, there are examples in which Mlabri has a long vowel and Tin a diphthong; these can be accounted for by a sound change, the underlying Khmuic vowel being long.

The analyst's expectation is, of course, that two out of the four monophthongal correspondence types can be accounted for by rule, i.e., lengthening in certain environments, or shortening in certain other environments, in one or another of the two branches of Tinic. A promising context in which to test that would be final nasals, since changes in vowel-length before nasals is a well-known type of sound-change globally. I shall, therefore, present the evidence in my data set; for completeness of exposition I include examples with a diphthong in Tin versus a long vowel in Mlabri. The relevant data then add up to 12 + 10 + 2 + 7 + 4 = 35 cognate pairs with final nasals:

Shortness in both Mlabri and Tin, 12 words:

Mla *boŋ* 'eat' ~ Tin *pɔŋ* 'id.'

Mla *gum* 'suck on something' ~ TinPrai *kəm* 'id.'

Mla *ɟeŋ* 'sew' ~ TinPrai *ciŋ* 'id.'

Mla *ɟun* 'candle' ~ TinMal *cən* 'bright'

MlaC *pr̩mpam (male lg., female lg.: pimpim)* and *r̩mpʌm* 'butterfly' (cf. the MlaA borrowing *tr̩lam pam* 'moth' from Khmu) ~ TinMalA *ᵐpʰam* in *ᵐpʰojʰ ᵐpʰam* 'butterfly'

MlaBC *ʔum* 'bathe' ~ Tin *ʔəm* 'id.'

Mla *lam* 'tree' ~ Tin *lam* 'id.'

MlaC *ʰmʌn* in *ɛːw ʰmʌn* 'little sister' ~ TinMalA *ʰmɔn* in *kʰwan ʰmɔn* 'daughter'

Mla *ɟen* 'be skilled'; TinPrai *ⁿɟen* 'teach'

MlaB *tʰapaɲ* 'small bug that bites at night' ~ TinMalA *ᵐpʰɤɲ* 'housefly'

MlaB *dɲrʌŋ*, MlaA *ɟn̩rʌŋ* 'housepole' ~ Tin *ⁿgraŋ* '(house)pole'

MlaAB *k̩lwʌŋ, kulwʌŋ* 'ball-shaped object' ~ TinMalA *kʰlwaŋ* 'seeds in fruit'

Length in Mlabri only, 10 words:

MlaB *boːm* 'bottle' ~ TinPrai *pom tʰeː* 'pottery jar'

Mla *ɟɤːɲ* 'foot' ~ TinMal *cəɲ*, TinPrai *ceɲ* 'foot' (cf. Khmu *ɟiəŋ* 'id.')

Mla *peːɲ* 'shoot' ~ Tin *pʰəɲ* 'id.' (cf. Khmu *piɲ*)

Mla *poːm* 'five' ~ Tin *pʰon* 'five' (Khmu has *puən* or *puːn*)

MlaC *plɤːm* 'landleech species' ~ Tin *pʰlɔm* 'id.'

Mla *tuːn* 'mole' ~ TinPrai *tʰun* 'big rodent species' (affinity to Tai but obviously very old in Tinic)

Mla *ɕoʔuːm* 'smell' ~ TinPrai *sijum* 'smell good'

Mla *wɤːɲ* 'chin' ~ TinMalA *wɔɲ*, TinPrai *waɲ* 'id.'

Mla *loːn* 'personal spirit' ~ TinPrai *silon* 'spirit; soul'

Mlabri *ʰr̩liːŋ* 'forget' ~ TinMalA *ʰliŋ com* 'id.'

Length in Tin only, 2 words:

MlaB *grɤŋ* 'sifter' ~ TinPrai *ⁿgraːŋ* 'sift'

Mla *ʰnʌm* 'year' ~ TinMalB *naːm* 'year'

Length in both Mlabri and Tin, 7 words:

MlaB *bungbɔːŋ* 'swell up; be swollen' ~ TinPrai *pɔːŋ* 'id.'

MlaA *bluːŋ* 'bamboo shoot' ~ TinPrai *poːŋ* 'sprout shoot'

MlaC *brɔːŋ* 'long since dead' ~ TinPrai *prɔːŋ* 'spirit of a deceased person; ghost'

Mla *keːŋ* 'carry in a strap across forehead' ~ TinPrai *kʰeŋ* 'stretch or
   hold up a rope'
Mla *taːɲ* 'weave' ~ TinMalA *tʰaːɲ*, TinPrai *tʰaɲ* 'weave' (cf. Khmu *taːɲ*
   'weave; plait')
Mla *tʰɤːŋ* 'five' ~ TinPrai *tᵉɔːŋ* 'id.'
Mla *dimrɛːŋ* 'bamboo flooring' ~ TinPrai *ᵐbrɛːŋ* 'id.'

Long vowel in Mlabri, diphthong in Tin, 4 words (disregarding Prai
monophthongization to short vowel in certain environments):
Mla *geːŋ* 'house' ~ TinMal *kiʌŋ*
Mla *tʰreːŋ* 'tooth' ~ TinMalA *tᵉiʌŋ*
Mla *mɛːm* 'blood' ~ Tin *miʌm* 'id.'
MlaBC *ɟiˀɛːŋ*, MlaA *ɟəˀɛːŋ* 'bone' ~ TinMalA *ˀiʌŋ*, Tin Pr *siˀiʌŋ* 'id.'

It will be apparent from this data set that with one exception, all corre-
spondence types are frequent. The exception is short vowel in Mlabri
versus long vowel in Tin. The overall occurrence of this correspondence
type is so limited that I assume that the words in question do not repre-
sent a regularity but owe either their shortness in Mlabri or their length
in Tin to sporadic lengthening for some (unknown) reason. The only
structural feature shared by the examples of this correspondence type in
my total data set is that the vowels are in all cases non-high, but then
there are numerous words with non-high vowels that do *not* exhibit such
a length alternation. Leaving that set, as well as the last correspondence
set involving a Tin diphthong, aside we have a residue of three seem-
ingly regular correspondences:

(i)   Short vowel in both Mlabri and Tin
(ii) Long vowel in both Mlabri and Tin
(iii) Long vowel in Mlabri, short vowel in Tin

It would seem safe to reconstruct Proto-Tinic with short vowels in type
(i) and long vowels in type (ii), assuming that type (iii) is due to a regu-
larity found in specific environments, be it a rule of vowel shortening in
Tin or a rule of vowel lengthening in Mlabri. It is, however, thought-
provoking that type (iii), with its alternating length, is slightly more fre-
quent in the data set above than type (ii). That is not a promising finding
if one is searching for specific environments. And in fact, looking at the
data above I cannot see how one could set up a reasonable rule of suffi-
cient generality by making reference to such features as the quality of

the onset, the quality (degree of aperture) of the vowel, or the point of articulation of the final.

My tentative conclusion is that there were two degrees of distinctive length in Proto-Tinic but that this does not account exhaustively for the behaviour of vowel quantity in Tinic. I do not want to engage in specu-lations beyond that conclusion, e.g. about former diphthongs that have vanished without other traces than a length alternation, or about a for-merly distinctive degree of half-length.

Vowel quality

Schematically, the vowel systems of Mlabri and Tin (disregarding length) can be confronted as in Table 3-6.

Table 3-6. Vowel charts of Mlabri and Tin

	Mlabri				Tin	
i	*ɯ*	*u*		*i*	*ɯ*	*u*
e	*ɤ*	*o*				
				e	*ə*	*o*
ɛ	*ʌ*	*ɔ*				
				ɛ	*a*	*ɔ*
	a					

The sounds [ɛ ɔ] in Tin are very open, approaching the openness of [a] (one can compare with Thai). The phonetic vowel [ʌ] occurs as the second component of diphthongs; Filbeck (1978) writes it as [-a]. Occasionally, I hear phonetic [ʌ] as a single vowel in Tin but I cannot establish a phonemic status for it.

Mlabri, in contrast, has an auditorily clear (more equidistant) separa-tion into four degrees of aperture, which are all mutually contrastive in the unrounded back series. The system is characterized phonetically by very high (narrow) [e ɤ o]-qualities unlike Tin (which again is more similar to Thai phonetically). As for the third step in Mlabri, *ɛ ʌ* are phonetically not very much more open than *e ə* in Tin, whereas Mlabri *ɔ* varies in openness, being sometimes a very open [ɒ] in affective MlaA speech (also in cases where *ɔ* has replaced *ʌ* in modern usage, e.g. *ʔoh*

ʔa wʌl 'I go home now!' > *ʔoh ʔa wɔl* [*wɒːl*]) . Altogether, the pattern of apertures in Mlabri seems to me more reminiscent of Burmese than of the closest neighbouring M-K or Tai languages.

This phonetic misfit between the degrees of aperture in Modern Mlabri and Modern Tin is significant because it might help to explain why there is a variation in terms of aperture in the etymological correspondences between Mlabri and Tin. By and large, the vowels in good cognate pairs differ at most by one degree of openness (e.g. *ɔ ~ o*) so the variation could well be due to vacillation in the auditory vowel identification.

As for the vowel system of Proto-Tinic I presented what I consider a plausible reconstruction already in Part 2, but it was motivated by comparative typology rather than evidence from cognate sets. In the data set above the multiple vowel correspondences between Mlabri and Tin were just defined as variables, and that applies to all vowel symbols used in the reconstructed forms, although the use of lower-case versus upper-case letters were meant to suggest a difference between straightforward and less straightforward vowel correspondences.

It seems useful at this point to present a null-hypothesis about the way in which Proto-Tinic vowels might be integrated into the vowel pattern of Early Mlabri. This is probably too simplistic, however. Instead, two competing hypotheses can be presented, both based on the assumption that the vowel systems at the time of the Tinic connection were not very different from those of the present:

(i) The vowel systems of Proto-Tinic and of Early Mlabri were practical identical from a topological perspective, i.e. the structurally corresponding entities in the two systems could be easily mapped onto each other. Therefore, when words from one language were integrated into the other, vowels were substituted according to their corresponding place in the system: /i/ by /i/, /e/ by /e/, etc.

(ii) There was a phonetic skewness between the vowel systems already at the time of the Tinic connection. Therefore, when words from one language were integrated into the other, the preferred choice was auditorily similar vowels even though they did not necessarily occupy the same structural positions in the vowel system as in the lending language. That might lead to such substitutions as [i] > [e], [e] > [ɛ].

Until these questions have been answered the whimsical vowel corre-
spondences between Mlabri and Tin spell trouble for the Tinic hypoth-
esis. In the following I shall look into the nature of the correspondences
themselves.

The simplest strategy is to take the correspondences pairwise, as
when an entity "x" in one language corresponds to two entities "y" and
"z" in the other, and look for possible regularities that could be formu-
lated as sound-laws. That might make it possible to reduce the number
of distinct units to be reconstructed for the common proto-language.
Since we are faced with a wickerwork of correspondences, however,
one must also take a unitary view of larger sections of the whole vowel
space. Unfortunately, my comparative data set is so small compared to
the number of different vowel correspondences that it is often impos-
sible to make safe generalizations.

For a start I shall consider the vowels *ɯ* and *ɤ*. There are five corre-
spondence sets; I disregard length here.

Set 1: Mla ɯ ~ Tin ɯ
Mla *guh* 'be ablaze'; TinMalA *kuh* 'id.'
MlaB *ruhruh* 'clean the field; remove weeds'; TinMalA *ruh* 'throw
   rubbish out'

Set 2: Mla ɯ ~ Tin ɯ/ɤ
Mla *kur* 'thunder'; most Tin *kʰur* or *kʰəːr* 'id.'

Set 3: Mla ɯ ~ Tin ɤ
MlaBC *tup* 'cover, v'; Tin *tʰəp* 'bury'
Mla *bul* 'die; dead'; Tin *pəl* 'id.'
Mla *blut* 'extinguished'; TinPrai *ᵐbət* 'id.'
Mla *jun* 'candle'; TinMal *cən* 'bright'
Mla *gur* 'shaft'; TinPrai *kəl* 'handle, n'
Mla *gum* 'suck on something'; TinPrai *kəm* 'id.'
MlaBC *ʔum* 'bathe'; Tin (MalA and Prai) *ʔəm* 'id.'

Set 4: Mla ɤ ~ Tin ɯ
Mla *glɤːʔ* 'head'; Tin *k(l)uʔ* 'id.'

Set 5: Mla ɤ ~ Tin ɤ
Mla *tɤl* 'cultivate; plant'; Tin (Mal and Prai) *tʰəl* 'id.'
Mla *jɤːŋ* 'foot'; TinMal *cəŋ* 'id.'

What one can get out of that is that there were presumably two Tinic vowels *u and *$ɤ$ involved. The data suggest that after Tinic time there happened a vowel lowering in Tin: *u > *$ɤ$ before an oral final (though not consistently), and conversely a vowel raising *$ɤ$ > *u before a laryngeal final (one single example). That would leave the two vowel qualities in complementary distibution in Tin, but that is not what one finds in Modern Tin, perhaps because of the introduction of new words. Going beyond the lexical material shared with Mlabri it is possible to find pairs such as TinMal *tup* 'beat, v' vs. *tɤp* 'mousetrap' with a minimal contrast.

We have established that Mlabri preserved the old vocalism whereas Tin underwent conditioned changes. Most of the sets above are very small, however. The lack of sufficient data detracts from the explanatory value of positing sound-laws since we cannot elimit variation in the Tin reflexes anyhow.

After having looked at the *u – *$ɤ$ pattern it is natural to throw a glance at the *i – *e pattern as well (the *u – *o pattern appears in a larger framework later). In this case there is a third degree of aperture that is well attested in both Mlabri and Tin: ε (long and short) so I include words that have ε either in Mlabri or in Tin, or in both.

Set 1: Mla i ~ Tin i
Mla *ti:ʔ* 'hand'; Tin *tʰi:* 'id.'
Mla *ti:ʔ* 'eight'; Tin *tʰiʔ* 'id.'
Mla *bi:ʔ* 'full'; satisfied'; Tin *piʔ* 'full'
Mla *ʰrli:ŋ* 'forget'; TinMalA *ʰliŋ com* 'forget'

Set 2: Mla e ~ Tin i
Mla *gehgeh* 'scratch'; Tin *kih* 'id.'
Mla *ɟeŋ* 'sew'; TinPrai *ciŋ* 'id.'

Set 3: Mla e ~ Tin e
Mla *prem* 'old; former; long ago'; TinMalB *pʰjeːm* 'id.'
MlaAB *preʔ* in *preʔ gem* 'chili'; Tin *pʰeʔ* 'spicy'
Mla *ɟen* 'be skilled'; TinPrai *ⁿɟen* 'teach'
MlaB *keh* 'greet; receive cordially'; TinPrai *kʰeh kʰaw* 'turn one's back on a visitor'

Set 4: Mla e ~ Tin ɛ
Mla *pleːʔ* 'fruit; nut; grain'; Tin *pʰlɛʔ* 'nut-shaped thing; fruit etc.'

Mla *kre:ʔ* 'hit the mark; correct'; TinPrai *kʰrɛʔ* 'correct'

Mla *reh* 'tear across; rip'; TinPrai *rɛh* 'id.'

Set 5: Mla ɛ ~ Tin ɛ

Mla *pɛʔ* 'three'; Tin *pʰɛʔ* 'id.'

MlaB *tɛ:k* 'snail species'; Tin *tʰɛ:k* in terms for various snail species

MlaA *kɛc* 'metal ring on shaft'; TinPrai *kʰrɛc* 'id.'

MlaA *kɛh* 'wild goat'; TinMal *kʰɛh* 'id.'

Mla *ʰmɛʔ* 'new'; TinMalA *ʰmɛʔ*

Mla *mrɛ:ŋ* in *diŋ mrɛ:ŋ*, *dimrɛ:ŋ* 'bamboo flooring'; TinPrai *ᵐbrɛ:ŋ* 'bamboo flooring'

Mla *mlɛ:t* 'pinch; strangulate'; TinPrai *ᵐblɛt* 'press (a button)'

MlaC *çak kṃrɛh* 'old person'; Tin *ᵐbrɛh* 'be old'

Mla *çŋkɛr* 'nail'; TinMalA *kʰɛr* 'id.'

Mla *çṃbɛp* 'lip region'; TinMal *ᵐbɛ:p* 'id.'

Mla *ʔjɛ:ʔ* 'far away'; TinMalA *ʔiʌʔ* (sic!), TinPrai *jɛʔ saʔ* 'do.'

Mla *gɲrɛ:* 'curry'; Tin (Mal and Prai) *ᵑgrɛʔ* 'pestle'

It may be due to the limited size of the data that there is not a single word with short *i* in both Mlabri and Tinic, considering that Mlabri long *i:* is well attested with a high vowel reflex in Tin. Set 2 might be construed to reflect a lowering rule *$*i > e$* in Mlabri but that is not plausible. If one looks beyond the Tinic data there are several Mlabri words that have short *i* in environments analogous to those of Set 2, e.g. *rih* 'here; this way', *bih* 'chop', (MlaB) *bia min* 'maggot', (MlaBC) *gliŋ*, 'clean; smooth', and it is hard to imagine that short [i] , one of the anchor vowels of the whole vowel system, was missing from the vowel inventory of Mlabri at any time. Thus one can assume *a priori* that if some words with short *e* in Mlabri reflect Tinic short *$*i$* that is most likely to be due to auditory identification because of phonetic skewness between the vowel systems of Early Mlabri and Proto-Tinic.

The sets which invite close consideration are 2 and 4, i.e. the ones that are defined by an offset in quality between Mlabri and Tin.

Sets 2, 3, and 4 at first glance suggest that Mlabri had a coalescence of three different Tinic vowels: *i*, *e*, and *ɛ* before laryngeal finals, but that analysis is contradicted by sets 1 and 5 and is derived of plausibility the moment one adduces external evidence.

The material in Set 2 is very meagre, the word for 'sew' being the only promising one from a comparative point of view. The question is whether it exists in Khmuic outside Tin-Mlabri at all; Phong, Khmu,

and Ksingmul have various unrelated words with this lexical meaning. Interestingly, however, Diffloth (1980) reconstructs Proto-Waic *ɟeŋ corresponding exactly to Mlabri ɟeŋ 'to sew'. That suggests (without, of course, proving it) that this word existed in Khmuic in a form with *e although it has been replaced by other words in most languages, and that the attested Tin form is due to vowel raising *e > i. That leaves the provenance of the form in Mlabri as an open question. It might be due to auditory identification of Tinic short /i/ as similar to Early Mlabri /e/; this is likely since Mlabri /e/ of today has a particularly narrow vowel quality before nasals (so narrow, in fact, that I used to mix Mlabri /i/ and /e/ up in this position). The other possibility is that it is an inheritance from Early Mlabri so that its connection with Tinic is indirect; I would consider that much less likely since this word is so specific to Tin and Mlabri.

Set 4, with three cognate pairs, is also illuminated by external evidence. The exemplar is the word for 'fruit': *ple:ʔ ~ pʰlɛʔ*, which is widespread in Northern Mon-Khmer. The vowel ɛ in Tin is corroborated by Phong: *plɛ*, suggesting that if there was a Phongic sub-branch of Khmuic (consisting of Tinic and Phong) it had *ɛ in this word. If one looks at other Khmuic evidence, however, Khmu has *pleʔ* and Ksingmul has *ple:* strongly suggesting that the higher vowel in Mlabri is the old vocalism. This is further corroborated by Palaungic: Mitani (1979: 143) reconstructs *ple* for Proto-Palaung-Riang, and Diffloth (1980) reconstructs *pliʔ* for Proto-Waic. Thus Phongic must have had an early lowering *e > *ɛ, whatever the conditions for that sound-shift. By being Phongic, it necessarily predates Proto-Tinic and hence the influx of Tinic into Mlabri. The inevitable conclusion is that the Mlabri word is unlikely to come from Tinic at all; it may be an inheritance from Early Mlabri or (less likely) a loanword from Khmu.

As for the word 'correct': *kre:ʔ ~ kʰrɛʔ* it occurs in Katuic with a diphthong [aj] and no final glottal stop. The same type of variation occurs within Mlabri in the word for 'head lice': MlaAB *ɕe:ʔ*, MlaC *ɕɛj*, so it is no criterion for casting doubt on the vowel quality of *kre:ʔ*.

I conclude that in the comparative framework of Proto-Tinic the *regular* correspondences show no variation over front vowels in Mlabri and Tin; these invariant vowels can thus be posited for Proto-Tinic:

PTinic *i > Mlabri *i* ~ Tin *i* (only attested with long vowels)
PTinic *e > Mlabri *e* ~ Tin *e* (attested with short and variable-length vowels)

PTinic *ɛ > Mlabri ɛ ~ Tin ɛ (attested with long, short, and variable-
   length vowels)

Another example mentioned in Part I involved Mlabri ɛ versus Tin *a*.
This time I widen the scope of comparison by including all Tin vowels
corresponding to Mlabri ɛ, and all Mlabri vowels corresponding to Tin
a but I consider only words that have a short vowel in both Mlabri and
Tin. There are five correspondence sets.

Set 1: Mla ɛ ~ Tin ɛ
Mla *pɛʔ* 'three'; Tin *pʰɛʔ* 'id.'
MlaA *kɛc* 'metal ring on shaft'; TinPrai *kʰrɛc* 'id.'
MlaA *kɛh* 'wild goat'; TinMal *kʰɛh* 'id.'
Mla *ʰmɛʔ* 'new'; TinMalA *ʰmɛʔ* 'id.'
Mla *çŋkɛr* 'nail'; TinMalA *kʰɛr* 'id.'
MlaC *çak kṃrɛh* 'old person'; Tin *ᵐbrɛh* 'be old'

Set 2: Mla ɛ ~ Tin a
Mla *mɛh* 'you, sg.'; Tin *mah* 'id.'

Set 3: Mla a ~ Tin a
Mla *kəplah* 'classifier for sheet-like objects'; Tin *pʰlah* 'classifier for
   big pieces of cloth'
MlaA *tar* 'string'; TinMal *tʰar* 'id.'
Mla *ʔa bah* 'it is early dawn'; Tin *pah saʔ* 'id.'
MlaAB *blah* 'run off in different directions'; TinMal *plah* 'escape; dis-
   appear'
Mla *ɟak* 'go'; Tin *cak* 'id.'
Mla *gajʰ* 'nine'; TinPrai *ᵑgat* 'id.'
Mla *mat* 'eye'; Tin *mat* 'do.'
MlaB *ŋaʔ* 'itch; it itches'; TinMalA *ŋaʔ* 'id.'
MlaA *-salʰ* in *bṛsalʰ* 'itching'; TinMalA *tᶜajʰ* 'itching'
Mla *lak* 'gourd; tuber'; Tin *lak* 'do.'
Mla *lṃbah* 'cabbage, etc.'; TinPrai *ᵐbah* 'component of terms for veg-
   etable species'
Mla *takat* 'be feverish; feel cold'; TinPrai *kʰat* 'fever'

Set 4: Mla a/ʌ ~ Tin a
MlaC *pṛmpam, rṃpʌm* 'butterfly'; TinMalA *ᵐpʰojʰ ᵐpʰam* 'id.'

Set 5: Mla ʌ ~ Tin a

Mla *pʌr* 'fly'; TinMalA *pʰar* 'id.'

Mla *kwʌŋ* 'something round, e.g. seeds of a fruit'; TinPrai *kʰwaŋ* 'seed (in fruit)'; also: MlaAB *k̩lwʌŋ* 'ball-shaped object'; TinMalA *kʰlwaŋ* 'seeds in fruit'

Mla *bʌt* 'dip water; scoop up water'; TinPrai *pat* 'id.'

Mla *wʌl* 'return home'; Tin *wal* 'return'

MlaB *dɲrʌŋ*, MlaA *ɟɲrʌŋ* 'housepole'; Tin *ᵑgraŋ* '(house)pole'

The well-attested correspondences are Mla *ɛ* ~ Tin *ɛ*, Mla *a* ~ Tin *a*, and Mla *ʌ* ~ Tin *a*. Proto-Tinic might have had all the three vowels **ɛ*, **a*, **ʌ* since they occur in some other Khmuic, but if so there is no trace of the distinction in any Tin. Since there are specific historical problems associated with Mlabri *ʌ* I shall take that up in a subsequent section.

As for other Mlabri counterparts of Tin *a* than *ʌ* there is just one example of Mlabri *ɛ* corresponding to *a* in Tin, namely the 2nd person singular pronoun. As mentioned in Part 1, external evidence suggests that Mlabri *ɛ* rather than Tin *a* is the old vowel, but no factor conditioning a vowel shift **ɛ > a* can be observed since an invariant Tin reflex *ɛ* is attested in other etyma, after the same category of initial and before the same final as in *mɛh* ~ *mah*. Once more, there is evidence that Mlabri faithfully reflects the state of affairs at the Tinic level whereas there have been vowel shifts in Tin.

Taking now a much more complex bulk of material I shall display the totality of correspondence sets involving *rounded back vowels* in Mlabri and Tin, this time sticking to a Tinic format of analysis. I shall meticulous exploit all possibilities of explanations in terms of context-dependent sound changes, though still with the whole exercise hanging in open air because the starting point, Tinic, is so far undefined in terms of its inventory of proto-vowels. Vowel-length being one of the factors that one might expect to be crucial to vowel qualities I subdivide each correspondence set according to the presence of distinctive length in one or the other language or in both.

Set 1: Mla u ~ Tin u

1a:

Mla *ɟruʔ* 'deep'; TinMalA *kruʔ*, TinMalB *kjuʔ*, TinMalC *kuʔ*, TinPrai & TinCuun *cuʔ* 'deep'; NB: Khmu *ɟruʔ*; Khabit *tɕruʔ*(!)

Mla *gul* 'seven'; TinPrai (conservative lg. in Laos) *ᵑgul* 'seven' (cf. Khmu ceremonial language: *guːl* 'seven')

Mla *puŋ* 'blow (produce an air stream with the lips); TinMalA *pʰuŋ* 'id.'

MlaAB *paluh* 'scold; abuse'; TinPrai *luh* 'offended; angry' (NB if this were from Tai one would expect -*ʰluː* in Mlabri)

1b:

Mla *çoʔuːm*, *çuʔuːm* '(A) have a bad smell; (BC) be fragrant'; TinPrai *sijum* 'smell good' (*2nd syllable*)

MlaBC *tʰuːl* 'meat and/or vegetables eaten with staple food'; TinMal *tᶜul*, *sul* 'food'

Mla *tuːn* 'mole'; TinPrai *tʰun* 'big rodent species' (affinity to Tai but obviously old in Tinic)

Mla *muːk* 'smell (with the nose)'; Tin *muk* 'id.'

Mla *ʰmuː* 'areca nut'; TinMal *ʰmuʔ* 'id.'

1c:

Mla *bluːʔ* 'thigh; upper part of leg'; TinMal(A) *pluː* 'leg' (NB cf. K: *bluʔ* 'thigh')

Mla *juːr* 'descend'; Tin MalA *cuː*, other Tin *cuːr*, *cuːl* or *cuː* 'descend' (cf. Khmu *juːr* 'descend')

Mla *duːc* 'tired'; TinPrai *ⁿduː* 'drunk'

Mla *tʰapuːl* 'stomach'; Tin variably *ᵐpʰul*, *pʰuːl*, *pʰul* 'id.'

Set 2: Mla u ~ Tin o

Mla *druːɲ* 'termite'; TinMla *kjoːn* 'id.'

MlaA *bluːŋ* 'bamboo shoot'; TinPrai *poːŋ* 'sprout shoot'

Mla *ʔuːlʰ* 'fireplace; firewood'; TinMal *ʔoːjʰ* 'id.'

Set 3: Mla o ~ Tin u

MlaAB *dok* 'poor'; TinPrai *tuk* 'poor'

MlaA *pol* 'crossbow'; TinPrai *pʰul* 'crossbow'

Set 4: Mla o ~ Tin o

4a:

Mla *broʔ* 'wound'; TinPrai *proʔ* 'skin lesion or infection'

Mla *poh* 'split'; Tin (Mal and Prai) *pʰoh* 'split'

MlaB *koh* 'cut apart'; Tin (Mal and Prai) *kʰoh* 'chop up' (Khmu has *koh* 'chop')

MlaB *bohboh* 'be boiling'; TinMalA *ᵐboh* 'be boiling'; TinPrai *ᵐboh* 'boil something'

Mla *pagoh* 'cause something to break'; Tin *ᵑgoh* 'strike something; knock on something'

Mla *loh* 'search for; hunt for'; TinPrai *loʔ* 'id.'

4b:

Mla *boːʔ* 'breast; breast feed'; Tin *poʔ* 'breast; breast feed' (Khabit and Khmu have *buʔ*)

MlaB *boːm* 'bottle'; TinPrai *pom* in *pom tʰeː* 'pottery jar'

Mla *groːʔ* 'make a sound; one hears a sound'; TinMalA *kroʔ* 'for an animal or bird to make its characteristic sound'; TinPrai *kroʔ* 'cry out; make a sound'

Mla *poːn* 'five'; Tin *pʰon* 'five' (Khmu has *puən*, *puːn* depending on dialect)

Mla *loːn* 'personal spirit'; TinPrai *silon* 'spirit; soul'

4c:

MlaAB *kot* 'encircle; bend around; hug', also as component of a redu-plicated form: MlaA *kotkwat* 'hug and carry in one's arms' ; TinPrai *kʰoːt* 'curl up (in resting position)'; NB affinity to Tai

4d:

MlaAB *broːc* 'spit liquid out of the mouth'; TinMalA *proːc* 'id.'

MlaAB *ɟoːk* 'suck'; TinMalA & TinPrai *coːk* 'suck (through a straw)'

MlaC *ʰn̩toːr*, MlaB *ʰn̩tor* 'orifice; 'duct'; TinPrai *ⁿtʰoːr* 'ear'

Set 5: Mla o ~ Tin ɔ
5a:

Mla *boŋ* 'eat' (AC: 'eat meat'); Tin *pɔŋ* 'eat'

Mla *toc* 'grasp; (B: 'take'); TinPrai *tʰɔc* 'grasp'

MlaB *toh* 'pull out (with a jerk)'; TinMalA *tʰɔh* 'pull out'

5b:

MlaC *koːlʰ* (idiolectally *koːjʰ*) 'wooden shaft of spear'; TinMal *kʰɔjʰ*, TinPrai *kʰɔt* 'spear'

5c:

Mla *poːlʰ*; TinMalA & TinCuun *pʰɔːjʰ*, TinPrai *pʰɔːt* 'barking-deer' (Khmu has *puas*)

Set 6: Mla ɔ ~ Tin o

6a:

MlaB *bɔ-* (presumably < *bɔː*) in *bɔbɛːp* 'whisper'; TinMalA *poː* 'whisper'

Mla *mɔh* 'nose'; TinMal *moh* 'id.' (Khmu has *muh* 'nose'; TinPrai *muh* may be a borrowing from Khmu)

6b:

Mla *mɔːj* 'one'; Tin *moːj* 'one' (Khmu sides with Tin)

Set 7: Mla ɔ ~ Tin ɔ

7a:

MlaB *ɟɔʔ* 'a few; a small quantity'; TinMalA *cɔʔ* 'short (of physical size)'

Mla *pleʔ klɔk* 'forest fruit species'; TinPrai *pʰɛʔ kʰɔk* 'id.'

MlaB *klɔt* 'knot; tie a knot'; TinPrai *silɔt* 'knot'

MlaB *kukɔʔ* 'neck'; TinMalA *ᵑkʰɔʔ* (TinPrai *ᵑkʰɔː*) 'id.'

7b:

MlaA *rɔc* 'rinse the interior of a slaughtered animal'; TinPrai *rɔːc*, Tin-MalA *rwac* 'id.'

MlaB *kukɔʔ* 'neck'; TinMalA *ᵑkʰɔʔ* (TinPrai *ᵑkʰɔː*) 'id.' (NB: connection to Tai spurious since one would then expect **g-* not **k-*)

7c:

MlaB *-bɔːŋ* in *bungbɔːŋ* 'swell up; be swollen'; TinPrai *pɔːŋ* 'swell up; be swollen'

MlaC *brɔːŋ* 'long since dead'; TinPrai *prɔːŋ* 'spirit of a deceased person; ghost'

Mla *krɔː* (MlaA) 'hope for', (MlaBC) 'ask for'; TinPrai *kʰrɔː* 'ask for' (NB: connection to Tai but the word is obviously old in Tinic; Khmu has *krɔː*)

Mla *pleʔ kɔːk* 'the thyroid cartilage; Adam's apple'; TinMalA *pʰlɛʔ kʰɔːk* 'id.' (NB: this is a Tai word but must have been borrowed very early into Tinic)

MlaA *ɲɔːk* 'plaited bag (like a net)'; Tin *ɲɔːk* 'trap (for fishing)'

This is the entire material. The question, then, is whether any clear-cut generalizations about context-dependency or about unconditional coalescences can be made.

The possible role of vowel-length. Most of the sets occur with both long and short vowels. The interesting exceptions are sets 2, Mla *u* ~Tin *o*, and 3, Mla *o* ~Tin *u*. Words in Set 2 have long vowels, whereas words in Set 3 have short vowels. This complementarity may, however, be accidental, i.e. due to the small size of the sample. In any case, one cannot get very far. To explain the pattern by rule one has to posit an additional and entirely hypothetical vowel (or diphthong) and claim that it has one or the other set of reflexes in the modern languages depending on whether it was long or short in Proto-Tinic. That hypothesis is unattractive since the vowel inventory becomes overloaded. If instead we stick to the vowel inventories defined by the modern languages, the vowels in Set 2 and Set 3 must somehow go back to either **u* or **o* (long versus short). Neither of those options works, however, since there is ample attestation of Mla *u* ~ Tin *u* and Mla *o* ~ Tin *o* with both long and short vowels (I would not hesitate to project vowel-length or shortness back to Proto-Tinic in the cases if Mlabri and Tin agree). Thus one would have to claim other hypothetical phonetic attributes for the proto-vowels in the alternation sets, e.g. that those in Set 2 were diphthongs and those in Set 3 lax vowels at the Tinic level. To conclude: no matter what vowel-lengths one posits for the words in Proto-Tinic, that does not in itself help to explain the erratic qualitative correspondences between Mlabri and Tin vowels.

The possible role of the initial. The following sets have so variegated initials that there is no basis for making interesting generalizations; 1, 3, 4, 5, and 7 (in 3 and 5 the initials happen to be all stops but that cannot be relevant). The remaining, potentially interesting, sets are 2 and 6. In Set 2: Mla *u* ~Tin *o* the initial is either a voiced stop or a glottal stop (the glottal stop preceding a vowel when there is not other onset); there are, however, also voiced initial stops both in Set 1: Mla *u* ~Tin *u* and in Set 3: Mla *o* ~Tin *u*, so that property of the initial does not define a special context. Switching now to Set 6: Mla *ɔ* ~Tin *o* we see that it stands out by having voiced labial stop or nasal initially in all the examples; there are, however, also voiced labial stops in Set 5: Mla *o* ~Tin *o* and in Set 7: Mla *ɔ* ~Tin *ɔ*. Again, the observed property of the initial does not define a special context. To conclude: no generalizations made on the basis of word onsets can explain aberrations in the vowel correspondences.

The possible role of the final. Sets 1, 3, 4, 5, and 7 all exhibit such a variety of types of finals (shared by the Mlabri and Tin cognates) that there is no possibility that the vowel correspondences in these sets are

conditioned by properties of the final. The remaining two sets: 2, Mla *u* ~Tin *o*, and 6, Mla *ɔ* ~Tin *o*, are small but potentially interesting in that they do not exhibit any final stops. Sets 1 (Mla *u* ~Tin *u*), 4 (Mla *o* ~Tin *o*), and 7 (Mla *ɔ* ~Tin *ɔ*), however, all exhibit examples with non-checked syllables as well, so one would have to arbitrarily postulate some extra vowels or diphthongs at the Tinic level anyway.

The possible combined role of onset and final. Two small sets, 2 (Mla *u* ~Tin *o*) and 6 (Mla *ɔ* ~Tin *o*), attract attention both when seen in relation to the onset and when seen in relation to the final. If we combine these two environments, there emerges a generalization to the effect that the two vowel sets occur in syllables that have a voiced onset and a non-checked final. Since they share the reflex *o* in Tin one might speculate that Tin has had a coalescence of Tinic **u* and **ɔ* with Tinic **o* in that combined environment. Off-hand (without the availability of further relevant information), one would posit Proto-Tinic **u* , **o* and **ɔ* as being reflected by the invariant sets Mla *u* ~Tin *u* (set 1), Mla *o* ~Tin *o* (set 4), and Mla *ɔ* ~Tin *ɔ* (set 7). In order to claim that **u* and **ɔ* changed into Tin **o* in non-checked syllables with a voiced onset one would have to ascertain that the two Proto-Tinic vowels did not remain unchanged in other instances in that type of combined environment. That is not the case, but it is conspicuous that there are few examples of invariant *u* and *ɔ* in the just-mentioned type of environment (they occur in subsets 1a, 1c, and 7c above). So far it seems just possible that we have at least a diachronic tendency here.

Tin, however, has thousands of words stemming from Proto-Tinic (though in most cases one cannot prove its provenance by comparing with Mlabri simply because the Tinic component in Mlabri is so limited). In order to make claims about restrictions on the occurrence of *u* and *ɔ* in non-checked syllables with a voiced onset one has to look at Tin vocabulary at large. A cursory glance at Prai reveals such examples as *tul* 'be fat', *pul mat* 'for eyes to reflect light' on the one side and *tɔŋ* 'egg', *kɔːm* 'termite' on the other. It should be remembered that plain voiceless initial stops in Tin are regular continuations of voiced stops so these words should have *ɔ* if the tentative sound-laws were valid. Consequently, the attempt to establish a pattern of regularities must be discarded.

Discussion: the predictability of vowel correspondences between Mlabri and Tin. As illustrated by the exercise above, the data sets with rounded vowels do not lend themselves readily to the establishment of

precise sound-laws operating between Mlabri and Tin. Some of the correspondence sets are very small. A possible explanation might be that only the large sets reflect the Tinic connection whereas the small sets reflect independent preservation of the etyma in question in both Mlabri and Tin. Another explanation might be that the small sets reflect irregularities in the establishment of a Mlabri norm, with competition between variant forms due to interference between the Tinic vowel pattern and the old Mlabri vowel pattern. Neither of these explanations, however, takes care of the apparent unpredictability of the difference between the fairly large sets 4 (Mla *o* ~ Tin *o*) and 5 (Mla *o* ~ Tin *ɔ*), both of which should be of Tinic provenance if the concept of Tinic has any sense at all.

Still, if one throws a side glance at the conclusions made for unrounded vowels above it is interesting that Mlabri *o* has different, well-attested correspondences just as was the case with Mlabri *e* of the same degree of aperture. Again, the smoothest explanation is that Tin for some reason vacillated in its transmission of high mid vowels. The most tempting solution, then, is that once more the topologically straightforward correspondences are the regular ones, i. e.

Mlabri *u* ~ Tin *u*
Mlabri *o* ~ Tin *o*
Mlabri *ɔ* ~ Tin *ɔ*

I do not want to overload the presentation with a similar exercise comprising correspondences among *all unrounded vowels*. The unrounded vowel inventory is considerably more complex than the rounded one both because it has one degree of openness more than the rounded series and because there are both front and back (or central) vowels which converge in the most open (low) part of the vowel space.

General conclusion about the transmission of vowel qualities from Proto-Tinic into Mlabri

My conclusion at this point is that there is evidence in favour of certain more or less irregular vowel shifts happening in Tin between Proto-Tinic and Proto-Tin, but that the basic correspondences are such that *both* Mlabri and Tin preserve the Proto-Tinic vocalism. Somewhat heavy-handedly, one can then reduce the 30 vowel correspondences

listed in Table 2-6 (in Part 2) to a Proto-Tinic vowel system of a moderate degree of complexity, i.e. some 9-10 vowel phonemes (long and short). That is, the reconstructed system displayed in Table 2-5 (in Part 2) is now to a moderate degree etymologically confirmed by comparison of Modern Mlabri with Modern Tin. It must be emphasized that this is strictly a reconstruction based on internal evidence. It has not been the task to see how the reconstruction fits in if one compares with other Khmuic languages.

The correspondence sets which exhibit a qualitative offset between Tin and Mlabri must also be taken into consideration. Accordingly, it requires three working-hypotheses about the Proto-Tinic point of departure and its reflexes in Mlabri and Tin to account (at least in principle) for the matches and mismatches in vowel aperture:

(i) It is the default case for Mlabri and Tin vowels in cognate pairs to occupy the same slot in the respective vowel systems. We must assume that the Mlabri not only imitated the phonetics of Proto-Tinic but had internalized the sound patterns of both languages and thus could automatically substitute vowels according to structural equivalence when borrowing lexical items from one language into the other. The vowel system of one language was mapped wholesale onto that of the other and whenever necessary, the phonetic vowel qualities of loanwords were adjusted to the phonetics of the borrowing language.

(ii) If Mlabri shows a higher vowel than Tin, it is *a priori* likely that Mlabri, with its altogether conservative phonology, shows Proto-Tinic vocalism. There must then have been a change in later Tin, either vowel lowering by rule (though it seems hard to formulate rules) or irregular substitution of a lower vowel. That must have happened some time between Proto-Tinic and Tin, and the irregularities probably had complex (and unknown) sociolinguistic causes involving competing norms.

(iii) If Mlabri shows a lower vowel than Tin, the explanation may be as outlined in (ii) above, but it is *a priori* also possible that Tin shows Proto-Tinic vocalism since Mlabri vowels are nowadays on the whole slightly higher than Tin vowels. The offset may then have a perceptual explanation: it may be due to occasional auditory identification by the Mlabri of the original vowel with a vowel that occupied a lower slot in their own language (e.g. matching /i/ in the Proto-Tinic system with /e/ in the Early Mlabri vowel system).

These are hypothetical scenarios about ways in which Tinic word-forms could be similar or become different in Mlabri and Tin. In order to have any substance they presuppose a working assumption: the cognate pairs presented in the beginning of Part 4 (or at least the great majority of them) are indeed Tinic. I have deliberately taken a rigid stand up to this point, using that assumption as my point of departure, because the wider search for the origin of Mlabri lexicon presupposes that one can form a realistic idea about the nature of the lexical inheritance from Early Mlabri. The more words one can identify as Tinic or as more recent loanwords from different sources, the safer conclusions can be drawn by looking at the residue from a comparative perspective.

To get a more balanced view, however, one must keep in mind that if Early Mlabri was a Mon-Khmer language or at least had a significant M-K component, some Mlabri-Tin cognate pairs presented as Tinic above may instead exhibit similarity due to distant relatedness. Such a scenario would furnish the basis for competition between Tinic and Early Mlabri forms in the speech of the newly founded ethnic group, and might explain offsets in vowel quality.

Unfortunately, one cannot get much further than just pointing to the explanatory potential of such a scenario as long as Early Mlabri is defined in hypothetical terms and entirely negatively: as the residue that remains after Tinic words and loanwords have been peeled off (and after later innovations in Mlabri have been identified as well). It becomes a very interesting hypothesis, however, the moment one has independent clues to the identity of Early Mlabri.

Excursus: ʌ versus *a* in Mlabri

In Mlabri, the vowel *a* is of course extremely frequent but the vowel ʌ is not at all rare either, although there is now a strong tendency among Mlabri speakers to substitute the rounded vowel ɔ for it in many words. A search for occurrences across the lexicon gave 155 entries with long or mostly short ʌ as a separate vowel, in addition to a modest number of occurrences as the second component of diphthongs (in this monograph I do not go into Mlabri diphthongs at all since their relationship to diphthongs in Tin is unclarified so far; many of them occur in loanwords). The only reasonable assumption is that this distinction goes back to Early Mlabri.

The occurrence of the difference between ʌ and *a* in Mlabri words for

which I have assumed a Tinic provenance is a crux, however. One has to scrutinize external comparative evidence for each single word with ʌ in order to hopefully explain of why, for instance, Mlabri has $t^hʌc$ 'meat; pulp' and *pʌr* 'fly, v' versus *tac* 'broken', *tar* 'string; rope'. I shall comment briefly on the two words that contain ʌ in these sets for exemplification of the situation.

The word for 'meat' occurs widely in Mon-Khmer but with *a* or reflexes of that vowel. I do not know of a Khmu cognate of $t^hʌc$ so one has to go beyond that language. Ksingmul has *sec*, showing that the vowel *a was prone to change in a palatal context in Khmuic. This is what we find in Tin: as said earlier, the Mal form *çəc* and the Prai form *sec* are due to raising and subsequent fronting of *a before palatals (incidentally, a similar change of *a into *ɛ* or even into *e* before palatal finals occurs as far away as in Aslian; Diffloth 1979: 56 gives this rule for Proto-Semai, a Central Aslian language, and Burenhult 2002 has forms such as *sɛc* 'meat' in Jahai, a North Aslian language). Tin seems to exemply a similar process of raising in the first person singular pronoun *ʔən*, presumably from *ʔaɲ. The Mlabri form $t^hʌc$ may simply reflect the first step in this process of raising: *sac > *sʌc (> $t^hʌc$). If so, its phonology is straightforward from a Tinic perspective if we assume that the raising to *ʌ had occurred already in Proto-Tinic. That leaves Mlabri *tac* 'broken' unexplained, however. It should also have had vowel raising to ʌ, cf. that there is the expected vowel shift in Prai t^hec 'broken'. An alternative possibility is that the process of raising observed in Tin had not started at Proto-Tinic time but was later, and that Mlabri $t^hʌc$ is a continuation of an Early Mlabri form rather than belonging to the Tinic component in Mlabri. If so, it is the origin of that Mlabri vowel in a word expected to have *a that is left unexplained.

As for the verb *pʌr* 'fly' I do not know of a cognate in Khmu but Phong has *par* and Ksingmul has *pal*, with the same vowel as Tin. Switching to Palaungic, however, Diffloth (1980) reconstructs *pər for Proto-Waic, and Mitani (1979: 147) reconstructs *pžr for Proto-Palaung. If, finally, one looks at Katuic, Brôu has *pʌr* (whereas Kui has *paːr*). Thus the Palaungic and Katuic evidence suggests that Mlabri *pʌr* is not just an aberration but has an old history in the language. It is hardly a Khmuic form but belongs to Early Mlabri. Thus it should be weeded out of the assumed Tinic component in Mlabri.

Assuming that Early Mlabri had a distinction between *ʌ and *a whereas Proto-Tinic just had *a (the process of raising and fronting before palatals having not yet taken place), an assignment of Proto-

Tinic words with this vowel to two different phonological categories in Mlabri may have happened when Mlabri speakers internalized Tinic vocabulary. Proto-Tinic *a was probably an [a] intermediate between Early Mlabri *ʌ and *a and perhaps closer to the former (cf. the vowel charts presented for Modern Mlabri and Modern Tin above). Thus, the alternative strategies presented above as hypothesis (i), i.e. system-oriented identification, and (ii), i.e. identification based on auditory similarity, might compete and cause inconsistency in the way in which Mlabri speakers assimilated Proto-Tinic words in *a, some words emerging with [ʌ], others with [ɑ].

It goes without saying that this attempt at an explanation is invalidated if external evidence throws doubt on the distinction between ʌ and a in Mlabri words that are old and *not* of Tinic provenance. In any case, we are left with another crux if early Mlabri did have such a distinction throughout its vocabulary, as I think it must have had. Where did the many occurrences of ʌ come from? The case study of *pʌr* above suggests a connection between Mlabri ʌ and Palaungic *ə in some words expected to have M-K *a. On the other hand, the overall similarity of the negative auxiliary ("negative imperative") *gʌm* with Khmer *kom* suggests that Mlabri ʌ at least in part originated as a rounded back vowel. Khmu has instances of a long vowel ʌː (different from both *aː* and *ɔː*), and its short vowel ɔ is of rather [ʌ]-like phonetic quality (Svantesson 1983: 14); I have not attempted to establish a correlation between the occurrence of these vowels across the lexicon in Khmu and Mlabri.

The most intriguing word pair with ʌ and a in Mlabri is *kwʌr* 'outsiders (cover term for people who are not Mla Bri)' versus *kwar* 'monkey species with a characteristic cry (apparently: slow loris)'. In the lack of etymologies one is tempted to speculate that these words arose by a differentiation of one word into two.

I have not gone further into these issues; in fact, it is very difficult to do so since the Tinic component in Mlabri has not been identified in its entirety so that it is still unclear what lexical items could be from Early Mlabri. Considering that the origin of Mlabri ʌ across the lexicon is an issue which may be of some general comparative interest, however, I present the complete material of contemporary Mlabri words in ʌ in Appendix I at the end of this monograph.

PART 4:

MLABRI BEFORE TINIC

The amount of Tinic in Mlabri

One of the main problems with my comparative Tinic data set is its modest size. I have so far managed to spot only some two hundred good Mlabri-Tin cognates that can qualify as Tinic. There are, however, thousands of words in my Mlabri files and new words have been emerging until I had to discontinue my fieldwork for health reasons. As for Tin, I had started an extensive campaign of lexical retrieval focussing on the most interesting dialect of Tin, MalA, but it was only in its infancy when I had to stop, and in fact I have more comparative data for Prai than for Mal. For most items in the Mlabri lexicon I never came around to asking about look-alikes in Tin, or I did not manage to retrieve them in any kind of Tin so far. There is every reason to assume that there may be more words from that source, be it from Mal or from Prai. Below, I refer to the unknown totality of Tinic words in Mlabri, as the *"real"* size of the Tinic component.

One way to attempt to assess the real size of the Tinic component in Mlabri is to look at a change happening within Tinic and see how strongly it is reflected in the total Mlabri vocabulary. We have such a sound change with the diphthongization $*a: > *ia$, which was explained in Part 2 as being due to "register" (possibly assisted by an influence from word-final nasals) and thus sensitive to voicing in the onsets.

Assuming that Khmuic $*a: >$ Tinic $*ia >$ Mlabri $\varepsilon:$ was conditioned by register one would expect the ratio of $\varepsilon:$ to $a:$ to be larger in words with voiced onsets than in words with voiceless onsets. As stated in Part 3, there are quite a few occurrences of $\varepsilon:$ in Mlabri which have a different source than Tinic $ia <$ Khmuic $*a:$ and which make the total number of words in $\varepsilon:$ disproportionately large in comparison with the total number of words in $a:$ in Mlabri. They ought, however, to be equally frequent after voiced and voiceless onsets, the only disturbance caused by them being that they raise the number of occurrences of $\varepsilon:$ in both of the two environments and make the difference between voiced and voiceless environments less clear the more numerous these occurrences from other sources are. The difference in proportionality between $\varepsilon:$ and $a:$ would, on the other hand, be the more conspicuous the greater the real proportion of Tinic vocabulary in Mlabri.

As referred to already in Part 3, I have approached this issue counting the number of lexical entries in my Mlabri files that contain either of these two vowels immediately after a voiced or a voiceless stop (leaving out palatal onsets for scarcity of data). Words in single stops are, of course, only a small fraction of the total number of words with these two long vowels in Mlabri but I wanted the difference in onset type to be as clear-cut as possible. The idea was to look non-selectively at the vocabulary, disregarding whether particular lexical items have an identifiable etymological source or not. The results were not very clear, unfortunately:

(i) The words with final nasal in the voiced-voiceless data set are so few that no conclusions can be drawn by looking specifically at them. One of those words is interesting in its own right, however: MlaA *ba:n* 'leaf fat'. If that had entered Mlabri from Proto-Tinic it would violate the register hypothesis; the reason for its failure to diphthongize is that it is not a Tinic word.

(ii) Looking at the whole set I found 12 examples of ε: versus 8 examples of a: in words with voiced single stops, and 20 examples of ε: versus 22 examples of a: in words with voiceless single stops. Although that agrees with the register hypothesis, the difference is modest and in particular, there is a disturbingly great number of occurrences of ε: after voiceless onset (if I weed out obvious loanwords from Tai there are actually *more* words in ε: than in a: after voiceless onset). This means that the words in which Mlabri ε: goes back to Khmuic *a: via Tinic diphthongization can hardly constitute the majority of words in ε:, as I would a priori have expected.

The interesting question is whether it is warranted to say something about the proportion of Tinic words in Mlabri on this basis. The evidence certainly does not suggest that most of the indigenous vocabulary in Mlabri is Tinic. Though inconclusive, it rather points in the opposite direction: *the non-Tinic proportion of Mlabri vocabulary seems to be so considerable that its properties tend to mask the pattern inherent in the specifically Tinic vocabulary.*

Although the real size of the Tinic component in Mlabri is undoubtedly larger than the set of cognates I have established by comparing Mlabri and Tin vocabularies, the conclusion above suggests that there are hardly thousands of Tinic words in Mlabri. The majority of indigenous lexical items are in fact likely to be non-Tinic. The next interesting

question, then, is to what extent that dominant residue is of Mon-Khmer origin, like the Tinic lexicon.

Tinic or non-Tinic?:
some etymological case studies

This section uses a different approach in addressing the question to what extent the Mon-Khmer appearance of Mlabri lexicon is due to the Tinic connection, and to what extent it reflects connections beyond Tinic. The various subsections below take up some Mlabri words which seem to have a strange history, as well as nomenclature within some narrow semantic fields that invite etymological speculations. In my view they illustrate that even in complicated cases one can arrive at suggestive etymologies by searching across the vocabularies of other languages more or less close to Mlabri and Tin.

Many of the suggestions below are very vulnerable to criticism; the last and longest essay (6, about water) shows how one is tempted to go further and further in the search for cognates, only to end up with mutually inconsistent etymologies.

More serious work in this difficult area would require a broader perspective and above all, a better understanding of the history and the Stammbaum of the Khmuic branch of Mon-Khmer.

By way of introduction, the complexity of addressing the origin of Mlabri lexicon can be illustrated by taking two particular words, the words for 'root' and for 'giant gecko' (A1-A2 below) plus three tiny semantic fields comprising a selection of small, useless or repulsive animals: leeches, some big or worm-like insects, and ants (A3-A5 below). Such terms are not likely to be borrowed by a forest-bound tribe during short encounters with villagers, but nevertheless, the vocabulary is tricky when it comes to provenance and reconstructions.

A1. The word for 'root' (of trees of other plants). The expected Tinic word, $*riaS_1$ (cf. Khmu $riəs$), occurs with straightforward phonology as Mla $re{:}l^h$ 'root' but has no direct reflex in Tin that I know of. Tin Prai has a word $ke{:}t$ for 'root' which might be the same etymon as MlaA $g\varepsilon j^h$ 'small twig' (both Tinic $*g{-}$ > Prai $k{-}$ and Tinic $*{-}S_2$ > Prai $-t$ are in accordance with general sound shifts). Though it looks strange off-hand, a semantic offset of this kind is not out of the question if the original denotation of $*g\varepsilon S_2/*ge{:}S_2$ was 'branching' > 'branched root' or 'twig'.

TinMalA, however, has $^{\eta}gri\Lambda j^h$ for 'root' ($^{\eta}gri\Lambda j^h$ $^{\eta}u\Lambda\eta$ 'root of tree'). That must be an indirect reflex of *riaS₁, the prenasalized stop suggesting an old prefix. That lost prefix may in fact have been a separate word, $^{\eta}gri\Lambda j^h$ being the remnants of an old compound denoting a particular type of root. In any case, judging from the history of other words with prenasalized clusters in Tin, $^{\eta}gr$- might reflect a former structure in which a dorsal stop (g or $\jmath$) was preceded or followed by some material before the stem-initial trill, all other material than stop+trill being deleted and compensated for by prenasalization.

There is no way to recover the original structure of the lost material in $^{\eta}gri\Lambda j^h$ since it is absent in Mlabri, so the full etymology of the Mal word remains obscure. Just to illustrate what degree of shrinking one would have to take into consideration in reconstruction, I shall venture two more or less plausible guesses: it may have been a compound that originally denoted the nucleus of a root (as on a tuber): *glĬ₁?-riaS₁, literally 'head of root', or it may have been a compound of the two Tinic words for 'root': *ge:S₂-riaS₁.

A2. The giant gecko.

A2. The giant gecko. The Mlabri term for 'giant gecko' has several alternative forms. Of these, (A) $t\partial g\varepsilon?$, (A) $tak\varepsilon:?$, (B) $t\partial k\varepsilon:?$, and (C) $tik\varepsilon:?$ suggest remote etymological ties to Tai, although one would rather expect *tukke: if the word was a recent borrowing. In MlaBC these Tai-like forms are reported to be female language. The corresponding term used in MlaBC male language is $t\partial k\jmath^h$, $tuk\jmath^h$, which is likely to be a cognate of the TinPrai word for 'small lizard': $^{\eta}k^h\jmath$ in spite of the semantic offset (I cannot decide on the evidence available to me whether the Prai word used to refer to one or another reptile species).

It is no valid objection against my claim about cognacy between the last-mentioned forms that Mlabri $t\partial k$- and Prai $^{\eta}k^h$- aberrate on the surface. Such correspondences are found in other cognate pairs and can be accounted for diachronically. There is, however, a difficulty with the final consonant: Prai would be expected to have -t not -j corresponding to Mlabri -j^h. The final -j in Prai is probably old so that it makes sense to reconstruct a Tinic form *$t\partial k\jmath$. If so, Mlabri may have had a substitution *-j > -j^h motivated by the occurrence of the unusual final -j^h in other words for "strange" animals such as $kro:j^h$ 'crocodile', $g\varepsilon j^h$ 'crab'. Apparently, this final had an expressive status in Mlabri.

A3. Leeches. Unfortunately, I did not retrieve any terms for leeches from MlaB (it is too late now), but in MlaA there are three terms for such creatures and from MlaC I retrieved one term:

MlaA *tʰaːk* 'land leech'
MlaC *plɤːm* 'land leech'
MlaA *glɤːw* 'water leech, big species'
MlaA *pliːn* 'water leech, small species'

Of these four terms, *plɤːm* and *pliːn* have Khmuic parallels. Tin has *pʰlɔm* and TinPrai has *piːn* (I do not have a Mal form); Khmu has *pliəm* and *pliːŋ*. Ksingmul has *ploːm* and *pliːn*, in good accordance with Mlabri.

The MlaC word for 'land leech' thus occurs in Khmuic but also beyond Khmuic; Thomas (1976) reconstructs **plʌːm* for Proto-East-Katuic. The vowel doublet *ɤː/oː* in this particular word may seem to go back to old Northern Mon-Khmer since Diffloth (1980) reconstructs a similar doublet **plim/plom* for this etymon in Proto-Waic. There are, however, several words that differ between Mlabri and Tin in terms of *ɤ(ː)* versus *ɔ(ː)* (see the words with "*Ë₂*" in Tables 3-1 to 3-3 in Part 3) so this is a very well-established phonological correspondence.

The lexical item meaning 'small water leech' is an "areal" word; it occurs in Tai with a final velar nasal, whereas I suppose the old form in Khmuic was as in Mlabri and Ksingmul: **pliːn*. The Khmu form *pliːŋ* is exactly what one would expect as an early loan from Tai (Central Thai now has a short vowel). The TinPrai form *piːn* betrays itself as a loan-word by not having consonant mutation, unlike *pʰlɔm*, and that also explains the absence of *-l-* (Lao has likewise lost the lateral), but the final *-n* suggests a contamination with Khmuic *pliːn*.

There remain two MlaA forms: *tʰaːk* 'land leech' and *glɤːw* 'water leech, big species'. These may reflect the non-Tinic and even non-Khmuic layer of vocabulary in Mlabri.

A4. Some insects. There are some insects or bugs with vaguely similar names, which tend to be mixed up:

MlaA *rəpʰɛp* (some young people say *bṛpʰɛp*) 'butterfly'
MlaBC *ruphɛp* 'cockroach'
MlaABC *kṇdeːp* 'centipede species (described by a MlaB speaker as being red)'

For 'butterfly' MlaB has a quite different word $^ʔuʌʔ$; if I look for butterfly or cockroach terms in neighbouring languages I do not come across anything similar to the forms above.

It is different with 'centipede': Khmu has a presyllable beginning with a velar: $k^ʔiːp$ 'centipede', which is reconcilable with the Mlabri form assuming that the latter has an older phonology with more material preserved (the added complexity of the Mlabri form does not look like infixation). Katuic also has forms with a presyllable beginning in an initial velar stop but the main syllable has initial h, cf. Kui $k^hheːp$ and Proto-East-Katuic $*kalheːp$ (following the reconstruction in Thomas 1976). Bahnaric also has an initial velar in the presyllable but agrees with Khmuic in having no h in the main syllable: Proto-North-Bahnaric had $*gaʔjìp$ (following the reconstruction in Smith 1972); Bahnar has $kaʔep$. Waic has an initial sibilant in the presyllable: Proto-Waic had $*sʔep$ (following the reconstruction in Diffloth 1980). The same sibilant occurs in Tin: MalC has $saʔep$ 'centipede', TinPrai $siʔep$ 'large centipede'; it is unusual for Tin to have preserved a bisyllabic form outside loanwords.

If all these forms are considered together they suggest to me that the basic word is $*ʔep$ and that the forms above are originally morphologically complex, with two different first components, one with the reflex $k(\ )-$, another with the reflex $s(V)-$. This is confirmed if we look at the most conservative Tin dialect. TinMalA has a form $rep\ ʔep$, clearly with a separate first component and at the same time with a conspicuous similarity to the Mlabri forms meaning 'butterfly' and 'cockroach'. Amazingly, however, it means 'centipede'. There is no doubt about the meanings; I checked with speakers of all these language variants what a $rəp^hɛp$, $rɯphɛp$, or $rep\ ʔep$ looked like.

Since the only meaning with some distribution across languages is 'centipede', one may suspect that the M-K term for this repulsive creature could be elaborated into a variety of compounds (which have later mostly shrunk into monosyllabic or sesquisyllabic simplex words). This suggests that the meanings 'butterfly' and 'cockroach' are strange innovations in Mlabri, which became available because Mlabri had settled on a particular form $kɳdeːp$ in the meaning of 'centipede'. The occurrence of a glottalized $-ɗ-$ in that form suggests that it is not Khmuic at all though it is clearly of M-K provenance.

A5. Ants. This presentation is only concerned with ants proper, not including termites, which the Mlabri, like other Mon-Khmer-speaking groups, designate by a quite different term $druːɲ$ (there is no cover term

for these two kinds of insects according to their native zoology, which on this point agrees with scientific zoology). For 'ant' Mlabri has at least six terms:

ʰŋlʰnul (MlaAB), a very big ant species (occurring in mythology)
lɤːŋ / lɤːŋ (MlaA), a red ant species making its mound in trees
micmɛc (MlaABC), a small, black ant species
mot (MlaABC), a big, black ant species; *mot ṭlceel* is a species that bites so that the skin swells
rm̩ram (MlaB), a small ant species (which irritates the skin)
tŋtʌŋ (MlaB), big, black ant species

Four of these items may be innovations within Mlabri. The two remaining terms, *micmɛc* and *mot*, have cognates in other Khmuic languages. I shall comment only on those below.

The form *mot* occurs identically in Tai. Several M-K languages have a word for 'ant' with the same vowel but a final palatal and mostly with a presyllable or a consonant before the nasal. This is true also of Khmuic. Phong has *s(ə)muⁱc*, Ksingmul has *səlmoːc*, Khmu has *ʰmuːc* reflecting Old Khmu **smuːc*.

TinPrai *simɔːc* and TinMalA *ʰmoc* likewise show relics of the material before the nasal; it is impossible to see whether the TinMalC form *mot* comes from Proto-Mal **ʰmoc* or is a loanword. In any case, Mlabri *mot* is likely to be a loanword from Tai, possibly via Tin.

The form *smuːc* has been analyzed as containing an agentive infix *-m-*, i.e. M-K **s-m-uːc* 'the one that stings'. The underived stem **suːc* is a M-K word for 'sting (with an insect's stinger)'; in Khmu the verb is preserved as *huːc*, and the stinging organ is called *hrnuːc* (dialectally also *prnuːc*, according to Premsrirat's *Thesaurus*), obviously with the instrumental infix *-rn-*. In Mlabri the expected form *tʰuːc* is a noun denoting the stinging organ of certain insects.

Returning to the word for 'ant', Diffloth (1980) presented the analysis involving an *-m*-infix in order to account for the relationship between Proto-Waic **hmic* 'ant' (i.e. **h-m-ic*) and **hic* 'sting'. The analysis, which also fits Khmu *ʰmuːc* 'ant' perfectly, was later presented for Monic as well by Diffloth (1984: 73), and he there characterizes the morphological formation as ancient (implying, of course, that it was inherited from Proto-Mon-Khmer into both Monic and Waic).

Diffloth's reconstructed Monic form for 'ant' is **smɔːc* (not the expected **smuːc*) with a discrepancy in vowel quality which he explains

as "probably due to the presence of *-m-*". That lowering effect of the nasal must be a comparatively late phenomenon applying only within Monic, for in the same monograph Diffloth assumes that infixal nasal could on the contrary have a raising effect at a very early time. In deriving *k( )muːn* 'nephew; niece' from *kɔːn* 'child' he states that "the difference in vowels, *uu* vs. *ɔɔ*, can be explained by the presence of *-m- and by the extremely early date of this morphological derivation, perhaps as early as Proto-Austro-Asiatic" (Diffloth 1980: 114).

It must be emphasized, however, that the alleged derivation *s-m-uːc* must also be extremely old in order to have become an areal word. According to Jerold K. Edmondson (personal communication) the etymon occurs everywhere in Tai, including Kam-Sui. There it has the segmental structure *mVt* (e.g. *mot* or *mət*).

The Mlabri evidence does not fit well with this etymological analysis of the word for 'ant'. Mlabri phonology is otherwise mostly extremely conservative, and I would have expected a trace of the initial *s- if *sm- had arosen by infixation. The expected reflex of *sm- in Mlabri would be *tʰm-, with a structurally non-permitted cluster which would change into a sesquisyllabic form (*tʰəmuːc) unless it was simplified into *ʰm-, since that is what has happened to all monosyllabic clusters of stop plus nasal in Mlabri. There is absolutely no such reflex in Mlabri. Moreover, if the word(s) for 'ant' derives from a word involving the notion of stinging, I would expect that the Mla Bri would preserve a native zoology according to which the ants sting, but in fact all Mla Bri always speak about the ants as "pissing" and thereby making one's skin itch (the word for pissing being *nɔːm*). In short: I would rather assume that the basic word for 'ant' in Mlabri is *mVc*, a word unrelated to 'sting': *tʰuːc* (< *suːc*).

A possible solution is that the Mlabri form was borrowed from another Northern Mon-Khmer language. Phong has *mec* 'ant' in one of its dialects but *moc* in another; that is not a likely source of borrowing. Bahnaric, however, provides two etyma: *smoːc* 'ant' and *( )-mac* for various mosquitoes and biting flies (Gérard Diffloth, personal communication). The latter etymon might be the source of Mlabri *micmɛc* in the meaning of 'ant'. Mlabri differs by having introduced both a semantic shift and reduplication in this animal name.

After these exercises in small semantic fields from nature I shall proceed to a few etymological case-studies of Mlabri lexicon pertaining to daily life and spiritual culture (B1-B5 below).

B1. Words for 'sweeping'. There are two Mlabri verbs meaning 'sweep with a broom'. One is found only in MlaB: *peːlʰ* (with the nominalization *prneːlʰ* 'broom'); this form is interesting by having the final aspirated lateral that is a reflex of M-K *-s. In Tin one would thus expect something like Mal *pʰeːjʰ, Prai *pʰeːt (although the vowel might be of any quality within a whole range of phonetic values), but I have not retrieved such a word in Tin.

The other, more widespread, Mlabri word for 'sweep with a broom' occurs both in MlaA and MlaB as *kwaːc* (with the nominalization *krwaːc* 'broom').

In MlaB, *peːlʰ* is said to be male language, *kwaːc* being the corresponding female term (MlaB has a further, "female" term for 'broom': *tʰup ɲuːl*; I have not looked into the provenance of that term).

MlaA, however, has another variant of the verb for 'sweep with a broom', namely *kwaːt* (with the nominalization *krwaːt* 'broom'). The doublet *kwaːc ~ kwaːt* is suggestive of a merger between two different (though etymologically related) verbs since Khmu has *kwaːc* 'scoop away' versus *kwaːt* 'sweep' (the latter with the nominalization *krwaːt* 'broom').

The form *kwaːt*, with its final dental, has ties to Tai, but the occurrence of a derivative with infixation suggests that it is of age in Khmuic. Tin would then be expected to have initial *kʰw- with mutation, but the form that actually occurs in Tin in the meaning of 'sweep', is *kwaːt*, indisputably a post-mutational borrowing. I would suppose that both Mlabri and Tin have borrowed from these forms from Khmu.

Looking now at MlaC, we encounter a form *kwaːlʰ* 'sweep'. That suggests a M-K reconstruction *kwaːs (with the same final as the synonymous MlaB-word *peːlʰ*, which has a direct cognate in Waic: Diffloth 1980 reconstructs *pes 'to sweep'). The form *kwaːlʰ* obviously belongs within the same etymological range as the forms *kwaːc/kwaːt*, but the MlaC form is unique within Mlabri and Tin in reflecting the old etymon rather than being a borrowing. It also clearly shows that sweeping the ground was part of the indigenous culture although both Mlabri and Tin borrowed words for that activity at a recent time.

It would obviously be rewarding to look at words for 'sweep' all across Khmuic.

B2. Words for 'star'. The common Mlabri word for 'star' is *çəmɔɲ*, with a rounded back vowel in the main syllable. That is unexpected in a (West) Khmuic context and suggests ties with more distant Mon-

Khmer. MlaA, however, traditionally makes a distinction between two words for 'star': *çəmɔɲ*, which is explained as denoting a 'big' star and *çəmeɲ*, a 'smaller' star. Although the "ablaut" relationship between these two forms is iconic by signalling size (i.e., brightness) by means of vowel quality, the front vowel in the form *çəmeɲ* clearly has a Khmuic provenance. Khmu has the front vowel: *simeɲ*, and a similar form occurs in Tin Prai. A third vowel quality occurs in Ksingmul *məɲ* 'star' (with loss of the presyllable).

Reconstructing Tinic *-*məɲ* with the same vowel quality as in Ksingmul is warranted by TinMalA *ʰməɲ* (elaborated as *tʰuʌʔ ʰməɲ*, literally 'star moon', *ʔojʰ ʰməɲ*, literally 'fire of star'), and TinMalC *(tʰuʌʔ) məɲ*, and it agrees with Proto-Waic *sʔməɲ* (Diffloth 1980). In Prai, *ə* is regularly fronted before palatals so *-*məɲ* > -*meɲ* is expected. As for the presyllable, it can be reconstructed at the Tinic level as *Sʒi-*, with Mal *h-* as an expected reflex (Prai would exhibit *si-* whatever the phonology of the old presyllable).

This makes it an open question where the rare Mlabri form *çəmeɲ* comes from. It is hardly the direct continuation of a Proto-Tinic form, for a correspondence between Mlabri -*e*- and Tin -*ə*- is not otherwise attested in my material (if vowels differ along the front-back dimension the correspondences are of the opposite kind: between Mlabri -*ɤ*- and Tin -*e*-). The vowel quality -*e*- suggests that *çəmeɲ* is a borrowing from either Khmu or Prai, with laxing of the vowel in the presyllable as often happens in Mlabri. It might, however, stem from an older, non-Tinic layer in Mlabri in which the fronting of *ə* before palatals likewise occurred early.

B3. Words for 'essence of life; personal spirit'. The common Khmuic word is *Sʒmaːl* 'personal spirit; soul' (Khmu *ʰmmaːl*, *maːl*). It also occurs in Tinic. In Tin proper, the initial complex has been altered into pseudo-prefixal Mal *h-* ~ Prai *si-* + -*m-* in accordance with the general pattern of change in Tin. The word is attested in TinMalB as *maːl* (< *ʰmaːl*) and in TinPrai as *simaːl*, its meaning in Tin being 'soul; essence of life'.

In Mlabri there are two possible reflexes of *Sʒm-* depending on whether the form was realized as sesquisyllabic, i.e. with initial *Sʒm-* or as monosyllabic, with an initial cluster *Sʒm-*. If the form had become sesquisyllabic, as in Ksingmul *həmaːl*, the regular Mlabri reflex of *Sʒ-* was *tʰ-*. The resulting form *tʰəmaːl* is actually attested in MlaC though in the specialized meaning of 'spirit of deceased person'.

For 'personal spirit', however, all varieties of Mlabri have ^{h}*ma:l*, with
*S₁- > h-, which shows that monosyllabic **S₁ma:l* was the main form
(as stated in Part 2, this very form is a confirmation that homosyllabic
onsets of sibilant + nasal underwent early weakening so as to appear in
Mlabri as preaspirated nasals, like onsets of stop + nasal did). I have no
explanation of why the word exists in both of these canonical forms.

For 'soul', MlaB has a form *çrma:l* (with a variant form *çrmu:l*)
'soul', which is clearly the same etymon with an *-r*-infix. This is hardly
an indigenous Mlabri form since *ç-* as a reflex of **S₁* before *r-* does not
otherwise occur in Mlabri (one expects **S₁r- >* Mlabri *tʰr-*, cf. **S₁riaŋ >*
tʰre:ŋ 'tooth'). That suggests that this form was borrowed into MlaB
from some other Khmuic language. Some Khmu actually has *hrma:l*,
which must reflect an older form **srma:l* , so that may be the source of
the 'soul'-term in Mlabri.

Another Tinic word for'personal spirit', *lo:n*, occurs all across
Mlabri. It denotes the spirit that resides in the whole body and ensures
one's well-being as long as it is united with the body. That may explain
its occurrence in the expression *du? lo:n* 'be suddenly frightened',
apparently involving the notion of being so startled that one departs
from one's personal spirit (for *du?* cf. the combination of this verb with
ɟak 'go' in MlaC *du? ɟak* 'escape'). In Tin I know of this etymon only in
the Prai form *silon* 'personal spirit; soul', which suggests that the word
was at some time augmented with a prefix in Tin.

It is unclear to me what used to be the semantic difference between
the two etyma discussed above.

B4. The words *kol* 'stick' and *kol* ^{h}*lek* 'traditional lighter'
There is an enigmatic relationship between two Mlabri words: *ko:lʰ*
meaning 'wooden spear-shaft' and *kol* meaning 'stick of wood'. The
form *ko:lʰ* is a direct cognate of the Tin word for 'spear': Mal *kʰɔ:jʰ*, Prai
kʰɔ:t (with a secondary phonological development; the latter form has
recently been borrowed into Mlabri in the meaning of 'spear'). The cor-
respondence between Mlabri *-lʰ* and Tin *-jʰ* is regular and points to a
Proto-Khmuic final which I have given the abstract symbol **-S₁* (see
Part 2 and Part 3). Within Mlabri, the form *ko:lʰ* seems now to exist only
in MlaC. Outside Khmuic it is corroborated by Katuic: Thomas (1976)
reconstructs Proto-East-Katuic **ko:s*.

As for 'stick', a form *kol* is unexpected. Both Tin and MlaB use the
Tinic word *lam* 'tree; stalk' for 'stick' (Khmu has other words). Thus
MlaB speakers would say *toc lam tek*, literally 'take stick beat', to refer

to the action of keeping a dog away with a stick (rather than just shouting *ɕeʔ* to it); in MlaAC the standard phrase for that is *ek kol tɛk*, likewise 'take stick beat'. One might guess that *kol* belongs to a non-Tinic stratum but its similarity with *koːlʰ* is conspicuous.

My assumption is that there was in Old Mabri a compound word **koːlʰ ʰlek* meaning 'spear-shaft fitted with an iron tip', i.e. the type of spear that required contact with villagers to obtain the iron tip (*ʰlek* 'iron' is a loanword from Ancient Thai). Vowel shortening would regularly occur in the first part of such a compound resulting in **kolʰ ʰlek*. In sandhi the consecutive aspirations of **kolʰ ʰlek* would merge, and misanalysis into unaspirated plus aspirated lateral could occur, resulting in *kol ʰlek*. Thus *kol* could emerge in the meaning of 'shaft', and so it did in MlaAC, with subsequent generalization of its meaning into 'stick'.

So far this is pure conjecture, but interestingly, MlaA does have a form *kol ʰlek* though in a different meaning, namely referring to a slim, tightly capped bamboo section that belongs to men's standard outfit (in MlaBC a bamboo section is simply referred to with the term for the bamboo species in question: *diːŋ* in B, *talaːw* in C). The *kol ʰlek* was formerly used to house the traditional lighter, consisting of a piece of iron, a stone, and some fibers though that was not its only use; it could also contain salt, for example. Associating *kol ʰlek* with 'stick' and 'iron' does not make any immediate sense.

My best suggestion is that in MlaA as well, *kol ʰlek* used to mean 'spear shaft with iron tip', but its meaning was eventually transferred – either jokingly or for euphemistic purposes – to that of 'bamboo section containing the iron for the lighter' (there are other instances of seemingly strange meaning transfer in Mlabri).

B5. Words for 'water', 'drink' and 'urinate'. It may be called for to warn the reader once more that this essay is not only longer than the preceding ones but speculative to a much higher degree. It can perhaps be characterized as an uncensured presentation of the kind of raw notes one would start with in the search for etymological connections. A more solid analysis will be possible only if the data are approached from a much broader comparative perspective, ideally on the basis of a much clearer picture of Proto-Khmuic and even of Proto-Northern-Mon-Khmer than I can envisage.

The Tinic word that is continued in Tin as *ʔɔːk*, in Mlabri as *wɤːk*, is both a noun meaning 'water' and a verb meaning 'consume a non-solid

substance', i.e. 'drink (a liquid)' or 'smoke (tobacco etc.)'. The existence of a verb whose meaning encompasses the actions of drinking and smoking, is attested elsewhere in Mon-Khmer. That includes Khmu, the verb in question being ʔiək , obviously a cognate of the Tinic form. The use of the same word to denote water as a substance and to denote the action of drinking may, however, be a characteristic of Tinic. In Khmu, 'water' is ʔom, as distinct from the verb ʔiək 'drink'.

The Khmu word for 'water' is in itself interesting in a Tinic context. Ksingmul has ʔuːm in the meaning of 'bathe', phonetically close to the Khmu word ʔom for 'water' but both phonetically and semantically suggestive of further cognacy with the word for 'bathe': ʔum, ʔɤm that we find in Phong, Tin and Mlabri. In any case, the notion of water as a liquid is involved. The vowel in Khmu agrees with Diffloth's (1980) reconstruction of Proto-Waic *om 'water' (Diffloth comments on the occurrence of this etymon across M-K, ibid.: 81-82).

It becomes more complicated when we proceed to the notion of 'urine' or 'urinate'. Mlabri has nɔːm but Tin has ʔoːm, at first sight an irregular etymological correspondence. Something must have happened in Tin, however. Mlabri nɔːm is undoubtedly the basic word for 'urine' or 'urinate' since it is a widespread M-K etymon in that meaning (found e.g. in North-Bahnaric, Smith 1972 reconstructs Proto-Hrê-Sedang *nɔːm; the Aslian language Jahai has knɔm). Khmu has nuːm 'urine'. Thus one would expect *noːm rather than ʔoːm in that meaning in Tin. That Tin has ʔoːm instead is suggestive of influence from the word ʔom 'water' that exists in Khmu, thus exemplifying a universal way of speaking euphemistically about urine. That, however, would be an old innovation since ʔoːm is found all across Tin Mal and Prai, referring to urine.

To return to 'water' and 'drink', the form wɤːk is shared by A- and C-Mlabri, whereas B-Mlabri has jrʌːk with exactly the same range of meaning. There is a TinPrai word sirɔːk meaning 'leaf cup for drinking water', which looks like a direct cognate of the Mlabri word jrʌːk (the correspondence Mla jr- ~ TinPrai sir- is perfect, as evidenced by such a word as Mla jraː 'skinny', TinPrai siraː). In TinMal the corresponding form would be expected to be *krɔːk or *kjɔːk but unfortunately I do not have any record of any such form in the meaning of 'leaf-cup'. In any case, the combined Mlabri and TinPrai evidence suggests a Tinic reconstruction *jrʌːk or *jrɔːk.

The next question is how the meaning of *jrʌːk or *jrɔːk could split into so different meanings as those in Mlabri and in TinPrai. Mlabri has

a verb *juk* 'scoop', which in B-Mlabri can enter a phrase *juk jrʌːk* 'scoop water'. If the second part of that was originally a Tinic name of the utensil, the phrase could eventually have been misinterpreted as to the concrete meaning of its second part, which was assigned the meaning 'water' instead (such a semantic shift would be parallel to the use in Mlabri of *gɲrɛː*, which seems to be etymologically the M-K word for a pestle used in pounding spices, to mean 'curry; soup'). It might then function as a synonym of the inherited word *wɤːk* / *ʔɔːk* 'water; drink, v', maybe as female versus male vocabulary. If so, the semantics of this noun would be extended accordingly so as to include the verb meaning. In other instances of where Mlabri has a distinction between male and female vocabulary the words of such a pair are normally perfect synonyms with the same semantic range, the choice between them simply serving the purpose of marking off sociolects.

Because of the similarity in rhyme between MlaB *jrʌːk*, Khmu *ʔiək*, and Tin *ʔɔːk* + MlaAC *wɤːk* it is tempting to speculate that these words for 'water' form an etymological triplet with an "ablaut"-like alternation: Early Mlabri *ʔiək* > *ʔjʌːk* (not attested without infixation), Khmu *ʔiək* (without change), and Tinic *ʔuək* with two developments, *ʔuək* > *ʔɔːk* in Tin and *ʔuək* > *wɤːk* in Mlabri. The form attested in MlaB, *jrʌːk*, has an initial stop instead of a glide. This in itself would be no major obstacle to the reconstruction assuming that an intermediate step *ʔj-r- arose by infixation: that cluster would then violate the constraints on possible onset clusters and therefore be likely to change into *jr-*, a well-attested cluster. Since, however, the vowel quality does not agree, and since infixation in a CVC form would require *-rn-* rather than *-n-*, the etymological association of the MlaB form with the other forms becomes more than dubious.

There is a stronger corroboration of the original association of *jrʌːk* with the handling of water in the form of an obviously related Khmu verb *jruək* 'wash rice and put in the steam pot' (the meaning is here cited from Premsrirat 2002b).

Final remarks about etymological hypotheses. The small essays above are admittedly different from the word comparisons that one typically encounters in comparative work on Mon-Khmer languages. For the purpose of establishing genetic groupings of languages and hopefully even sound-laws the normal approach is to use carefully selected lists of words known to be amenable to cross-language comparison. As every linguist knows, however, that is not all there is to etymology. The

histories of individual words and their meanings, also of the much less
amenable ones, are interesting in themselves and important as input to
the study of cultural history, and that applies in particular when it is dif-
ficult or impossible to retrieve the past history, as is the case with the
Mla Bri.

The problem is that etymologizing on individual words often leads to
situations in which the choice between alternative explanations has no
rational basis, and two or more options must simply be left open (one
can assure oneself of that by looking at entries in etymological diction-
aries of some of the best studied Indo-European languages).

Beyond Tinic

There is no way to make sure that words shared in by Mlabri and Tin
are indeed Tinic words in Mlabri, not even in cases where they show
strict correspondences. Some of the words may be shared simply
because they belong to the core vocabulary of Mon-Khmer and are of a
shape making the words less prone to phonological changes. This is evi-
dently true of a word such as *mat* 'eye'. Words may also be of that kind
even if they do not seem that "basic" from a semantic perspective. An
example is Mlabri *blah* 'run off in different directions', which I regard
as Tinic because Tin has the expected form *plah* (< *blah*); that etymon
also shows invariance over widely separated branches, cf. Monic *blah*
'escape; go free' (Diffloth 1984).

A rare complication in determining the provenance of Mlabri words
is variation within the Mlabri language itself. The Mlabri word for
'head lice' is *çeʔ*; it could be old in the language or it could be a bor-
rowing from Khmu *seʔ* 'head louse'. It could also be Tinic; TinPrai has
ⁿseʔ for 'lice', with a trace of some lost material at the front end of the
word. Intriguingly, MlaC has a deviant form *çɛj* which cannot be recon-
ciled with the forms cited above because it differs both in absence of the
final glottal stop and in its diphthong instead of a short vowel. That is
also Khmuic; it agrees with the Ksingmul word for 'lice': *ceːj* (the cor-
respondence *ç-* ~ *c-* is expected). Thus we must recognize the existence
of two different but related forms of the same etymon, and it is signifi-
cant that the form that probably occurred in Early Mlabri form sides not
with Khmu or Tin but with the distant language Ksingmul. The variant
form with no glottal stop has a wider distribution, cf. that the word for
parasites such as lice in Monic is *caj* (Diffloth 1984) and that 'head

lice' is *ɲcɛː* in Kui (the initial nasal component in Tin and Kui is likely
to be a relic of a word meaning 'a multitude of –'; Khmu has two-word
phrases ending in *seʔ* used in the collective sense of 'lice').

Old M-K vocabulary in Mlabri that may be Khmuic but does not agree with Tin

Looking at content words (as against grammatical function words) in
Mlabri, the confrontation with Tin certainly does not exhaust the possi-
bilities of old Khmuic lexicon in the language. Several lexical items
have rather convincing cognates in other Khmuic languages (although I
have not established rigid sound correspondences) but fail to occur in
the data on Tin available to me, cf. Table 4-1. Still, the Khmuic but pos-
sibly non-Tinic words I have spotted so far (disregarding recent loan-
words) are considerably fewer than the cognates shared with Tin.

Table 4-1: Words which I have retrieved in some Khmuic but so far not in Tin

Mla *tʰɛːŋ* 'macaque (monkey)', cf. Khmu *chaːŋ* 'macaque' (the
phonology does not fit well)

Mla *tʰuːc* 'stinger (of insect)', cf. Khmu *huːc* (< *suːc*) 'sting, v'

MlaAC *tʰrːoc* 'snail species', cf. Khmu **ʰroːc* (dcumented as *roːc*)
'snail species'

MlaBC *pɤːr* 'lightning', cf. Khabit *pir* 'id.'

MlaAC *pɤj* 'eat fruit, vegetables, or sweets', cf. Phong *pɔj* 'eat'

Mla *poːlʰ* 'barking-deer', cf. Khmu *puas* 'muntjak', Phong *pwas* 'roe
deer'

MlaAB *plʌlʰ* 'flick; (B also:) strunt', cf. Khmu *plih, plis* 'flick'

MlaB *prijʰprɛjʰ*, MlaC *prɔjʰprɛjʰ* 'crisp', cf. Phong *pras* 'dry'

Mla *tel* 'listen', cf. Khabit *tal* 'id.'

MlaBC *tʌp* 'cover; thatch', cf. Phong *tɔp* 'bury; inhume'

Mla *tuːl* 'pointed tip', cf. Khmu *tuəl* 'end or tip of thing'

MlaAC *trɤːk* 'side of torso; rib', cf. Khmu *tʰriək* 'id.' (Tin has unrelated
terms within the semantic range: Mal *niʌr*, Tin *ᵐbən laː*)

Mla *ciaːk* '(sambhar) deer', cf. Khmu *tjak* 'deer species' (a widespread
etymon; Diffloth, 1980, reconstructs **cak* 'sambhar deer' for Waic)

Mla *kaːl* in *l�socⁱ ŋkaːl, lŋkaːl* 'forthcoming' (Mlabri *luŋ* means 'in the di-

rection of; together with'), cf. Khmu *ka:l* 'first' (the morphosyntactic and semantic differences make it unlikely that the Mlabri form is a borrowing from Khmu)

MlaB *kap* 'song', cf. Khabit *kʰap* 'id.'

MlaA *ka:p* 'duck', cf. Khabit *ka:p* 'id.'

Mla *ke:t'* 'ear', cf. Phong *kət* 'id.'

Mla *ki:ʔ* 'the moon; month'; in certain expressions also referring to the sun'; Phong has *ki*, but the word is scarce in Khmuic (TinMalA has *tʰuʌʔ* (sic!), TinPrai *suaʔ*; Khmu has *moŋ*); it is shared with Waic, however, the Proto-Waic form being *$kʰiʔ$ (Diffloth 1980)

Mla *kiaŋ* 'bird species', cf. Khmu *kiaŋ kiaŋ* 'hornbill'

Mla *çukkɔʔ* 'wet', cf. Khmu *skɔʔ* 'id.'

Mla *kuh* '(for liquid to) spill over', cf. Khmu *kəh* 'pour from one container into another'

MlaAB *braŋ* 'horse', cf. Khmu *hmbraŋ* 'id.'

Mla *bri:ʔ* 'forest', cf. Khmu *briʔ* 'sky; weather' (occurs in identical form in distant M-K languages so hardly a borrowing from Khmu)

Mla *dɔ:k* 'put away for later use', cf. Phong *doʔ* 'put; place' – the final velar is attested outside Khmuic, cf. Katu *dɔ́:k* 'put away for later use', but has changed into -ʔ in Phong (note: the Phong change *-k > -ʔ is attested in (loan)words such as *paʔ* 'stick; thrust', *lam baʔ* 'difficult', also cf. Phong *ci maʔ* 'hate' vs. Mlabri *mak* 'love')

Mla *kokdrɔ:j* 'pidgeon species', cf. Khmu *ku:k tru:l* 'peaceful dove'

Mla *gal* 'ten', cf. Khmu (ceremonial lg.) *gal* 'ten' (Mlabri and Khmu share the proto-form found as far away as Munda; the other Mlabri numerals below ten are not borrowings from Khmu so this is hardly a borrowing either)

Mla *klʔa:k* 'crow', cf. Khmu *klʔa:k* 'id.' (occurs in identical form also in distant M-K languages so unlikely to be borrowing from Khmu)

Mla *ʔɛm* 'lie in resting position; sleep', cf. Ksingmul *ʔe:m* 'lie'

Mla *ʔih* in *çiʔih* '(AB) be talkative in a high-pitched voice; (C) sing in native voice using falsetto' (*çi* is a preverb meaning 'prone to'), cf. Khmu *ʔih* 'talk'

Mla *çrɔ:ɲ* 'dry', cf. Khmu *tçrɔ:ɲ*, Khabit *ʰrɔ:ɲ* 'id.'

MlaAB *gm̩hɤːjʰ* 'gasp for breath', cf. Khmu *km̩ʰmɔːc* 'id.'

Mla *mak* 'love, v', cf. Phong *maʔ* 'id.'

Mla *mʌːr* 'crawl; creep', cf. Phong *mur* 'walk; creep'

MlaA *ɲrɲar* in *tar ɲrɲar* 'spinal chord', cf. Khmu *lɲar* 'spinal fluid'

MlaAB *ʰnɛʔ* 'respectful third person pronoun; marker of plurality', only with a remote cognate (if at all related): Khmu *sna:* 'they, dual'

Mla ^{h}nʌr 'wing', cf. Khmu *pnir* 'id.'

Mla *laːc* 'finish' (mostly as *a laːc* 'finished'), cf. Khmu *laːc* 'disappear'

Mla *dalaːw* 'bamboo species; bamboo tube' (with *l*-infix?), cf. Khmu *taːw taːw* 'bamboo tube (musical instrument)'

MlaBC *lɔŋ/loŋ klʏ(ː)ŋ* 'behind', cf. Phong *(taʔ) luəŋ kləŋ* '(lie) on the back'

Mla ^{h}*luːt* in *keːt* ^{h}*luːt* 'with a hearing impairment; deaf', cf. Khmu *sluːt*, Ksingmul *luːt* 'deaf'

Then there are some words which do occur in Tin but with so deviant phonology that they cannot have a Tinic provenance. They might have been borrowed from some other Khmuic language, but they do not agree with Khmu, and the semantic category of some of the words makes it altogether less likely that they can all be loanwords. Instead, they can be taken as indicative of a non-Tinic but Khmuic or at least Mon-Khmer component in Mlabri lexicon.

The existence of such words thus constitutes crucial evidence for the existence of Early Mlabri lexicon. Table 4-2 gives a list of examples.

Table 4-2: Mon-Khmer words in Mlabri
with irregular phonological offsets from Tin

MlaAC *pum* 'rigid rattan basket (used to store one's belongings)'; Tin-Prai *p^{h}lum* 'do., Mlabri style'; Khmu *plum* 'basket for keeping clothes'

MlaB *pjɔːl^h* 'stretch oneself', MlaC *pjal^h* 'stretch one's legs'; TinMalA *p^{h}ral*, TinMalC *p^{h}jaw*,TinPrai *p^{h}rɔl* 'stretch one's legs'

MlaAB *plʌk* 'hiccough'; TinPrai ^{m}*blʌʔ* 'hiccough'

Mla *taːl* (MlaB has *t^{h}aːl*) 'six'; TinPrai *t^{h}uʌl* 'six' and Khmu *tɔːl* 'six in ritual language)

MlaAB ^{h}*ntaːʔ* 'tail', cf. TinPrai ^{n}*t^{h}al* 'do.' (but a perfect fit with Khmu: *hntaʔ* 'id.')

MlaB *gʁtɔr* in *gʁtɔr lɔt* 'low part of the throat', cf. TinMalA ^{ŋ}*groːr* 'throat'

Mla *lʔmbaːʔ* 'cow', cf. TinMalA ^{m}*buʌʔ* 'cow'

Mla *bluut* 'extinguished'; TinPrai ^{m}*p^{h}ət* 'id.'

MlC ^{h}*ndɛh* in ^{h}*ndɛh mɛːm* 'breathe' (*mɛːm* means 'blood'), cf. TinMalA ^{n}*t^{h}uh* in *p^{h}lah ^{n}t^{h}ih* 'breathe'

Mla *gɛj^h* 'crab', cf. TinPrai ^{ŋ}*gaːc* 'crab claws'

Mla *glaŋ* 'husband'; TinPrai *kʰləŋ* 'id.'

MlaBC *gləːc* 'whistle (with the lips)', cf. Tin MalA *ᵑkʰɔc* 'whistle (with the lips)'; TinMalA *kʰrɔc*, TinMalB *kʰlɔːt*, TinPrai = TinCuun *silɔːc* 'whisper'

MlaA *kamɓɤːt* '(elder generation's word for) aeroplane, cf. Tin (Mal & Prai) *ᵐbuʌt* 'big, bee-like insect'

MlaC *ɕɔːʔ* 'hoe', cf. TinMalA *cɔː* 'hoe' (apparently contamination in Mlabri of words for 'hoe' and 'axe', see **ɕoːʔ* in H1b; there is a homonym *ɕɔːʔ* 'dog' in H1b)

MlaBC *ɕrɛːt* 'comb, V' (plus the nominalization *ɕɲrɛːt* 'comb, N'), cf. TinPrai *saːc* 'comb', Khmu *tɕriəs* 'comb'

Mla *ɕwal* 'light (i.e. not heavy)', cf. Tin (Mal and Prai) *sul* 'light'; TinMal also has variant of this etymon as a loanword: *cɔ(ː)l*, Khabit has *sɯəˑl* 'id.'

Mla *mʌːʔ* 'creeper species (with a very thick stem); python', cf. Mla *mʌːr* 'creep', Tin *maːr* 'snake' in D1

MlaAB *ɲaːm* 'season (of the yearly cycle)'; TinMalA *ɲum* 'id.'

Mla *tʰaɲaːp* 'yawn', cf. TinPrai *wah ᵑgiʌp* 'do.' and Khmu *hɲaːp* 'do.'

MlaB *ɕraɲɔːk* 'bird species'; TinMalA *ᵑgok* 'hornbill (bird species)'

MlaA *ʰnɛl* (<*knɛl) 'rat', cf. TinPrai *sinɛː* 'rat' (NB: Khmu has *kn-* but like Tin and all other M-K languages possessing this etymon that I know of it has no final lateral)

Mla *wɤːŋ* 'chin', cf. TinPrai *waŋ* 'id.' (better fit with Phong: *voŋ* 'id.')

Mla *grwɛc* 'finger; toe', cf. TinMalA *ᵑgrajʰ* 'do.' (the finals do not fit, whereas Mla *grw-* ~ Tin *ᵑgr-* is in accordance with general correspondence rules)

Mla *jɤ(ː)h*, *ʔjɤ(ː)h* 'hey! (calling somebody', cf. TinMalA *ⁿjah* in *ⁿjah swaʔ* 'call a dog'; TinMalC and TinPrai *cah* in *cah swaː* 'call a dog'

Mla *ʔjɔh* 'spit', v; TinPrai *cuah* 'spit' (NB: same vowel discrepancy as in Mla *bɔh*, Tin *puah* 'ashes')

The non-Khmuic component in Mlabri lexicon

In Khmuic, as apparently in Mon-Khmer in general, there are certain lexical items that occur in many different languages, also in languages that are located far apart. There is otherwise an enormous lexical differentiation within Mon-Khmer, and also within Khmuic, so the recurrence of certain etyma is rather remarkable.

The absence of an expected etymon in one or another language is, of

course, insignificant in itself, but if terms for certain very basic concepts, e.g. relating to body-parts or important phenomena in nature, differ it is nevertheless interesting. Table 4-3 gives a few examples of Khmuic lexicon that one might have expected in Mlabri (the Mlabri cognates have no counterparts in Khmuic languages that I know of).

Table 4-3: Examples of Khmuic basic vocabulary with unrelated counterparts in Mlabri

Tin (Mal and Prai) *ⁿtʰaːk* 'tongue'; Khmu *ʰntaːk* 'tongue'; this is a widespread M-K word but Mlabri has *r̩tlat*, in C mostly *lutlat*, a derivation from the verb *lat* 'lick'

Tin (Mal and Prai) *kʰwan* 'child'; Khmu *kɔːn* 'child'; this is a very widespread M-K word but Mlabri has *ʔɛːw* 'child'

Tin (Mal and Prai) *ⁿdah* 'forehead'; Khmu *kdah* 'forehead'; Mabri has *buːk* 'forehead; face'

Tin *se(ː)m* 'bird'; Khmu *siːm* 'bird'; this is a widespread M-K word but Mlabri has a different word: *aːc* as the generic term for 'bird'

Tin Mal *lwaŋ*, TinPrai *waːŋ* 'sky'; Khmu *lwaːŋ* 'sky'; Mlabi has *klaːr*, *kluːr* 'sky'

Looking at such discrepancies from the other side it does not take much study of Mlabri lexicon to see that it contains innumerable lexical items with culturally salient meanings which are unfamiliar in a Khmuic context. Some occur in distant languages in more or less deviating form (and sometimes meaning).

Occasionally there are even synonyms with non-Khmuic and Khmuic connections. An interesting case is the Mlabri terminology for 'father'. One word is *mɤm*, which I do not know from other Khmuic but which may be the same etymon as Monic *maːm* (Diffloth 1980) and Stieng *mɔːm*, for example. The shared Monic and Stieng meaning is 'father's brother' but Stieng also has the meaning 'father' (Haupers and Haupers 1991: 148). Then there are two mutually related Mlabri words, *joːŋ* and *jjoːŋ*, which are of Khmuic but not Tinic provenance, and which basically mean 'man; male' though they can also mean 'father'. The Mlabri word *joːŋ* is mostly used about men opposed to women: *ʔuːj joːŋ* 'woman and man', 'wife and husband'. It may have acquired the additional meaning of father under the influence of Khmu *joŋ* 'father' (Tin has other terms for 'father': *ʔaw* and 'man': *kʰlɔŋ*). The Mlabri

word *jjoːŋ* has a broader range of meanings associated with the male gender, including the transferred meaning of 'blacksmith's hammer' (*ʔuːj* being the anvil; these terms for male and female are also used about the Morning Star and the Moon, of which the former is considered to be chasing the latter). It does, however, refer to the paternal role e.g. in *ʔot jjoːŋ = ʔot mɤm* 'my father'. Mlabri *jjoːŋ* is indeed Khmuic as evidenced by its cognate *zuŋ* 'man' in Ksingmul.

In several other instances there is only one term, or one primary term, for a concept, and it is then often difficult to retrieve in other Mon-Khmer. If a word only exists in one variety of Mlabri there is every reason to suspect that it may be a neologism, although it might of course be an old word that has been lost in the other varieties. If a word occurs in all three varieties of Mlabri one would suspect that its cognates exist somewhere else in Mon-Khmer. Table 4-4 lists a small selection of such words. I have limited the sample to words which are not known to me in similar phonological shape either from Khmuic or from other branches of Mon-Khmer, but which have so general meanings that other terms for the same concepts are listed in a variety of M-K vocabularies.

Admittedly, one could take probably any language of this family and find that a substantial part of the vocabulary does not seem to appear in neighbouring languages, but the appearance of idiosyncratic lexicon (with or without opaque presyllables) is indeed a conspicuous characteristic of Mlabri. That is true in particular if one includes words with sound-symbolism but it is true even without them (I have included just one such word in the list, the word for 'elephant').

I certainly do not claim that the words in Table 4-4 are unique to Mlabri. Some I know to have vaguely similar cunterparts in remote Mon-Khmer languages ("look-alikes", which may then be true cognates or spuriously similar and due to chance, cf. that *mʌc* 'see' is vaguely similar to Phong *muh* and to Monic *-maj* of a similar meaning; as for the first component of *kobliːŋ* 'alive' it is thought-provoking that Phong has a two-word phrase *kɔ sjom* in that meaning). It is, however, my impression that considered as a whole, the Mlabri words listed in Table 4-4 deviate from the core of M-K etyma that occur widely and are typically considered in cross-language comparisons within this large language family.

**Table 4-4: A random selection of Mlabri words which
seem to lack cognates in neighbouring languages
and may not be "mainstream" Mon-Khmer either**

(a) nature
'sky': *kla:r*
'cloud': *ku:p*
'wind': *r̥mut*
'soil; ground': *bɛ:ʔ*
'stone': *kɛp*

(b) hunted animals
'elephant': *pm̥po:*
'rabbit': *twɛr*
'squirrel sp.': *katʰe:r*
'another squirrel sp.': *cakdar*
'monitor lizard': *pje:*
'crab': *gɛjʰ*
'fowl': *ʔjoc*
'bird': *ac*

(c) other important animals
'cobra': *tm̩ʔo:ʔ*
'bee': *ʔjek*

(d) staple food
'tuber; taro': *ʔe:ʔ*
'rice' (boiled wet rice): *ju:k*

(e) kinship terms for minors; group terms
'younger sibling': *ro:j*
'child': *ʔɛ:w*
'grandchild; nephew': *nɔ:ʔ*
'person; man; Mla Bri: *mla:ʔ*
'outsider': *kwʌr*

(f) terms related to the body
'face; forehead': *bu:k*
'hair': *-mu(:)j* in MlaA *k̩lmu:j*, MlaB *mujmuj* (MlaC *ŋwujŋwuj*)
'penis': *doŋ*

'vulva': *plil*
'blood': *ʰmʌŋ* (this is blood inside the body versus blood from a wound,
    which is *mɛːm*, a widespread M-K etymon)
'wound, n': *broːʔ*

(g) verbs with rather "general" meanings
'walk': *teːn*
'chase': *raːp*
'grasp': *toc*
'evade': *niːʔ*
'enter': *blʌk*
'put': *lʌp*
'swallow': *lʌːr*
'(be) alive': *bliːŋ, kobliːŋ*
'ache': *greːt*
'cry': *beːt, beːc* ('the baby cries': *ʔɛːw beːt/beːc*)
'speak': MlaAC *tʌɲ*, MlaB *glaʔ*
'see': *mʌc*

The phonology of the words in Table 4-4 permits an important general-
ization, however. In spite of the lack or scarcity of cognates the phono-
logical shape of each and every form is consistent with its being of
Mon-Khmer origin. Considered from that perspective the Mlabri lex-
icon of content words does *not* suggest that the language has an origin
outside Mon-Khmer.

Another way to approach lexical differences between Mlabri and Tin
is to look at particular semantic fields. There is one semantic field in
particular that is of that kind: kinship terminology. The Mlabri kinship
system is complicated and still imperfectly understood, and the kinship
terminology differs across the three varieties of Mlabri (for A-Mlabri
see Suebsaeng 1992; for B-Mlabri see Rischel 1995: 116-132; as for C-
Mlabri some kinship terminology has been recorded but no overall
analysis has been attempted so far). Still, there is a common core of
concepts relating to kinship and family with a terminology that is rather
well understood. A few such terms were included in Table 4-4 above. It
is intriguing to compare the kinship terminology of Mlabri with that of
Tin and with that of Khmu.

The kinship terminology that I know of for Tin is altogether different
from Mlabri. It is not just that the individual lexical items are different;
Tin exploit a different strategy from both Khmu and Mlabri in that it

uses phrasal expressions to specify basic types of family relation of a higher order. The words for 'father', 'mother' and 'child' are *ʔaw*, *meː* and *kʰwan*, hence one refers to the corresponding second-order relatives: grandfather, grandmother and grandchild by the same terms plus an epithet.

Mlabri has several complex terms in its kinship terminology but they have different functions: some give a specific definition of the nature of the relatedness between two family members, others are collective terms such as *km̩bɛːr km̩tiːʔ* meaning something like 'the whole group of kinspeople'. Most expressions can be analyzed morphologically; in this particular example *bɛːr* is the numeral 'two' and *tiːʔ* the numeral 'eight', the expression *bɛːr – tiːʔ* referring to an unspecific number like *bɛr tʰɤːŋ* 'two five' used about a plurality of persons (the presyllable *km̩-* is probably a contraction of *kumɔm*, a word referring to children and thus originally reflecting the conceptualization of the Mla Bri as defenseless vis-à-vis nature and other people).

There is some, but not much shared terminology between Mlabri and Khmu. It is apparent from the survey in Premsrirat's Thesaurus (2002b: 517) that Khmu has a device of "ablaut-like" differentiation between ascending or descending family relationship of different order (*ɟeʔ* 'grandchildren', *ɟɔʔ* 'great grandchild', *ɟɛʔ* 'great great grandchild'; *jaʔ* 'grandmother', *ɟeʔ* 'great grandmother'; *taʔ* 'grandfather', *tɛʔ* 'great grandfather'); Mlabri has nothing of the kind; the most conspicuous overall characteristic of Mlabri kinship terminology is that it does not exceed relations of the second order and that it has cover terms for grandfather + uncle: *taːʔ*, grandmother + aunt: *jaːʔ*, and grandchild + nephew or niece: *nɔːʔ*, in addition to more specialized terms defining the relationship of one's elder relative to one's father or mother.

The conclusion must be that the Mla Bri had an altogether different conceptualization of family relationships than the Tin, and it is not very similar to that of the Khmu either. Even the terms themselves are in most cases alien to Khmuic. All of this must go back to Early Mlabri, understood as the language spoken by remnants of a once sizeable Mla Bri population before the assumed new founder event involving an influx of Proto-Tin.

The difficulty with the provenance of seemingly unique Mlabri lexicon is that one can posit two kinds of origin. Some of the words may be due to lexical innovation after the Tinic period; the lexical discrepancies among the three varieties of Mlabri makes it likely that lexical inventiveness has played a great role in the shaping of this language. It is, on

the other hand, tempting to speculate that before the obvious contact with Khmuic, Early Mlabri may have belonged to a different branch of Khmuic or even of Northern Mon-Khmer. If that branch cannot be identified, the working-hypothesis must be that Early Mlabri constituted a branch all by itself.

Most of the Mlabri lexicon has so far no known provenance; the null hypothesis would be that it is all Mon-Khmer. I must concede that it is a major drawback that most of the Mlabri lexicon I have gathered and edited over twenty-five years is still totally unexplored with respect to historical provenance, since my emphasis has been on the Khmuic connection.

Within Mon-Khmer, it is my impression that the number of obvious correspondences even in the closest neighbouring languages outside Khmuic is modest. The most interesting neighbouring branch is Palaungic with its Palaung and Waic sub-branches. Some Mlabri words are closer to Waic than to Khmuic, apparently. An example is MlaB *mʌːlʰ* 'name', which agrees well with Diffloth's (1980) reconstruction Proto-Waic **mis* 'name' (the West-Katuic language Kui has *muh* because *-s > -h there), whereas it only has a remote – if at all related – Khmuic cognate, namely Khmu *meh* 'to be called', *pnmeh* 'to name'.

I made a quick search for cognates in the Proto-Waic lexicon of Diffloth (1980), which according to my count contains 516 entries (including some words that are shared with Tai but of age in Waic; if those etyma occur as fully integrated words in Mlabri I included them as well). Off-hand I spotted 92 entries with cognates in Mlabri, mostly close cognates but some rather aberrant. That gives some idea of the degree to which there is shared lexicon but *not* of the extent to which Mlabri has old lexical affinities to Waic. To approach that issue one must leave out of consideration all the cognates of Mlabri words that have been listed in Part 1 and Part 3 as Tinic or as recent borrowings from Tin or Khmu, viz. 34 entries in total. One must also leave out 15 words which are shared with Tai plus a widespread old loanword for 'gun' (Proto-Waic **snat*, Mlabri *çənaːt*), all of which might have entered Mlabri from anywhere.

There remain 42 entries. Of these, about half are shared between Waic and Mlabri but have clear parallels in Khmuic languages. As for the remaining entries I have not so far spotted close Khmuic parallels but I may have overlooked some e.g. in Ksingmul, and I am totally in lack of data from some of the "small" Khmuic languages. Five of these remaining words occur in Katuic, however, so they are likely to be part

of the traditional Khmuic vocabulary as well, since Khmuic is "in between" Waic and Katuic, as it were. Leaving those words aside as well, there remains a residue of 17 words that are potentially interesting by possibly pointing to a specific Mlabri-Waic (or rather: Mlabri-Palaungic) connection, though I must admit that additional Khmuic data may show that even they are shared by Khmuic. For the convenience of future comparative study I list them in Table 4-5.

Table 4-5
Mlabri-Waic cognates without so far recognized,
obvious Khmuic parallels

(NB!) PW *pɒ, *pɔ 'wake up' – Mla *puh* 'get up'; C also: 'wake up'

(NB!) PW *kʰiʔ 'wood; firewood' – MlaAB ʰŋkeːʔ 'firewood'

(NB!) PW *kʰiʔ 'moon; month' – Mla *kiːʔ* 'do.'

(NB!) PW *siʔ 'pain; disease' – MlaAB *ɕi* (with complement) 'do.'

(NB!) PW *hwɛk* 'dark' – MlaAC *tʰukʔwɛk* (C: *tʰukʔwɛk*) 'dark'

PW *bik 'tie, v' – MlaAB *buk* 'put a garment on (on torso)'

(NB!) PW *mək 'cow' – MlaA *ləmɔkmɛk* 'deer species (extinct)'

(NB!) PW *ɲit 'extinguish' – MlaAB *kaɲit* 'dark'

(NB!) PWL *cip 'put on; wear' – MlaAC *ɕup* 'put close-fitting garments on the head or the extremities'

PW *( )nim 'urine; urinate' – Mla *nɔːm* 'urinate'

(NB!) PW *kdiŋ 'big' – Mla *ʔdiŋ* 'id.'

PW *diŋ* 'paternal uncle; parent's elder brother; father's elder sister' – Mla *diŋ* 'close elder relative; elder brother'

(NB!) PW *ʔjoŋ 'small village' – MlaA *jʌŋ* 'marketplace; town'

PW *pes 'sweep' – MlaB *peːlʰ* 'id.'

(NB!) PW *kdəl 'stomach' – Mla *kṇdul* 'buttocks'

PW *joh 'call' – Mla *ʔʃh* 'id.'

PW *klaw* 'testicle' – Mla *klaw* (A: *klaːw*) 'id.' (Khmu shares with Kui a distant cognate: *klaː* as the 2nd component of terms for 'testicles').

Some cognates in Table 4-5 are fairly straightforward; a very few show minor semantic deviations but many more show phonological discrepancies. Many of these phonological discrepancies consist in presence or absence of a final laryngeal, differences in vowel length, or differences in vowel quality, i.e., exactly the kinds of discrepancies that make it difficult to establish regularities between Mlabri and Tin in putative Tinic words.

Mlabri has t^h- corresponding to Waic *h-, and both predictably go back to Mon-Khmer *s-; similarly, Mlabri has -l^h corresponding to Waic *-s, and both predicably go back to Mon-Khmer *-s.

In some cases Mlabri has presyllabic material or other phonological discrepancies from Waic that are informative for deeper reconstruction; I put "(NB!)" in front of word pairs with such more important or more intriguing deviations. As for PW *$kdiŋ$ 'big', Diffloth (1980) cites the Khmu form because in his view, Khmu exhibits more ancient phonology than Palaungic in such a case. The Mlabri form $diŋ$, however, supports the vowel quality of Waic vis-à-vis Khmu. At the same time, the correspondence kd- ~ d- opens up for an explanation of the origin of glottalized voiced stops in Mlabri: some of them, at least, are reflexes of various M-K clusters of voiceless stop + voiced stop.

If one goes outside Table 4-5 there are numerous other interesting correspondences although they are not evidence for a specific bond with Waic. The cognate of PW *$kdim$ 'ripe' is Mlabri $^h ndo:m$, suggesting a reconstruction with initial *kn+d. It deserves mention also that the vowel quality in Mlabri and Khmu gal 'ten' is the original one, as seen by comparison with Mundaic; I wonder whether Waic *kol 'ten' can be due to interference between the old Austroasiatic numerals *gul 'seven' and *gal 'ten'.

Within Palaungic, Waic constitutes a sister branch beside Palaung-Riang. It would have been interesting to include Proto-Palaung in the comparative considerations. Unfortunatly, I have not been able to do that except with a few etyma quoted from Short (1963) and Mitani (1979) in earlier parts of this monograph.

If one turns to the south-eastern side of the Mlabri area the closest non-Khmuic languages of the M-K family are the Katuic ones. As mentioned already in Part 1, I unfortunately did not have access to the recent Katuic etymological dictionary (Pejros 1996a) so I had to use sources with a more fragmentary coverage. Searching for cognates in the West-Katuic language Kui (Sriwises 1978) I did find some but not an impressive amount (a few have been cited in this monograph). As for East-Katuic, the languages in question are spoken further to the east, inside Vietnam, so their geographical distance from the present Mlabri area is considerable. On the other hand, there is the advantage that Dorothy Thomas (1976) gives a list of 667 etyma each of which is attested in at least two out of three East-Katuic languages. Thus the list may also be fairly representative of the general Katuic lexicon. The languages are Brôu (data from Mrs. John Miller), Pacőh (data from Mr. and Mrs.

Richard Watson), and Katu (data from Judith Wallace and Nancy Costello; I have otherwise been using the dictionary of the last-mentioned scholar).

By cursory inspection of the East-Katuic list (as with the Waic data), I found fifty entries that have Mlabri cognates that are familiar to me, i.e. roughly half the number of etyma on the Proto-Waic list that are shared with Mlabri (note that the Waic and East-Katuic lists are of comparable size, the latter actually being the largest, and that my approach was the same in both cases). The lexical closeness to East-Katuic is apparently considerably smaller than the lexical closeness to Waic. I therefore did not go into detail such as checking the provenance of the Mlabri words that have cognates on the East-Katuic list.

As for phonological similarity, my overall impression is that the words that are shared, occur in fairly similar basic shape across the Khmuic-Katuic divide, at least as far as the consonants are concerned, though some cognates are very deviant. I renounce on giving a list of the fifty cognates here because of difficulties in representing the vowel qualities of the different languages correctly (the PEK reconstructions do not always give a sufficient coverage of these word-forms from the perspective of Mlabri-Katuic or Khmuic-Katuic comparison).

My conclusion from the above is that there is a certain amount of lexical cognacy between Khmuic and Katuic, as one knows already, but there is no evidence of affinity between Mlabri in particular and Katuic. By comparison, the affinity of Mlabri to Waic, or more generally to Palaungic, becomes even more interesting.

The word for 'cave': evidence from a remote past? It is thought-provoking that although most Mlabri of the A- and B- groups refer to a house or lean-to as *gɛːŋ*, the small group in Laos (speaking C-Mlabri) uses the word *krʔuŋ*, which basically means 'cavity' in Mlabri, to refer to any kind of dwelling or shelter or camp. This usage, which makes perfect sense if it referred originally to caves of some kind, is attested already in the earliest word list of Mlabri (Bernatzik 1938).

I would assume that the word that basically means 'cave' is the original Mlabri term for 'native dwelling', although the A- and B-varieties have extended the use of the word for 'house', *gɛːŋ*, to refer to the lean-to as well.

The cave-word has an apparent counterpart in TinPrai, *grɔŋ* or *ᵑgrɔŋ*, which means 'hole; cavity' (some speakers claim that there is a difference in meaning between the two pronunciations, but I have a suspicion

that this difference is spurious). The voiced stop and the optional pre-nasalization in Tin show that some consonantal material was lost in the beginning of the word. Still, the Mlabri and Tin forms are difficult to reconcile in their phonological shapes. If one compares with other Tin words of similar structure the recent history of which can be reconstructed, $^{v}grɔŋ$ looks more like the opaque reflex of a deverbal noun formed by nasal infixation after an initial lingual stop. If that is what it is, it has nothing to do with the Mlabri word for 'cavity; cave'. I do not know whether the Mlabri etymon exists in other Khmuic; Khmu and Phong have a loanword from Tai in the meaning of 'cave'.

One must look at more distant languages to retrieve safe cognates of the Mlabri form. I have noticed that it is attested as far away as in Stieng (in southern Vietnam) in the form *rhuŋ*, in Jahai (on the Malayan Peninsula) in the form *gahuŋ*, and in Monic in the reconstructed form **sruŋ* (Diffloth 1984). There may have been different prefixes such as $*k(r)(V)$- and $* s(r)$- on the same main syllable *-uŋ*. In the various sources the meaning is given as 'hole', 'cavity', etc.; in Jahai it means 'cave'.

By a stretch of the imagination, the cave-metaphor could be seen as reflecting a former use of caves as shelters. The existence of prehistoric cave-dwellers is well-attested in Northern SEA, and local archeologists have in fact speculated (independently of the just-mentioned lexical usage) that the Mlabri might have a connection to the prehistoric Hoabinhian culture.

Is the old Mlabri lexicon all of Mon-Khmer origin?

No matter from what angle I approach the relationship of Mlabri to other languages I arrive at the conclusion that I can account for only a minority of its words. Close to half of the entire lexicon is made up of words that are vaguely of Mon-Khmer appearance but do not have specific look-alikes that I know of.

Considering that there are considerable lexical differences among the three varieties of Mlabri there is every reason to assume that recent neologisms have played a great role in shaping the lexicon of this language. The same may have been true in earlier time. One would, however, assume that there must be a sizeable residue of "original" words. It would take a very comprehensive search across Mon-Khmer languages to retrieve a substantial amount of additional M-K material, and then one must be aware of the danger of similarity by chance. After all, most

words in these languages are made up of one syllable of very simple structure so there are limitations on the combinatory possibilities which make chance similarities more likely to occur.

The next, and no less important, step is to look outside the Mon-Khmer language family, i.e. to include languages that belong to the Sino-Tibetan phylum. Such languages are nowadays spoken not too far from the present Mlabri area. If one looks at northernmost Laos, the language specimens (phonological introductions and short vocabularies) in Kingsada and Shintani (1999) comprise several languages classified as Lolo-Burmese: six languages of the "Phunoi" group, five languages of the Akha group, and a non-classified language Sida (Sila). In addition, there is one Sinitic language (Ho). West of the Mlabri area there is Akha and Lahu and Lisu (Lolo-Burmese), for example. Even Burmese presence around the Golden Triangle is of some age. Given the time depth, however, I have no clue as to how one could identify specific languages as being particularly likely to have been in ancient contact with Mlabri.

Moreover, there is so far no evidence for the extensive presence of non-Mon-Khmer, non-Tai lexicon in Mlabri. A cursory glance at some of the Lolo-Burmese languages in Kingsada and Shintani (1999) has *not* given me the impression that they can provide the key to vocabulary of unknown provenance in Mlabri (whereas their lexical relatedness to Burmese is very obvious), but then the questionnaire used in that survey comprised only a very limited and very basic vocabulary: concepts for which Mlabri has Mon-Khmer vocabulary.

Another way of approaching the issue was to look through Bradley's (1979) glossary of 886 numbered entries (in fact some 900 entries in total) with reconstructed Proto-Loloish Vocabulary and attested reflexes for each gloss. Again, the result was singularly negative. The look-alikes (e.g. *tek^H 'kick' ~ Mlabri *tɛk* 'hit', *go*[1] 'that (near)' ~ Mlabri *gʌh* 'this; here') are so conspicuously few that they might be due to chance similarity. Then Mlabri has a word *kaːp* for 'duck' that does not occur in Tin and Khmu and seems unusual in a wider M-K context but is similar to *kap* 'duck' in two of the Loloish languages of northernmost Laos, Phunoi and Bisu. That might be discarded as trivial similarity between onomatopoetic words, but there is in fact evidence for it being due to borrowing the *opposite* way, i.e. from Khmuic into those languages, since Khabit and Ksingmul side with Mlabri in having *kaːp*.

In order to possibly retrieve a significant amount of Loloish vocabulary in Mlabri one would have to perform a difficult search for special-

ized vocabulary, not only terms related to hunting, zoology and botany, and the like, but verbs in highly differentiated semantic fields such as posture, movement, hitting, cutting, etc. I am not in a position to pursue such studies.

Final remarks about lexicon. Hopefully, it is clear from the preceding sections that my whole approach to Mlabri lexicon is extremely tentative as soon as the scope moves beyond Tin and Khmu. Firstly, one cannot draw conclusions about the absence of a word in a certain language or branch because one fails to retrieve it in the sources, although one is on somewhat safer ground if there is positive evidence that the language(s) in question use an unrelated synonym with seemingly the same semantic range. Secondly, the unfortunate conditions under which this monograph had to be completed prevented me from having much access to comparative data outside the field I am primarily working in.

I might have stopped the lexical exercises within the bonds of Khmuic, but I hope that my more or less carefully phrased suggestions about the wider scenario will be taken as suggestive or at least provocative.

The origin of Mlabri grammar

As shown in Part 2, the morphosyntax of Mlabri has a considerable number of peculiarities. These add up to giving Mlabri a profile which is quite different from that of Tin on the one hand, and from that of Khmu on the other. This has not influenced the placement of Mlabri as Khmuic in all modern reference works. Nothing was known about Mlabri grammar until rather recently and in addition, the classification of Northern Mon-Khmer languages has been based on lexical similarities not on grammar, as I understand it.

I shall now summarize the features that make Mlabri grammar inexplicable in a Khmuic context.

Features affecting the canonical shape of wordforms
(1) *all* transparent pre- and infixation is accompanied by obligatory syllabicity associated with the affix itself
(2) both in verbs and nouns there is an abundance of word-formation by means of reduplication
(3) there are numerous sesquisyllabic words in Mlabri; the majority of these are not morphologically transparent

The word shapes created by these characteristics make Mlabri look like a full-fledged Mon-Khmer language. That does not, however, give Mlabri a Khmuic appearance. The abundance of obscure presyllables sets Mlabri off from its geographical neighbours. It contrast both with the fairly rich and transparent morphology of Khmu and with the near-absence of morphology and of sesquisyllabicity in Tin or Phong.

The strict constraint on sesquisyllabicity in derived word-forms sets Mlabri off from much other M-K as well. If one looks at the two most common nominalization affixes, -*r̩*- and -*r̩n*-, it makes sense to consider syllabicity a phonemic property of the affixes, as indicated by the transcriptions, since that readily accounts for their distribution. If MlaB *glaʔ* 'speak' is nominalized one gets *gr̩laʔ* 'conversation' but if MlaB *peːlʰ* 'sweep' is nominalized one gets *pr̩neːlʰ* 'broom' (not **preːlʰ* since that is ill-formed as a Mlabri nominalization). In any case, the crucial characteristic of infixation in Mlabri is that it regularly creates minimal phonological contrasts between non-derived and derived words, e.g. wordforms beginning in *grw*- (underived) versus *gr̩w*- (derived).

Although four morphological processes are operative in Mlabri: prefixation, infixation, reduplication and compounding, these do not account for the abundance of sesquisyllabic (or disyllabic) word-forms. Most of these do not lend themselves to a morphological analysis in terms of any of the four morphological processes.

It would be a reasonable expectation that some of the presyllables in this large residue of sesquisyllables had identifiable ranges of meaning, i.e. that they functioned as quasi-morphemes signalling that the words carrying them belonged to a particular semantic area. I have not been able to establish revealing generalizations of that kind so far, however. The vast majority of sesquisyllabic wordforms (e.g. in *kl̩*-, *cŋ*-, *rm̩*- and numerous others of similar kind) defy morphological decomposition; they simply have that structure as lexical items.

This suggests that the abundance of sesquisyllabicity and its (morphological) origins must be a very old feature of the language. Some of the presyllables that are now semantically and morphologically non-transparent may have had well-defined meanings but if so, these meanings are long since lost (one would probably have to perform wide-embracing comparisons along the lines of Daladier 2002 and Anderson 2006 in order to posit reconstructed meanings of some persuasive force). That the word-forms themselves have been preserved faithfully is not due to morphological conservatism but to phonological conservatism.

Affixes and particles

(4) The only general prefix is causative *pa ~ ba* (by comparison, Khmu has several prefixes; Tin has no productive prefixes left)

This prefix is conspicuous by containing an obligatory full vowel (which is not the case in the neighbouring M-K languages for which I have data) and by having voicing dissimilation between the onset of the prefix and the onset of the main syllable (which is not the case in surrounding languages either).

(5) There is only one morphologically functional infix, a nominalizing infix with three alternants -*ŗ*-, -*ņ*-, -*ŗn*-

Again, Mlabri differs from the surrounding M-K languages I know of.

(6) Mlabri has more or less clear patterns involving suffixal -*h*, -*ʔ* and -*t* in the pronominal system

The suffix -*t* has a possessive meaning, making Mlabri stand out as a language with possessive pronouns; as for the other suffixes, the exact meaning is obscure.

(7) Mlabri has a Definite Article (most likely as a consequence of the emergence of possessive pronouns)

This is makes for the encoding of referentiality in a manner which to my knowledge is unusual in a Northern M-K context.

(8) Mlabri has a particle *di* used to encode possession: POSSESSOR + *di* + POSSESSUM, the possessor having the same possibility of explicit referentiality as nouns headed by the definite article

I do not know that this syntactic device exists elsewhere in Northern Mon-Khmer.

(9) There is a tendency toward systematic exploitation of the vowels *a* and *i* in prefixes and particles

As for the vowel *a*, it occurs in the morphemes that are most salient in verb morphology and syntax at the single-verb level: causative *pa ~ ba*

and perfective *a*, *ma*, *ta*, *tama*. As for the vowel *i*, it occurs in particles that connect clauses or predicates: *i*, *di*, *ni*. This exploitation of the vowels within rather well-defined categories is hardly accidental; I know of no parallels in Khmuic but something vaguely similar has been reported for Palaungic.

Conclusion about morphosyntax. As a sweeping conclusion from the above I would say that morphology and syntax point in the same direction as the non-Tinic part of Mlabri lexicon: if Early Mlabri was at all Khmuic it must have constituted a separate branch, but the basis for arguing in favour of that classification is slim. It is worth considering that Mlabri may be the remnant of a separate branch of Northern Mon-Khmer although it has been influenced by Khmuic.

General conclusions based on linguistic evidence (I-II)

(I) Tinic, Khmuic and deeper layers in Mlabri

The irregular vowel offsets between Mlabri and Tin that were discussed extensively in Part 3, can be seen in a radically different light if one assumes that Mlabri before the Tinic connection was a Northern Mon-Khmer or even Khmuic language forming a separate branch. In that case there would be a substantial number of etyma occurring both in Early Mlabri and in Proto-Tinic. If a new ethnic group was founded by a combination of speakers of early Mlabri and Proto-Tinic it could well create a clash between the forms occurring in one and the other language. One thing that strikes the observer when comparing different Khmuic languages is how much and seemingly unpredictably they differ on the vowels in individual words. This is not just a problem to do with the relationship between Mlabri and Tin; it is all over this branch of Northern Mon-Khmer.

It is interesting to envisage the social situation shortly after the founder event assumed by the biologists (Oota et al. 2005), with children being raised hearing words pronounced in two different though related languages. All evidence suggests that Mlabri is a conservative language in terms of its phonology; Tin, on the other hand, must have undergone various changes already in the interval from Khmuic to Proto-Tinic. Thus in the Mlabri-Tin symbiosis there might be a compe-

tition between two norms, possibly with the result that some wordforms survived with Mlabri vocalism, others with Tin vocalism.

This is consistent with the finding that the vowel offsets are not chaotic but mostly minimal differences. To recapitulate, it is typically so that pronunciations in contemporary Mlabri and Tin differ by only one step of aperture or by another, phonologically binary, articulatory feature (front-back, less often rounding). This is exactly the kind of variation over pronunciation one finds in languages which for one or another reason have coexisting and competing norms.

There is, however, another explanation of some of the offsets between Mlabri and Tin forms in the shared vocabulary, namely that Tin underwent changes *after* Tinic time, some of these changes happening already while Tin was still essentially one unit, before the bifurcation into Mal and Prai. That explanation is fully confirmed by the vowel shifts in Tin that were identified in Part 3. On the other hand it does not account for some other offsets between Mlabri and Tin.

Methodologically, this leaves the researcher in an impossible situation: one is faced with an equation which cannot be solved because there are too many variables. There is a scenario of vowel shifts in Proto-Tinic, a scenario of competing norms in the language resulting from the founder event, and a scenario of vowel shifts in later Tin. All of this predates the chronological period treated in Filbeck's historical study of Tin (1978), his reconstruction of Proto-Tin being essentially a common denominator for the two branches of Modern Tin. When it comes to the chronological period before that, the degree of progress in the understanding of Mlabri and Tin language history depends on the degree to which well-defined sound-laws can be spotted in the vowel correspondences and shown to have operated in Tin after the Tinic period.

In any case, the hypothetical scenario of Mlabri-Tin confrontation sketched above would suggest some mutual intelligibility between the two "ancestors" of present-day Mlabri. That might be the case even if some of the shared words differed in more awkward ways, as exemplified in Table 4-2 above.

The words listed in Table 4-2 show offsets in the consonants as well. They do not present any clear picture since the instances of this in the material under consideration are too few, so it is difficult to draw safe conclusions from them. Some of the discrepancies are, however, of such a nature that it is tempting to assume that they go back to very early Tinic. If so, Mlabri did not get these words from Tinic but had them inde-

pendently. That would make sense if Early Mlabri constituted a separate branch of Khmuic, as against the Tinic or perhaps Phongic branch.

(II) What happened to Tin?

The Tin villagers probably formed a small group leading a very isolated life for a long time in a rather small area filled with high mountains speckled by their villages. They may not have been in much contact with outsiders until fairly recently; I do not know of solid evidence for old loanwords in Tin. More recently, they clearly got into closer contact with other local ethnic groups and with the lowlanders; there are numerous modern loanwords from Northern Thai or Lao in Tin. Again, that may have something to do with economy; in some of the poorest Tin villages they can still make a supplementary income by producing such cheap items as brooms and thatching.

At a fairly recent time, the Tin language split up into Mal and Prai dialects, suggesting that there were subgroups who settled apart from each other. That has led to hampered inter-communication between the subgroups and a gradual ethnic separation between them.

At some point before the bifurcation into Mal and Prai commenced, the pronunciation norm of the Proto-Tin language must have undergone considerable changes. A possible reason is a mixing of norms. That could happen if the ethnic Tin fused with another ethnic group speaking a very closely related language or perhaps just a very aberrant and now lost, early dialect of Tin.

Nothing is known about the extent to which small Khmuic languages of the area have vanished before our time, and we do not know the original ethnonym or ethnonyms of the Tin; the Mlabri call them *grɛh*, which would be **krɛh* in contemporary Tin if it existed. Maybe that term denoted a distinct subgroup. In any case, a disturbance of the pronunciation norm before or around the splitting into Mal and Prai could account for some of the strange vowel mismatches which we can now observe between the Mlabri and Tin pronunciations of shared words.

A hypothetical historical scenario

The apparently mixed character of the Mlabri language even before the Tinic era is suggestive of old, extensive bilingualism involving a

Northern Mon-Khmer language and one or more languages belonging somewhere in the Sino-Tibetan phylum (I deliberately keep the definition so vague rather than talking about Tibeto-Burman, since one of the most pervasive features of Mlabri syntax, the use of *di*, is Sinitic). Such a connection outside Mon-Khmer cannot so far be proven in the general lexicon of "content words".

In spite of lack of evidence it makes sense now to sketch a fairly plausible history of the Mla Bri and their language. Let us assume that more than a millennium ago there was a fairly numerous and active ethnic group, the ancestors of the Mla Bri (on the male side), who lived partly by fouraging in the forest, partly by activities in the open land such as trade of meat, wax, etc. for other goods. They would have to be at least bilingual; many, at least among the grown-up men, would be trilingual. The relevant languages in northernmost Southeast Asia would belong to Sino-Tibetan and to Mon-Khmer. There might even be intermarriages across ethnic and cultural barriers. The eventual outcome was a language whose lexicon was of general Mon-Khmer appearance, whereas its syntax on some points resembled Sino-Tibetan.

If several Mlabri words had by now been identified as Sino-Tibetan I would consider the hypothesis that the language of the ethnic group originally belonged to that phylum, and that in a certain time period there happened a massive replacement of the original lexicon by Mon-Khmer lexicon. That would have happened because the ethnic group had come to live in a Mon-Khmer speaking environment, with even the local peer language or the lingua franca belonging to that family. In short, the hypothetical scenario would be that of an ethnic group speaking Mon-Khmer on a non-Mon-Khmer substratum.

The problem with this assumption, apart from the lack of lexical evidence so far, is that the tangible evidence for a Sino-Tibetan substratum would be syntax. Preservation of syntax while the lexicon is gradually replaced is *not* what one expects to happen in situations of language shift. That is normally a wholesale thing happening after the native language has lost more and more domains of use.

The situation may have been different, however. The alien features of Mlabri syntax may represent an *adstratum*; such diffusion of syntactic phenomena across language barriers is well-known from areas with a so-called "Sprachbund". Assuming that the "original" mother-tongue of the Mla Bri was a Mon-Khmer language, the impact from Sino-Tibetan might be due to extensive use of another language for communication outside the domestic environment, i.e. extensive bilingualism. Again,

however, one would expect lexical not just syntactic impact. There may have been a strong social bond within the group which was instrumental in the preservation of native vocabulary without mixing it up extensively with the extraneous vocabulary.

The conclusion is that Early Mlabri was a Northern Mon-Khmer language, possibly constituting a branch of its own but with lexical and morphological affinities to Palaungic; it seems to have had some syntactic affinity to Sino-Tibetan.

Later, the Mla Bri experienced another early encounter with outsiders: Tai-speaking newcomers to the area. They continued their external contacts including barter trade, and that readily explains the presence of such words as *bɛːŋ* 'costly' in Ancient Thai form in Mlabri.

Eventually, the Mla Bri must have been considered an unwanted section of the population. Maybe the number of individuals roaming the northern Thailand-Laos border area was indeed decimated by the Northern Thais, as suggested by a Mla Bri myth, and a few survivors took refuge in the deep forest. That was where they encountered a few ethnic Tin, including one or at most two girls according to the biologists' findings (as said earlier, the Tin myth suggests that there was just one girl and also a boy).

These few individuals then founded the ethnic group anew, with the result that the language had a lexical influx of Tin words, mostly belonging to domestic vocabulary rather than being related to hunting. There was no perceptible influence on Mlabri grammar; it was taken over wholesale from Early Mlabri.

In the first centuries after that the Mla Bri population grew to more or less its present size but they stuck to a minimum subsistence as hunter-gatherers, perhaps because the collective memory of previous near-eradication transformed into a constant fear of outsiders, particular of lowlanders (a fear put into system by developing a belief into the spirits strictly banning interaction with outsiders).

For centuries, probably, there has been intermittent interaction with highlanders living in villages such as the Khmu and in particular the Tin; in some places also Tai villagers choosing high slopes for their farming. That has left its stamp in the form of recent loanwords, in particular from the common lingua franca, which was Lao or Northern Thai depending on the location. In terms of material culture that has stimulated the development of high skills in producing such things as baskets and mats made of plant fibres and weeds. Locally, the Mla Bri were renowned for such handicrafts which they exchanged for goods.

The present contact with the Hmong and the Mien, on the contrary, is quite recent and has not affected the indigenous culture and language. The interest of those ethnic groups in the Mla Bri has been to have access to a cheap labour force in clearing the mountain slopes for agricultural use.

The historical progression in the account above is admittedly speculative but to the best of my conviction not implausible. It accounts for all the linguistic observations and for the myths as well, both those of the Mla Bri and those of the Tin. It suggests that the hunter-gatherer lifestyle of the Mlabri was a means of survival, and that is consistent with the finding that their material culture is on the one hand minimalistic and on the other hand characterized by tools and techniques that are used also by other locals if they camp in nature or go hunting.

Future areas of study:
specialized lexicon and spiritual culture

One of the challenges for future historical and comparative study of Mlabri is to look at the specific hunter-gatherer terminology and find a way to assess its age in the language. It is a particularly interesting question to what extent names of wild animal species are ancient and to what extent they are on the contrary recent innovations coined in order to avoid mentioning the game by its proper names. I have recorded a fairly small number of animal names which seem to be associated with hunting and have synonyms used in domestic life, but I have not been able to determine the extent or age of such practices.

Another challenge is to look at the spiritual culture of the Mla Bri: their rituals, myths, epic singing, and general narrative tradition. Much of that has now been preserved in audio or video recorded form, but it is a huge task to process the texts, which are often difficult to understand and translate. Just scratching the surface of that material gives one the impression that the Mla Bri have a distinctive spiritual culture of their own. Still, the vocabulary of ritual texts is sometimes strangely infected by words that are hardly old in Mlabri. This is true of the very first sentence used in invocation of one's own personal spirit: *ku:k* ^h^*ma:l mu:*, literally something like: 'call spirit return'. The first word is a Lao word shared with Kui but not otherwise used in Mlabri; in MlaB it is invariably spoken in a (subdued) falsetto voice, apparently imitating a call. The third word is likewise a Lao word; in MlaC it is used in daily

speech as the word for 'return (home); go home' whereas all other Mlabri say *wʌl* (in MlaB also *brarul^h*) in that general meaning. Another example is the elaborated MlaA expression *gʌm jak gulgut wɤːk gulgut naːm*, literally 'don't go spring water spring water', a warning against the dangerous place where a well springs. Here, *wɤːk* is the Mlabri word for 'water', and *naːm* is its Tai equivalent, which is not otherwise used.

The basic cosmology and some rituals are shared by all three Mla Bri groups but I have no indication so far that the general narrative tradition is shared. Fairy-tales in the A-group seem to be fairly short and deal with concrete events. The fairy-tales in the B-group, which all belong to the repertory of one (now deceased) man, are on the contrary long and eventful and more or less phantastic or grotesque (I have retrieved no such material from the C-group). It is not my impression from talking to people with familiarity with Khmu and Tin narratives that there is any particular similarity.

The Mla Bri may have brought their spiritual tradition with them from elsewhere; e.g. their myths about cataclysms and resurrection of life on earth through the intervention of a cultural hero are too specific for it to be likely that it evolved over a short period of time. Maybe we shall never know their origin.

Appendix I:
Mlabri words containing the vowel /ʌ/

This list contains all the Mlabri words with short or long ʌ [ʌ, ʌ:] that I have retrieved in three varieties of Mlabri. There are 154 entries.

Loanwords from Tai are included; the occurrence of the vowel ʌ in a few loanwords in which one expects *a* is in itself interesting.

The order of arrangement is according to the final consonant. Under each final, words are listed according the first consonant of their main syllable, then according to presyllabic material, and finally according to consonantal material after the first consonant of the main syllable.

It will be seen that the vowel ʌ occurs before all possible finals in Mlabri (before -*j* only long *ʌ:* occurs in the data, however). As for onsets, the vowel occurs after all oral obstruents whereas there are some accidental holes in the occurrence after sonorants.

Occasionally, the vowel ʌ varies with *a(:)*; this is true of the MlaC word *trʌl* below which is the same etymon as *tra:l* 'bright' in MlaA (i.e., the MlaC entry means '(I) look (and it is) not bright').

The coverage is limited to occurrences of the vowel *ʌ(:)* as a syllable nucleus by itself. Diphthongs containing a vowel of this phonetic quality (i.e. *iʌ, iʌ:, uʌ, uʌ:*) are disregarded because the vowel has a quite different status and a different history as a component of those diphthongs. The MlaC word *pabjʌ:c* below is marginal in this respect since it would be *pabiʌ:c* in some other Mlabri; the word *jjʌl* also occurs in the pronunciation *jiʌl*.

ʰ*mʌ*- MlaAB: component of kinship terms
jʌ: MlaA(C) 'which?'

tʰʌp 'roast; bake'
pʌp MlaB 'break'
pʌ:p MlaC 'rest; relax'
takʌp MlaC 'weave'
çikʌp MlaC 'weave a Lao emblem'
krʌp MlaAC '(for an insect to) bite'
grʌp MlaA 'cut off (in a biting manner)'
jipjʌp/jiupjʌp MlaAC 'scissors'

gɲrʌp MlaA 'scissors'
lʌp 'put'

tʰʌt MlaC in *cuʌk tʰʌt* 'dig deeply'
cʌt MlaBC 'crouch; nest'
bʌt MlaABC 'scoop up'
ɟʌt MlaB 'while'
ʰŋʌt 'for the rain to stop'
rmɕʌt MlaB in *rmɕʌt toc* 'pick up'
wʌt 'tie; bind around'
tunwʌt/tɳwʌt 'lashing'
kutʔwʌt MlaC 'bamboo species'

tʰʌc MlaB 'flesh; pulp'
pabjʌːc MlaC 'cook until it is soft'
mʌc 'see; know'
mʌc MlaA: a personal name
ʰmʌc MlaC 'sticky; sticky rice'
rʌc MlaB 'bird species'
ruɕrʌc MlaB 'sift; sickle down'

pʌk MlaB 'be on a hilltop'
gmpʌk MlaB 'perform a string ritual'
plʌk MlaAB 'hiccough'
tʌk MlaBC 'put in resting position'
tʌːk MlaB 'pour'
ɟrʌːk MlaB 'drink; water'
mʌːk MlaB 'be happy'
ɲrʌk MlaB in expression about teasing
ɲrʌːk MlaB 'complete a trajectory'
tʰarʌk MlaA 'slide downhill'
jʌk MlaAB 'thus; like this'
rukʔjʌk MlaC '(sit) on one's heels'

tuʔtʌʔ MlaA 'open mountain slope'
crʌʔ MlaA 'bird sp.: great hornbill'
drʌːʔ MlaA 'belch'
digʌʔ MlaA 'over there'
ɕugwʌʔ 'down there'
mʌʔ MlaB in *mʌʔ lɛʔ* 'malleolus(?)'

mʌʔ MlaA in *luʔ mʌʔ* 'delirious'
mʌːʔ 'creeper (vine) species; python'
ɲʌʔ 'that one; there'
dɲɲʌʔ MlaB '(to) that place'
gɲɲʌːʔ MlaA 'chew'
lɱɲʌʔ MlaC 'just over there'
ɕʌʔ in *grok ɕʌʔ* MlaB 'the core family'
ɕʌʔ MlaA: postnominal expressive
ɕʌːʔ Mla B 'that's it!'; 'that's it??'
jʌʔ MlaA: component of kinship terms

rɱpʌm MlaC 'butterfly'
plʌm 'lump of bees' wax'
tʌm MlaBC in *mak tʌm* 'button'
dʌm 'brain'
dʌm MlaAB 'continuously'
ɕʌm 'stumble'
ɕʌm MlaA 'again'
kʌmʌm MlaC 'dumb; cannot speak'
ʰnʌm 'long time; a year'
tʰalʌm MlaB 'clench one's fists'

pʰjʌn (sic!) MlaA 'friend'
ɓʌn MlaA 'thick (layer)'
ɲʌn (B also *ɲɔn*) MlaBC 'because'
ʰmʌn MlaC: kinship term

tʌɲ MlaAC 'speak'
kətʌɲ MlaA 'however'
cʌɲ 'tooth; sharp edge'
pɾgʌɲ 'whet (a knife); polish'
tɱnʌɲ MlaAC 'talking'

pjʌŋ MlaC 'size' (cf. *jʌŋ*)
tŋtʌŋ MlaB 'ant species'
kwʌŋ MlaBC 'something round' (cf. *ḵlwʌŋ*)
jʌŋ MlaA 'sift by shaking'
ɲʌŋ MlaC 'safely at home'
ʰmʌŋ MlaA 'pouring blood; a bleeding'
ɲɲrʌŋ MlaA 'house pole; tripod'

dɲrʌŋ MlaB 'house pole'

kɽwʌŋ MlaB in *kɽwʌŋ juːt* 'elastic band'

kḷwʌŋ MlaAB 'something round' (cf. *kwʌŋ*)

jʌŋ MlaC 'size' (cf. *pjʌŋ*)

jʌ(ː)ŋ MlaA in *jak jʌŋ* 'go to the market'

r²ʌh MlaB 'soon'

crʌh MlaC 'clear one's throat'

təkʌh MlaB 'over there, far away'

çikʌh MlaC 'up there(?)'

lakʌh MlaB 'up there (in the air)'

blʌh MlaAC 'split wood, v'

jrʌh MlaB 'clear the throat' (cf. *jərʌh*)

gʌh 'this one; here'

tʰugʌh/tʰugʌh MlaA 'up/down to here'

təgʌh/tugʌh MlaAC 'over there'

dəgʌh MlaAC 'around here'

ʰŋgʌh MlaC 'this; that'

çugwʌh MlaAC 'down on the other side'

ɲʌh MlaA 'there'

jərʌh (also: *jərah*) MlaA 'cough' (cf. *jrʌh*)

pʌr 'fly, v'

tʌr MlaAC 'down(wards)'

cʌr 'fruit species'

kʌr-kɛr MlaB in *jraː kʌr-kɛr* 'very skinny'

kwʌr term for non-Mlabri ethnic groups

bʌr MlaB: *kiʔ a bʌr* 'full moon'

bʌr (~ *bɽ-*) 'property of – '

tɽbʌr MlaC 'cigarette'

jʌr MlaBC 'fly through the air'

ɓʌr MlaC 'size'

çʌr MlaBC 'stand upright; rise' (cf. *jʌr*)

mʌːr 'crawl; creep'

ʰnʌr MlaB 'wing'

lʌːr 'swallow, v'

tawʌr MlaC 'tortoise species'

jʌr 'rise up; stand up' (cf. *çʌr*)

jʌr in *het jʌr* MlaA 'mushroom species'

dm̩pʌl MlaB 'tree species'
dm̩pʌl MlaB 'soldier'
tʌl MlaB 'let something slide along'
gr̩tʌl MlaAB 'upper arm (A), elbow (B)'
gltʌl MlaC 'elbow; heel'
trʌl MlaC in *dʒŋ met trʌl* 'cannot see'
ç̩rkʌl MlaB 'for water to be muddy'
dʌl 'flat end'
jjʌl 'dusk'
wʌl MlaAB 'return to home base'
kʰʌʔjʌl MlaA 'stamp with the feet'

plʌlʰ MlaAB 'flick (away)'
cʌlʰ MlaBC 'go off for a purpose'
kʌlʰ MlaB 'fall apart'
dukʌlʰ MlaBC 'sneeze'
dʌlʰ MlaB 'matured'
dʌlʰ MlaBC 'amount'
mʌːlʰ MlaBC 'name; term'
tr̩nʌlʰ MlaAB 'collar bone'
tr̩nʌlʰ MlaB 'house pole'
tl̩ʰnʌlʰ MlaC 'oblique rafter in lean-to'
kr̩lʌlʰ MlaC: 'crooked pin on crossbow'

tʰʌw MlaC 'roast; bake' (cf. *tʰʌp*)
pʌw MlaC 'blow at something'
pupʌw MlaC 'used up; all gone'
krʌw 'feel scared'
gʌw in *b(l)aːj gʌw* 'big'
palʌw MlaC 'roll something'
jʌw MlaB 'like (this); because'

pʌːj MlaB 'edible bamboo shoots'
kujkwʌj MlaB 'wrap'
gŋ̍jʌj MlaA in *gŋ̍jʌj jʌːɲ* 'chew'

jʌjʰ MlaAC 'delicious'
luwjʰgwʌːjʰ '(a branch) sticking up in the air'
gljʌjʰ MlaB 'turn'

Appendix II:
Transcribing Mlabri
and other Mon-Khmer languages

The transcriptions in this monograph are in a broad version of the transcription system of the International Phonetic Association: the IPA notation. All transcriptions are in italics to make them stand out from ordinary text, also when I cite from sources that do not use italics for that purpose (the only potential problem with that is that the two "a"-qualities [a] and [ɑ] come out the same way in my IPA italics font: *a*; I have not found that that could cause any confusion since I nowhere cite forms with an /a/ that contrasts with /ɑ/, or vice versa). Occasionally, more detailed phonetic notation is used, and then in ordinary type in square brackets.

In rendering forms from different Mon-Khmer languages I expose myself to criticism by subjecting different notational practices to standardization. I use a broad phonetic notation and attempt to use it in uniform manner in order for the quoted forms to be easy to compare phonetically.

In transcribing consonants in Mlabri and Tin I use raised symbols to show concomitant articulatory features such as aspiration or partial devoicing (*ʰn = n̥n, pʰ = ph, jʰ = j̊j*, etc.). In my analysis, these count synchronically as single phonemes. As for all other M-K languages, however, I write full letters to represent pre-or postaspiration in order to avoid misrepresenting clusters as single phonemes. As for all other M-K languages, including Khmu, there are different ways of representing aspirated stops but I write them as *ph*, *th*, etc. because they count morphologically as clusters (in some languages, infixation can happen between an intial stop and /h/, for example, which is not the case in Mlabri or Tin).

I mark the glottal narrowing or closure ("glottal catch") before syllable-initial vowels by a raised symbol for glottal stop: *ˀV*, also when citing forms from sources that put a full symbol before the vowel: /ʔV/. This is meant to be a hopefully acceptable compromise (in M-K sources one encounters a variety of usages: some use a full stop symbol: /ʔV/, others put no symbol in front of initial vowels: /V/, and still others have a hyphen before a vowel in hiatus: /V'V/ but not elsewhere). I have felt

uncomfortable, however, with such transliteration if the source has sesquisyllabic wordforms of the structure /C?V(C)/, but I have chosen to render it as $C^?V(C)$ if the source otherwise renders word-initial vowels as /?V/ (if the opposite applies, such a form should be interpreted as consisting of a presyllable /C?/ [C°?] plus a main syllable with an initial vowel /V(C)/, i.e. $C?^?V$ in my notation).

Final glottal stops are probably written with a full letter by all authors writing on M-K; some use /q/ for the glottal stop (I myself do that in my Mlabri romanization). In this monograph I use the notation ? throughout in final position, transliterating /q/ as ?. The sequence $V:?$ occurs in Mlabri; phonetically it is realized as a long vowel with laryngealization (much like the Danish stød on long vowels followed by a sonorant) but structurally it is a long vowel plus /?/.

The languages under consideration here have a distinction between what can be broadly defined as labial versus dental (alveolar) versus (alveolo)palatal versus velar. The most pervasive problem in transcribing consonants in Mon-Khmer languages is the representation of the palatal series. There is no problem with the nasal; it is transcribed /ɲ/ or /ñ/ in the sources, and I generalize the symbol ɲ in the forms I cite. The palatal glide and the corresponding stop are often rendered as /y/ vs. /j/; I normalize such transcriptions to j vs. $ɟ$ in accordance with IPA practice.As for the voiced and unaspirated voiceless palatal stops, many authors make a point of stating that the point of articulation is alveolopatal or palatoalveolar rather than palatal, and that they are affricates. At least for Mlabri I find that a broad transcription $ɟ$, c is less confusing than the use of complex symbols which exaggerate the fricative component. As for aspirates and sibilants with a palatal component there is one such entity in Mlabri which I have here chosen to render as $ç$; its phonetic realizations range all the way from a retracted sibilant [s̠] to an alveolopalatal affricate, which some would render as [tɕʰ] (in previous work I have rendered it as /ch/ because it can be seen as patterning with aspirated stops). Tin likewise has a variable affricate with a predominant sibilant component which can vary from alveolar to alveolopalatal; since there is no contrast with a plain sibilant I render this phoneme as s. Occasionally I use the symbol $ç$ because some speakers of the conservative MalA dialect jointly insist that certain words have an alveolopalatal, *not* an alveolar affricate in their speech (I do not know how to interpret that, however, since it does not seem to be consistent over the speech community; there is a need for much more lexical data on this particular dialect).

In my Mlabri transcriptions, there are numerous instances of presyllables in which a sonorous consonant is marked as syllabic by means of a vertical stroke under the consonant symbol. Words given in that form have phonetic realizations with a more or less audible epenthetic (parasitic) vowel before the sonorant consonant, so notations with a syllabicity mark are in fact broad notations. The vocalic realizations are highly variable across the three varieties of Mlabri, and there are often individual preferences or even idiosyncracies in the realizations. Some words have a full, non-predictable vowel in the first syllable if spoken distinctly (at least in the pronunciation of some speakers) but it may reduce to a parasitic [ə]-vowel in more casual speech. I have not found it practical to render this mass of variation consistently here (in a forthcoming Mlabri dictionary I go into detail with vowel variation in presyllables). What is important from a comparative perspective is to indicate syllabicity as such because it is distinctive in Mlabri.

This insistence on stating syllabicity is at variance with the practice of perhaps most scholars writing on Mon-Khmer languages, who in their broad notation represent word-forms with presyllables as if they were monosyllables with an initial cluster. There are often very good reasons for the latter usage, either because the syllabicity is predictable from the nature of the cluster or because there is variation between syllabic and non-syllabic realizations in the language in question. Kui (Katuic), for example, has no rigid distinction between words with an initial cluster and sequisyllabic words, and there is variation over Kui dialects as to whether certain preconsonantal consonants are syllabic or not (Theraphan and Gainey in Sriwises 1978; ix and xxii-xxiii). In citing forms from languages other than Mlabri and Tin I use broad notation if that is available in dictionaries and other reference works.

As for presyllables in Mlabri which do *not* contain a sonorant, I always transcribe them with a vowel after the initial, choosing the vowel quality that seems most widespread among speakers. Other languages are transcribed in broad notation if that is given in the sources, although it occasionally leads to under-differentiation. For example, Suwilai Prensrirat in her *Dictionary of Khmu in Laos* (2002): 284 has two apparently homophonous entries **ské:p**[1] and **ské:p**[2] with different meanings, but if one consults her fine phonetic notation they have vowels of different quality between the first two consonants: [sɨkĕ:p] vs. [sakĕ:p]. My position is that if experts on particular languages can live with such occasional under-differentiation in their broad notations, so can I when citing forms from those languages in order to throw light

on Mlabri-Tin comparisons. There is no single instance in the present monograph where the vowel quality of presyllables in other languages is crucial.

As for vowels in main syllables, the most important question is what happens in the front-back dimension and how to represent it in phonetic notation. In main syllables, Mon-Khmer languages have unrounded front vowels and rounded back vowels, and then there is a category "in between". It is most customary to represent non-low vowels of the third series as unrounded central vowels: /ɨ, ə/, and some authors argue explicitly that they are central. Some, however, represent the high vowel as back: /ɯ/, or that representation is extended to both vowels: /ɯ, ɤ/. If there is a phoneme between [ɤ/ə] and [a] it is mostly rendered as /ʌ/, which in IPA means a back vowel but is often taken to mean a central vowel. Transcriptions of the Katuic language Brôu use the symbol /â/ to denote the vowel quality [ʌ]; I transliterate it accordingly as ʌ.

The unrounded, non-front vowel series in Mlabri is back not central, and thus I render it as ɯ, ɤ, ʌ. I also hear the high vowel in Tin as back [ɯ], and I render it accordingly, whereas I am less certain about the non-high vowel: [ɤ] or [ə]. In this monograph I have chosen to render the latter vowel as ə when transcribing Tin forms.

My main principle in transliterating word-forms from sources on other Mon-Khmer languages is to follow the conventions of the literature I use as long as they are compatible with IPA-type phonetic notation. Otherwise I transform the symbols into IPA, at the risk of introducing phonetic inaccuracies (since I engage in detailed phonetic-phonological comparison only when it comes to the Mlabri-Tin connection, that is hardly a serious problem). Khmu has a diphthong which some render as /ia/ others as /iə/ or /ɯə/; I use the notation iə. As for the unrounded, non-front vowels many authors writing on languages in Vietnam used "hooked u" and "hooked o", which I transliterate as ɯ and ɤ. In transcriptions of languages spoken in Vietnam one finds vowel aperture distinctions such as /ê/ versus /e/, /ô/ versus /o/; I transliterate those into the IPA-counterparts e versus ɛ, o versus ɔ.

Vowel-length is very often rendered by doubling the vowel symbol; that notation is here transformed into use of the length mark (a: for /aa/). In some transcription systems used for languages in Vietnam it is the short vowels that are marked (by a breve mark) and the long ones that are unmarked. Here, as well, a conversion into the IPA usage is made (a for /ă/, a: for /a/). I cite some forms from Phong and from Palaung in which a distinction is likewise made between vowels with

and without the breve mark; if a main syllable ends in an unmarked vowel I permit myself to add a length mark (*ka:* for *ka*, etc.) so as to facilitate comparisons, whereas I leave the formulatic reconstructions of vowels in proto-languages as they are. Diacritic marks (breve, acute, etc.) used in transcriptions of Katuic languages with their rich vowel inventories are particularly complicated for outsiders (both aperture and vowel shortness are involved); I have tried to avoid citing such forms with diacritics.

Some Mon-Khmer languages have live voice quality contrasts ("register"), which may appear in notations as diacritics on vowels (e.g. grave accent); I have rarely had occasion to cite such forms and I then preserve the accent mark, with or without explanation. Then again some M-K languages are tonal. Tonal distinctions have no direct relevance to my study since there are generally reconstructions of pre-tonal stages available (I do not cite forms from Vietnamese). Khmu has tonal and non-tonal dialects; although the current research focusses much on tonal Khmu I consistently cite forms in the conservative, non-tonal type of Khmu (preserving of old differences in the voicing state of initial consonants), which are more easily compared with Mlabri and Tin forms.

In word-forms of more than one syllable stress falls predictably on the main syllable, i.e. the last syllable, all across Mon-Khmer. This is not indicated in the notation of such word-forms. Mlabri has a few words which some speakers tend to pronounce with high prominence on a non-final syllable, even on a presyllable with a syllabic consonant (e.g. *kṇ* in *kṇde:p* 'centipede'). This feature is disregarded in the present comparative work since I cannot make any useful generalizations about it.

I wish to round this section off by briefly mentioning that a variety of other conventions for transcribing Mlabri have been used or are now available.

In work from the 1980es both my deceased colleague Søren Egerod and I used a hybrid IPA-notation in which the vowels *ɯ, ɤ* were rendered as *ï, ë*. In my more recent work on a comprehensive Mlabri dictionary I use a typographically easy but phonologically adequate romanization with symbols such as *ï, ë, ä* for central vowels, *è, ò* for low mid vowels, *q* for final glottal stop, and digraphs for some of the consonants. Syllabicity is marked consistently by placing a dot at all syllable boundaries. Beside that romanization I have proposed a Thai-script based Mlabri orthography, which may or may not come into use.

Bibliography

The selection of entries in this bibliography requires comment because it goes far beyond a listing of references from the text, and at the same time falls short of a complete bibliography of publications relevant to the Mlabri–Tin issue. Needless to say, the emphasis of the bibliography is on the descriptive and comparative study of Mon-Khmer languages and on Mlabri in particular.

As for the Mla Bri ethnic group and their language, publications that represent the early research history up to the Second World War are listed rather generously. From the last fifty years, however, there are so many publications containing some kind of information on this ethnic group that only linguistically relevant literature and literature that presents extensive demographical data (be it reliable or perhaps less reliable) is included.

As for Tin (Mal and Prai) and Kammu, the bibliography includes items that are relevant to the study of lexicon or language history, or which give important information on demography. The only extensive word lists of Tin available to me, except for my own field notes, are unpublished dictionary files worked out several years ago by Dr. David Filbeck and the Rev. David Jordan. As for Kammu, there are excellent, published dictionaries, which are entered below.

As for other languages relevant to the present study, a number of handbooks and papers with a comparative linguistic scope are included. Besides, several word lists and dictionaries have been consulted during the preparation of this monograph. The majority of these tools were published by SIL (Summer Institute of Linguistics), Chulalongkorn University and the University of Mahidol in Thailand. Such standard literature is as a rule *not* included in the bibliography unless it contains detailed information on the structure of the target language.

(JSS = Journal of the Siam Society, Bangkok)

Anderson, Gregory D. S. 2004. "Recent advances in Proto-Munda reconstruction". *Mon-Khmer Studies* 34: 159-184.

Anderson, Gregory D. S. 2006. "Proto-Munda and Proto-Austroasiatic: problems in diachronic linguistics". Handout for a paper given at the workshop on Historical linguistics and hunter-gatherer populations in global perspective. Leipzig: MPI-EVA, August 2006.

Bauer, Christian. 1990. "Reanalyzing reanalyses in Katuic and Bahnaric", *Mon-Khmer Studies* 16-17: 143-154.

Benjamin, Geoffrey. 1976. "Austroasiatic subgroupings and prehistory in the Malay Peninsula", *Austroasiatic Studies* I: 37-128. The University Press of Hawaii.

Bernatzik, Hugo A. (*unter Mitarbeit von* Emmy Bernatzik), *Die Geister der gelben Blätter: Forschungsreisen in Hinterindien* (F. Brugmann Verlag, München, 1938; 2nd ed., Koehler und Voigtländer Verlag, Leipzig, 1941).

Bernatzik, Hugo A. 2005. *The Spirits of the Yellow Leaves. The Enigmatic Hunter-Gatherers of Northern Thailand.* (Originally 1938 and 1958; new edition by Jørgen Rischel). Bangkok: White Lotus Press.

Boeles, J. J. 1963. "Second Expedition to the Mrabri of North Thailand", *JSS* 51,2: 133-160.

Bourke-Borrowes, D. 1926-27. "Further Notes on the Phi Tong Lu'ang", *JSS* 20,2: 167-168.

Burenhult, Niclas. 2002. *A Grammar of Jahai.* Lund: Department of Linguistics and Phonetics, Lund University.

Chazée, Laurent, *The Mrabri in Laos. A World Under the Canopy* (White Lotus Press, Bangkok, 2001).

Davidson, Jeremy H.C.S., ed. 1991. *Austroasiatic Languages: Essays in Honour of H.L. Shorto.* London: School of Oriental and African Studies.

Daladier, Anne. 2002. "Kinship terms grammaticalized as Classifiers in *K(V)- for 'Animate' nouns in Austroasiatic and renewed again as Gender Pronoun-Determiners in War-Khāsic or Animate, Non-Animate, Dual suffixes in Munda". Paper given at the BLS 28 Conference, Feb. 15, 2002.

Dessaint, William Y. 1973. "The Mal of Thailand and Laos", *Bulletin of the International Committee on Urgent Anthropological and Ethnographic Research* 15: 9-25.

Dessaint, William Y. 1981. "The T'in (Mal), dry rice cultivators of Northern Thailand", *JSS* 69: 107-137.

Diffloth, Gérard. 1976a. *Proto-Mon-Khmer Final Spirants.* Kyoto: Kyoto University Center of Southeast Asian Studies.

Diffloth, Gérard. 1976b. "Expressives in Semai", *Austroasiatic Studies* I: 229-247. The University Press of Hawaii.

Diffloth, Gérard. 1977. "Mon-Khmer initial palatals and 'substratumized' Austro-Thai", *Mon-Khmer Studies* VI: 39-57.

Diffloth, Gérard. 1979. "Expressive phonology and prosaic phonology in Mon-Khmer". In Theraphan L. Tungkum, Pranee Kullavanijaya, Vichin Panupong and M.R. Kalaya Tingsabadh eds.: *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie J.A. Henderson*: 49-59.

Diffloth, Gérard. 1980. *The Wa Languages.* Linguistics of the Tibeto-Burman Area Vol. 5, Number 2.

Diffloth, Gérard. 1984. *The Dvaravati Old Mon Language and Nyah Kur.* Bangkok: Chulalongkorn Printing House.

Diffloth, Gérard. 1991. "Palaungic vowels in Mon-Khmer perspective", in Davidson ed.: 13-27.

Downer, Gordon B. 1992. "The Thai element in Khmu'", *Mon-Khmer Studies* 18-19: 44-51.

Egerod, Søren. 1982. "An English-Mlabri basic vocabulary", *Annual Newsletter of the Scandinavian Institute of Asian Studies* 16: 14-20.

Egerod, Søren and Jørgen Rischel. 1987. "A Mlabri-English Dictionary", *Acta Orientalia* 48: 35-88.

Elliott, Julia and members of the Lua' community. 1992. "The Lua' (T'in): remembering life in Laos", in: Lewis, ed.: 73-123.

Ferlus, Michel. 1964. Word list from a Mlabri dialect, taken down in Laos in 1964 (unpublished).

Ferlus, Michel. 1974. "Les langues du groupe austroasiatique-nord", *ASEMI* V.1: 39-68. Paris: CNRS.

Ferlus, Michel. 1977. "L'infixe instrumental RN en Khamou et sa trace en Viet-namien", *Cahiers de Linguistique Asie Orientale* 2: 51-55.

Ferlus, Michel. 1978. "Reconstruction de /TS/ et /TŠ/ en Mon-Khmer". *Mon-Khmer Studies* 7: 1-38.

Ferlus, Michel. 1980. "Formation des registres et mutations consonantiques dans les langues Mon-Khmer", *Mon-Khmer Studies* 8: 1-76.

Filbeck, David. 1975. "A Grammar of Verb Serialization", in Harris and Cham-berlain 1975, p. 112-129.

Filbeck, David L. 1976. "On */r/ in T'in", *Austroasiatic Studies* I: 265-283. The University Press of Hawaii.

Filbeck, David. 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", *JSS* 75: 129-138.

Filbeck, David 1991: "Keeping things up front: Aspects of information pro-cessing in Mal discourse structure", in Davidson ed.: 161-182.

Filbeck, David. unpubl. *Mal dictionary files.*

Flatz, Gebhard. 1963. "The Mrabri: anthropometric, genetic, and medical examinations", *JSS* 51,2: 161-177 plus Appendices.

Haudricourt, André G. 1965. "Mutation consonantique en mon-khmer", *BSL* 60: 160-172.

Haupers, Lorraine and Ralph Haupers. 1991. *Stieng-English Dictionary.* Ma-nila: Summer Institute of Linguistics.

Huffmann, Franklin E. 1976a. "The register problem in fifteen Mon-Khmer lan-guages", *Austroasiatic Studies* I: 539-574. The University Press of Hawaii.

Huffmann, Franklin E. 1976b. "The relevance of lexicostatistics to Mon-Khmer languages", *Austroasiatic Studies* I: 575-589. The University Press of Ha-waii.

Jacob, Judith M. 1963. "Prefixation and infixation in Old Mon, Old Khmer, and Modern Khmer". *Linguistic Comparison in South-East Asia and the Pacific*: 62-70. London: School of Oriental and African Studies.

Janzen, Hermann. 1976a. "The system of verb-aspect words in Pale", *Austro-asiatic Studies* I: 659-667. The University Press of Hawaii.

Janzen, Hermann. 1976b. "Structure and function of clauses and phrases in Pale", *Austroasiatic Studies* I: 669-691. The University Press of Hawaii.

Jordan, David. unpubl. *Prai dictionary files.*

Kingsada, Thongpheth and Ryuichi Kosaka. 1999a. "Khabit", in: Kingsada, Thongpheth and Tadahiko Shintani, eds.: 297-317.

Kingsada, Thongpheth and Ryuichi Kosaka. 1999b. "Khmu", in: Kingsada, Thongpheth and Tadahiko Shintani, eds.: 318-338.

Kingsada, Thongpheth and Tadahiko Shintani. 1999. "Lü", in: Kingsada, Thongpheth and Tadahiko Shintani, eds.: 1-21.

Kingsada, Thongpheth and Tadahiko Shintani, eds. 1999. *Basic Vocabularies of the Languages spoken in Phongxalt, Lao P. D. R.* = Lao-Japan Project, Data Paper 1. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA).

Kraisri Nimmanhaeminda. 1963. "The Mrabri Language", *JSS* 51,2: 179-184 plus Appendices (1963).

Lewis, Judy ed. 1992. *Minority Cultures of Laos: Kammu, Lua', Lahu, Hmong, and Iu-Mien*. Folsom Cordova Unified School District, Rancho Cordova (California): Southeast Asia Community Resource Center.

Lindell, Kristina. 1974. "A vocabulary of the Yùan dialect of the Kammu language", *Acta Orientalia* 39: 191-207.

Lindell, Kristina, Jan-Olof Svantesson and Damrong Tayanin. 1978. "Two dialects of the Rmeet (Lamet) language", *Cahiers de Linguistique Asie Orientale* 4: 5-22.

Lindell, Kristina, Jan-Olof Svantesson and Damrong Tayanin. 1980. "Phonology of Kammu dialects", *Cahiers de Linguistique Asie Orientale* 7: 59-67.

Lindell, Kristina, Jan-Olof Svantesson and Damrong Tayanin. 1981. "Phonology of Kammu dialects", *Cahiers de Linguistique Asie Orientale* 9: 5-22.

Matisoff, James A. 1991. *Endangered Languages of Mainland Southeast Asia*. In Robins & Uhlenbeck, eds.: 189-228.

Mingkwan Malapol. 1989. *Pray Grammar at Ban Pae Klang, Thung Chang District, Nan Province*. Unpublished MA-thesis. Mahidol University.

Mitani, Yasuyuki. 1979. "Vowel correspondences between Riang and Palaung", in Theraphan L. Tungkum, Pranee Kullavanijaya, Vichin Panupong and M.R. Kalaya Tingsabadh eds.: *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie J.A. Henderson*: 142-150.

Oota, H., B. Pakendorf, G Weiss, A. von Haeseler, S. Pookajorn et al. 2005. "Recent origin and cultural reversion of a hunter-gatherer group", *PLoS Biol* 3(3): *e*71.

Paulsen, D. A. 1992. "Phonological reconstruction of Proto-Plang". *Mon-Khmer Studies* 18-19: 160-222.

Pejros, Ilia. 1996a. *Katuic Comparative Dictionary*. Pacific Linguistics C-132. Canberra: Australian National University.

Pejros, Ilia. 1996b. "The Vietnamese etymological dictionary and 'new' language families (Project in progress)", in: *Pan-Asiatic Linguistics, Proceedings of the Fourth International Symposium on Languages and Linguistics*, Volume III: 883-894. Salaya: Institute of Language and Culture for Rural Development, Mahidol University.

Phra Winit Wanadorn. 1926-27. "Some information concerning the Phi Tong Luang obtained from a few residents in a village in the Nam Wa District east of Nan", *JSS* 20,2: 171-174.

Pinnow, Heinz-Jürgen. 1959. *Versuch einter historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Otto Harrassowitz.

Pinnow, Heinz-Jürgen. 1963. "The position of the Munda languages within the Austroasiatic language family", *Linguistic Comparison in South-East Asia and the Pacific*: 140-152. London: School of Oriental and African Studies.

Polibyenko, T. G. and Buy Kxany Txe. 1990. *Jazyk Ksingmul*. In Solntsev et al. eds.: 9-419.

Prasert Sriwises. 1978. *Kui (Suai) – Thai – English Dictionary*. Bangkok: Indigenous Languages of Thailand Research Project, Chulalongkorn University Language Institute.

Premsrirat *see* Suwilai Premsrirat

Proschan, Frank.1992. "Ethnonymy, with special reference to the Kmhmu", paper for the Mon-Khmer Workshop at the International Conference on Sino-Tibetan Languages and Linguistics, University of Berkeley, 1992 (67 pp. plus appendix).

Proschan, Frank. 1996. "A survey of Khmuic and Palaungic languages in Laos and Vietnam", in: *Pan-Asiatic Linguistics, Proceedings of the Fourth International Symposium on Languages and Linguistics,* Volume III: 895-919. Salaya: Institute of Language and Culture for Rural Development, Mahidol University.

Rischel, J. 1982. "Fieldwork on the Mlabri language: a preliminary sketch of its phonetics", *Annual Report of the Institute of Phonetics, University of Copenhagen* 16: 247-255.

Rischel, J. 1989a. "Fifty years of research on the Mlabri language: A reappraisal of old and recent fieldwork data", *Acta Orientalia* 50: 49-78.

Rischel, J. 1989b. "Can the Khmuic component in Mlabri ('Phi Tong Luang') be identified as Old T'in?", *Acta Orientalia* 50: 79-115.

Rischel, J. 1989c. "What language do 'The Spirits of the Yellow Leaves' speak?: a case of conflicting lexical and phonological evidence", *Annual Report of the Institute of Phonetics, University of Copenhagen* 23: 87-118.

Rischel, J. 1992. "Isolation, contact, and lexical variation in a tribal setting". In Ernst H. Jahr, ed.: *Language Contact* = Trends in Linguistics, Studies and Monographs 60: 149-177. Mouton / de Gruyter.

Rischel, J. 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.

Rischel, J. 1995. *Minor Mlabri: A Hunter-Gatherer Language of Northern Indochina*. Copenhagen: Museum Tusculanum Publishers.

Rischel, J. 1997a. "Typology and reconstruction of numeral systems: The case of Austroasiatic". In Jacek Fisiak, ed.: *Linguistic Reconstruction and Typology*, pp. 273-312. Berlin – New York: Mouton de Gruyter.

Rischel, J. 1997b. *Language Contact and Language Maintenance in Two Low-Prestige Minority Groups of Northern Thailand*. Research Report submitted to DANIDA (of the Danish Ministry for Foreign Affairs). 92 pp. + Appendices. Copenhagen.

Rischel, Jørgen. 1999. "'Small' languages and their speakers", *European Review* 7,2: 191-218. Cambridge University Press.

Rischel, J. 2000. "The dialect of Bernatzik's (1938) 'Yumbri' refound?", *Mon-Khmer Studies* 30: 115-122.

Rischel, J. 2002. "Language and Ethnicity in and around the Golden Triangle", *The Royal Danish Academy of Sciences and Letters, Annual Report 2000-2001* (Det Kongelige Danske Videnskabernes Selskab: Oversigt over Selskabets Virksomhed 2000-2001): 248-249. Copenhagen: C. A. Reitzel.

Rischel, J. 2003a. "The role of a 'mixed' language in linguistic reconstruction", *Proceedings of the 17th International Congress of Linguists, Prague* (CD-ROM, ISBN 80-86732-21-5). Prague: Matfyzpress.

Rischel, J. 2003b. "Culture and language in a three-dimensional area", *Language and Culture = Copenhagen Studies in Language* 29: 11-27.

Rischel, Jørgen. 2004a. "In what sense is Mlabri a West Khmuic language?". Paper given at the 37th ICSTLL Mon-Khmer workshop on 29th September 2004, organized by Dept. of East Asian Languages, Lund University. A summary of this paper is on the net, arranged by the organizers.

Rischel, Jørgen. 2004b. "Pan-dialectal databases: Mlabri, an oral Mon-Khmer language". Paper given at the *Lexicography Conference* 24-26 May, 1984 at Payap University, Chiangmai (Thailand); subsequently made *available on the net* by the organizers Brian Migliazzi and Dough Cooper.

Rischel, J. 2005. Preface (p. xi), Introduction (p. xiii-xxxviii), Appendix (pp. 135-172, i.e. a detailed analysis of Bernatzik's Mlabri word list of 1938), and Selected Bibliography (pp. 173-175) to Bernatzik 2005.

Rischel, J. 2006. "The 'definite article' in Mlabri". *Mon-Khmer Studies* 36: 61-102.

Rischel, J. and S. Egerod. 1987. "'Yumbri' (Phi Tong Luang) and Mlabri", Acta Orientalia 48: 19-33.

Robins, R. H. and E. M. Uhlenbeck, eds. 1991. *Endangered Languages*. Oxford/New York: Berg Publishers.

Sakamoto, Hinako, ed. 2005. *Mlabri Text = ELPR (Endangered Languages of the Pacific Rim)* A3-17.

Sataporn Boonprasert. 1989. *A Study of Language Choice for Social Communication among Hmong, Khmu, and Prai at Ban Namsord, Thung Chang District, Nan Province*. Unpublished MA-thesis, Mahidol University.

Schumacher, W. W. 1969. "Zur Klassifizierung der Phi Tong Luang-Sprache. – Ein kommunikationsteoretischer Versuch", *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 22,4: 360-363.

Seidenfaden, Erik. 1919. "Further notes about Chaubun, etc.", *JSS* 13,3: 47-53.

Seidenfaden, Erik. 1927. "The Khā Tong Lu'ang", *JSS* 20,1: 41-48.

Shorto, Harry L. 1963. "The structural patterns of Northern Mon-Khmer languages". *Linguistic Comparison in South-East Asia and the Pacific*: 45-61. London: School of Oriental and African Studies.

Shorto, Harry L. 1976. "The vocalism of Proto-Mon-Khmer", *Austroasiatic Studies* II: 1041-1067. The University Press of Hawaii.

Sidwell, Paul. 2000. *Proto-South-Bahnaric*. Pacific Linguistics. Canberra: Australian National University.

Smalley, William A. 1961. *Outline of Khmu' Structure*. New Haven: American Oriental Society.

Smalley, William A. 1963. "Notes on Kraisri's and Bernatzik's word lists", *JSS* 51,2: 189-201.

Smalley, William A. 1992. Review of Suwilai Premsrirat 1987, *Mon-Khmer Studies* 18-18: 262-272.

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

Smith, Kenneth D. 1972. *A Phonological Reconstruction of Proto-North-Bahnaric* = Asian Pacific Series, Number 2. SIL. Santa Ana, California.

Smith, Kenneth D. 2000. *Sedang Dictionary*. Mon-Khmer Studies, Special Volume No. 1.

Solntsev, V. M. et al. eds. 1990. *Materialy sovetsko-vetnamskoi lingvisticheskoi ekspeditsii 1979 goda, Jazyk ksingmul*. Moscow: Nauk.

Sriwises *see* Prasert Sriwises

Suebsaeng, Nipatwet. 1992. "The Mlabri family and kinship system", Chapter 3 of Pookajorn, ed.

Sujaritlak Wajanarat. 1979. "Classifiers in Mal (Thin)", *Mon-Khmer Studies* 8: 295-303.

Surin Pookajorn, ed. 1992. *The Phi Tong Luang (Mlabri): a Hunter-Gatherer Group in Thailand*. Odeon Store Publishers, Bangkok.

Suwilai Premsrirat. 1987. *Khmu, a Minority Language of Thailand*. Pacific Linguistics A-75. Canberra: Australian National University.

Suwilai Premsrirat. 1991. "Aspects of inter-clausal relations in Khmu", in Davidson ed.: 123-139.

Suwilai Premsrirat. 2002a. *Dictionary of Khmu in Laos*. Salaya: Institute of Language and Culture for Rural Development, Mahidol University.

Suwilai Premsrirat. 2002b. *Thesaurus of Khmu Dialects in Southeast Asia*. Salaya: Institute of Language and Culture for Rural Development, Mahidol University.

Svantesson, Jan-Olof. 1983. *Kammu Phonology and Morphology* (= Travaux de l'Institut de Linguistique de Lund XVIII). Lund: CWK Gleerup.

Svantesson, Jan-Olof, Damrong Tayanin, and Kristina Lindell. 1994. *Kammu-Lao Dictionary* [in Lao]. Vientiane.

Svantesson, Jan-Olof, T. Tayanin, Kr. Lindell and H. Lundström. 2002. *Kammu-Yùan – English Dictionary, Preliminary Edition*.

Tayanin, Tamrong (Kàm Ràw) and Lue Vang. 1992. "From the village to the city: the changing life of the Kammu", in: Lewis, ed.: 1-72.

Thê´, Bùi Khánh. 2000. "The Phong language of the ethnic Phong which live near the Melhir Muong Pon Megalith in Laos (field work notes)", in: *Pan-Asiatic Linguistics, The Fifth International Symposium on Languages and Linguistics, Proceedings,* Vol.1: 199-253. National University Hochiminh City: University of Social Sciences & Humanities.

Thomas, David. 1964. "A survey of Austroasiatic and Mon-Khmer comparative studies", *Mon-Khmer Studies* 1: 149-163.

Thomas, David and Robert K. Headley, Jr. 1970. "More on Mon-Khmer sub-groupings", *Lingua* 25: 398-418.

Thomas, Dorothy. 1976. *A Phonological Reconstruction of Proto-East-Katuic* = *Working Papers* vol XX, Suppl. 4. Grand Forks, North Dakota: Summer Institute of Linguistics.

Thongkum, Theraphan L. 1983. "'To eat' in the Mlabri (Phi Tong Luang) language" (in Thai; mimeographed, 5pp.).

Thongkum, Theraphan L. 1984. *Bibliography of Minority Languages of Thailand*. Faculty of Arts, Chulalongkorn University.

Thongkum, Theraphan L. 1985. "Minority Languages of Thailand", *Science of Language Papers, Fasc. 5, August 1985: Languages and Dialects*: 29-69. Bangkok: Faculty of Arts, Chulalongkorn University.

Tongkum, Theraphan L. 1992. "The Language of the Mlabri (Phi Tong Luang)", Chapter 2 of Surin Pookajorn, ed. 1992.

Thongkum, Theraphan L. and Jerry W. Gainey. 1978. Editors' Introduction to Sriwises 1978: ii-xxvii.

Trier, Jesper. 1981. "The Khon Pa of Northern Thailand, an enigma", *Current Anthropology* 22.3: 291-293.

Unchalee Singnoi. 1988. *A Comparative Study of Pray and Mal Phonology*. Unpublished MA-thesis, Mahidol University.

Watson, Richard and Saundra Watson. 1979. *Pacoh Dictionary*. Manila: Summer Institute of Linguistics.

Zide, Arlene R. K. and Norman H. Zide. 1976. "Proto-Munda cultural vocabulary: Evidence for early agriculture", *Austroasiatic Studies* II: 1295-1334. The University Press of Hawaii.

Zide, Norman H. and Gregory D. S. Anderson. 2000. "Issues in Proto-Munda and Proto-Austroasiatic nominal derivation. The bimoraic constraint", *Proceedings of the SEAsia Linguistic Society Conference*. Madison.

Manuscript

General. – Manuscripts and illustrations must comply with the details given below. The original manuscript and illustrations plus one clear copy of both should be sent to the editor of the series. Manuscripts on digital media are accepted; contact the editor in advance, giving technical specifications.

A manuscript should not contain less than 48 printed pages. This also applies to the Mat.fys.Medd. where contributions to the history of science are welcome.

Language. – Manuscripts in Danish, English, German and French are accepted; in special cases other languages too. Linguistic revision may be made a condition of final acceptance.

Title. – Titles should be kept as short as possible, preferring words useful for indexing and information retrieval.

Abstract, Summary. – An abstract in English is required. It should be of 10-15 lines, outline main features, stress novel information and conclusions, and end with the author's name, title, and institutional and/or private postal address. – Papers in Danish must be provided with a summary in another language as agreed between author and editor.

Manuscript. – Page 1 should contain title, author's name and the name of the Academy. Page 2: Abstract, author's name and address. Page 3: Table of contents if necessary. Consult a recent issue of the series for general layout. Indicate the position of illustrations and tables. A printout must accompany manuscripts submitted electronically.

Figures. – All illustrations submitted must be marked with the author's name. It is important that the illustrations are of the highest possible quality. Foldout figures and tables should be avoided.

References. – In general, the editor expects all references to be formally consistent and in accordance with accepted practice within the particular field of research. Bibliographical references should be given in a way that avoids ambiguity.

Correspondence

Manuscripts are to be sent to

The Editor
Det Kongelige Danske Videnskabernes Selskab
H.C. Andersens Boulevard 35
DK-1553 Copenhagen V, Denmark
Tel: +45 33 43 53 00   Fax: +45 33 43 53 01
E-mail: kdvs@royalacademy.dk.
www.royalacademy.dk

Questions concerning subscription to the series should be directed to the Academy

Det Kongelige Danske Videnskabernes Selskab
The Royal Danish Academy of Sciences and Letters

Historisk-filosofiske Skrifter

Historisk-filosofiske Meddelelser

Priser ekskl. moms / Prices excl. VAT