AUSTROASIATIC LANGUAGES

Essays in honour of H.L. Shorto

Edited by J.H.C.S. Davidson

School of Oriental and African Studies
University of London
1991
Collected Papers in Oriental and African Studies

AUSTROASIATIC LANGUAGES
Essays in honour of H. L. Shorto

Edited by
Jeremy H. C. S. Davidson
Formerly Lecturer in Vietnamese
School of Oriental and African Studies

SCHOOL OF ORIENTAL AND AFRICAN STUDIES
UNIVERSITY OF LONDON
1991
CONTENTS

CONTRIBUTORS..................................................................................................................... vi

PREFACE.............................................................................................................................. vii

ACKNOWLEDGEMENTS.......................................................................................................... viii

H. L. SHORTO: A BIOGRAPHICAL NOTE J. H. C. S. Davidson................................. 1

PUBLICATIONS OF H. L. SHORTO Helen Cordell....................................................... 3

AUSTRIC: AN ‘EXTINCT’ PROTO-LANGUAGE Paul K. Benedict ......................... 7

PALAUNGIC VOWELS IN MON-KHMER PERSPECTIVE G. Diffloth........... 13

COMMUNICATIVES, EXISTIVES, AND STATIVES IN PROTO-SOUTH-BAHNARIC David Thomas................................................................. 29

A LINGUISTIC ANALYSIS OF SOME SOUTH MUNDA KINSHIP TERMS, I Norman H. Zide & Arlene R. K. Zide ........................................ 43

PROBLEMS AND PITFALLS IN THE PHONETIC INTERPRETATION OF KHASI ORTHOGRAPHY Eugenie J. A. Henderson ..................... 61

HU — A LANGUAGE WITH UNORTHODOX TONOGENESIS Jan-Olof Svantesson................................................................. 67

ON AUSTRONESIAN LEXICON IN VIETNAMESE Kenneth Gregerson ........ 81

SEVENTEENTH-CENTURY VIETNAMESE LEXICON: PRELIMINARY GLEANINGS FROM ALEXANDRE DE RHODES' WRITINGS Nguyễn Đình-Hoà ................................................................. 95

THE PHONOLOGY OF KOMPONG THOM CHAM Robert K. Headley ............. 105

ASPECTS OF INTER-CLAUSAL RELATIONS IN KHMU Suwilai Premsrirat ................................................................. 123

AN INSTRUMENTAL STUDY OF CHONG REGISTERS Theraphan L. Thongkum ................................................................. 141

KEEPING THINGS UP FRONT: ASPECTS OF INFORMATION PROCESSING IN MAL DISCOURSE STRUCTURE David Filbeck ............... 161

LES DÉRIVÉS DÉSIDÉRATIFS EN KHMER — Saveros Pou ......................... 183

A DIACHRONIC SURVEY OF SOME KHMER PARTICLES (7th to 17th CENTURIES) Judith M. Jacob ......................................................... 193

THE FORM syah IN ANGKORIAN KHMER Philip N. Jenner ......................... 227

OLD MON s- Christian Bauer ........................................................................ 241

Frontispiece — Professor H. L. Shorto
CONTRIBUTORS

J. H. C. S. Davidson, formerly Lecturer in Vietnamese, School of Oriental and African Studies, University of London

Helen Cordell, Sub-Librarian, School of Oriental and African Studies, University of London

P. K. Benedict, Adjunct Professor of Anthropology, Arizona State University, Tempe, Arizona, USA

G. Diffloth, Professor, Modern Languages, Cornell University, Ithaca, New York, USA

D. Thomas, Summer Institute of Linguistics, International Linguistics Center, Dallas, Texas, USA

N. H. Zide, Professor, Department of South Asian Languages and Civilizations, University of Chicago, Chicago, Illinois, USA

Arlene R. K. Zide, Department of Foreign Languages, Harold Walsh College, Chicago, Illinois, USA

†Eugénie J. A. Henderson, Professor Emeritus of Phonetics in the University of London

J.-O. Svantesson, Professor, Department of Linguistics and Phonetics, University of Lund, Lund, Sweden

K. J. Gregerson, President, Summer Institute of Linguistics, International Linguistics Center, Dallas, Texas, USA

Nguyễn Đình-Hoà, Professor, Vietnamese Studies Program, San Jose State University, San Jose, California, USA

R. K. Headley, Senior Research Linguist, US Government, Washington, DC, USA

Suwilai Premsrirat, Institute of Language and Culture for Rural Development, Mahidol University, Salaya, Nakhorn Pathom, Thailand

Theraphan L. Thongkum, Department of Linguistics, Chulalongkorn University, Bangkok, Thailand

D. Filbeck, Christian Mission to the Orient, Chiang Mai, Thailand

Saveros Pou, Directeur de Recherche, CNRS, Paris, France

Judith M. Jacob, formerly Senior Lecturer in Cambodian, School of Oriental and African Studies, University of London (rtd.)

P. N. Jenner, Director, Department of Indo-Pacific Languages, University of Hawaii, Honolulu, Hawaii, USA

C. Bauer, Lecturer in Linguistics and Mon, Institute of Language and Culture for Rural Development, Mahidol University, Salaya, Nakhorn Pathom, Thailand
PREFACE

The present volume is the fourth in this series to be dedicated to a retired scholar of the School. Here we honour Professor H. L. Shorto, whose standing in the field of Austroasiatic—especially Mon-Khmer—Studies is well reflected in the range and quality of the papers offered as a tribute in this volume.

The arrangement is, I hope, obvious—opening with a questioning paper by Paul Benedict, then moving from the languages of the rims of the area to focus on the Mon-Khmer heartland of Harry's major interest. The papers do not call for any comment from me as Editor, nor do their authors need introduction. Their contributions speak for themselves and will undoubtedly stimulate further research in the field of Austroasiatic Studies in coming years.

Jeremy H. C. S. DAVIDSON
ACKNOWLEDGEMENTS

I should like to thank all those who have helped make the task of editing these seventeen contributions lighter. In particular I wish to thank the Publications Committee (especially its Chairman, Professor Shackle, and its Secretary, Mr Martin Daly) for financial support in this publication of the volume. Thanks are due to the Editorial Secretary, Miss Diana Matias and to the SOAS photographer, Mr Paul Fox, for help with its production.

The photograph on the cover is by Hans Hinz and is reproduced by permission from Thai painting published by the Office du Livre, Fribourg.
HARRY LEONARD SHORTO: A BIOGRAPHICAL NOTE

Professor Harry Shorto is one of the founding scholars of Mon-Khmer studies, the field which his knowledge and insight helped to make the active and expanding area of scholarship that it is today.

It is now some forty five years since Harry Shorto completed his MA in Modern and Mediaeval Languages at Cambridge before moving to SOAS to take up a training appointment as Lecturer in Linguistics in 1948. From that point he was to pursue the study of the Austroasiatic and Austronesian language families of his choice: in 1952 he was appointed Lecturer in Mon; in 1964 Reader in the Languages and Literatures of South-East Asia; in 1971 to the Chair of Mon-Khmer Studies in the University of London. During this period he produced his two magisterial dictionaries, *A dictionary of Modern Spoken Mon* (1962) and *A dictionary of the Mon inscriptions from the sixth to the sixteenth centuries* (1971), to which he added numerous other stimulating publications listed in the bibliography of this volume.

Harry Shorto is a private man, given to weighing his words carefully between frequent re-lightings of his ever-present pipe. His Common Room colleagues may often have found awesome the ease with which he negotiated the depths and further reaches of his unfamiliar linguistic territory, but the learning was always leavened with humour, the delight in language for its own sake infectious.

The lightness of touch and passion for detail which Harry brought to the investigation of languages, their histories, constructions, and interconnections, are in complete harmony with his other absorbing interest: early musical instruments, in particular the flute, gamba, cornet and virginal. This interest extends to performance and he was for a time a member of a consort of viols. Today he continues to perfect his technique on his 1960 Morley virginal.

To Harry Shorto's many colleagues and former students it must be a source of great satisfaction that retirement has not put an end to his encyclopaedic researches, notably his project for a comparative phonology of the Mon-Khmer languages. We wish him well and offer this volume in appreciation of his contribution to scholarship over the years.

Jeremy H. C. S. DAVIDSON
PUBLICATIONS OF H. L. SHORTO
Helen Cordell


The devatā plaques of the Ananda basement. In, Essays offered to G. H. Luce by his colleagues and friends in honour of his seventy-fifth birthday, (eds.) Ba Shin et al.. Ascona: Artibus Asiae. 2, 156-65.


A dictionary of the Mon inscriptions from the sixth to the sixteenth centuries, incorporating materials collected by the late C. O. Blagden. London: Oxford University Press. (London oriental series; 24), xlii, 406.


1976 In defense of Austric. Computational Analyses of Asian and African Languages 6, 96-104.


1978 The planets, the days of the week and the points of the compass: orientation symbolism in ‘Burma’. In, *Natural symbols of South East Asia*, (ed.) G. B. Milner. London: School of Oriental and African Studies (Collected papers in Oriental and African studies), 152-64.


A leading Mon-Khmerist (Diffloth 1985) has recently raised the question: what happened to Austric? One can dilly-dally over this at great length but the answer is very brief: it became extinct.

The Austric hypothesis was developed shortly after the turn of the century by Wilhelm Schmidt (1906), a pioneer in the field of comparative Mon-Khmer linguistics. He had noted a general similarity in morphology, with even some resemblances in prefixes and infixes, between Austronesian (AN) and Mon-Khmer (MK), the latter grouped with the Munda languages of India to make up the Austroasiatic (AA) stock. True, he had found nothing in MK to compare with the elaborate suffixial apparatus of AN but, not surprisingly, he skipped lightly over this discrepancy. And what of the anticipated common stock of 'core' (basic) lexical elements? Linguists of that period, and well up into the middle of this century, laboured under the mistaken notion that lexical elements are pretty much the last place to look in setting up linguistic relationships; furthermore, in those pre-Swadesh days, that even when one does get around to inspecting these elements he need only apply an aphorism that Gertrude Stein might have expounded: a word is a word is a word. Henri Maspero, perhaps linguistically the most sophisticated of all the French sinologists of the first half of the century, predictably followed along these lines of thought in wrenching Vietnamese out of its native MK setting and misidentifying it as kindred to the Tai languages. It would have been equally predictable, given the circumstances, that a linguist such as Schmidt would seek to establish a genetic linkage between AN and AA. Thus it can be said that the birth of ‘Austric’ was expected; rather less expected has been a certain continuing enchantment with it on the part of some scholars, a reluctance to accept its demise. A review of the matter, as attempted in this paper, is thus in order.

The writer, when not busy with actual (Sino-Tibetan [ST], Austro-Tai [AT]) rather than fantasied language stocks, has from time to time given some consideration to the basic question here. He at first (Benedict 1942) en passant expressed a willingness to go along with Schmidt’s hypothesis on the assumption that continued research in that still largely unexplored field would turn up a respectable body of AN/MK cognate sets for ‘core’ lexical elements, thereby ‘fleshing out’ Austric, so to speak. Much later, after a sizeable amount of data in the field had become available, he read a paper on the subject at the First International Conference on Austroasiatic Linguistics at Honolulu (Benedict 1973). Here he played the ‘devil’s advocate’ role, first setting up as good a case as possible for
Austric, tying in both phonological and lexical evidence, before finally—and with some reluctance—reaching the conclusion:

AT and AA [the earlier levels of AN and MK] do not have a core vocabulary in common, despite the morphological similarity of the two language stocks, hence the idea of an ‘Austric’ superstock must be abandoned.

The homelands of both stocks must be assigned to adjacent areas in South-East Asia (SEA) and the similarities in overall patterning can be ascribed to areal factors. This corner of the Asiatic mainland has, in fact, over the past decades acquired a certain fame among linguists for the vast scope of its areal influences, involving even highly specific processes such as vocalic transfer (Benedict 1979), and there is no reason to suppose that similar influences were not also operative at an early period.

In the paper presented in 1973 the writer went rather further in suggesting that a handful of lexical similarities in basic vocabulary between AT and AA are indicative of an early relationship of substratum type between the two stocks. It now seems evident that he overplayed his ‘devil’s advocate’ role in this respect, paying insufficient attention to the alternative explanation in terms of ‘look-alikes’ or ‘comparabilia’ (Matisoff 1976), along with areal influences or ‘border phenomena’. Thus, the classical pair: PAN *mata (as earlier reconstructed); PMK *mat ‘eye’ are attention-getting, to be sure, but the syllable reduction is basically unmotivated (*d- as well as mono-syllabic roots occur in MK); PAN *maCa (with *C a cover symbol for reflexes representing proto-level palatal or consonant cluster), as now reconstructed by Austronesians, weakens the comparison (final *-c is a prominent feature of PMK) while PAT *mapra (for the earlier cited *mapla; see Benedict 1990) makes matters even worse: finally, PAA *mat, the likely reconstruction on the basis of the Munda cognates (see Pinnow 1959), takes us even further away from PAT *mapra at the early time level at which these proto-comparisons must be made. We can, in fact, do much better by comparing the PST root for ‘eye’: *mygk (as now reconstructed; see Benedict 1976a), with final *-k > *-t shift after the medial *-y- (lacking in AA), even obviating the problem of syllabic reduction!

The above is an excellent illustration of the useful linguistic rule of thumb: ‘look-alikes’ look less and less alike as we attain more precise reconstructions at earlier and earlier levels. Thus, the suggested ‘twin/two’ set of the Benedict (1973) paper, labelled as ‘doubtful’ even by Shorto (1976), must now be discarded in the light of recent MK evidence re the medial vocalism while the discussion there of the suggested set for ‘dog’ overlooked Monic final *-r, which vitiates the comparison. Both comparisons are also burdened with the problem of unmotivated syllabic

---

1. The recently reported Lai language of Southern China (Guǎngxi) has /bi/ for ‘two’, supporting a PMK reconstruction of *?biar or *?biar type, in line with Shorto’s suggested *bi?aar > *?bi(a)r rather than the *?baar cited in Benedict (1973). See Benedict forthcoming, n.6.
Austric: an 'extinct' proto-language

reduction, found also in 'eye' (above) as well as in a fourth set: PAN *busuk; PMK *sok *hair'.2 A fifth, often cited set: PAN *ikan; PMK *ka (or *ka') 'fish' has the same syllabic discrepancy along with 'problems' in both final and initial (the Munda languages point to *q- as the PAA initial, as reported in Benedict 1973). These difficulties led the writer to reject the comparison in 1973 and the present reconstruction and analysis of the PAN form (Benedict 1990) supports that decision: *śikan (for the cited *śisikan), from *śi-*ka-n, an archaic derivative (> Japanese ika 'squid') of the ubiquitous AT 'core' root: *ka/ka* 'eat'.

It is entirely possible, of course, that the borrowing of one or more basic roots such as 'hair' (above) will eventually be assignable to the AT/AA 'border phenomena'. If so, the syllabic reduction factor requires that one set up the direction of the loan(s) as AT > AA. This direction is also required for the early AT > AA culture-word loans, notable those for 'copper' (> Munda group 'iron') and 'sugarcane', as pointed out in the 1973 paper. The Aslian languages maintain an archaic final *s (generally > -h = -0 in non-Formosan AN languages) in the latter loan and an extensive corpus of Aslian material of this kind recently collected by Geoffrey Benjamin (pers. comm.) includes what appear to be parallel examples of early loans that preserve other archaic phonological features, including even a final *-l (maintained as a lateral in Formosan languages only). The analysis of this material is still in an early phase and it is unclear how many, if any, of these loans can be established as roots at the PMK (possibly even PAA) level; in any event, they supply additional evidence for the priority of the AT-speaking peoples in the early emergence of 'high culture' in South-East Asia (Benedict 1975).

At the 1976 Toronto symposium on AT, Harry Shorto made a spirited defence of Austric, prompting a 'Comment' by the writer (1976b). The argumentation there need not be repeated in this paper; much of it concerns lack of agreement re lexical items. Shorto proposed a number of new AN/MK cognate sets, falling for the most part in what the writer termed the ketketbongbong class, involving lexical areas well outside 'core' vocabulary, with a striking attenuation or gap in this key aspect. As summed up by Benedict (1976b: 106):

...the point to be stressed is that a score of ketketbongbong do not one language stock make, since correspondences of this kind have significant value for comparative purposes only when found in association with a corpus of core vocabulary items.

2. The Formosan evidence, strongly supported by the Japanese (see Benedict 1990), points to an underlying trisyllabic root: *busukas or *bosokas (PA medial *o and *u were merged in PAN *u), which yielded Malayo-Polynesian forms of *busuk > buhuk type. The mid or mid-low height of the PMK vowel—the cited form reflects the Proto-Semai vocalism—that can readily be explained as a retention rather than an innovation while the length can be interpreted through vocalic transfer (Benedict 1979). This root appears to be unrepresented in the mainland AT languages (Kadai, Miao-Yao); the indicated line of development, with syllabic reduction on-the-left along with vocalic transfer, is of the sort that characterises the Kadai family as a whole. PAT *o is also maintained in this family, which may well have been the source of an early loan (as an alternative to a 'look-alike') for this lexical item.

9
The final clause was italicised lest the point be overlooked, as apparently has occurred in the more recent paper by Diffloth (1985). The nub of the writer’s argument here lies in the fact that Shorto, a MK specialist with good access to a wealth of comparative AN data, failed to turn up an even halfway decent corpus of ‘core’ vocabulary correspondences. This is indeed the name of the game, at least in the South-East Asia region, on the basis of an apparent linguistic consensus at the present time, and on this basis Austric surely emerged as the loser. And the writer (1976b: 105), even ‘at the risk of appearing whimsical’, could quite honestly report his belief that Shorto’s paper had served only to strengthen the case against Austric!

Almost a decade passed before the appearance of another defence of Austric, this also by a leading Mon-Khmerist, Gérard Diffloth (1985). Unlike Shorto, however, who proposed a sizeable number of new AN/MK cognate sets, Diffloth included only two ‘possible Austric etymologies’: for ‘wood’ (‘tree’) and ‘bone’, apparently the prize members of his collection, preferable to certain ‘runts’ which he modestly kept to himself. The pair that he does offer are in fact ‘look-alikes’ or ‘comparabilia’, with Diffloth admitting to serious difficulties in reconstruction even at the PMK, let alone a PAA or ‘Austric’, level. It must also be noted that he has unparalleled MK comparative material at his disposal along with the most recent advances in the AN field, all of which he has made use of, hence his failure to do any better than Shorto simply further strengthens the null case here. Given the present circumstances, the notion that a presentable corpus of ‘core’ AN/MK (or AT/AA) cognate sets lies ‘out there somewhere’ waiting to be discovered strikes one as quixotic in the extreme.

Diffloth’s citation of ‘look-alikes’ in an undertaking of this nature conforms to general linguistic practice, as does his ‘sharpening the reconstructions of promising cases’ (Schmidt had proposed the ‘wood’ etymology). Quite novel, however, is his attack upon what he labels ‘Swadesh’s “basic” notions’, which is no less than a frontal assault upon the whole idea of utilising basic or ‘core’ vocabulary in comparative work. Granted a certain feeling of frustration on his part in being unable to uncover ‘core’ AN/MK cognate sets, one can hardly follow him in this rash line of thinking. After pointing out that given items of ‘core’ type can on occasion be replaced through factors such as ‘taboo or euphemistic passe-passe’, he makes the following summary statement:

Over the millennia, the recurrence of factors like the above [cit. supra] could easily wipe out 100 such ‘basic words’, or even 200 [i.e. both the standard Swadesh-type word lists], while protecting for us many hidden gems, such as ‘to whittle bamboo strips’ or ‘scruff’, not to mention ‘smegma’ and the like. Benedict may detest it, but the Austric hypothesis is still very much alive. (Diffloth 1985)

The writer does not ‘detest’ Austric; in fact, as an old supporter of the hypothesis he must admit to a certain sentimental attachment to it and a
Austric: an 'extinct' proto-language

sadness at its demise. He must confess, however, that he is dismayed at the thought of a proto-language with a corpus of reconstructed roots made up exclusively of 'scruff', 'smegma', and the like. One can hardly visualise comparatists recognising a Proto-Indo-European language put together in such a fashion, and there is no reason for them to accept an 'Austric' of this sort either. South-East Asia has its novelties as a research field, to be sure, but it is hardly extra-linguistic. The cardinal point, as emphasised above, is that roots of this kind, whether called ketketbongbong or not (they need not be reduplicated), can be used in establishing a genetic relationship only in conjunction with 'core' vocabulary. On theoretical grounds, of course, all 100 or even 200 'Swadesh list' roots could have been replaced in a given instance, but how would one prove it, especially in the case of proto-languages, such as PAT and PAA, that occupy contiguous territories? Perhaps, in the dim South-East Asian past, AT and AA not only influenced each other structurally but also exchanged a few scruffy or smegmatic words in the bargain, along with some 'high culture' items after AT had made the early advance towards 'civilisation'. None of this, however, justifies the setting up of an 'Austric' superstock. As reported at the outset of this paper, Austric is 'extinct'.

REFERENCES


Thirteen years ago, H.L. Shorto pointed to the vowel system of Proto-Mon-Khmer as being the 'crux' in the historical phonology of this family (Shorto 1976). His assessment remains as valid today as it was then, even though some advances in reconstruction have been made; our data base has considerably expanded and improved, but the mirage of Proto-Mon-Khmer vowels continues to recede, even as we penetrate further into the past.

The solution proposed then: vowel variation in the proto-language, is consistent with certain facts which can be observed in several Mon-Khmer languages spoken today. In Bahnar, Sre, Khmu and Semai, to select but a few, whole families of Expressives (Diffloth 1979; in press) are often built on vowel permutations, and such Expressives occasionally find their way into the prosaic (Non-Expressive) vocabulary; conversely, prosaic words often serve as a starting point for building families of Expressives which differ, for example, only by their major vowel. This has surely contributed to the formation of word-families such as those identified in Shorto (1973). This explanation, however, has its limits: presumably, these processes would have affected a word here and a word there, at different times, but it is difficult to see how it could have pervaded the thousands of items which form the non-expressive lexicon of one language, not to speak of an entire family. Other factors are needed in order to account for the numerous vowel correspondences which have been detected so far.

For example, it may well be that the Proto-Mon-Khmer vowel system reconstructed until now, although sizable, is not rich enough for the purpose, and that we need to expand it with some additional phonological dimensions.

Tone has been practically ruled out for Proto-Mon-Khmer since the simple tone systems of Büläng and Riang, and the tone-cum-register system of Nyah Kur (Diffloth 1980, 1984) can all be explained as innovations; but the newly recorded Angkuic languages U and Man Met (see below) have four-tone and six-tone systems respectively, the origins of which remain partly unknown for the moment. Then again, Haudricourt's account of Vietnamese (VN) tonogenesis has generally been accepted, but it leaves out, as tonally irregular or unexplained, a large number of words which do belong to the indigenous Mon-Khmer stratum of the language. And the recently discovered Palyu language, called Lai in Chinese, also has six tones which may, or may not, turn out to be
recalcitrant—Palyu is apparently Mon-Khmer (Liang 1986, Benedict, in press), but its position in the family is still undecided.

Register is a better candidate. This is, typologically, a well-established feature of Mon-Khmer languages (Huffman 1976). The general consensus is that Register is a relatively recent phenomenon, and Shorto, accordingly, does not reconstruct it for Proto-Mon-Khmer. Ferlus (1979) described all Mon-Khmer register systems found until then as being the result of one type of evolution: devoicing of initial consonants. This explanation has long been accepted in the case of Spoken Mon (Blagden 1910) and of Modern Khmer (which ironically has now lost phonation-type distinctions), and it does account for the registers of several other newly recorded languages (e.g. Kuy, Bruu, Phalok). But it is inadequate in certain other languages: the Pacoh register system, admittedly an innovation, has nothing to do with the process of devoicing which has independently taken place in this language. It is also inapplicable to the North Bahnaric languages where no devoicing has taken place, except in Sedang. In Pacoh, the genesis of register is due to changes in vowel quality, namely, the fronting or backing of certain proto-central vowels (Diffloth 1982), and a similar innovation has apparently also taken place in North Bahnaric (Diffloth 1983). So, we do not have yet a case of reconstructing register as being ancient in Mon-Khmer.1

However, the Pearic branch might force us to do that: recently, Huffman (1985) has shown that Chong, a language of the Pearic branch, had a Clear vs. Breathy distinction, criss-crossing a Plain vs. Glottalised contrast, giving rise in effect to a four-register system. Theraphan (this volume) describes in detail the complex bundle of phonetic features these four registers contain. This phenomenon has no historical explanation, and Headley himself (1985) has abandoned to the sagacity of future historical linguists any attempt in this regard.

Gage (1985) has pointed out that certain unexpected occurrences of the śćăc tone in Vietnamese seem to find an echo in the register system of Pearic. The tonogenesis of Vietnamese requires that the śćăc tone occur with final proto-stops, and indeed cannot explain the tones of many VN words which have excellent Mon-Khmer etymologies, such as: bôn ‘four’, chín ‘cooked’, gió ‘wind’, or: ngái ‘far’. Cognates to all four of these words happen to have glottalisation in Pearic. In Chong, as I have recorded it, the first three have the ‘tight’ register:2 /phọơn/ ‘four’, /chijın/ ‘cooked’, /kəjəŋy/ ‘wind’; the fourth word has a ‘breathy-creaky’ register3: /ŋąy/ ‘far’. Other examples can be found, e.g. VN: câm, Chong /kəŋəm/ ‘rice-husk’, but there are counter-examples as well, e.g.: VN: chitm, Chong /cjiım/ ‘bird’. Since the Pearic and the Việt-Muông branches are only distantly related, the implications of this fact could go back directly to Proto-Mon-Khmer.

1. Smith’s opinion to this effect (Smith 1972) was not based on the establishment of sound correspondences, but on statistical tendencies within a very small set of possible Mon-Khmer cognates, which a more thorough comparison does not confirm.
2. Clear voice plus glottalisation in Huffman’s (1985) analysis.
3. Huffman’s (1985) breathy voice plus glottalisation.
Palaungic vowels in Mon-Khmer perspective

There are also less exotic vowel features which have not been fully used in Mon-Khmer reconstruction. Diphthong systems can be much richer than the simple *ia and *ua usually proposed; I have reconstructed Proto-Katuic with five proto-diphthongs (Diffloth 1982), and Nancowry Nicobar (Radakrishnan 1981:25) is described even today as also having five diphthongs: /iá/, /iá/, uá/, /uá/ and /uá/, which seem to correspond with what we can reconstruct for Proto-Aslian.

Some of these phonological features may have to be reconstructed back to Proto-Mon-Khmer, and could well explain a number of Shorto’s variations as being regular outcomes of a much richer proto-vowel system. But then, the number of proposed Proto-Mon-Khmer etyma becomes a relevant issue, and what has been published so far can be said to represent only a sample.

In this paper, I will not explore these possibilities, but only prepare the comparative ground to do so; I will try to clarify some points in the history of vowel systems in the Palaungic branch, where recently recorded material allows us to make systematic reconstructions. This may seem at first to be somewhat irrelevant: if Waic and Palaung are notable for one thing, it is precisely the poverty of their vowel systems. The old vowel-length contrast was already lost in Proto-Waic, and the best source of information on Milne’s Palaung (= Ta-ang) appears not to have a phonemic contrast of this kind.

But the Palaungic branch has an important role to play in reconstruction: it belongs to a distinct division of the family, the Northern Division, and it provides us with an independent testimony for the reconstruction of Proto-Mon-Khmer vowels. Besides, as I will try to show, Palaungic vowel systems are not as poor as they first seem to be.

1. Proto-Waic

The term ‘Waic’ covers (1) several Wa languages, e.g. Paraok, Avüa, La (Zhōu & Yān 1984) and their dialects; (2) the Phalok language, formerly referred to as Khalo or Mae Rim Lawa (Flatz 1970); (3) Lawa and its dialects (Mitani 1972); and (4) the Bulang-Phang complex with its many dialects (Diffloth 1980). Certain Waic languages, Lawa and Paraok in particular, currently have rich and complicated vowel systems but this is

---

4. Professor Shorto has let me use his own notes from Riang and from the same Palaung language, Ta-ang, as described in (Milne: 1931); this is the source of the ‘Ta-ang’ and ‘Riang’ words quoted here. Let the present article be a small token of appreciation for his kindness.
5. I collected the information on Phalok included here in two separate field trips, one in April 1981 with the help of Theraphan L. Thonkum, and the second by myself in July of the same year. This research was financed by a grant from the National Science Foundation (NSF) entitled ‘An etymological lexicon of Mon-Khmer’. More information on Phalok will be made available in the forthcoming volume ‘Wa-Lawa-Bulang’.
6. In that study, I called ‘Samtao’ a language which later turned out to be identical to that spoken by the Bulàng National Minority in Yunnan, China. No linguistic information was available on Bulàng at the time, as Zhōu & Yān (1983) had not yet appeared.
due, in part, to the influx of Tai borrowings\(^8\) and in part to recent processes of vowel warp, conditioned by Registers and final consonants. Only nine proto-vowels are needed at the Proto-Waic stage\(^9\)

**Proto-Waic Vowel system**

\[
i \\
e \\
\varepsilon \\
a \\
\circ
\]

This maximum system is found with most final consonants, but there are certain distributional gaps; for example, with final *-ʔ, only eight proto-vowels are found (all the above except *o); with final *-h, only seven are found (*y and *o are excluded); and there are no open final syllables in Proto-Waic.

2. **Proto-Palaung-Rumai**

The term ‘Palaung-Rumai’ also covers several languages, the best known of which is Ta-ang, i.e. the Palaung of Nam Hsan described by Milne (1931). The Rumai language and its dialects, also belongs here\(^10\), as well as the dialects of Riang;\(^11\) it also includes another distinct group sometimes called ‘Palé’, which contains at least Da-ang and Na-ang; and several other languages, like Ka-ang and Ra-ang\(^12\). Other Palaung-Rumai languages surely await description in Burma or Yunnan, and they may or may not belong to one of the seven groups mentioned here.

There is considerable diversity within Palaung-Rumai, but this is not the place to present all the phonological innovations which can now be documented. Mitani (1977, 1979) has already reconstructed the Palaung-

---

8. Lawa has borrowed from Northern Thai and more recently from Standard Thai; Paraok has borrowed from Shan (referred to in China as ‘Dehong Dai’), Bulang has borrowed from Lü (referred to in China as ‘Xishuangbannah Dai’, or Xi Dai for short).

9. This reconstruction was presented in (Diffloth 1980); since then, a Paraok-Chinese dictionary has been published (Yán et al. 1981), and, in 1984, I was fortunate to collect Waic linguistic material in China, with the help of Zhōu, Z.-Z., under another NSF grant entitled ‘Comparison of the Mon-Khmer languages of China with other languages of the Mon-Khmer family’; with this new material, the number of reconstructed Proto-Waic words has now more than doubled, and the reconstruction of a few etyma given in Diffloth (1980) has been modified; this new information confirmed the nature of the Proto-Waic system I had reconstructed in 1980 with the help of Y. Mitani.

10. All the Rumai examples quoted here were collected in May 1981, with the help of a family of Rumai emigrants living in Chiang Mai at the time.

11. See note 4 above.

12. The information on Na-ang included here was kindly given to me by Yán, Q.-X. as part of a research programme in China (see n.9). She is the author of a valuable sketch on the Běnglóng language(s) (Yán 1983). I collected myself the Ka-ang data from a native speaker in Künming, in the course of the same research programme. The Da-ang and Ra-ang data were collected in 1981 (see n.5), and 1984, during my stays in Thailand. It was not possible for me to determine with precision the geographic spread of these languages, especially for those spoken in the Shan States (Burma) where there seems to be a lot of small-scale migrations. For China, Svantesson et al. (1981), Yán (1983), Zhōu & Yán (1983, 1984) provide geographic and demographic information.
Rumai vowel system, using older material; and even though none of the sources used had indicated the vowel length contrasts which are clearly present today in most of these languages, he did reconstruct vowel length contrasts at the Proto-Pal-Ru stage. Remarkably, his reconstructions, which he termed as tentative, are confirmed by the better material we now have; this is true at least for the presence of a proto-length contrast, and for the number of proto-vowels: 10 (although this represents a minimum); only the reconstructed quality of some of these vowels can now be improved upon.

Briefly, with only a few rare gaps and notation problems, we find the following correspondences in words having a proto final Velar Nasal:

<table>
<thead>
<tr>
<th>Rumai</th>
<th>Na-ang</th>
<th>Da-ang</th>
<th>Ra-ang</th>
<th>Ta-ang</th>
<th>Ka-ang</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>-aŋ</td>
<td>-aŋ</td>
<td>-aŋ</td>
<td>-aŋ</td>
<td>-aŋ</td>
</tr>
<tr>
<td>(2)</td>
<td>-ōŋ</td>
<td>-ōŋ (-œŋ)</td>
<td>-ňŋ</td>
<td>-ňŋ</td>
<td>-ňŋ</td>
</tr>
<tr>
<td>(3)</td>
<td>-oön</td>
<td>-œŋ (-œŋ)</td>
<td>(?)</td>
<td>-œŋ</td>
<td>-œŋ</td>
</tr>
<tr>
<td>(4)</td>
<td>-ŋ</td>
<td>-ŋ</td>
<td>-ŋ</td>
<td>-ŋ</td>
<td>-ŋ</td>
</tr>
<tr>
<td>(5)</td>
<td>-oŋ</td>
<td>-œŋ (-œŋ)</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
</tr>
<tr>
<td>(6)</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
</tr>
<tr>
<td>(7)</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
</tr>
<tr>
<td>(8)</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
</tr>
<tr>
<td>(9)</td>
<td>-oön</td>
<td>-œŋ (-œŋ)</td>
<td>-œŋ</td>
<td>-œŋ</td>
<td>-œŋ</td>
</tr>
<tr>
<td>(10)</td>
<td>-yọŋ</td>
<td>-iañ</td>
<td>-iañ</td>
<td>-iañ</td>
<td>-iañ</td>
</tr>
</tbody>
</table>

Phonetic values for reconstruction of these proto-rimes do not come to mind immediately, to say the least. But the first four are relatively easy: Mitani (1977) reconstructs *aa, *a, *unu, and *ur respectively, for (1), (2), (3) and (4).

Correspondence (1), includes items such as:
- ‘bone’ Ru, Na, Da, Ra: /kaʔaŋ/, Ta: /kaŋaŋ/, Ka: /kaʔaŋ/
- ‘house’ Ru: /gaŋ/, Na, Da, Ra: /kaŋ/, Ta: /gaŋ/, Ka: /gaŋ/.
- ‘torch, lamp’ Ru: /raŋ/, Ta: /raŋ/.

Riang-Lang cognates have a back /a/ in these words:
- /prún/ ‘sour, acid, rancid’.

Waic cognates always have Proto-Waic *a:

13. Evidently, some vowels have caused PPal-Ru *ŋ to palatalise to -n, and sometimes further merge with *-n; but both *-p and *-n must be reconstructed at the proto-Palaung stage, in contrast with *-ŋ.
Lamet cognates have a long /aa/:


And in the Angkuic branch of Palaungic, Man Met\(^\text{14}\) has /-aːn/ finals with a low-falling tone, and U\(^\text{14}\), which does not have vowel length contrasts, shows finals in /-ən/ with a high-falling tone:


Outside Palaungic, Mitani’s reconstruction of *-aːn is confirmed, for instance, by Khmu:


Note that we will not be concerned here with the history of initial stops in these languages, interesting as that may be. In a nutshell, Rumai and Ta-ang have preserved the original values of proto-voiced and proto-voiceless stops; the same state of affairs also exists in Ka-ang, except that initial *p- and *t- have become implosive b- and d- respectively; in Ra-ang, *p- and *t- have followed the same evolution, and *k- has become a voiced stop g-, while all proto-voiceless stops have become voiceless; Da-ang and Na-ang have followed the same course as Ra-ang, and, in addition, have lost the implosion of b- and d-, which then become ordinary b- and d-; Da-ang and Na-ang therefore show a total reversal of the voicing values of PPal-Ru stops: another illusion of ‘flip-flop’, with implosiveness as the point of transition in this game of musical chairs. Angkuic, on the other hand, has undergone a completely regular ‘Germanic’-type of sound change where proto-voiceless stops are now aspirated, and proto-voiced stops are now voiceless. But note that Angkuic tono-genesis is mainly due to vowel qualities and quantities, with some influence from the final consonant; it is unrelated to the earlier or the present voice features of initials.

Correspondence (2) includes:

‘bitter, gall’ Ru: /sɔn/, Na, Da: /sɔŋ/, Ra: /sɔŋ/, Ta: /sɔŋ/, Ka: /səŋ/
‘thatch-grass’ Ru: /plɔŋ/, Na: /blɔŋ/, Da: /blɔŋ/, Ra: /blɔŋ/, Ta: /pləŋ/
‘bamboo-shoot’ Ru: /bəŋ/, Na: /pəŋ/, Ta: /bəŋ/, Ka: /bəŋ/
‘house-pole’ Ru: /rɔŋ/, Na, Da: /rɔŋ/, Ka: /ʁŋ/
‘horse’ Ru: /mbrɔŋ/, Da: /mprɔŋ/, Ta: /brɔŋ/, Ka: /braŋ/ 

Riang-Lang cognates have a front /a/ in these words:


\(^{14}\) I recorded Man Met and U from native speakers, in Yunnan, during the research project mentioned above (n.9); they were introduced to me as speaking ‘dialects’, or more exactly ‘fāngyān’, of the Būlāng language. The Chinese term ‘fāngyān’ corresponds most of the time to what Western linguists consider to be different languages; this was true of these ‘fāngyān’ of Būlāng which do not even belong to the Waic branch of Palaungic, but to the little known Angkuic branch (Diffloth, 1974). The location of the U language is given in (Zhōu & Yān 1973), Man Met is spoken a few miles from Jinghong, Xīshuāngbānnà, Yūnnān.

18
Proto-Waic cognates have *o:


Lamet has a short /a/:

/can/ ‘bitter’; /plan/ ‘thatch grass’; /tpan/ ‘bamboo-shoot’; /mpan/ ‘horse’.

And in the Angkuic branch, Man Met has /-an/ rimes with a high-rising tone, and U has /-aÊ/ rimes with mid-rising tone:


Here again, Mitani’s reconstruction of a short *a is confirmed, outside Palaungic, by Khmu which regularly shows cognates with short /a/:


The contrast between correspondences (1) and (2) shows that length vs. shortness has been preserved everywhere in the Palaung-Rumai languages, at least for this pair of vowels, except in Ta-ang which does not show any length distinctions anywhere in its system.

It also shows that in the Angkuic branch, tonogenesis is directly due to vowel-length, and has nothing to do with the proto-voice feature of initials; this kind of tonogenesis is unique in the Mon-Khmer family, but, annoyingly, it accounts for only some of the tonal contrasts found in Angkuic languages. It should also be noted that Man Met does undergo this kind of tonogenesis even though the older length distinction is retained; in U however, the length contrast, before disappearing, leaves another trace in the final consonant; it de-nasalises final nasals after short vowels.

Another remark: if we only had the Waic and the Riang-Lang material at our disposal, it would appear that a so-called flip-flop has taken place: Waic has a front *-a- where Riang-Lang has a back /a/; and vice-versa: Waic has a back *-o- for Riang-Lang’s front /a/. The former presence of a length contrast shows this apparent flip-flop to be nothing but a synchronic illusion.

Correspondences (3) includes:

‘high’ Na: /leep/, Ta: /hlunj/, Ka: /hlon/
‘to dig (a hole)’ Ta: /kunj/

Riang-Lang has an /o/ reflex:

/kunj/ ‘to dig’, /tsunj/ ‘foot’

and Proto-Waic shows *-o-:

*hlon (N77) ‘high’, *kon (N9) ‘to dig’, *jon (N18) ‘foot’

Lamet has /ee/:

In Angkuic, Man Met has long /ee/ with a low falling tone, U has /e/ with a high-falling tone and a final nasal:

\[ \text{MM: /leə/ 'high, long'; U: /kʰeə/ 'to make with a dibble-stick'; MM: /ceə/, U: /cəə/ 'foot'.} \]

Outside Palaungic, Khmu has cognates with /ia/:

Khmu Yuan (Svantesson, personal notes): /kʰian/ (aspiration unexplained) 'to dig up', Southern Khmu: /jiaŋ/ 'foot'.

**Correspondence (4):**

- ‘bed-bug’ Ru, Na, Ra: khanj/, Ta (Milne, 1931): kõŋ
- ‘stalk, trunk, post’ Ru, Na: /taŋ/, Ta: /taŋ/ (Milne: tõŋ, tōŋ), Ka: /bōŋ/.

The Riang reflex is /a/:


Proto-Waic has *-y-:

*hrŋ (Paraok: /hun/, Drage’s Wa: hōng, Phalok: /hun/, Phang: /hun/, ‘bed-bug’).

**Lamet has a short /a:/**:

/raŋ/ ‘bamboo’, /hōŋ/ ‘bed-bug’.

In Angkuic, Man Met has a short /a/ with a high-rising tone (but the tone of ‘bed-bug’ seems to be low-rising), while U has a /a/ with a mid-rising tone and de-nasalisation of the final:

\[ \text{MM: /hōŋ/, U: /hrōk/ 'bamboo'; MM: /sōŋ/ (tone?), U: /sōk/ 'bed-bug'} \]

Outside Palaungic, Khmu has a short /ui:/

/huŋ/ ‘bed-bug’.

Another etymon with the same proto-vowel, unfortunately without Palaung-Rumai attestations, is:


Mitani’s reconstruction of a length contrast, *uru vs. *u, for these two proto-vowels is confirmed independently by the evidence of Lamet and Angkuic (both Man Met and U); in Palaung-Rumai, Na-ang and Rumai also display the same contrast, which appears to be a retention, not only from Proto-Palaungic, but from still earlier periods, as the Khmu evidence indicates.

For all four proto-vowels *aa, *a, *uru and *u, Na-ang and Rumai have thus preserved the older length feature. I depart slightly from Mitani in terms of vowel-qualities: (3) and (4) were probably central mid-vowels, *ə and *ə. This would explain how (3) became *ee in both Lamet and Angkuic and *o in Proto-Waic, and why (4) has very open vowel quality reflexes throughout Palaung-Rumai, except in Ta-ang where (3) and (4) merged and were pushed higher to /u/ by the shift of *a to /a/. The original value of (4) is preserved in Riang, Lamet and Angkuic.
For the remaining vowel correspondences, Mitani does not reconstruct any length contrast. For (5), (6) and (7) he proposes a back-rounded series: *u *o and *ɔ respectively.

Correspondence (5) includes:
‘big village, country’ Ru: /kunŋ/, Ta: /kunŋ/
‘drum’ Na: /grvŋ/, Ra: /grvŋ/, Ta: /kruŋ/, Ka: /kruŋ/
‘bamboo-strip mat’ Ru: /bluŋ/, Ra: /plvŋ/, Ta: /bluŋ/
‘love, like’ Ru, Da: /ʔωυŋ/, Ra: /ʔvŋ/, Ta: /ʔυŋ/
‘to bury’ Ta: /kpruŋ/, Ka: /kɔðuŋ/ (Vowel ?)
The corresponding PWaic vowel is *i: 
*kıŋ (N10) ‘wet field, country’, *kıŋ (N63) ‘drum’, *kripsi (N41) ‘to bury’

In Angkuic, Man Met has a short /u/ with high tone, and U has /u/ with a mid-rising tone and de-nasalisation of the final:
MM: /khuŋ/, U: /khɛk/ ‘wet rice-field’
And outside Palaungic, Khmu has /u/: /kuŋ/ ‘village’

Lamet cognates are missing, although there are other etyma with apparently the same proto vowel, but without Palaung-Rumai attestations; these show the Lamet reflex to be short /u/:
‘to blow’: Lamet /pʊŋ/, PWaic *piŋ (N40), Man Met: /phʊŋ/, Khmu /punŋ/ ‘a sprout’: Lamet: /plʊŋ/ (tone ?), PWaic *biŋ (Paraok /plaŋ/), Khmu: /bluŋ/

For this correspondence, the evidence for proto-shortness is clear: in Palaung-Rumai, only Na-Ang has a long vowel (in a single item which could have been misrecorded); all other languages where a length contrast exists, Rumai, Da-ang, Ra-ang, Ka-ang, Lamet, Angkuic, Khmu, have a short reflex.

The other two correspondences (6) and (7), seem, by contrast, to be on the long side:

Correspondence (6)
‘male bird’ Ru: /kʊŋ/, Ta: /ʔo-kʊŋ/, Ka: /kʊŋ/
‘Classifier: round objects’ Ru: /poŋ/, Ta: /pʊŋ/, Ka: /boŋ/
‘buttocks’ Na: /səpoŋ/, Da: /səpoŋ/, Ra: /səpoŋ/, Ta: /səboŋ/ (my own recording, cf. Milne sə-bong which would indicate /səboŋ/)
‘far, long’ Ru: /dʊŋ/-/ndoŋ/, Na: /toʊŋ/, Ra: /tʊŋ/, Ka: dʊŋ/
The Proto-Waic reflex is *o: 
*kʊŋ (N61) ‘knee-cap’, *kʊŋ (N8) ‘peacock’
Lamet has a long /oʊ/ reflex:
/kxɔoŋ/ ‘knee’

In Angkuic, Man Met has a long /oʊ/ with a low-falling tone, and U has an /o/ with high-falling tone but no denasalisation of the final:
Correspondence (7) is well documented and contains well-known etyma; it also includes some surprising Front reflexes for what is certainly a proto Back vowel:

- 'hornet' Ru: /ɔeŋ/, Da: /ɛeŋ/, Ta: /ɔoŋ/, Ka: /ɔoŋ/
- 'rainbow' Ta: /ɔɛγŋ/, Ka: /ɔɛγŋ/
- 'back (of body, of knife)' Ru: /kɛeŋ/, Ra: /ɡɛeŋ/, Ta: /krɔŋ/. Ka: /krɔŋ/
- 'stairs' Ru: /ndeŋ/, Na: /nteŋ/, Ka: /doŋ/

In Riang, the normal reflex is a diphthong /ua/ or / uɔ/:

- /uən/ 'hornet', /prɪuəŋ/ ‘rainbow’, /rənðuəŋ/ ‘stairs’

Proto-Waic regularly shows *o:

- /ɔoŋ/ (N6) ‘hornet’, *pryɔŋ (N84) ‘rainbow’, *krɔŋ (N60) ‘back’

Lamet has a long /ɔo/ reflex:

- /uəŋ/ ‘hornet’, /pxyɔuŋ/ ‘rainbow’, /kɔɔŋ/ ‘back’

In Angkuic, the Man Met reflex is a long /ɔo/ with a low falling tone, and the U reflex is a diphthong /ua/, with a high-falling tone and no denasalisation of the final:


Outside Palaungic, the Khmu reflex is regularly a long /ɔo/:


Mitani was certainly justified, on phonemic grounds, in leaving the feature of length unspecified in the back vowels series of Proto-Palaung-Rumai; but we can afford to be more precise now, and say that *u was probably short, while *o and *ə were probably long *oo and *ɔ. This will allow for an easier description of the gradual collapse of the older vowel-length system in languages like Ta-ang and Proto-Waic. And for Front vowels, the same line of reasoning will help us to solve a curious problem.

The proto Front Vowel system was also left unspecified as to length by Mitani, who reconstructed *i, *e, and *e for our correspondences (8), (9) and (10) respectively.

Correspondence (8):

- 'to sew' Ru: /jɛn/, Na: /ceep/, Ra: /ceŋ/, Ta, Ka: /jɛŋ/
- 'head' Ru: /kɛn/, Na: /geŋ/, Da: /giŋ/, Ra: /geŋ/, Ta, Ka: /kɛŋ/
- 'bamboo water-container' Ru: /dɛn/, Ra: /teŋ/, Ta, Ka: /dɛŋ/
- 'husband' Ru: /mɛŋ/

Riang has an /i/ reflex for this correspondence:

- /kɛn/ ‘head’, /kɛndɛŋ/ ‘navel’

But PWaic has an *e:


Lamet normally has a short /i/, except for one case of long /ii/:

This last item may actually be one of the many words Lamet has borrowed from the neighbouring Khmu; unfortunately, I have not yet found the expected word /hmiiɪŋ/ in any dialect of it. But the Khmu reflex of correspondence (8) is indeed a long /ii/:

Khmu Yuan (Svantesson, personal notes): /kńiiɪŋ/ ‘navel’, /tiɪŋ/ ‘bamboo water-container’

In Angkuic, Man Met has a short /i/ with a high-rising tone, and U has /i/ with a mid-rising tone and denasalisation of the final Nasal:


The Palaung-Rumai evidence, except for Na-ang, points to a proto-short vowel *i, and this is confirmed at the Proto-Palaungic level by Lamet and Angkuic.

Correspondence (9):


‘wheel’ Ru: /kəlōon/, Ra: /kōeŋ/, Ta: /kənleŋ/

‘equal amount’ Ta: /kpręŋ/

Riang cognates have an /e/ vowel:


And Proto-Waic, surprisingly, has an *i:

*kliŋ (Paraok: /kliŋ/) ‘to spin (yarn)’; *mriŋ (Bo Luang and Umphai Lawa: /mbręŋ/) ‘to compare quantities, to match’

The Lamet evidence is, unfortunately, limited to a single item:

/tampliɪŋ/ ‘sky’

In Angkuic, I do not have Man Met cognates, but U has the same reflex as in correspondence (8): /i/ with mid-rising tone and denasalisation:

U: /phliŋ/ ‘sky’

This merger of (8) and (9) is specific to U and not general in Angkuic: there are other cognate sets, unfortunately without attestations in Palaung-Rumai, where this U rime corresponds to Man Met /-eeŋ/ with middle tone, (and to Proto-Waic *-iŋ, as correspondence (9) requires):

‘to return home’ U: /?ıkk/, PWaic *?iŋ (N1)

‘wall, partition’ U: /ndhik/, Mok (a close relative of Man Met spoken in Thailand, Wenk’s (1965) ‘Ya Ang Lawa’): /theen/, PWaic *ntiŋ (N21)

I propose to reconstruct PPal-Ru *ee for correspondence (9), and for PPalaungic as well. The Palung-Rumai reflexes indicate a long vowel, even in Rumai, where the strange reflex, /-oon/, represents a merger with *-oəŋ (correspondence 3), itself a proto-long vowel.

This reconstruction also provides a simple explanation for what would appear to be yet another case of flip-flop: if we kept Mitani’s reconstruction, without length, the correspondences would be:

(8) *i-: PPal-Ru *-i- = PWaic *-e-
(9) *e-: PPal-Ru *-e- = PWaic *-i-

If, however, we reconstruct PPal-Ru *ee for (9), head-on collisions are easily averted, and gradual phonetic change can proceed smoothly.

In addition, we can also explain the Lamet and Angkuic reflexes: in those two sub-branches, Proto-Palaungic *əə (correspondence 3) was fronted to a long /ee/, pushing the older *ee (corr. 9) out of the way: in Lamet, this *ee was raised to /ii/, keeping its long feature and filling a gap in the vowel system; but in Angkuic, Man Met and U evolved differently: in Man Met *ee and *əə merged in terms of qualities and length, but seem to have acquired different tones, whereas in U, *ee simply merged with *i. If this is correct, the movement of *əə to /ee/ is probably not an innovation shared by Lamet and Angkuic; it may have happened independently in these two sub-branches which do not appear to be especially closely related to each other.

The last correspondence, (10), poses a special problem in that it is difficult to decide if *EE or *iə should be reconstructed, for either PPal-Ru or PPalaungic. In any event, a short *e seems a very unlikely value for this correspondence in view of its reflexes:


The Riang reflex is a diphthong, usually transcribed /ie/:


In PWaic, the reflex is *e:

*?əŋ (N3) ‘excrement’

The normal Lamet reflex is a long /ee/, although ‘excrement’ has unexpected reflexes, possibly due to euphemistic deformation:


In Angkuic, the Man Met reflex is a short /e/, usually with a high tone, whereas in U, we find a diphthong /ia/ with a high-falling tone and a Nasal final; ‘excrement’ is irregular only in U, having what seems to be an otherwise unattested proto-short diphthong:

Palaungic vowels in Mon-Khmer perspective

Only one cognate has been found so far in Khmu\textsuperscript{16}: /kian/ 'elbow'.

For the sake of consistency with Mitani's notation, I will arbitrarily reconstruct \(*ee\) for this correspondence at the PPal-Ru level.

The vowel system reconstructed for PPal-Ru now appears as follows:

\begin{center}
\textit{Proto-Palaung-Rumai Vowel System}
\begin{tabular}{ c c c c c c c c c }
\hline
 & *i & *e & *\=e & *\=a & *\=o & *o & *\=oo & *\=aa & *\=a \\
\hline
\textit{ee} & *e & *\=e & *\=o & *\=aa & *\=a \\
\textit{\=ee} & *\=e & *o & *\=a & & \\
\hline
\end{tabular}
\end{center}

In contrast to the 9-vowel system shown above for Proto-Waic, this 10-vowel system shows the vowel length contrast to be still operating, but with a small functional load. I have not attempted here to reconstruct the Proto-Palaungic state of affairs, because Lamet and Angkuic are still too poorly known; but it seems likely that the vowel system will be, if anything, richer at that stage, probably filling some of the gaps in the long vs. short contrasts.

The Proto-Palaung-Rumai vowel system appears to be, typologically, half-way between a full South-East Asian system, as found for example in Khmu or Standard Thai, and a more contracted system where vowel length has been lost, as found for example in Proto-Waic, and Modern Mon. This vowel-system contraction seems typical of the Burma-Yunnan linguistic sub-area. But what Palaung-Rumai shows is that areal pressure does not work like a stream-roller: Ta-ang did lose the length contrast, but Rumai, even with two mergers, maintains it systematically and even innovates in this respect with complete disregard for its more forceful neighbours.

Something similar can be said of U: while it did lose vowel-length and acquired tones, thus conforming to its neighbours, it did so in a way which is completely original since the tono-genesis of U is partly due to vowel length.

In the perspective of Proto-Mon-Khmer, we may have to cope with as many upheavals, during that long stretch of history which separates us from these ancient times, as we can see in the relatively short adventure of Palaungic vowels.

REFERENCES


\textsuperscript{16} In spite of its conservative phonology, Khmu is not as useful for reconstructing Proto-Palaungic as it might appear at first glance; it is true that the Khmuic branch is closer to Palaungic than to other branches of Mon-Khmer, but the Khmu lexicon has undergone a great many replacements.

1980. The Wa languages. (= Ling. Tibeto-Burman Area 5(2), 1-182.)


Flatz, G. 1970. The Khalo or Mae Rim Lawa, a remnant of the Lawa population of Northern Thailand. J. Siam Soc. 58(2), 87-104.


Palaungic vowels in Mon-Khmer perspective


COMMUNICATIVES, EXISTIVES, AND STATIVES IN PROTO-SOUTH-BAHNARIC

David Thomas

0. Introduction
This paper is a study of some communicative, existive, and stative clause types in South Bahnaric languages, comparing them, and postulating some Proto-South-Bahnaric (PSB) forms. The data from the various languages, as will be apparent from the discussion, are uneven in both quantity and quality, so that the present paper must be considered preliminary.

The South Bahnaric languages are the southern section of the Bahnaric branch of Mon-Khmer (Thomas & Headley 1970), located mostly in southern Vietnam, with some spilling over into Cambodia. I draw most heavily on Chrau, Eastern Mnong (Rlâm), and Stieng, as representative languages of the group, with additional data from Köho Sre and Central Mnong (Bunâr and Preh). The three main languages above are respectively at the south-eastern, north-eastern, and western edges of the South Bahnaric area, so should give a fairly good picture of the range of diversity.

In the examples, words whose main significance seems to be as functors rather than as content words are underlined. Vocabulary items, mostly nouns and adjectives, whose meaning is not basic to the structure of the clause, are glossed beside their first occurrence. Functors and central verbs, i.e. elements that are basic to the clause structure, are listed and glossed below the set of examples. It would be desirable to list other verbs that take the same structure, but in most cases I am limited to the published data sources.

In the reconstructions, an agreement of Chrau, Rlâm, and Stieng is taken as sufficient evidence to reconstruct it for Proto-South-Bahnaric. An agreement of just Chrau and Stieng is also considered sufficient if there is no contrary evidence.

1. Communicatives
The talking – perceiving – quoting – informing group of clause types have a basic Speaker—V—Addressee—Information order in all the daughter languages.

1. This is a companion article to “Some Proto-South-Bahnaric Clause Grammar”, paper delivered to the 18th Sino-Tibetan Conference, Bangkok, 1985. (Mon-Khmer Stud. 15, 1989, 111-24). That article dealt with transitivity and locational clause types. The numbering of the examples here follows on the previous numbering.

The clause presentation is based on my clause components outline (Thomas 1983: 137-42). It is semantically based, looking for and comparing the forms which manifest the desired meanings.
DAVID THOMAS

Mnong Bunär (= B)

B3a:  göp ngööi ngach ‘I speak fast’ (Phillips 1963 = MLC. 1.1) (ngach ‘fast’)

B3k:  göp ntaay an naao BuNoong ma khön ay ‘I will teach the Mnong language to you’ (MLC.4.2) (naao BuNoong ‘Mnong language’, khön ay ‘you f.pl.’)

ma ‘to’
ngööi ‘speak’
ntay an ‘teach’

Chrau (= C)²

C3a:  aŋ ṭaai (yuur yuurr) ‘I speak (slowly)’
C3b:  aŋ ṭaai siq neh ‘I talk about him’ (neh ‘him’) aŋ ṭaai sinlō ńi heq ‘I talk about this house’ (ńi heeq ‘this house’) aŋ naai siq sinlō ńi heeq ‘id.’
C3c:  aŋ ṭaai bay neh ‘I spoke to him’
C3d:  aŋ chiith neh ‘I scolded him’
C3e:  aŋ ṭaai paān neh saaq ‘I said he went/I said “He went”’
C3f:  aŋ paān neh saaq ‘I said he went’ aŋ paān, neh saaq ‘I said, “He went”’

2. The Chrau data are my own (see Thomas 1971); the Köho Sre data are from Evans & Bowen n.d. (indicated as KLC) and Manley (OSS); the Mnong Bunär data are from Phillips Ms. (MLC); the Mnong Preh data are from Phillips & Kem (1974; CMLL); the Mnong Rläm data are from Tang (1976; MLLL), plus personal communications from Evangeline Blood (1985; unmarked), and the Stieng data are from Miller 1976 (OSG), plus personal communications from Ralph Haupers (1985; unmarked). I was not able to recheck any of the data with native speakers.

Because of varied spelling conventions used in the different sources, I have standardised the writing of length as VV, the voiceless velar stop as k, and the final glottal stop as q. The ‘whiskered’ œ and œ are rendered here as ō and ū. The Köho o with the lowered dot is rendered ō.

Four different spelling systems have been used for Köho in the past (Manley 1972:39), so I have converted the data from the different sources to the so-called SIL system, as it more closely matches the spelling of the other South Bahnaric languages.

In some cases I have taken the liberty of replacing nouns and place names with other nouns and places names for reader ease.

In the original sources shortness/length is marked as follows:

<table>
<thead>
<tr>
<th></th>
<th>Usual Markings</th>
<th>Other Markings</th>
<th>Environments</th>
</tr>
</thead>
<tbody>
<tr>
<td>short</td>
<td>long</td>
<td></td>
<td>only short</td>
</tr>
<tr>
<td>Bunär</td>
<td>v</td>
<td>v,e,ö</td>
<td>-i/y, u/o, ū/ö</td>
</tr>
<tr>
<td>Chrau</td>
<td>v, ñ, ü</td>
<td>v, ê, ö, ō</td>
<td>-i/y, u/o</td>
</tr>
<tr>
<td>Köho KLC</td>
<td>vN, vT</td>
<td>vN, vT</td>
<td>-i/y, u/o</td>
</tr>
<tr>
<td>OSS</td>
<td>v</td>
<td>v=vq</td>
<td>-</td>
</tr>
<tr>
<td>Preh</td>
<td>v, ñ</td>
<td>v, ê, ü, ö, ō</td>
<td>-i/y, u/o</td>
</tr>
<tr>
<td>Rläm</td>
<td>v, ü</td>
<td>v</td>
<td>-i/y, u/o</td>
</tr>
<tr>
<td>Stieng</td>
<td>v</td>
<td>vv</td>
<td></td>
</tr>
</tbody>
</table>

A combined phonetic chart of the vowels would be:

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>ü</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>ê</td>
<td>ō</td>
<td>ì, üö, uö are</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>ā</td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td>ŋ</td>
<td>-</td>
</tr>
</tbody>
</table>

30
Clause types in Proto-South-Bahnaric

C3g: añ ñāai bay neh paañ añ saaq 'I told him that I was going'
C3h: añ chiih neh paañ neh saaq 'I scolded him, saying that he went'
C3i: añ ñāai aan neh göt (paañ) añ saaq 'I spoke letting him know I was going'
C3j: añ paañ neh saaq 'I invited him to go'
añ sīr neh saaq 'I invited him to go'
añ aan neh saaq 'I allowed him to go'
C3k: añ padau neh göt troong Chrau 'I'll teach him the Chrau language'

aan 'permit, allow'

aan...göt 'let...know'
bay 'with, to'
chiih 'speak, scold'
göt 'know'
ñāai 'speak'
padau 'teach'
paañ 'say, saying, invite'
sīr 'invite'
siq 'concerning (lit. returning)'

Köho Sre (= K)
K3a: añ (qqq) dōs 'I (don't) speak' (Manley 1972 = OSS.217)
añ dōs adaar adaar 'I speak slowly' (Evans & Bowen, n.d. = KLC.2) (adaar 'slow')
K3j: añ jaaq mē saao 'I invite you to eat' (KLC.63) (mē saao 'you eat')
dōs 'speak'
jaaq 'invite'

Mnong Preh (= P)
P3a: gāp ngōoi ngach 'I speak fast' (Phillips & Kem 1974 = CMLL.18) (ngach 'fast')
P3j: gāp jaq may sōöng sa 'I invite you to eat' (CMLL.2 (sōöng sa 'eat')
P3l: gāp nti aan an may git nau BuNoong 'I will teach you Mnong'

an...git 'inform, let know'
jaq 'invite'
ngōoi 'speak'
nti aan 'teach'

Mnong Rlām (= R)
R3a: añ ngōoi brōq brōq 'I am speaking slowly' (brōq brōq 'slowly')
R3b, c: añ ngōoi ta kan 'I spoke to/about him' (kan 'him')
R3f: añ lāh kan saak 'I said he went/ I said "He went"' (saak 'go')
R3g: (laai) añ lāh ta kan añ saak 'I told him I was going'
R3i: añ ngōoi aan kan göt añ saak 'id.'
R3j: añ ndōöm kan saak 'I invited him to go'

31
R3k:  aañ aan kan saak 'I allowed him to go'
aan  'allow, give, let'
lah  'say'
laai 'past' Preferred in R3g.
ndööm  'invite'
ngööi  'speak'
ta  'to, for, etc.'

Stieng (= S)
S3a:  hêy mor (dreet dreet) 'I speak (slowly)'
S3b:  hêy chhuôr baak bu 'I spoke about him' (bu 'him')
S3c:  hêy lah a bu 'I spoke to him'
S3d:  hêy lah bu 'I scolded him'
S3e:  hêy chhuôr lah bu han 'I said that he went/I said "He went'' (han 'go')
S3f:  hêy lah bu han 'I said that he went'
S3g:  hêy lah a bu, lah hêy han 'I told him that I was going'
S3i:  hêy lah aan bu güt (lah) hêy han 'I spoke informing him that I was going'
S3j:  hêy maan/sièr bu han 'I invited him to go'
S3k:  hêy aan bu han 'I let him go'
S3l:  hêy tti bu güt mor Sôdiêng 'I taught him to speak Stieng'
a  'to, for, from, etc.'
aan  'allow, let'
baak  'matter, concerning'
chhuôr  'relate, tell'
güt  'know'
lah  'say, scold, tell'
maañ  'command'
mor  'speak'
sièr  'invite'
tti  'teach'

From the foregoing data one can immediately reconstruct an intransitive talking clause (3a) as Proto-South-Bahnaric *S—V—(Adv.), attested in all six daughter languages. The presence of an adverb or a negative with this construction seems to be preferred.

For 'talking about' clauses (3b), all three attested languages have S—V—Link—Content, but the type of Link varies from a preposition (Rlâm) to a generic noun (Stieng, Chrau) or a motion verb (Chrau). The first Chrau form, with a verbal Link, may possibly be a calque on Vietnamese vê, and it is not attested in my data from the other languages. The second Chrau form, with a nominal Link, is matched by Stieng; it could of course, also be a calque on Vietnamese viêc, but it seems to be a more general South-East Asian pattern, and Stieng has been under less Vietnamese influence than Chrau. The Rlâm form is ambivalent for 3b and 3c. The wide variation in South Bahnaric forms here might seem to indicate some
Clause types in Proto-South-Bahnaric

instability or ambivalence at the proto stage, probably not matching any of the attested current forms.

Addressee clauses (3c in C, R, S) all have the structure S–V–Prep.–Addr., so that structure should be posited for PSB. For the preposition Stieng and Rlâm use their broad-spectrum prepositions a and ta, but Chrau, lacking such a broad-spectrum preposition uses bay ‘with’. Perhaps a broad-spectrum preposition (*ta?) should be posited for PSB.

In Chrau, the communicative verbs have been split into at least four classes. V1 verbs, like ṭaa ‘speak’, dōōm ‘converse’, are used in 3a, b, c, etc., and require a Preposition before an Addressee, and require a Quote Introducer before a Quote. V2 verbs, like chiīh ‘scold’, lah ‘scold’, payoom ‘praise’, are used in 3d, h, and require a Quote Introducer before a Quote. The V3 verb paañ ‘say’ cannot take an Addressee, and it takes no Quote Introducer since that would be homophonous with it. And paañ replaced PSB *lah as the Quote Introducer. V4 verbs, like siēr ‘invite’, paañ ‘say, invite’, are used in 3j.

A transitive talking clause (3d in C, S) *S–V–O should be reconstructed for verbs like ‘scold’ or ‘praise’.

There was apparently no distinction between direct and indirect quotatives (3e, f) in Proto-South-Bahnaric. Form 3e S–V1–QuotInt.–Quot. is attested in Chrau and Stieng. The dropping of V1 gives a simpler form (3f in C, R, S) in which the Quote Introducer (C. paañ, R, S. lah) becomes the main verb *S–VQuot.–Quot. The verb lah ‘speak, scold’ is attested in C, R, S and probably served as both a VQuot. and a QuotInt. in PSB, but its QuotInt. function dropped out in Chrau. All these languages have many verbs that can function as V1, but only one that can function as QuotInt.

Simple quotative addressee clauses (3g) in Chrau and Stieng have the form S–V1–Prep.–Addr.–QuotInt.–Quot., in which V1 and QuotInt. are the same as in the direct and indirect quotatives (3e, f), and the Prep. is that in 3c. Rlâm has S–V1–Addr.–Quot. The modern forms could be accounted for by positing a PSB *S–V1:v/VQuot.: /lah–Prep.–Addr.–QuotInt. :lah–Quot. Stieng seems to have this structure. Mnong Rlâm dropped the redundant second lah. Chrau dropped lah completely out of this construction by substituting paañ in the QuotInt. slot and by putting lah into the V2 class, which does not occur in the 3g construction.

The 3h form, used with Chrau class V2 verbs, does not have a Preposition before the Addressee. Data from other languages are lacking, so no conclusions can be drawn for PSB.

The longer quotative addressee form (3i) has the form *S–V1–aan–Addr.–git–Quot., as attested in C, S, R, with aan...git ‘let know’ functioning as a benefactive marker. In C there is an optional QuotInt. before the quotation; this may be a Chrau idiosyncrasy on the analogy of 3e.

Imperative clauses, with verbs like ‘invite, command, permit’ (3j) have the same form as 3f, i.e. S–V–Addr.–Quot., attested in C, S, R, so may be reconstructed for PSB in that form. In Chrau, the 3f and 3j verbs are
homophonous in one instance, paañ, yielding an ambiguous clause. The causative verb aan ‘give, allow’ can also be used here (C, R, S).

Teaching clauses (3k) have the form S–V–aan–Addr.–git–Content (=3i) in C and P, perhaps reconstructable for PSB. Stieng lacks the *aan. Only Bunar has S–V–Content–ma–Addr.

2. Existives
The simple existence – identified existence – transitive existence (existive possession) – naming – becoming group of clause types take a variety of forms, as may be seen from the following:

_Mnong Bunâr_

B4b: `geh ngîh ta ri ‘There is a house there’ (MLC.2.11) (ngîh ‘house’, ri ‘there’)
B4c: `ta ti jêëng ngîh ‘Over there are houses’ (MLC.2.11)
B4e: `buum jêëng du ntiil ndô sa ‘A tuber is a kind of food’ (MLC.3.2) (buum ‘tuber’, du ntiil ndô sa ‘one kind of food’)
B4f: `göp geh du play ngîh ‘I have a house’ (MLC.2.7) (du play ‘one cl.’)
B4g: `göp NDjréët ‘I am Djréët’ (MLC.4.1)
B4k: `amoh göp NDaan ‘My name is Dan’ (MLC.4.1)

_mnong_ ‘name’

geh ‘have’

jêëng ‘be’

ta ‘at’

_Chrau_

C4a: `geh nî ‘There are houses’ (nî ‘house’)
C4b: `geh nî u heeq ‘There are houses here’ (heeq ‘this, here’)
C4c: `u heeq geh nî ‘id.’
C4d: `nî heeq ‘This is a house’
C4e: `heeq la nî ‘id.’ (rare < Vietnamese)
C4f: `añ geh nî ‘I have a house’
C4g: `GaPe añ ‘I am GaPe’
C4h: `añ heeq GaPe ‘id.’
C4j: `añ saq GaPe ‘I am named GaPe’
C4k: `saq añ GaPe ‘My name is GaPe’
C4m: `añ tan’hya saq neh GaPe ‘I named him GaPe’ (neh ‘him’
C4n: `neh jêëng yaw ‘He became a tiger’ (yaw ‘tiger’)
C4o: `añ ðöp vadaai jêëng/luh nî ‘I made the lean-to into a house’ (vadaai ‘lean-to’)

geh ‘be, have’

heeq ‘here, this one’

jêëng ‘become’

la ‘is’ (< Vietnamese)

luh ‘appear, become’

ðöp ‘make, do’

saq ‘name, be named’
Clause types in Proto-South-Bahnaric

tan'hya

'so to name'

Koho Sre

K4a: gōs hiu 'There are houses' (hiu 'house')
K4b: gōs hiu téeng dō 'There are houses here' (KLC.9) (dō 'here'
K4e: chi dō la hiu 'This is a house' (KLC.26) (chi dō 'this')

khay lahjēeng caw mih 'He is an American' (OSS.189)

K4f: aī gōs kroac 'I have oranges' (KLC.14a) (kroac 'orange')
K4h: chi dō kroac 'This is an orange' (KLC.14)
K4l: sōndan aī la K'Poh 'My name is K'Poh' (KLC.14)
K4n: khay jēeng/gōs kliu 'He became a tiger' (OSS.189-190) (khay

'he', kliu 'tiger')

gōs 'be, have, become'

jēeng 'become'

la, lah 'is' (< Vietnamese)

sōndan 'name'

tēeng 'at'

Mnong Preh

P4c: tām bri geh ne 'In the forest there are rats' (CMLL.28) (bri

'forest', ne 'rat'

P4e: gap jēeng BuNoong 'I am a Mnong' (CMLL.16)

P4f: gap geh pe nuyh koon 'I have three children' (CMLL.8) (pe nuyh

koon 'three children')

P4k: moh sak gap NDoong 'My name is Dong' (CMLL.7)

geh 'have'

jēeng 'be'

moh sak 'name'

tām 'in, at'

Mnong Rlām

R4a: geh hih 'There are houses' (hih 'house')

R4b: geh hih ta han 'There are houses there' (han 'there')

R4c: to car Mriik mau hih 'In America there are houses' (MLLL.36)

(car Mriik 'America')

R4d: hih hō 'This is a house' (hō 'this')

R4e: hō eh jēeng hih 'This is a house' (eh 'this')

R4f: aī geh hih 'I have a house'

R4g: Jhang aī 'I am Jhang'

R4i: aī jēeng Jhang 'id.'

R4k: nan aī Jhang 'My name is Jhang'

R4m: aī nan to kan Jhang 'I named him Jhang'

R4o: aī mhoq tüm hin njeeng hih 'I made that shelter into a house'

(tüm hin 'shelter')

geh 'have'

jēeng 'be'
A simple existence clause *ExistV–S (4a) may be posited for Proto-South-Bahnaric on the evidence of C, K, R, S. And a PSB existence verb *geh is attested by B, C, (K?), R, S.

A located existence clause ExistV–S–Prep.–Loc. (4b) is attested by B, C, K, R, S with a demonstrative Loc. The locative preposition, however, is different in all five languages, leading one to suspect that perhaps PSB had no preposition there; but the need for a preposition (in a preposing language group) became felt, perhaps to avoid ambiguity with a homophonous N-Dem. noun phrase. Bunăr, Chrau and Preh data (4c) also include a transposed Prep.–Loc.–ExistV–S form, emphasising the Location rather than the Subject, which should probably be posited for PSB. Bunăr has different ExistV in 4b and 4c.

The simplest identification clause (4d) is Ident.–S, attested in C, R(?), or its reverse S–Ident. attested in S. It is not clear what should be reconstructed for PSB.

A copula-linked identification clause (4e) S–Cop.–Ident. is found in B, K, P, R. It is absent from Stieng and only borrowed in Chrau, the two most reliably attested languages in the sample. (Calques from Vietnamese or English are very possible in the language lesson books, the sources for most of the other language data.) The copula jeeng used in B, K, P, R, however, is a verb of becoming in Chrau (4n) and Rlâm (4o), so perhaps
PSB should be reconstructed without a copular identification clause form, then the northern tier of daughter languages expanded the use of jëëng ‘become’ to provide a copula.

A possession clause (4f) *S–Poss.–V–Item is clearly reconstructable from B, C, K, P, R, S, and the Possessive Verb in all of them is the existence verb *geh/gös of 4a, b. In PSB, apparently, existence and possession were parts of a single semantic category; perhaps possession should be viewed as transitive existence. (This use of have/be is paralleled by Vietnamese có, Thai mii, Khmer mii-an, and many other South-East Asian languages.)

The personal identification clause (4g) is like the simple identification clause (4d), with Name³-S attested in C, R, and S–Name in B, S. Reconstruction is not clear. A topicalized form *S–Dem.–Name (4h) is more widely attested in C, K, S and should be reconstructed for PSB. Râm has also a copular form (4i) S–Cop.–Name, not to be reconstructed for PSB (see 4e).

A name clause (4j, k) is semantically very close to the personal identification clause. A form *S–NameN–Name (4j) with sak ‘body, name’ as the Name Noun can be reconstructed from C and S. And an alternate form *NameN–S–Name (4k) can be reconstructed from B, C, P, R, S. The nan ‘name’ in R is probably a loan from Rade. Köho also has a copular form (4i) with the borrowed Vietnamese copula là.

The naming clause (4m) has different forms in C, R, and S. The history of these is not clear.

A becoming clause (4n) *S–jëëng–O is probably reconstructable from C, K (see 4e). There is no outside support for the Köho variant S–gös–O, which is homophonous with the possession clause (4f).

The transforming clause (4o) has the form *S–V₁–Form₁–V₂–Form₂. C and R have jëëng as V₂, probably reconstructable for PSB. C can also have !uh as V₂, but S has !uh as V₁. PSB *luh ‘go out’ must have been part of the PSB semantic field of transforming, but its syntactic use is not clear.

3. States
The ambient – stative – comparative – superlative – evaluative group of clause types tends to be S–V in South Bahnaric languages, though V–S state clauses are quite common.

Mnong Bunär
B5a:  
naar aaö geh mih ‘There will be rain today’ (MLC.4.2) (naar aaö ‘today’, mih ‘rain’)

B5b: 
klaang boôk naar jëh ‘It is noon already’ (MLC.2.2) (klaang boôk naar ‘noon’)

B5c: 
naar aaö ji kat ‘Today it is cold’ (MLC.4.2) (ji kat ‘cold’)

B5d: 
göp ji ngoot ‘I’m hungry’ (MLC.2.2) (ji ngoot ‘hungry’)

geh ‘have’

jëh ‘already, now’

3. Structurally ‘name’ is the head of the Subject noun phrase, but semantically ‘I’ is still the Subject, as in the two preceding forms.
DAVID THOMAS

_Chrau_
C5a:  kô mi ‘It is raining’ (kô ‘sky’, mi ‘rain’)
C5b:  nar tamvôôp een ‘It’s noon now’ (nar ‘day’, tamvôôp ‘middle’)
C5c:  nar heeq takat ‘Today is cold’ (heeq ‘this’, takat ‘cold’)
C5d:  añ takat ‘I’m cold’
C5e:  takat añ ‘I’m cold’ (more common)
C5f:  añ takat añ ‘id.’ (emphasising ‘I’)
C5g:  añ luh takat ‘I became cold’
C5h:  añ takat dóong neh ‘I’m colder than him’
C5i: neh êeq takat ka añ ‘He’s not as cold as me’
C5j:  añ takat dóong leq ‘I’m the coldest of all’

dóong  ‘more than’
een  ‘already, now’
êeq  ‘not’
ka  ‘like, as’
leq  ‘all’
luh  ‘become, appear’

_Köho Sre_
K5b:  guul ngai rau ‘It is noon already’ (KLC.11) (guul ngai ‘noon’)
K5c:  ngai dô noat ‘Today is cold’ (KLC.38) (noat ‘cold’)
K5d:  añ kqopp ‘I’m sick’ (KLC.22) (kqopp ‘sick’)
K5g:  khay gôlôh koop ‘He became sick’ (OSS.190)
K5i:  gö qqq niam be chi dô ‘They are not as good as this one’
      (KLC.33) (niam ‘good’, gö ‘they’)
K5j:  chi dô buôn rlaü jôh ‘This one is cheapest’ (KLC.33) (buôn ‘sell’)
      be  ‘as, like’
gôlôh  ‘become’
qqq  ‘not’
rau  ‘already, now’
rlau jôh  ‘most, superlatively’

_Mnong Preh_
P5a:  bri mih ‘it is raining’ (CMLL.24) (bri ‘jungle’, mih ‘rain’)
P5c:  naar aâô duh (ngan) ‘Today is (very) hot’ (CMLL.23) (nar aâô
      ‘today’, duh ‘hot’)
P5d:  rpual prah joong ‘A melon is long’ (CMLL.29) (rpual prah
      ‘melon’, joong ‘long’)
P5h:  gâp kataang lôön ma may ‘I am stronger than you’ (CMLL.23)
      (kataang ‘strong’) lôön ma  ‘more than’

_Mnong Rlâm_
R5a:  mih ‘It is raining’ (mih ‘rain’)
R5c:  naar o kôkât ‘Today is cold’ (naar o ‘today’, kôkât ‘cold’)
R5d:  añ kôkât ‘I’m cold’
Clause types in Proto-South-Bahnaric

R5e: kökat a̱i ‘id.’
R5g: a̱i jèéŋ kökat ‘I became cold’
R5h: an kökat hin ta kan ‘I’m colder than him’
R5i: kan han ay so kökat blah a̱i ‘He is not as cold as me’
R5j: a̱i kökat hin ta leq möt nih ‘I am the coldest of all’

ay so ‘not’
blah ‘as, like’
han ‘not’ (?)
hin ta ‘more than’
jèéŋ ‘become’
leq möt
nih ‘all, everyone’
sò ‘see, perceive’ (?)

Stieng

S5a: mi ‘It is raining’ (mi ‘rain’)

mi lôh ‘id.’
S5c: ‘lèék ‘It is cold’ (‘lèék ‘cold’)
nar ‘lèék ‘id.’ (nar ‘day, sun’)
S5d: hêy ‘lèék ‘I am cold’
S5e: ‘lèék, hêy aq! ‘Cold, indeed I am!’
S5f: hêy ‘lèék hêy ‘I am cold’
S5g: hêy lôh ‘lèék ‘I became cold’
S5h: hêy ‘lèék huôs bu ‘am colder than him’

hêy teq a bu ‘lèék hêy ‘I am colder than him/Beside him I’m cold’
S5j: let pal nêy, ‘lèék hêy ‘I am the coldest of all/Of all of them I’m coldest’

stative = ± S + P:Vi (OSG.9)
aq ‘exclamation’
biit ‘become’
huôs ‘than, more than’
leet pal
nêy ‘completely, all of them’
lôh ‘appear, become’
teq a ‘place beside, compare with’

There is no agreement on the form of the simple ambient clause (5a). Köhö, Rliam, and Stieng have a simple Amb. structure. Chrau and Preh use a dummy subject DumS–Amb. structure. Stieng can use a dummy verb Amb.–DumV structure. And Bunær uses a geh–Amb. existive (4a) structure. This variety could perhaps be explained by positing a PSB simple *Amb. as in K, R, S. Chrau (which has been in close geographical proximity to Vietnamese) and Preh adopted a Vietnamese-like structure, treating the Ambient as an intransitive verb. The Bunær form and the alternate Stieng form treat the Ambient as a noun. Positing a simple proto *Ambient would most easily account for the verb and noun developments. Nominal use of
the Ambient does appear also in Chrau in the form *mi sa neh ‘he was heavily rained on’.

A time clause *Time–Adv. form (5b) can be reconstructed on the evidence of B, C, K.

A time-located ambient *Time–Amb. form (5c) is reconstructable from B, C, K, P, R.

A stative *S–State (5d), which is similar in form to the intransitive (1a), is reconstructable from B, C, K, P, R, S.

A reversed stative form *State–S (5e), emphasising the state, is reconstructable from C, R, S. The Stieng form seems to be more emphatic than the Chrau, and is normally accompanied by an emphatic final particle. An echo form *S–State–S (5f) is reconstructable from C, S, giving mild emphasis to the Subject.

An inceptive state *S–IncepV–State (5g) is reconstructable from C, K, R, S. The Inceptive Verb is *uh in C, *jeeng in R, and *gölös (from *gös ?) in K.

The Stieng form seems to be more emphatic than the Chrau, and is normally accompanied by an emphatic final particle. An echo form *S–State–S (5f) is reconstructable from C, S, giving mild emphasis to the Subject.

4. Summary of reconstructions
The reconstructed Proto-South-Bahnaric forms may be summarised, with sample glosses, as below. Binomial slot : filler formulations are given when both the functional slot and the actual filler were discussed.

Communicatives

3a: *S–V–(Adv.) ‘I speak (slowly)’
3b: *S–V–Link:?–Content ‘I spoke about him’
3c: *S–V–Prep.:ta–Addr. ‘I spoke to him’
3d: *S–V–O ‘I scolded him’
3e: *S–V1–QuotInt.:lah–Quot. ‘I spoke saying he went’
3f: *S–VQuot.–Quot. ‘I said he went’
3g: *S–V1:v/VQuot.:lah–Prep.–Addr.–QuotInt.–Quot. ‘I told him that I was going’
3h: ??
3i, k?: *S–V1–BenV:aan–Addr.–BenV:git–Quot. ‘I let him know that I was going’
3j: *S–V:v4/aan–Addr.–Quot. ‘I invited him to go’

Existives

4a: *ExistV:geh–S ‘There are houses’
Clause types in Proto-South-Bahnaric

4b: *ExistV:geh-S-(Prep.:?)–Loc. ‘There are houses there’
4c: *Prep.–Loc.–ExistV–S ‘Over there are houses’
4e: (PNSB *S–Cop.:jēēng–Ident. ‘That is a house’)
4f: *S–PossV:geh–Item ‘I have a house’
4g: ??
4h: *S–Dem.–Name ‘I here am GaPe’
4i: –
4j: *S–NameN:sak–Name ‘I am named GaPe’
4k: *NameN:sak–S–Name ‘My name is GaPe’
4l: –
4m: ??
4n: *S–V:jēēng–O ‘He became a tiger’
4o: *S–V1–Form1–V2–Form2 ‘He made the lean-to into a house’

Statives
5a: *Amb. ‘It is raining’
5b: *Time-Adv. ‘It is noon now’
5c: *Time-Amb. ‘Yesterday it rained’
5d: *S–State ‘I am cold’
5e: *State–S ‘I am cold’
5f: *S–State–S ‘I am cold’
5g: *S–IncepV:?–State ‘I became cold’
5h: *S1–State–CompMk:?–S2
5i: *S1–Neg.–State–CompMk–S2 ‘I am not as cold as he’
5k: ??

REFERENCES


41
A LINGUISTIC ANALYSIS OF SOME SOUTH MUNDA KINSHIP TERMS, I

Norman H. Zide & Arlene R. K. Zide

Part 1

1. In this paper we present etymologies for a number of South Munda kinship terms. South Munda (SM), one branch of Proto-Munda (PM), branches into Kharia-Juang (KJ) or Central Munda, and Koraput Munda (KM); KM branches into Sora-Juray-Gorum (SJG) and Gutob-Remo-Gta? (GRG). While we have not done anthropological analyses of the Munda kinship systems, we have made use of the anthropological sources in defining and relating kin terms and kin-term sets of the languages and proto-languages.

Our objective has been to provide a linguistic analysis of the SM kin-terms, with a view to reconstructing as much of the original (i.e. non-borrowed) terminology as possible, and to integrate the results into as coherent and plausible a system as we can. Since KM and SM noun morphology have not been analysed and described, we will present a description of SM noun morphology with particular reference to the derivational morphology of full forms (FF), and combining or compositional forms (CF) of Munda kin terms. This will come in the second paper in this series, along with the first full sets of kin-term etyma.

In the first section of this paper we take up and criticise the work of Bhattacharya (1970) and Parkin (1985) on Munda kin terms, and in the second section we discuss how relationship and reciprocity are shown in Munda kin terms.

We have taken our data, the SM kin terms and definitions, from a wide range of existing sources, and from our own fieldnotes. The anthropological sources are fuller in the coverage of the terms and their uses, but are poorer in linguistic transcription, and lack morphological analysis. The linguistic sources provide better linguistic data, but are incomplete and, for purposes of kinship term analysis, poorly organised. A few publications (e.g. Deeney, 1975, on Ho) are exceptional in presenting the kinship system and the kin terms fully and perspicuously in linguistically well-analysed form.

---

1. We use the following abbreviations in describing kin terms: M, 'mother'; F, 'father'; P, 'parent'; Br, 'brother'; Si, 'sister'; Sib, 'sibling'; Y, 'younger'; O, 'older'; Hu, 'husband'; Wi, 'wife'; Sp, 'spouse'; So, 'son'; Da, 'daughter'; Ch, 'child'.

The abbreviations for the modern language names are: Sa, 'Santali'; Mu, 'Mundari'; Kher, 'Kherwarian'; Kh, 'Kharia'; Ju, 'Juang'; So, 'Sora'; Go, 'Gorum'; Jr, 'Juray'; Gu, 'Gutob'; Re, 'Remo', and Ga, 'Gta?'. The more common ethnonyms for the KM languages are: Sora: Saora, Savara; Juray: Juray Sora; Gorum: Pareng, Parenag, Parenji; Gutob: Gad(a)ba, Gutob Gadba; Remo: Bonda, Bondo; Gta?: Diayi, Diayei, Dire.
There has been widespread borrowing of kinship terms into SM: from languages as diverse as the Dravidian Ollari Gadba, and Indo-Aryan Kotia Oriya, standard Oriya, and standard and dialectal Hindi, Bhojpuri and Bengali, and English. Undoubtedly earlier borrowings, particularly from Dravidian, have been missed by us. We hope, for a later paper in this series, to collaborate with a Dravidianist on Dravidian borrowings, early and late, of both kinship behaviour and kinship words. This paper does not deal with kinship terms that are not genetically old in Munda, and that are not likely to have cognates in Mon-Khmer (MK). We propose to present the full set of Munda kin terms with Mon-Khmer (and, possibly, Austronesian) cognates in a later paper.

The only published general treatments of Munda kin terms are those of Bhattacharya (1970) and Parkin (1985), and the relevant sections of Pinnnow’s Versuch (1959), as updated by him in 1960 in his unpublished monograph on Juang. We show that Bhattacharya’s weighting of semantic similarity at the expense of sound correspondences leads him to lump together forms that are semantically very close, but not cognate. Parkin, an anthropologist who has written a dissertation on Austroasiatic kinship, needs to use linguistic evidence. He makes use of Bhattacharya’s material, but is sometimes misled by Bhattacharya’s methods of analysis and presentation of data. Parkin also takes rough phonetic similarity as indication of genetic relationship, and uses the—usually spurious—‘cognition’ to support anthropological arguments about kinship. He sometimes ignores Bhattacharya’s conclusions (e.g. about bare (Bhattacharya’s Set 66. (1970: 455), Ga bare; ZZ *boHre, ‘a woman’s brother’) and mistakenly connects reflexes of this *boHre (e.g. bo?re, bok’re) with *bokɔ, YSib. By doing this he misses the importance of PM *boHre and the existence in PM of terms restricted to male or female speakers, e.g. GRG *bo(b)re, ‘a woman’s brother’, *bVlon, ‘a woman’s sister’, *bVyan, ‘a man’s brother’, and *tonan, ‘a man’s sister’. We need the anthropologist to make kinship sense of the meanings of cognate sets where we cannot reconstruct a properly precise meaning for the PM term, much less account for the change in meaning in NM and SM and the modern languages. The example of Northern Munda (NM) *hili, OBrWi, and KM *hVli SpYBr will be discussed in some detail later in the paper. Bhattacharya has confused the issues by putting NM *hili and KM *hVli into different semantic-cognate sets because their meanings are not close enough.

We should note that anthropological studies of Munda kinship have flourished in the past fifteen years; we are thinking of the work of S. Bouez, Deeney, Parkin, Pfeffer and Vitebsky. However, most of these...
analyses do not bear directly on the linguistic problems with which we are concerned, and no further reference will be made to most of these.

Bhattacharya’s survey paper is very useful in that it presents and begins to organise his own rich field data. He is less thorough in abstracting the published literature. His analyses, his semantic-cognate sets, observations on borrowings, etc., are usually perceptive. In some cases he is more conservative than he needed to be, e.g. in putting Ho *haam*1 ‘old man, husband’ in one set, and the reflexes of GRG *hV-n-čam* (e.g. Ga *handa*, etc.) in another. Had he noted Mundari *haram* (same meaning as the Ho and the GRG), he would have been compelled to put all these forms in one semantic-cognate set. In other etyma, particularly where he brings in possible cognates in Mon-Khmer, he is too indiscriminately accepting. (But then, he is the only Indian Munda scholar who has paid any attention to Mon-Khmer.) Bhattacharya knows several of the Munda languages, and he knows Indo-Aryan and some Dravidian. Parkin, when he disregards Bhattacharya, usually goes wrong.

Parkin is conscientious in examining long lists of kinship terms, and he turns up interesting forms not earlier appreciated (e.g., Remo *N-kwi*, YSi, which Bhattacharya also records). He can be perceptive about borrowings, e.g. noting that KM *mama*, MBr (in some languages SpF), is borrowed (along with kin behaviour) from Dravidian, and not directly from Indo-Aryan. But his use of rough phonetic similarity as evidence of cognation is self-defeating. We do not quarrel with Parkin’s anthropology. It is true, as Parkin says (1985: 705) that the absence of studies of comparative Munda kinship constitutes ‘a major gap in south Asian studies’, and that Parkin has done a considerable amount of serious work in filling that gap. We give three examples of the sorts of things we object to:

(1) his Table of ‘Basic forms of typical NM and SM terminologies’;
(2) *boko* and *bare*; and
(3) *e1Jga*, ‘mother’.

After this we take up the confused (mostly by Bhattacharya) case of Juang *ini-bo*, HuYBr.

In his Table, Parkin gives ‘typical’ NM and SM terminologies, and includes the terms for SpF, SpM, ChCh, same sex SibCh (of the same sex as Ego), and FSibCh of same sex as Ego, opposite sex SibCh (of opposite sex to Ego) and of PSibCh, BrWi, HuSib, SiHu and WiSib, etc. What is a ‘typical’ terminology, and how and why does one compare typical terminologies? ‘Typical’ would seem to mean ‘well-described’, or ‘well-known’. The NM and SM forms given in the table are of little value. To a linguist, the forms to be compared would be the reconstructed NM and SM forms. These turn out to be closer to each other than Parkin’s typical forms from typical modern languages—when they are cognate. NM

---

4. We have—without marking these forms—retranscribed some of Bhattacharya’s transcriptions in order to indicate morpheme boundaries, and to prune dubious and excessive phonetic detail. We will give Bhattacharya’s transcriptions and comment on them in the full data sets in later papers in this series. The form transcribed by us *haam* was transcribed *ham* by Bhattacharya, and *hām* by Deeney.
*kankar and NM *kinkar, SpM, are fairly similar whereas Parkin’s Mu hanar and (KM ?) *kinar are less so—perhaps because Korku, Kharia and Juang, which reserve PM *k, are not typical enough. Parkin’s jai ‘grandchild’ (in Mundari-Ho) was selected as the representative NM form because he wanted to connect it with jia, ‘grandmother’. The connection is dubious. A better word for NM ‘grandchild’ than jai, found only in the Kherwarian branch of NM, would be kVrar/kVrat, found in both branches (Korku kurar, Santali korar). The data can be found in Bhattacharya.

2. Both in the table and elsewhere Parkin notes the important distinction between sibling terms where one must know whether the speaker is of the same or opposite sex to the kin-term referent. What seems important—to generalise the case—is not same or opposite sex, but whether the set of terms is sensitive to speaker’s and referent’s sex. That is, we have in KM a set of four terms where this feature is marked (‘a woman’s brother’, ‘a woman’s sister’, ‘a man’s brother’, ‘a man’s sister’). This set contrasts with another set of terms where this feature is absent, but where we mark relative age of the speaker and referent, i.e. OBr, YBr, OSi, and YSi. Parkin’s boko belongs to this second set (and perhaps should be defined not YBr, but YSib) and the others of the sibling terms in his table tonan, bokre (from *boHre, Bhattacharya’s bare set) and misi belong in the first set. It is possible (as the NM data suggest) that there were only two sex-sensitive terms in PM in this first set, and these were both opposite-marking terms, but this needs to be demonstrated. Parkin takes bokre as cognate with boko, and not with bare, *boHre which leads him to miss the one cognate in PM of the sex-sensitive set. Since such terms go back to PM (however many may have to be reconstructed), we want to know what the distinctive functions of these two sets of sibling terms were—in earlier times and protos—and what they are now. None of the anthropologists we have read takes up this important problem.

If one wants to use linguistic evidence in kinship arguments, then it is necessary to be able to recognise the historical depth of an etymology. *boHre can be reconstructed for PM; this is noteworthy. It is possible that *boHre at some more distant level—perhaps PAA—can be shown to have a morpheme in common with *bokə, but good MK evidence of this would be needed to make such a case. One has to be able to recognise that *V?(g)-leŋ is a good KM etymology for ‘grandchild’, but not a good SM or PM etymology.

3. Lastly, Parkin’s treatment of ‘mother’. Parkin writes that ‘the standard NM and KM term for “mother” is enga... ’. enga is only standard in one branch of NM, Kherwarian. If one wants an etymological formula for PM ‘mother’, it would be V-ya-N, V-yə-N. Korku has ayom in one dialect, the more archaic ayan in others; the reduced form of this before -te? is an-; Santali has enga and ayo/ayə; Mundari-Ho has enga, but
**South Munda kinship terms**, I

e(y)an in the vocative (we do not find much morphological alternation of this sort in Munda); Juang has bwi-N, Kharia has ayo/ayo-N, Sora has yan, Juray ayan and yon, Gorum has yan, Remo yon, and iyon, and Gta? yan. The basic form is the V-ya-N/V-yo-N. The question is how to relate enga and e(y)an (vocative) in Ho-Mundari. Deeney has anticipated us in pointing out (Bhattacharya and Pinnow have made partly similar observations) that the vocative e(y)an resembles enga-ii/eya-ii, i.e. the non-vocative stem with the first person singular pronoun -ii, i.e. ‘my mother’. This suggests that most of the Munda forms for ‘mother’ were originally vocatives with first person pronominal enclitic (of possession) -ii-ŋ, and that enga represents (how accurately?) the basic form of ‘mother’. How we get from enga to eya/ena where these forms precede the enclitic -ii-ŋ has just been shown in Deeney’s data. One could suggest metathesis, common in Munda, but we have no (other) cases of eya/eya, y/n, metathesis. We reconstruct V-ya which becomes V-ŋa before final nasal in the vocative. To take enga, ayo, ean, etc. as obviously cognate is risky (it was not wrong). In any case, ‘enga is not the standard form of “mother” in NM and KM’. The point of these criticisms is not that anthropologists should reconstruct linguistic proto-forms, but that they should be less free in identifying putative cognates, and using these—largely spurious—cognates to support other arguments.

 ini-bo or ini-bau/-boi. Our first comment on Parkin’s rejection of Bhattacharya’s connecting ini-bo and hVli was that Bhattacharya’s judgment on cognition was better than Parkin’s, and whether the semantics agreed with Parkin or not, ini and hVli were almost certainly cognate. A re-examination of Bhattacharya’s sets 81. and 83. (1970: 457) shows the situation to be more complicated. Bhattacharya misleads his readers by setting up two semantic-cognate sets, and putting NM *hili in the second (with Juang kuli), and KM *hVli in the first. (Bhattacharya does not think with or use reconstructed form; we have constructed ‘Bhattacharya’s reconstructions’ for him.)

The facts are these: Bhattacharya has made two semantic-cognate sets, 81. and 83., these coming in his section of ‘Terms for Brother, Sister, Brother-in-Law, Sister-in-law’, one of his more complicated—and important—sets of terms. In 83. he puts Juang kuli/koli, OBrWi, and Kherwarian (there is no Korku cognate), *hili, OBrWi. In 81. there are three sets of words: 5 (a) KM *hVli(-boj) (ZZ reconstruction), Vli-boi, HuYSi, SpYSi; (b) KM *Vrel/*Vrer, HuYBr, SpYBr, and (c) NM *erwel/*Vrel, HuYBr. Bhattacharya does not sort these into sets, so that it is not clear if he thinks Juang ini-bo goes in set (a), or with sets (b) and (c). The latter two sets clearly reconstruct to PM *Vrel.6 The meanings of the 83. forms are fairly close: (a) HuYSi, SPYSi, and (b) – (c) HuYBr, SpYBr. ini-bo means HuYBr. Since 83. (a) is clearly cognate with 81. which includes Juang kuli, to relate ini to the augmented 83. we either have to disqualified and reject kuli, somehow to relate kuli and ini, or to connect ini with 81. Can we—as Bhattacharya’s array of data suggests—

---

5. See the Appendix for fuller presentation of the data.
show *ini* to be a cognate of PM *Vrwel, HuBYr, SpYBr? We need to show that the sound correspondences are possible: that Juang *n* corresponds with *r* (or something like *r*); this is possible; that the first vowel *i* corresponds to *V*, where the following, stressed, vowel is *e* and *i* in Juang: this is possible; the correspondence of Juang stressed *i* and PM *e*: this is possible; and that a final consonant, Juang *l*, can be lost compound-medially: this also fits into the set of expectable, regular, correspondences. We accept, tentatively, *ini* as cognate with PM *Vrwel, and rearrange Bhattacharya’s sets as follows: 81. (b), (c) plus *ini*, HuYBr, SpYBr, and 83. plus 81 (a), *kuli(-boj), OBrWi, HuYSi, SpYSi.

**Part II**

This second section takes up:

1. (once more) the two different sets of sibling terms in a number of the Munda languages;
2. the old system of pronominal enclitics (ProP) marking inalienable possession. The kin terms of Kherwarian that do not take ProP are: (i) name-like kin terms, and (ii) conjoined pairs of terms whose referents are related to each other, not to the speaker or some other named or pronominally indicated person. The latter, paired reciprocal terms, are common in Munda, both north and south. Reciprocal infixes in kin terms (the infix is the same one found as the verbal reciprocal marker) are commonly used where each of the referents of a hypothetical pair refers to the other by the same kin term. The various meanings of this reciprocal infix, NM -*p*, SM -*m*, are examined in several Munda languages.
3. reciprocal (kinship) terms of address are examined in Santali, Juang and Sora.
4. Some kin-term affixes that look like but are not reciprocal -*p/-m* are described, and traced to PAA.

We noted earlier that in some Munda languages—Remo is the best example of one with two complete sets—there are two different sets of sibling terms, one marking relative age (e.g. ‘younger brother’, ‘older sister’), and the other marking ‘relative sex’ of speaker and referent (i.e. ‘a man’s brother’ versus ‘a woman’s brother’). Relative age is commonly marked in the Indian area; relative sex is not. So far as we know, none of the anthropologists who have worked on Munda groups with two (complete or incomplete) sets of sibling terms has described the distinctive functions of these sets. Parkin does note these terms, but in his table he distinguishes between same-sex (of speaker and referent) and opposite-sex terms, whereas we see the basic difference as between sex-marked (same or different sex) terms, and age-marked terms. The GRG languages have two full (four-term) sets, but it is not clear that cognate forms of the sex-marked set in Sora (e.g. GRG *bVyan, So boñap, GRG *tonan, So tonan, these meaning ‘a man’s brother’ and ‘a man’s sister’ respectively in GRG)
have the same meanings and sort into similar sets in Sora. In Kherwarian
the sex-marked set has only two terms, both marking opposite sex kin.
The history of these sex-marked terms is still largely obscure. But in any
case the distinction is noteworthy, and an understanding of the functions
of the two sets should be important to the anthropologist studying
Munda kinship.

When studying the Munda languages that preserve the old system of
pronominal enclitics for marking inalienable possession (e.g. Ho, Santali,
Juang, Kharia), we find different definitions of grammatical inalien-
ability. The minimal set of inalienably possessible nouns would be the kin
terms, the next minimal set would add words for body parts. The
pronominal enclitics, Pro\textsuperscript{IP}, are required with almost all inalienably
possessible nouns, with a few exceptions. The common exceptions are: (a)
‘name-like’ terms; and (b) paired reciprocal terms where the relationship
obtains between the referents of the pair. These paired reciprocal kin
terms are common in Munda.

One Mundari example from Hoffmann (1950: 1303) of a name-like
term has to do with the speech of siblings (i.e. those who have a common
referent for ‘mother’ and ‘father’). In this context—siblings speaking to
each other—one cannot say ‘my mother’ without the (erroneous and
offensive) implication that she is not also ‘your mother’. What is said—
and is preferable to using a Pro\textsuperscript{IP} for first person dual inclusive—is simply
‘mother’ (perhaps ‘Mother’ would better represent this), i.e. the form
\textit{en\textgreek{a}} without a Pro\textsuperscript{IP}.

Paired reciprocals of one form or another are characteristic of several of
the Munda languages, North and South. The kinds of pairs found are:

(i) the senior term occurs followed by the ‘collective’ suffix -ya/-ea; in
Santali, e.g. \textit{hili-ya} (\textit{hili}, OBr\text{Wi}); we discuss the meaning of such
forms below;

(ii) the senior term of the reciprocal relationship (if there is one) is
mentioned followed by an echo word; in Santali, e.g. \textit{hili+hal\textgreek{i}}; we
discuss the meaning of these forms below;

(iii) both members of the pair are mentioned, the senior one first; in
\textit{Gta'}, e.g. \textit{nta\textgreek{p}+ag\text{\textgreek{a}}} ‘grandfather and grandchild’; this means (we,
they etc.) are grandfather and grandchild. Whether, as for similar
forms of Santali (i and ii above), the compound can also mean the
relationship (in this case the grandfather–grandchild relationship) is
questionable.

(iv) the construction with \textit{nu-} plus (usually junior) kin term in Remo;
e.g. \textit{nu-girin} ‘I and my wife’s younger brother’, and probably also
‘we two are OSi\text{Hu} and Wi\text{YBr};

(v) where both members of the (hypothetical) pair would refer to each
other by the same term, e.g. GRG *b\text{Vya}n, ‘a man’s brother’; in the
GRG languages, e.g. GRG *b-Vm-lon (*b\text{Vlon, ‘a woman’s sister’})
'each other’s sisters', Gutob bumulon, Remo bunlu, Gta? bumlug.7 The -m- infix (which becomes -n- in Remo) marks this relationship. There are two seemingly irregular cases in Remo: t-un-una, and gVN-riy ([giirdiy]), where tuna is ‘a man’s sister’ (younger sister according to Bhattacharya) and giriy is WiYBr. These have the same meanings as the forms nu-tuna and nu-giriy, i.e. ‘I and my younger sister’; ‘we are (older) brother and (younger) sister’ and the giriy form with -n- would be glossed the way the nu- form was. These are reciprocal pairs, but unless Remo has a common term of reference (or of address?) for brother and sister, and for OSiHu and WiYBr, these are not terms for referents who refer to each other by a common kin term, and we conclude that Remo -n- has been extended to a few such pairs.

The first definition (in the third volume of Bodding’s Santal dictionary (1935: 116)) of hili+h9li is ‘a man’s wife and younger brothers’, then (that) ‘relationship’, and then, more literally, ‘a hili and those who call her so’, i.e. her erwel(s), HuYBr. That is, hili+h9li means ‘the hili-erwel relationship’; hili-and-erwel: the erwel(s)’ hili, and the hill’s erwel(s). Some echo words can be defined as ‘what goes along with (the preceding N1, which the echo word echoes)’ or N1 ‘and the like’. What goes with hili, OBWi, here is its reciprocal. We find echo words of kin terms used comparatively infrequently in this way, but the synonymous construction with -ya/-ea (Bodding’s ‘collective’) is more common. (hili-ya, according to Bodding, means exactly what hili+h9li means.) Bodding in his Materials (1929: II 41-2) lists more than twenty such kin-term pairs (or collectives), most of them in common use.

So far as we know, such reciprocals are not common elsewhere in the Indian area. There are in Indo-Aryan and Dravidian pairs like Hindi maa+baap ‘mother and father’ (in more formal speech maataa+pitaa), and bhaaIi+bahan, ‘brother and sister’, but not twenty other pairs of kin-term reciprocals in common use.

The other examples of SM -m- are found in Juang, and of NM *-p- in North Munda fairly widely (we discuss forms in Ho, Santali and Korku). Pinnow gives two Juang examples of -m-: semelay (from selan, ‘grown-up girl’), ‘young woman, woman’, and komanger (from kanger, ‘young man, youth’), ‘master, lord, husband’. In the first case, the affixed form is slightly more general; in the second the infixed form is a kind of honorific. The only feature of the reciprocal evident in the second example is its use of plurality to mark the honorific, something quite common in the Indian area.

Ho—the third of Deeney’s small 1975 dictionary that we sampled—has few examples of -p- in kin terms. It is interesting that the two we found have precisely the meanings of the Juang forms but are not cognate with

---

7 A common Indian areal construction repeats the noun (kin term) for this reciprocal—‘distributive’—meaning. Thus, Hindi ham bhau bhau hai. (bhau, ‘brother’) ‘We are (like) brothers’, ‘buddy buddy’. Similar constructions have been noted in some of the Munda languages.
South Munda kinship terms, I

them: \textit{h-Vp-anum} ‘young woman’ (there is no \textit{hanum}), and \textit{s-Vp-eed} (Mundari \textit{sepered}) ‘young man’. There is no free \textit{seed}, but it occurs bound in the ‘nephew’ term \textit{hon-seed} (NM *\textit{kon-sered}, *\textit{kon-serej}, FBBrSo).

The Santali examples—we used as a sample the letter H in the third volume of Bodding’s dictionary (1935: 1-184)—are \textit{hapamün}, \textit{hapamüb} ‘grown-up girl’ (there is no \textit{hamün} or \textit{hamū}); \textit{hapram-ko/haprum-ko} ‘ancestors, forefathers, collection of old men’ (\textit{haram}, \textit{harum} ‘an old man’); \textit{hepel}, along with \textit{herel}, ‘man, male’ (no \textit{el}); and \textit{kopon} along with \textit{hon}. He translates \textit{kopon} ‘offspring, young, child, son daughter’; adj. ‘small, little’, and \textit{hon} ‘a son, child’, now used only in compounds.

For Korku, N. Zide recorded along with \textit{kon}, \textit{kon-jei}, \textit{kurar}, \textit{kosered}/\textit{koserej} and some others, \textit{kopon} (with the dual or plural), \textit{koponjei}, \textit{kupərər}, \textit{kopsered}/\textit{kopserej}. The simple forms were much more common. It never became clear what the meanings of the dual and plural \textit{-p-} infixed forms were. The meaning of \textit{-p-} in NM is still obscure; the plural aspect of the ‘reciprocal’ (and its development into an honorific), and the notion of a set or collection can be seen in some of the infixed forms. The forms with \textit{-m/-p-} for ‘young girl’ and ‘young man’ in Juang, Santali and Ho have been noted, but what the function of the reciprocal is here is unknown.

Of mutual appellatives (Bodding’s term, i.e. reciprocal terms of address—in our examples, of persons not of the same generation), Parkin notes Santali \textit{gərəm}, which Bodding translates ‘namesake’, used by grandfather and grandson as terms of address. Parkin also notes the existence of such forms in Juang. McDougal (1963: 141) writing of generation sets and the extension of kin terms, says that ‘any male of an odd-numbered ascending or descending generation, regardless of kin type, may be addressed with the term for “father”, and any female with the term for “mother”... Any male of Ego’s own or any even-numbered generation may be addressed with the term for “brother”, any female with the term of “sister”. For example, a man may address his brother, father’s father, and son’s son with the term for “brother”...’. In Vitebsky’s as yet unpublished notes on Sora kinship there are also examples of cross-generation kin addressing each other by the same term. The term \textit{tata}, FEBr, is used in address reciprocally, i.e. by YBrSo. Similarly \textit{entalaij}, MOSi, can be used reciprocally in address by YoSiDa.

It is likely that more Munda languages have mutual appellatives, but that they have not been recorded.

There are two more possible infixes found with kin terms: the \textit{-m-} in NM \textit{k-Vm-on} ‘nephew, niece’, from the simplex \textit{kən}, ‘child’, also attested in SM: Sora \textit{amən-sij} ‘nephew’, and \textit{amən-sil}, ‘niece’. This is the only example of this PM \textit{*m-}. G. Diffloth tells us that there is a cognate affix, \textit{-m-}, old but unproductive in Mon-Khmer, with cognate forms in a number of MK families (Khmer, Bahnaric) meaning ‘child’ for the simplex, and ‘nephew/niece’ for the infixed form.

The other case of a kin term that might be analysed as containing an infix is PM \textit{*kin-} or \textit{kin-}, found only in \textit{*kinkar/kinkar} ‘mother-in-law’
(and perhaps Juang kan-đae, 'old woman'), the *kin having a cognate in MK meaning 'woman'. There is another kin term in Munda, *kimin, ‘SoWi, YBrWi’, which might be related morphologically at an early stage to *kin. Diffloth informed us of an MK -mp- infix, that would account for this derivation. If we take the *-mp- as going to *mh- in PM, this helps clarify a previously baffling set of correspondences. We find *kVmin everywhere in Munda except in SJG, where the cognates—if they are cognates—are unexpected: Sora kə'ën, (Vi: koin, kaon), Juray kaun, [kə'əun], and Gorum konun for YBrWi. We reconstruct *k-Vmh-in/k-Vmh-in, and begin to understand how these SJG forms (and presumably others) developed.

It is difficult to say what in the Munda kinship (and kinship term) system comes from Austroasiatic and what does not. Of the topics we have discussed, those peculiar to Munda (in the Indian linguistic area) are: the sex-marked sibling terms; the use of cross-generational reciprocal terms of address with some frequency (fuller data on more Munda groups would probably show more evidence of this); the use of paired reciprocal terms commonly; and the various infixes and their uses. Indo-Aryan and Dravidian do have age-marked terms, as has Mon-Khmer. Munda has some, but a number of the distinctions and terms look like borrowings. It would take a closer examination of Munda and Mon-Khmer kinship to speak with authority on this. That age distinctions in kin terms are old in MK does not, of course, mean they are old in Proto-Austro-Asiatic. Munda is not the only language family to be influenced by its neighbours, although it may be true that it is easier to perceive areal influences on Munda than on Mon-Khmer. The usual assumption, that MK preserves much more of PAA than does Munda, is probably true, but what is and is not PAA needs to be demonstrated after sufficient MK and Munda evidence is in, examined, and evaluated, and the outlines of PAA are clearer than they are now.8

8. Some of the fieldwork reported in this paper was done under a grant from the National Science Foundation.
APPENDIX

*kuli

Bhattacharya 83. (abridged) Juang koli, kuli, OBrWi; Kherwarian *hili, OBrWi

We noted earlier that adding KM *hVli makes the SM evidence more convincing. Pinnow in his Juang monograph also connected Juang kuli and Kher *hili, but took the etymology no further. We note the form kulaya-sini (only in McDougal) which is, apparently, a derivative of kuli-sini (elsewhere -sel, -sen) is the combining form of (McD.) kon-chalan (Pi. kon-selan) ‘young woman’. kulayasini means HuYSi, and can be roughly glossed ‘kuli-junior’, i.e. the ‘junior’ of OBrWi is HuYSi. Note the use of -sini in the ‘grandchild’ terms boko-lap (McD.), ‘grandson’, and boko-sini, ‘granddaughter’. The grandchild terms probably derive from boko-X-lab/sini, the X perhaps being the -du? found in Kharia (cognate with words for ‘child’ in Gorum (a)ud), and for ‘young man’ in Gotub (o-rag)).

SM *hVli(-boi), PM *Vrel, SM *Vrel/*Vrer, NM *erwel

Bhattacharya 81. (abridged, rearranged, and provided with additional data, the latter in parenthesis)

Set (a). Ga ili-boy (MZ hli-bwe’?), HuYSi; Re liboi, HuYSi; So aliboi, HuYSi (Vi. aliboi, HuYSi, WiYSi); ?? Ju ini-bo (McD. ini-bou, MM ini-bo), HuYBr. -boi/-boy/-b:J mean ‘woman’ in these compounds; -sij/sij (full form in So pasij, ‘child’, in compounds with terms meaning nephew/niece -sij means ‘male person’. This -sij is probably cognate with the -serje, -seresi/-sered, -sered mentioned earlier for NM and meaning ‘young man, nephew’. Bhattacharya seems to be suggesting that the -bo in ini-bo is a combining form of boko, YBr. We have no other examples of this—or anything else—as a combining form of boko.

Set (b). Ga urive, HuYBr (MZ wrwe’?); Re ere (ZZ ere(l)), HuYBr; Gu erel, HuYBr; Go ilil, HuYBr; So erer-sij-an, HuYBr, erer-sij-an, WiYBr (we take erer, and ere to be the same), erel-boi, WiYSi (Vi eri-sij, WiYBr, HuYBr; Sur erer-sij, WiYBr, HuYBr; ali-boy, arrel-boy, WiYSi, HuYSi).

One complication in Sora that needs comment; we find in both Vitebsky’s and Suryanarayana’s data that along with ali-boy for SpYSi, we also find eri- (Vi.) and arrel-(Sur.)-boi. The eri- and arrel- are both from KM *Vrel, although the eri looks like a portmanteau of ali and Vrel. Since the marking of ali as feminine, and of Vrel as masculine in Sora is being lost (this sex-marking role has been taken over by -boi and -sij) the meanings of Vii and Vrel have come closer. Vitebsky suggests that eri-
means ‘younger’, i.e., ‘younger SpSib’, since it can now be used with either wife’s younger sister (eri-boy along with ali-boy) or younger brother (eri-sij), and similarly for a female ego for HuYBr and HuYSi. For Vitebsky’s Soras this would work, but not for Bhattacharya’s. For Suryanarayana SpYBr is erer-sij; *Vrel has been generalised, but then split into arrel, which commutes with ali-, i.e., goes with -boy. The other form *Vrel has split into is erer-, and this is used only with the masculine -sij.

We are grateful to R. J. Parkin for copies of his papers, and to Piers Vitebsky for his notes on Sora, and for copies of the extract from Suryanarayana’s dissertation.

Set (c). Sa erwel, HuYBr (Bodding era + el ?); Mu iril, irul, iriul, HuYBr; Korwa irvil, HuYBr; Koku, ilur, HuYBr. (WiYBr in NM is *teýa (i.e. [tḗyā]).

For set (a)—putting aside ini-bo for the moment—we reconstruct *hVli(-b:Jj), HuYSi, SpYSi; using Bhattacharya’s data only it would be *Vli-boi. We find this to be cognate with Bhattacharya’s 83., *kuli/ *kVli, OBWi. The definition of the new PM *kuli would be ‘female affine of ego’s generation’ (i.e. English ‘sister-in-law’), OBWi, HuYSi, SpYSi. An alternative interpretation would reject Juang kuli as cognate with either NM *hili or KM *hVli, and include ini and NM *hili and KM *hVli in the reconstruction.

Bhattacharya presents sets of forms that are semantically similar. He would like these semantic sets to be cognate sets as well. But he wants his semantic-cognate sets to be closer in meaning—for him to accept them—than, we think, such cognate sets in PM will often be. It is true that some of these kin terms’ semantic-cognate sets (e.g. Bhattacharya’s bare, ZZ bɔHrɛ) show forms that have remained remarkably close in meaning and in phonetic form in the modern languages. But given the chronological separation of the Munda languages (at least twenty-five hundred years), there is no reason to expect this degree of closeness. Just as we have no hesitation in calling Ga swa and So aŋɛl, both ‘fire, firewood’, cognates—because we can see how both developed regularly from KM *sVAHl—we should have no trouble in accepting NM *hili and KM *hVli as cognate despite the excessive (to Bhattacharya) difference between the meanings of the two forms. We need the anthropologist to reconstruct a more precise meaning for the proto of these, i.e.—if we accept Juang kuli—*kVli, and to account for the changes in meaning between the PM form, and the forms in the modern languages.

From the forms in set (b) we reconstruct KM *Vrel/*Vrer, HuYBr, SpYBr. From this and NM *erwel (reconstructed from the forms in set (c)), we reconstruct PM *Vrwel/*erwel, SpYBr, HuYBr.

It is possible, as Bhattacharya suggests, that *Vrwel is bimorphemic, and that the second morpheme (of *Vr( )-) is -hVl(i) or -kVl(i), but we see no evidence of this.
i,:zi-bo, i,:zi-bou/-bou

What then of i,:zi-bo? Since bau/bou is the reciprocal in Sora of KM *hVli and KM *Vrel derivatives, and a similar situation may obtain in Juang, let us look at Bhattacharya’s Set 62., bau/bou.

PM *bao/*bau

Bhattacharya 62. (edited, and with additional material): So bao-n, ‘brother-in-law’, (Vi. baon OSiHu, Sur. boung, OSiHu, kinar-boung SpOBr); (ZZ Go bao, ‘brother-in-law’); ZZ Gu ū-ban, m-ban, OSiHu); Remo (Bhattacharya, 1968, 111) m-ban, OSiHu; (ZZ Ga m-bia, OSiHu); Kh bau HOBr, Roy bao-tang, WiOBr (VU boutang, WiOBr); Ju bōv-kar, HuOBr (Pi. bau/bou, OSiHu, MM bau, SiHu, McD. bou, ESiHu, bokar (probably from bau-kar), HuOBr); NM: Ho bavo, OBr, SpOSiHu; bau honjar, SpOBr (Deeney bau, OBr or cousin; bau honyar, SpBr); Mu bau, OBr (in address); bau honjar, SpOBr; Sa bahn-jhar, SpOBr; Korku bao, WiOBr, (Girard bao, WOBr). The final nasals in some of these words are frozen first person prenominal enclitics used in terms of address, i.e. literally, ‘my OSiHu’, etc.

To return to i,:zi-bo (or -bou), we stated in the body of the paper that with regard to sound correspondences and semantics, a case can be made for i,:zi being cognate with PM *Vrel, SpYBr, HuYBr, if not the strongest case. What of the -bo or -bou? i,:zi-bo or -bou is an inalienably possessed noun in Juang. It is true that reciprocals—both terms and the relationship between then—are often maintained through many linguistic changes. But that a pair of reciprocals meaning (an inalienably possessible) HYBr, this deriving from the usual paired reciprocal meanings, could have developed in Juang is, for semantic reasons, highly unlikely. -bo, although we have seen no (other) examples of it as a combining form of boko seems more likely. Why does i,:zi need a second morpheme at all to have the meaning HYBr? As far as we can see, it does not. The spouses of both kulayasini and i,:zibo are both some kind of boko: the i,:zibo’s wife is (McD.) boko-ray (YSi), and kulayasini’s husband is boko (YBr). HYBr in McDougal is boko-ger (from boko-ker?); elsewhere he is just boko. Apparently i,:zi too required a second morpheme. -bo from boko makes better sense than bau/bou.

-kar, ‘in-law’

On bau-kar and the use of the affine-marking -kar in Kharia-Juang. The affinal -kar, ‘in-law’, is found in PM in *kikar/*kankar, SpM (or perhaps, SpM, WiOSi), and *ku(X)nkar, SpF (or SpF, HuOBr). The -kar may be related to what has become the NM word for ‘man, person, member of the community’, kpr. The Kharia -kar marks the agent in forms like rema-kar, ‘call-er’, rema- ‘to call’ (Malhotra, 1982: 311). The kar is also used in forming singular pronominal stems from demonstrative
bases (e.g., ho-, u-), thus ho-kar, u-kar, 'he, she'; i.e. -kar means 'one'. In KJ we find the following forms: Juang keeps the term for 'father-in-law', kuńkar/kwińkar/kwińkar, and also has bao-kar, HuOBr, and aji-kar, SpOSi. Kharia keeps the old word for 'mother-in-law', kinkar, and has, in addition, boker (probably from boko-ker), SPYBr, and (Roy) aichkar, WiOSi. (We noted earlier that Juang has boko-ger—with the ger perhaps from -ker and not from koN-ger, 'young man'.)

In some but not all cases the -kar can be translated simply as 'in-law', i.e. if X means YSi, then X-kar means YSi-in-law, i.e. SpYSi. This presupposes that X is not an affinal term. The examples of this in KJ are aji, ESi, and boko, YBr. One could, simplistically, try to derive the remaining terms from hypothetical simple terms, not in KJ, but in Kherwarian. Thus Ho bau, OBr, Juang bau-kar, HuOBr. This would miss the fact that Juang bau/bou is an affinal term, OSiHu, and that kar derivatives of affinal terms are different from derivatives of simplex terms. Where the kar-derivative is formed on an affinal term, the simplex refers to a sibling's spouse, and the kar-derivative to a spouse's sibling. The other example of this in KJ is (Roy) aji, OBrWi, and aich-kar, WiOSi. The Juang boko-ger, HuYBr, if regular, should come from bok(V)-(k)er-ker.

The term bok-sel in Kharia, SpYSi (boker is SPYBr) is probably to be derived from bok-ker-sel. And the Juang 'grandchild' terms, boko-lab, 'grandson', and boko-sini, 'granddaughter' are probably to be derived from boko-X-lab/-sini, the X perhaps being cognate with the *ui? found in Kharia in bok-(ui?, 'grandchild'. The reduction of similar compounds in Sora from morphemes 1-2-3 to 1-3 was observed by N. Zide in regard to the Sora numerals.

The Kherwarian uses of *-kar (which, if it were reduced the way it is in KJ would be *-har): PM *kinkar/*kankar (Kher *hanhar) and *ku(X)ńkar (Kher *hńnar): these occur in compounds of the structure X-hanhar/hńyar, which have the meaning hanhar's or hńyar's X, i.e. kaka-honyar in Ho means honyar's kaka, i.e. SpFYBr. For X as aji and bau this is not the case. The construction with aji-hanar, and bau-honyar in Ho, is like the construction in KJ with non-affinal simplex plus -kar. Note that (unlike KJ) aji in Ho is OSi, and bau is OBr; aji-hanar means SpOSi, and bau-honyar means SpOB. Note that these are 'spouse's sibling' terms, which, in KJ, are those derived from affinal simplex terms, not, as here, non-affinal simplex terms.

Santali is much like Ho in its SpF (hńhar) and SpM (hanhar) terms: most of them have the X-hanhar/hńhar construction where this means, e.g. hanhar's X (where, for instance, X would be gongo, OBr). For the two forms cognate with the two just distinguished in Ho, Santali bahńhar and ajnihar, the meanings are parallel to those in Ho: SpOB, and SpOSi. Note that bahń does not occur as a free form (as does the Ho cognate bau, OBr), but aji, OSi, does. Note the reduction of hńhar to -har in the first form, and that of hanhar to -hanar in the second.
It is not surprising that Parkin (with Bhattacharya's help) sees forms related to *boko almost everywhere. He is right to connect Ga *tabo (MZ ta-bo?) and Re *tabuk', FYBr, MYSi, stepF (Bh. Set 31.) to *boko. Cognate forms in Set 33. should also be included: Bh. *mbu, (FF mbu?, Elwin umbsuk-boi), MYSi, FYBrWi, stepM. Bhattacharya is probably wrong in bringing in *boho (he translates the Gta? form, wrongly, we think, as YBr). In KJ *boko and possibly related forms cover a good deal of semantic territory: apart from the basic—YSib—meanings there is the Juang word for 'clan' (Pi. *bok, McD. *bok), the 'grandchild' terms we have talked about (Kh (Roy) *bok-du(?), Juang (McD.) *bokolap and *bokosini). Pinnow records the Juang term *bok-rad, "Verwandtschaft" (relation), Bedeutung und Form unklar'. The form found in Kherwarian, *bok2, is unexpected, since one would expect the intervocalic k to go to h before o. Bhattacharya gives a form, Ho *boko, which is not found in Deeney (but see the other forms in Bhattacharya's Set 64. (1970: 455)). There do seem to be related forms in Kherwarian with h, e.g., in Bodding's Santali Materials II (1929: 21) there are *bok boe ha and *bok bo ha boe ha, 'cousins any number of times removed, descendants in the male line'. One wants to know not only the nouns, the kin terms, but the rest of the kinship vocabulary—the verbs and the constructions. This sort of information can be found in Bodding and Vitebsky.

As to the rest of the terms with initial b that might be etymologically connected with *boko, a better knowledge of MK cognates would help. For instance, we would tentatively connect the first piece (etymologically, presumably, a morpheme) of *vyay, boya, 'brother, a man's brother' and GRG *vlo 'sister, a woman's sister', Sora (Vi.) bui-mai, (bui-mai? *vi-mai?), 'true sister, usually used by women'. Bhattacharya wants to connect *vyay with Sora, YBr, ubba-n-ubba- (Vi. u'baa).- We think it is probably cognate with *boko. To make a case for either—or, conceivably, both—considerably more synchronic phonological and morphological analysis needs to be done. For instance, by Bhattacharya's hypothesis, one would take -boj as a combining form of *vyay, odd but not impossible, but it does not fit in with assumption of a morpheme *vii-/ *vii- shared by *vlo and KM *vyaN. And we have other forms: Gorum *ing-ger, 'bloodbrothers', Sora birinda, 'clan', i.e. (Vi.) 'exogamous agnatic localised virilocal kin group' (Vitebsky translates *vyay as 'male member of own birinda'), and Sora FaYBr kim-bom, Vi. kin-bom, kin-boj that come into the picture. The sibling terms and related vocabulary will be taken up in a later paper, which will take MK cognates into consideration, and do more synchronic phonological and morphological evidence than we have done here.
REFERENCES


South Munda kinship terms, I


Vitebsky, P. 1979. Field notes on Sora kinship.


The growing interest in Khasi in recent years, as the lone representative of the Austroasiatic family in an area surrounded by Indo-Aryan and Tibeto-Burman languages, has resulted in a welcome number of serious studies of the language. A number of otherwise valuable studies have, however, been flawed by misunderstanding arising from the Khasi spelling system. This looks relatively simple on the surface but there are pitfalls for the unwary. Scholars with an English or Indo-Aryan linguistic background may be led astray if they are tempted to take the spelling at its face value and to trust their eyes rather than their ears. The sections devoted to phonetics and phonology in some studies I have seen have contained phonetic or phonemic ‘transcriptions’ which are little more than letter-for-letter replicas of the spelling. In these circumstances, it is worthwhile taking a closer look at problem areas within the spelling system itself.

There have, I believe, been certain attempts during the last few years at minor revisions of the official orthography, and I cannot claim to have up-to-date knowledge of all of these. As far as I am aware, however, the system is substantially still the same as that used in Diengdoh (1965), Kharkongngor (1968), Bars (1973), and Blah (1974). I shall also refer in this paper to the older and probably better-known dictionary of U Nissor Singh (1906), and occasionally to even older forms. The pronunciations cited are from Cherrapunji speakers, since this variety of Khasi is accepted as ‘standard’.

Vowel length and final consonants
What has to be borne in mind is that the Khasi roman spelling was devised in the first half of the nineteenth century, not by English or Scots or American missionaries, as in some other parts of India and Burma, but by Welsh missionaries, who introduced certain Welsh orthographic conventions unfamiliar to most English speakers.

To the non-Welsh, one of the most confusing of these conventions relates to the use of the letters -p, -b, -t, -d, -c, -j in word final position. Linguists accustomed to the phonemic differentiation of voiced and voiceless stops, in such languages as French and English for example, find it natural to suppose that Khasi has this sort of phonemic distinction also,

1. This does not apply to Rabel (1961), whose phonetic observations are beyond reproach.
2. On the absence of final -g in Khasi, and on the special distribution of final -k, see Henderson (1965).
when they are confronted for the first time with spellings such as those in (1) below:

1. Khasi:  
   - `dab` ‘bullock’  
   - `kad` ‘to tear’  
   - `mad` ‘to taste, try’  
   - `sngab` ‘gill’  
   - `dap` ‘full’  
   - `kat` ‘so much, as much as’  
   - `mat` ‘joint’  
   - `sngap` ‘to hear’

For those whose work has hitherto been mainly with Indo-European languages, there is nothing very remarkable about this kind of final alternation, but those who have specialised in the languages of South-East Asia to which Khasi is related, it would be not a little surprising to find a phonemic distinction between voiced and voiceless final stops. The characteristic typological pattern found all over continental South-East Asia is for the voice : voiceless contrast, which is frequently found in initial position, to be neutralised in final position. What we usually get is an unexploded voiceless stop, often accompanied by simultaneous glottal closure. Listening to the pronunciation of the Khasi words in question soon makes it clear that Khasi is no exception to the general rule—the final stops are voiceless, no matter what letter they are written with. What we also observe is that, in the great majority of cases, where the final letter in the spelling is `b` or `d` the preceding vowel is always long, while where it is `p` or `t` it is almost always short. The words in (1) above are regularly pronounced:

```
2.  
   - `dab` [da:p']  
   - `kad` [ka:t']  
   - `mad` [ma:t']  
   - `sngab` [sIJa:p']  
   - `dap` [dap']  
   - `kat` [kaC]  
   - `mat` [maC]  
   - `sngap` [sIJap']
```

This seems very curious until it is realised that Welsh has just such a convention: vowels before the letters `b`, `d` and `g` are almost always long, whereas before `p`, `t` and `c` are almost always short. (For this rule, and for the exceptions to it, see Wells (1979).)

3. Welsh:  
   - `mab` [ma:h] ‘son’  
   - `tad` [ta:c] ‘father’  
   - `log` [lo:ɡ] ‘interest’  
   - `map` [map] ‘map’  
   - `at` [at] ‘to, towards’  
   - `llac` [lak] ‘slack’

Phoneticians who have worked on Welsh claim that while the sounds written `-b`, `-d`, `-g` are voiceless finally, they are also lenis, or lax, as compared with those written `-p`, `-t`, `-c`, which are voiceless and fortis, or tense. Wells thus prefers to describe the lenis set as ‘devoiced’, here symbolised `[l;>, cj, g]`, following the IPA convention. In Khasi there is no such tense/lax distinction, any more than there is a voice/voiceless distinction in final stops: they are all voiceless. It seems clear that the early Welsh missionaries made use of the their own spelling conventions in

---

3. See note 2 above.
order to mark distinctions of vowel length in Khasi.

This convention works well with stops, where there are two sets of letters available, but it does not work for words ending in final nasals or in final r, which are always voiced. In these circumstances, in Welsh, a circumflex accent is used to mark the long vowel:

4. Welsh:  
- tân [ta:n] ‘fire’  
- tan [tan] ‘under’  
- tôn [to:n] ‘tune’  
- ton [ton] ‘wave’  
- glân [gla:n] ‘clean’  
- glan [glan] ‘bank’

Earlier spellings of Khasi also tried this device. In Nissor Singh (1906) the circumflex accent is often used to mark vowel length before nasals and r; alternatively, a grave accent may be used to show length, or an acute accent to show shortness, but this is not at all regular, e.g.:

5. Khasi:  
- bām [ba:m] ‘to eat’  
- sbāi [sba:i] ‘cowry’  
- kēr [kē:r] ‘to enclose’  
- kpēr [kpē:r] ‘garden’  
- hēr [hē:r] ‘to fly’  
- ēr [ēr] ‘to slip, fall’  
- tur [tur] ‘to butt’

It appears probable that, in these words, there is some variation in usage among Khasi speakers, which may account for some of the seeming irregularities. Thus, we find ēr ‘to slip, fall’, marked with a short vowel in Nissor Singh, but tur ‘to butt’ written without an accent, although also pronounced with a short vowel by my Khasi informants. Similarly, we find the spelling kēr [kē:r] ‘to enclose’, but kpēr [kpē:r] for the derived form meaning ‘garden, enclosure’. Hēr, marked short, was pronounced with a long vowel by my informants, but it may be that some speakers use a short vowel here. This irregularity and uncertainty in the use of accents may be one of the reasons why they seem to have been abandoned completely in most of the recent dictionaries. This leaves the problem, for the foreigner, of being unable to tell from the spelling whether the vowel is long or short in such words. Father Bars’ (1973) dictionary, for example, has to resort to such entries as ‘bam (long a)’ and ‘bām (short a)’.

The vowel ‘y’

Another problem for linguists has been the vowel spelt y in Khasi. It is almost always unstressed and behaves in much the same way at the shwa vowel [ə] in English, as in ‘about, again, collapse, suppose’, etc. When we look at Welsh, we find this similarity is no accident, since y is the Welsh way of spelling the vowel [ə]. Depending upon its position in the word, y may in Welsh also represent a vowel which is pronounced either [i] or [i], according to dialect, e.g.: mynydd [ˈmɔnɪd], ‘mountain’. y, with the [ə] value, is also the vowel that is used where a cushion vowel or epenthetic vowel is felt to be needed to break up awkward or alien consonant clusters, as in the forms ysgôr ~ sgôr [asɡɔːr, sɡɔːr] ‘score’. The rather old-fashioned Khasi spellings kypa ‘father’, bydi ‘twenty’, bynnai ‘moon’, month’, etc., show the same process at work. Since clusters like [kp, bd,
EUGÉNIE HENDERSON

bn], and many others, are difficult for Welsh or English speakers to pronounce without a short vowel in between, the y vowel was inserted. More recently the convention on the whole is not to write y between 2-consonant clusters, but to insert it where there are clusters of more than two consonants, e.g. bna 'to hear', bdi 'twenty', but byndi 'to mortgage'. There seems still to be some uncertainty about this, however, as Father Bars still gives spellings like byna. Phonologically speaking, there would be much to be said for dropping the y in the 3-consonant clusters as well, simply writing bndi, which is perfectly clear. This is the solution proposed by Dr. Rabel in her phonemic analysis of Khasi (Rabel 1961).

Final palatals
Khasi has final palatal consonants; Welsh does not. What appears to have struck the ears of the missionaries who first listened to Khasi was the i-like glide that precedes the final palatals. We thus have early spellings such as skain ‘fly n., bysein ‘snake, python’, etc., and it seems clear that the final consonant was thought of as plain n in such words, though its palatal quality was later recognised by the use of the letter ŋ: skain, bseiii, etc. Since there is no accepted roman equivalent for the final palatal stop, -d and -t continued to be used, except after ie, when we find j, e.g.:

         kshaíd [kʃəc] ‘waterfall’  kaic [kəkic] ‘plantain’
         bțuid [bución] ‘slippery’  bizt [buc] ‘cunning’
         miej [mi:ɕ] ‘table’

From the phonological point of view, the vocalic segments preceding these consonants should be interpreted as plain vowels (or in some cases as centring diphthongs), with an automatic palatal on-glide to the final consonant, rather than as phonemes /ei, ai ui/, etc. Father Schmidt, writing on Khasi as early as 1904, recognised this; his knowledge of the Austroasiatic family led him to expect that final palatals might be present and he interpreted the missionaries’ spelling in this light.

Representation of the glottal stop
The representation of the glottal stop, which is a consonant phoneme in Khasi, does not derive from a Welsh convention, but rather from a much more general missionary usage, found also in the roman orthographies of Tibeto-Burman languages such as Lushai and Chin. Word-initially, the glottal stop is not written at all; word-finally it is spelt with -h; which is not otherwise needed in final position:

         soh [soʔ] ‘fruit’  kseh [kseʔ] ‘pine’

Problems arise when [ʔ] occurs medially, or when it is the second consonant of an initial cluster. Devices used in such contexts include hyphens, apostrophes, or the vowel letter y. In the last case, the y
Problems in Khasi orthography

presumably represents the slight vocalic sound that can be heard as the glottal stop is released:

8. Khasi  
   ha-oid, ho-oid, h’oid  
   sh’iap, shyiap  
   l’er, lyer  
   s’iar, syiar  
   lyoh  
   syyiem

   [haʔɔɛt]  ‘Yes’;
   [ʃʔiap]  ‘sand’;
   [lʔer]  ‘wind’;
   [ʃʔiar]  ‘fowl’;
   [lʔoʔ]  ‘cloud’;
   [sʔiːm, sʔeːm]  ‘king, chief’

In the word for ‘cloud’ above, we find the glottal stop represented by h finally, and by y in the initial cluster.

Dental/alveolar distinction

It is common for Western linguists to expect that t, th, d, n will form a natural class, united by a common place of articulation. In Khasi, however, as in certain other languages in the area, this turns out not to be the case. [t] and [th] are dental, [d] and [n] are alveolar. This is not simply a matter of phonetic detail; it is important in understanding the rules for permitted sequences of plosives and nasals in Khasi initial clusters. At first sight it appears that a sequence of virtually any two consonants may form an initial cluster, but while the possibilities are very rich, there is a constraint upon sequences of plosives and nasals whereby those with the same place of articulation may not form a cluster together. For example, bn, bt, bth, pd, pn, bn, thŋ, dp, km, kt, etc., are permitted; pb, pm, bm, kg, khŋ, are not; td, tn and thŋ are permitted, and seem to be exceptions until it is recognised that the sequences here are also hetero-organic, and hence perfectly regular.

Ambiguous vowel spellings

The letters u, o, and e, and the digraph ie are sometimes phonetically ambiguous.

Finally and before a glottal stop, ie represents a pure vowel, a very close [e]. Before other consonants, pronunciation varies between [iː] and [e:], depending upon the speaker:

9.   khlieh [khleʔ]  ‘head’  
      sdiə [sde]  ‘axe’
      dieng [diːŋ, deːŋ]  ‘tree, wood’

   lieh [leʔ]  ‘white’
   kmie [kme]  ‘mother’
   ktien [ktiːŋ, kteːŋ]  ‘word’

In open syllables, the spelling o represents a very close [ə] in some words, [ɔ] in others. There seems to be variation in usage between speakers for some words:

10.   ro [ro  or ro]  ‘quicksilver’;
       -to [to]  (indicating assent);

   to [tɔ]  ‘that (enclitic)’.

In closed syllables, the phonetic value is always [ɔ].

4. For a more detailed account of Khasi vowels, see Henderson (1967).
Before a final glottal stop is \([u']\) in some words, \([o']\) in others:

11. `ruh` \([ru'\text{?}]\) ‘also, indeed’  `ruh` \([ro'\text{?}]\) ‘cage’
    `duh` \([do'\text{?}]\) ‘lose’  `shniuh` \([\text{[po'}\text{?]})\) ‘hair’

Before other consonants, \(u\) may represent long \([u:\text{;};i]\). \([o:\text{;};i]\) or short \([u]\):

12. `bud` \([bu:a'\text{;};i', bo:a'\text{;};i']\) ‘follow’  `thub` \([\text{thu:a'}\text{;};i', \text{tho:a'}\text{;};i']\) ‘reduced in size’
    `tup` \([\text{tu: }\text{p'}]\) ‘cannon’  `lum` \([\text{lu: }\text{m}, \text{lo:m}]\) ‘hill’
    `lüm` \([\text{lu:m}]\) ‘gather’

In open syllables, \(e\) may represent either close \([e]\) or open \([e]\):

13. `te` \([te'\text{;};i]\) ‘then, but’  `de` \([de'\text{;};i]\) ‘also’,
    `me` \([me'\text{;};i]\) ‘thou’  `re` \([re'\text{;};i]\) ‘or’

In closed syllables the value is always \([e]\).

To sum up, with a proper understanding of its origins and background, Khasi spelling has much to recommend it. Once mastered, the conventions regarding the use of the letter \(y\), the use of `-b, -d, -j` to indicate length of the preceding vowel, and the conventions for the representation of the glottal stop, are in the main clear and unambiguous. A reasonable complaint might be that vowel length before nasals is commonly not shown. However, in my view, the traditional spelling, properly understood, represents Khasi usage far more accurately than the misinterpretations and attempted ‘improvements’ that have been suggested by some linguists with an Indo-Aryan linguistic background, and little or no understanding of the strong Mon-Khmer legacy that still survives in Khasi phonological structure.

REFERENCES


HU – A LANGUAGE WITH UNORTHODOX TONOGENESIS

Jan-Olof Svantesson

In this article I will describe and analyse a small vocabulary I collected in September 1984 from a native speaker of Hu /xú?/, a Mon-Khmer language spoken by about 1000 persons in a few villages in the Xiāo Měngyáng area in Jīnghóng county, Sipsong Panna (Xīshuāng Bānnà), Yúnnán province, China. The Hu are known among the local Chinese as Kōnggē.

Hu belongs to the little-known Angkuic group of the Palaungic branch of the Mon-Khmer languages. Small Angkuic populations are scattered over south-western Yúnnán province, and in another article (Svantesson 1988) I describe the language of another of these, U, spoken in the village Pă̄ Xẹp (Bāngxié) in Shuāngjiāng county.

The place of Angkuic within the Palaungic branch is shown in the following table according to Diffloth (1982a):

<table>
<thead>
<tr>
<th>Palaungic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>East Palaungic</td>
<td></td>
</tr>
<tr>
<td>Waic</td>
<td></td>
</tr>
<tr>
<td>Angkuic</td>
<td></td>
</tr>
<tr>
<td>Lamet</td>
<td></td>
</tr>
<tr>
<td>West Palaungic</td>
<td></td>
</tr>
<tr>
<td>Danaw</td>
<td></td>
</tr>
<tr>
<td>Riang</td>
<td></td>
</tr>
<tr>
<td>Rumai</td>
<td></td>
</tr>
</tbody>
</table>

Although my material is too small to allow a complete synchronic phonemic analysis, the most important historical developments in Hu can be inferred from it. From a general phonological point of view, the most interesting phenomenon is the development of a two-tone system where the tones are not the reflexes of voiced/voiceless proto-initials, as is most often the case in Mon-Khmer two-tone (or two-register) languages. Instead, the tones are the reflexes of the long/short vowel opposition which existed in Proto-Palaungic (inherited from Proto-Mon-Khmer). As far as I know, no language with this kind of tonogenesis has been described before.
Initial consonants

The Angkuic languages are characterised by a ‘Germanic’ development of the initial stops—that is to say, voiceless stops have become aspirated, and voiced stops have become voiceless unaspirated. This is illustrated by the following examples:

<table>
<thead>
<tr>
<th>Language</th>
<th>Hu</th>
<th>U</th>
<th>Lamet</th>
<th>N. Kammu</th>
<th>S. Kammu</th>
</tr>
</thead>
<tbody>
<tr>
<td>*voiceless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phiŋ</td>
<td>phet</td>
<td>pin</td>
<td>pin</td>
<td>pin</td>
<td>‘to shoot’</td>
</tr>
<tr>
<td>than</td>
<td>than</td>
<td>tsan</td>
<td>taoŋ</td>
<td>taŋ</td>
<td>‘to weave’</td>
</tr>
<tr>
<td>khaŋ</td>
<td>khaŋ</td>
<td>kaap</td>
<td>kaap</td>
<td>kaap</td>
<td>‘jaw’</td>
</tr>
<tr>
<td>*voiced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pan</td>
<td>pan</td>
<td>pan</td>
<td>—</td>
<td>—</td>
<td>‘white’</td>
</tr>
<tr>
<td>priŋ</td>
<td>priŋ</td>
<td>priŋ</td>
<td>briŋ</td>
<td>‘forest’</td>
<td></td>
</tr>
<tr>
<td>phlakaŋ</td>
<td>phlakaŋ</td>
<td>ktaak</td>
<td>ktaak</td>
<td>kdaak</td>
<td>‘palm (of hand)’</td>
</tr>
<tr>
<td>kaaŋ</td>
<td>kaaŋ</td>
<td>kaaŋ</td>
<td>gaaŋ</td>
<td>‘house’</td>
<td></td>
</tr>
<tr>
<td>kak</td>
<td>kaaŋ</td>
<td>kaaŋ</td>
<td>‘to bite’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Southern Kammu (as recognised from Lindell et al. 1981), the original Proto-Mon-Khmer voicing contrast is retained. The unaspirated stops in the Angkuic languages Hu and (Paŋ Xep) U correspond to voiced stops in Southern Kammu, and the aspirated stops correspond to voiceless unaspirated. Original voiced and voiceless stops have merged in Lamet (Rméeet; from Lindell et al. 1978) and Northern (Yuan) Kammu, giving rise to lax and tense register in Lamet, and low and high tone in Northern Kammu, as is the case in Mon-Khmer and other languages with ‘orthodox’ register or tone development*. The examples also show that the Hu tones are not the result of orthodox tonogenesis.

Hu has a contrast between initial s- (with the allophone [ɔ] before i, and in the word sêt ‘red’) and θ-, an opposition which is not present in U or other Angkuic languages, but which is found in Danaw (Luce 1965), as a contrast between ts- and θ-. In Lamet, and in the rest of Palaungic (and in Kammuic), Hu s- and θ- correspond to s- and h-, respectively. Diffloth (1977) and Ferlus (1978) reconstruct these as Proto-Mon-Khmer *c- and *s-, respectively. According to Diffloth (1977), *c- became *ts- in Proto-Palaungic. Examples are:

<table>
<thead>
<tr>
<th>Language</th>
<th>Hu</th>
<th>U</th>
<th>Lamet</th>
<th>Danaw</th>
</tr>
</thead>
<tbody>
<tr>
<td>*c-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sêt</td>
<td>so</td>
<td>sêt</td>
<td>tso</td>
<td>‘dog’</td>
</tr>
<tr>
<td>salêt</td>
<td>salè</td>
<td>slèet</td>
<td>kālê</td>
<td>‘rain’</td>
</tr>
<tr>
<td>nṣît</td>
<td>nchi</td>
<td>sî</td>
<td>tsi</td>
<td>‘louse’</td>
</tr>
</tbody>
</table>

*In the Northern Kammu form ktaak, the tone is determined by the voiceless k. (Ed.)
Hu – A language with unorthodox tonogenesis

<table>
<thead>
<tr>
<th>*s-</th>
<th>θūm</th>
<th>sūp</th>
<th>hūum</th>
<th>θən⁴</th>
<th>‘to bathe’</th>
</tr>
</thead>
<tbody>
<tr>
<td>paṭēn</td>
<td>sēt</td>
<td>phīn</td>
<td>pāṭēn⁴</td>
<td>‘snake’</td>
<td></td>
</tr>
<tr>
<td>θəʔən</td>
<td>sʔən</td>
<td>‘dry’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θəmāʔ</td>
<td>samāʔ</td>
<td>‘wind’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

U has the regular reflex ch- of *c- after minor syllables, as in the example ‘louse’. Proto-Palaungic also had *h-, which is retained as such in all languages but, unfortunately, I failed to elicit any word with *h- in Hu.

There are some words in which Palaungic s- corresponds to c- in Kammu—they include Hu sān ‘bitter’, Kammu cāy. Here, Ferlus reconstructs *tš- in Palaungic, Kammuic and Viet-Muong (corresponding to two Proto-Mon-Khmer initials, *ts- and *tš-).

Minor syllables

As in other Northern Mon-Khmer languages, most non-compound words are monosyllabic or sesquisyllabic, i.e. consisting of a major syllable preceded by an unstressed minor syllable (cf. Shorto 1960). Minor syllables have a syllabic sonorant (such as m in θmphūp ‘lung’) or a (probably) non-contrastive vowel which I have written as a (kaʔa ‘two’).

There is also a contrast between θ- and s- minor syllable initial in Hu (and in one word, ts- is attested). This opposition is not maintained in this position in U, Lamet or Proto-Waic (= PW; Diffloth 1980), where the most common corresponding initial is s-:

<table>
<thead>
<tr>
<th>Hu</th>
<th>U</th>
<th>Lamet</th>
<th>PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>θanāt</td>
<td>nāt</td>
<td>snāat</td>
<td>*snat</td>
</tr>
<tr>
<td>θaʔaw</td>
<td>saʔa</td>
<td>sʔāar</td>
<td>‘sour’</td>
</tr>
<tr>
<td>θaʔən</td>
<td>sʔən</td>
<td>‘dry’</td>
<td></td>
</tr>
<tr>
<td>θavāŋ</td>
<td>savāa</td>
<td>‘to ask’</td>
<td></td>
</tr>
<tr>
<td>θathāŋ</td>
<td>sathāt</td>
<td>‘old’</td>
<td></td>
</tr>
<tr>
<td>θathāʔ</td>
<td>sathā</td>
<td>ntāʔaʔ</td>
<td>*sntaʔ</td>
</tr>
<tr>
<td>θamāʔ</td>
<td>samāʔ</td>
<td>*maʔ</td>
<td>‘wind’</td>
</tr>
<tr>
<td>θmphūp</td>
<td>saphop</td>
<td>‘lung’</td>
<td></td>
</tr>
<tr>
<td>samōʔ</td>
<td>samō</td>
<td>*smō</td>
<td>‘stone’</td>
</tr>
<tr>
<td>salēʔ</td>
<td>salē</td>
<td>slēʔ</td>
<td>*hleʔ</td>
</tr>
<tr>
<td>saplāŋ</td>
<td>saʔāa</td>
<td>smplāʔ</td>
<td>‘shoulder’</td>
</tr>
<tr>
<td>saŋāy</td>
<td>nāy</td>
<td>nāay</td>
<td>*ŋay</td>
</tr>
<tr>
<td>sŋkhōʔ</td>
<td>khū</td>
<td>*ŋkoʔ</td>
<td>‘yesterday’</td>
</tr>
<tr>
<td>tsanāl</td>
<td>sanēn</td>
<td>snāal</td>
<td>*ŋal</td>
</tr>
</tbody>
</table>

Presumably, θ- and s- are the reflexes of Proto-Palaungic *s- and *ts- (< *c-), respectively, which have merged into s- in U, as usual. It may be noted that *s- has (at least in some cases) been retained in this position in
Lamet and Waic. As mentioned above, word initial *s- has usually become *h- in these languages. There are irregularities, which may be due to more complex initial consonant clusters.

The occurrence of minor syllable initial *ts- in *tsañål ‘blue’ suggests a different Proto-Palaungic consonant, presumably the reflex of Ferlus’ (1978) *tš-, since the Kammu cognate is *cňāar with initial c-.

Hu has also retained the contrast between s- (< *ts- < *c-) and θ- (< *s-) in major syllable initial position after a nasal minor syllable:

<table>
<thead>
<tr>
<th>Hu</th>
<th>U</th>
<th>Lamet</th>
<th>PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>*nsi?</td>
<td>*nchi</td>
<td>si?</td>
<td><em>si?</em> ‘louse’</td>
</tr>
<tr>
<td>*nasök</td>
<td>*sūr</td>
<td>yŏok</td>
<td>*hyok ‘ear’</td>
</tr>
<tr>
<td>*nīćac</td>
<td>*ntsăt</td>
<td>màac</td>
<td>*hmac ‘sand’</td>
</tr>
<tr>
<td>*nīćim</td>
<td>*ntsip</td>
<td>lmhiim</td>
<td>*mhem ‘claw’</td>
</tr>
</tbody>
</table>

Here, U has s in ‘ear’, where the minor syllable has disappeared, otherwise ch or tš (which are probably allophones of a single phoneme). After a nasal minor syllable, s never occurs in U, but has developed into tš/ch, which accounts for the unexpected occurrence of these consonants in the words for ‘sand’ and ‘claw’.

Hu also has a number of words which have a nasal minor syllable followed by a voiceless nasal major syllable initial:

<table>
<thead>
<tr>
<th>Hu</th>
<th>U</th>
<th>Lamet</th>
<th>PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>*mām</td>
<td>sanām</td>
<td>nāam</td>
<td>*hnam ‘blood’</td>
</tr>
<tr>
<td>*nīm</td>
<td>sanāp</td>
<td>nīm</td>
<td>*nvm ‘year’</td>
</tr>
<tr>
<td>*nāt</td>
<td>sanā</td>
<td>*nāt ‘comb’</td>
<td></td>
</tr>
<tr>
<td>*nēʔ</td>
<td>nē</td>
<td>*nēʔneʔ ‘meat’</td>
<td></td>
</tr>
<tr>
<td>*mūl</td>
<td>mūn</td>
<td>kmūul</td>
<td>*kmil ‘silver’</td>
</tr>
<tr>
<td>*nāt</td>
<td>nār</td>
<td>knāas</td>
<td>*kpas ‘to laugh’</td>
</tr>
</tbody>
</table>

The Hu forms suggest a *nasal + h initial cluster, while Lamet and Waic in some cases have clusters with a stop and a nasal. Taken together, this implies proto-forms with *stop + nasal + h clusters. Medial h has been lost in most of Palaungic (cf. Diffloth 1977), and in Hu, where h is retained, the initial stop has disappeared.

U often loses cluster initial stops (see Svantesson 1988), as is the case in the last three examples above. In the preceding list, the minor syllable sa in the first three words in U, taken together with evidence from outside Palaungic, suggests Proto-Mon-Khmer *j- (palatal voiced stop). This tallies perfectly with Diffloth’s (1980:175) reconstruction of ‘blood’ as Proto-Mon-Khmer *jnhaam. For the other two words, *jn- clusters are also supported by evidence outside Palaungic: Kammu criās (Southern Kammu jriās) ‘comb’, Mon cnām ‘year’.

70
Hu — A language with unorthodox tonogenesis

Final consonants
Hu has retained the Proto-Palaungic (and even Proto-Mon-Khmer) system of final consonants rather intact. An exception is final *-s, which has developed into -t (merging with original *-t), another innovation shared by Hu and Danaw (and several other Mon-Khmer languages as well):

<table>
<thead>
<tr>
<th>Hu</th>
<th>Danaw</th>
<th>Lamet</th>
</tr>
</thead>
<tbody>
<tr>
<td>?axêt</td>
<td>k'r'et</td>
<td>kriis</td>
</tr>
<tr>
<td>phôt</td>
<td>pyt²</td>
<td>pôos</td>
</tr>
</tbody>
</table>

The final palatal stop *-c sometimes appears as -t in Hu (cf. Hu pêt ‘to spit’, Lamet pèc; Hu ?amût ‘mosquito’, Lamet mûuc ‘ant’), but my data are too limited to reveal the exact circumstances under which this has taken place.

Initial *r- is realised as a uvular [k] (as is also the case in some Lamet dialects, in Bûlûngshăn Blang, and in some other languages of the area). In final position, [k] is pronounced as a rather vocalic uvular glide which can be written [ś] (incidentally a development which has also taken place in my own southern Swedish dialect! See Lindau 1985 for different kinds of r). Thus ?śr ‘fowl’ is pronounced [śśā]. After the vowel a, final *-r has disappeared, or is retained as -w:

<table>
<thead>
<tr>
<th>Hu</th>
<th>Lamet</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka?ą</td>
<td>?lāar</td>
</tr>
<tr>
<td>mà</td>
<td>màar</td>
</tr>
<tr>
<td>ða?aw</td>
<td>sʔār</td>
</tr>
<tr>
<td>kāw</td>
<td>kāar</td>
</tr>
</tbody>
</table>

Tones and vowels
There are two tones in Hu, namely, high (denoted by ’ over the vowel) and low (’). The co-occurrence of tones, vowels and final consonants is restricted, as shown in this table:

<table>
<thead>
<tr>
<th>Vowel:</th>
<th>i u i</th>
<th>e ə a ɔ</th>
<th>o ɔ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final:</td>
<td>ø ʔ k</td>
<td>others</td>
<td></td>
</tr>
</tbody>
</table>

As mentioned above, the general rule is that originally long vowels have conditioned low tone, and short vowels have conditioned high tone, and the co-occurrence restrictions probably reflect restrictions on the co-occurrence of long and short vowels with final consonants by the time
that the tone system was formed. Since Lamet is a Palaungic language which retains the Proto-Palaungic (and even Proto-Mon-Khmer) length distinction, I will use Lamet examples for showing the relationship between Proto-Palaungic vowel length and Hu tones. For non-high vowels, the relationship is quite clear-cut:

<table>
<thead>
<tr>
<th>Hu</th>
<th>Lamet</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>*short vowels:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yám</td>
<td>yám</td>
<td>'to die'</td>
</tr>
<tr>
<td>paθán</td>
<td>phán</td>
<td>'five'</td>
</tr>
<tr>
<td>mën</td>
<td>krmën</td>
<td>'star'</td>
</tr>
<tr>
<td>ncën</td>
<td>kcën</td>
<td>'heavy'</td>
</tr>
<tr>
<td>*long vowels:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yám</td>
<td>yáam</td>
<td>'to cry'</td>
</tr>
<tr>
<td>lèk</td>
<td>liik</td>
<td>'pig'</td>
</tr>
<tr>
<td>?3m</td>
<td>?6om</td>
<td>'water'</td>
</tr>
<tr>
<td>nasök</td>
<td>yōök</td>
<td>'ear'</td>
</tr>
</tbody>
</table>

Before a final glottal stop, the tone is always high, probably due to shortening of the vowel in this position before the development of tones. For comparison, forms from (Northern) Kammu are given. This language, which has orthodox tonogenesis, is another language where the length distinction in vowels is lost before a glottal stop. Examples:

<table>
<thead>
<tr>
<th>Hu</th>
<th>Lamet</th>
<th>Kammu</th>
</tr>
</thead>
<tbody>
<tr>
<td>s5ʔ</td>
<td>s5ʔ</td>
<td>s5ʔ</td>
</tr>
<tr>
<td>kathéʔ</td>
<td>ktáʔ</td>
<td>ptéʔ</td>
</tr>
<tr>
<td>?5ʔ</td>
<td>?6ʔ</td>
<td>?6ʔ</td>
</tr>
<tr>
<td>phléʔ</td>
<td>pléeʔ</td>
<td>pléʔ</td>
</tr>
</tbody>
</table>

The high vowels i and u always have high tone, except before k or in open syllables, where both tones occur. Compare the following examples with long proto-vowels (I have no examples with long *ii):

<table>
<thead>
<tr>
<th>Hu</th>
<th>Lamet</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>?asim</td>
<td>siim</td>
<td>'bird'</td>
</tr>
<tr>
<td>phám</td>
<td>priim</td>
<td>'old'</td>
</tr>
<tr>
<td>mmúul</td>
<td>kmúul</td>
<td>'silver'</td>
</tr>
<tr>
<td>?úp</td>
<td>?úup</td>
<td>'cooked rice'</td>
</tr>
<tr>
<td>Ṝum</td>
<td>húum</td>
<td>'to bathe'</td>
</tr>
<tr>
<td>But:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>?iœ</td>
<td>?œër</td>
<td>'fowl'</td>
</tr>
</tbody>
</table>

One possible explanation for the absence of a tone contrast in the high vowels is that they have higher intrinsic pitch than low vowels, as has been shown for many different languages (see, e.g., Lehiste 1970: 68-71), which might have conditioned high tone for both long and short high vowels when the Hu tone system developed. This explanation does not, however, account for the occurrence of high vowels with low tone on open syllables.
Hu – A language with unorthodox tonogenesis

On the other hand, there are indications that the length contrast was already lost in the high vowels in Proto-Angkuic, i.e. before the development of the Hu tone system. Thus, judging from the words given in Diffloth (1982a), there is no length contrast for \( u \) and \( i \) in Mok, while the contrast is retained for non-high vowels. In U, the vowel length contrast has disappeared, although it has left traces in final nasals which are retained after originally long vowels but have become stops after originally short vowels. Denasalisation has, however, taken place after both *long and *short \( i \) and \( u \), which are thus treated as if they were short.

The following examples show this development. Lamet cognates are given because they retain Proto-Palaungic vowel length:

<table>
<thead>
<tr>
<th>Lamet</th>
<th>Hu</th>
<th>U</th>
<th>Mok</th>
</tr>
</thead>
<tbody>
<tr>
<td>*non-high long:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yäm</td>
<td>yäm</td>
<td>yäm</td>
<td>naam</td>
</tr>
<tr>
<td>?öm</td>
<td>?öm</td>
<td>?öm</td>
<td>?oöm</td>
</tr>
<tr>
<td>pön</td>
<td>pöphon</td>
<td>phon</td>
<td>phoon</td>
</tr>
<tr>
<td>*non-high short:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yäm</td>
<td>yäm</td>
<td>yap</td>
<td>yem</td>
</tr>
<tr>
<td>ntäm</td>
<td>nthäm</td>
<td>ntháp</td>
<td>tham</td>
</tr>
<tr>
<td>kcën</td>
<td>kcën</td>
<td>kêt</td>
<td>kacen</td>
</tr>
<tr>
<td>*high long:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>siöm</td>
<td>?asim</td>
<td>pachip</td>
<td>?a-sim</td>
</tr>
<tr>
<td>kuaŋ</td>
<td>khpŋ</td>
<td>khit</td>
<td>khun</td>
</tr>
<tr>
<td>priim</td>
<td>phkim</td>
<td>xip</td>
<td>phim</td>
</tr>
<tr>
<td>*high short:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kʰŋ</td>
<td>(kʰŋ)</td>
<td>kʰŋ</td>
<td>kʰŋ</td>
</tr>
</tbody>
</table>

(The Hu word kʰŋ, which occurs in kʰŋ kʰŋ ‘knee’ may be cognate to the words meaning ‘head’ in the other languages.)

The reason for the loss of vowel length in the high, but not in the other, vowels may be their shorter intrinsic length, something which has been attested for various languages (see Lehiste 1970: 18-19). Furthermore, the length contrast in the high vowels seems to have carried a rather small functional load.

As mentioned above, final -ŋ is more or less vocalic, which may explain why low tone can occur on high vowels before this final, as it does in open syllables.

The reason why \( o \) and \( ø \) occur only with low tone (except before \( ? \)) in my data is probably that there were relatively few words with short *\( o \) and *\( ø \). This is the case in Kammu, where short and long \( ø \) do not contrast (see Svantesson 1983).

---

1. Proto-Angkuic did not have tones, as is proved by the absence of tones in Mok (Diffloth 1982a). U has a tone system, which is different from that in Hu, and for the other Angkuic languages it is difficult to know whether they have tones or not, since they are known only from older and not very reliable sources, which do not give any tones.

73
Vowel length or tones?

Figure 1 below shows some typical examples of the fundamental frequency ($F_0$) contours of the two tones on different types of syllables:

![Graph of fundamental frequency contours for 'yam', 'yam', 'pap', and 'khap'.]

In these cases, as in many others, there is a co-variation between fundamental frequency and vowel duration, so that vowels carrying low tone have longer duration than vowels carrying high tone. One might, therefore ask whether an analysis in terms of vowel length rather than tones is possible.

To investigate this, the duration of the vowel and the mean value of the fundamental frequency over the vowel were computed (using the ILS* program package) for some words with high and low tone. The words were said in isolation by the female informant. Since the tone contours are rather flat, the average frequency value can be used to characterise the tones. The results are given in Table 1.

As seen in this table, vowels carrying low tone often have longer duration than those with high tone, as seen in the only recorded minimal pair, 'yám' 'to die' and 'yàm' 'to cry'. On the other hand, some vowels with high tone are longer than some with low tone, and this overlapping of duration can be taken as evidence against treating vowel length as distinctive. Each of the analysed high-tone words also has higher fundamental frequency than each of the low-tone words.

Thus, fundamental frequency is definitely a consistent phonetic correlate of the investigated opposition, while vowel length may be regarded as a concomitant factor. It seems, therefore, reasonable to analyse the opposition as one consisting of two tones.

* ILS = Interactive Laboratory System (Ed.).
Hu – A language with unorthodox tonogenesis

Table 1

<table>
<thead>
<tr>
<th></th>
<th>( F_0 ) (Hz)</th>
<th>Duration (ms)</th>
<th>Mean values:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High tone:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yâm</td>
<td>269</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>247</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>263</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>263</td>
<td>120</td>
<td>260</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>pâp</td>
<td>253</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>249</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>95</td>
<td>248</td>
</tr>
<tr>
<td></td>
<td>102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kâk</td>
<td>253</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>252</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>258</td>
<td>120</td>
<td>254</td>
</tr>
<tr>
<td></td>
<td>117</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low tone:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yâm</td>
<td>214</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>215</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td></td>
<td>215</td>
<td>225</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>khâp</td>
<td>204</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>203</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>208</td>
<td>115</td>
<td>205</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>?àk</td>
<td>205</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>201</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>201</td>
<td>180</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>188</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The \( F_0 \) ranges are 242-269 Hz for high tone and 201-215 for low tone.

**Conclusion**

The Hu data raise a number of intriguing questions, two of which will be discussed here.

One concerns the classification of Palaungic. There are some striking similarities between Hu and Danaw. In particular, both have \( \theta- \) where Proto-Mon-Khmer has *s-, and this is found nowhere else in Palaungic. In the sub-classification of Palaungic given in Diffloth (1977), Danaw is close to Angkuic, but in Diffloth (1982a), a classification (shown on p. 67), which puts Danaw and Angkuic rather far from each other is given. Mitani (n.d.), using lexicostatistical methods for classifying Palaungic, also finds that Danaw and Angkuic are rather distant from each other. One might ask if my data from Hu—being an Angkuic language which shares the innovation *s- > \( \theta- \) with Danaw—necessitates a revision of this. Not necessarily, since it is quite possible that the development *s- > h-, which has taken place in all Palaungic languages except in Angkuic and Danaw, is the final result of two different processes: *s- > \( \theta- \) and \( \theta- > h- \). If that is the case, Hu and Danaw are not languages which share an
early innovation, but rather languages where a phonological rule \((\theta - > h- )\) has not applied, and there is no reason to assume that they have branched off together from the rest of Palaungic at some early time.

Another question concerns tonogenesis. In almost all cases where a Mon-Khmer language has acquired tones (or registers), the development of the tone (or register) system is the result of a loss of contrasts—usually voicing contrasts—in the consonants, but in Hu, the tones have developed in connection with the loss of vowel length. The only other case known to me, in which tonogenesis of a similar kind may have taken place, is in Estonian, where a tonal distinction (different from that in Hu) has possibly developed from an earlier vowel length opposition (see Lehiste 1978). One somewhat similar case in Mon-Khmer is Pacoh which, according to Diffloth (1982b), has an unorthodox regiogenesis, where a register difference has replaced an earlier difference in vowel quality.

Both the acquisition of tones and the loss of vowel length are ongoing processes in the area where Hu is spoken, so it is perhaps not surprising to find a language that combines both. There might be a phonetic explanation as to why long vowels have acquired low tone and short vowels high tone, since there seems to be a general tendency for vowel duration and fundamental frequency to vary inversely with each other. As already mentioned, it has been shown for many languages that high vowels have intrinsically shorter duration and higher pitch than low vowels. For instance, measurements of the intrinsic pitch and duration in Standard Chinese vowels (pǔtōnghuà) have shown that, if other factors are constant, vowels with relatively high intrinsic pitch also have relatively short duration (Shi Bō, pers. comm.). Naturally, pitch differences of 40 Hz or more, as are found in Hu, are not the result of automatic adjustments, but an originally non-distinctive pitch difference could have taken over some of the functional load carried by vowel length, eventually acquiring phonemic status.

Vocabulary

The vocabulary is presented in reverse alphabetical order.

<table>
<thead>
<tr>
<th>Dictionary entry</th>
<th>Pinyin</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka?á</td>
<td>‘two’</td>
<td>mà ‘dry field’</td>
</tr>
<tr>
<td>masà</td>
<td>‘thing’</td>
<td>pà ‘not’</td>
</tr>
<tr>
<td>ści</td>
<td>‘to go’</td>
<td>?amò ‘one’</td>
</tr>
<tr>
<td>?a?ś</td>
<td>‘monkey’</td>
<td>mè ‘axe’</td>
</tr>
<tr>
<td>lû</td>
<td>‘bad’</td>
<td></td>
</tr>
<tr>
<td>lā?</td>
<td>‘leaf’</td>
<td>θamá? ‘wind’</td>
</tr>
<tr>
<td>ṇā?</td>
<td>‘to itch’</td>
<td>masá? ‘to steal’</td>
</tr>
<tr>
<td>θathā?</td>
<td>‘tail’</td>
<td>?ē? ‘we (pl.)’</td>
</tr>
<tr>
<td>ké?</td>
<td>‘they (pl.)’</td>
<td>salè? ‘rain’</td>
</tr>
<tr>
<td>phlé?</td>
<td>‘fruit’</td>
<td>mè? ‘you (sg.)’</td>
</tr>
<tr>
<td>Hu</td>
<td>A language with unorthodox tonogenesis</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----------------------------------------</td>
<td></td>
</tr>
<tr>
<td>kamé?</td>
<td>‘dream’</td>
<td>thamé?</td>
</tr>
<tr>
<td>nhé?</td>
<td>‘meat’</td>
<td>phé?</td>
</tr>
<tr>
<td>thé?</td>
<td>‘tree’</td>
<td>thé?</td>
</tr>
<tr>
<td>kathé?</td>
<td>‘earth’</td>
<td>vé?</td>
</tr>
<tr>
<td>ē?</td>
<td>‘day’</td>
<td>kē?</td>
</tr>
<tr>
<td>ē?</td>
<td>‘pestle’</td>
<td>pē?</td>
</tr>
<tr>
<td>pasi?</td>
<td>‘string’</td>
<td>nŚ?</td>
</tr>
<tr>
<td>paś?</td>
<td>‘blood-vessel’</td>
<td>thē?</td>
</tr>
<tr>
<td>ēkhō?</td>
<td>‘rice (husked)’</td>
<td>snēkhō?</td>
</tr>
<tr>
<td>samō?</td>
<td>‘stone’</td>
<td>pō?</td>
</tr>
<tr>
<td>ē?</td>
<td>‘I’</td>
<td>ṭaklŚ?</td>
</tr>
<tr>
<td>sō?</td>
<td>‘dog’</td>
<td>katś?</td>
</tr>
<tr>
<td>palū?</td>
<td>‘salt’</td>
<td>xu?</td>
</tr>
<tr>
<td>nθāc</td>
<td>‘sand’</td>
<td></td>
</tr>
</tbody>
</table>

| lih | ‘to go down’ |
|??? | ‘bow’ | kāk | ‘to bite’ |
| thēk| ‘buffalo’ | ṭāk | ‘rice (plant)’ |
| phlēk | ‘palm (of hand)’ | nthēk | ‘tongue’ |
| lēk | ‘pig’ | phēk | ‘ribs’ |
| ēθēk | ‘rat’ | tēk | ‘small’ |
| ōk | ‘flea’ | nasēk | ‘ear’ |
| nthēk | ‘head’ | ōk | ‘hair’ |
| thēk | ‘to hang’ |

| ŋāl | ‘fire’ | tsāŋāl | ‘blue’ |
| mphläl | ‘mortar’ | phēl | ‘wing’ |
| ntōl | ‘wine’ | mŋūl | ‘silver’ |
| katūl | ‘belly’ |

| nŋūm | ‘blood’ | nthūm | ‘egg’ |
| yām | ‘to die’ | yām | ‘to cry’ |
| nēm | ‘younger brother’ | ?aŋēm | ‘right (side)’ |
| ōm | ‘to live’ | kūm | ‘village’ |
| phēm | ‘old (of things)’ | ?asēm | ‘bird’ |
| nōm | ‘claw’ | nŋūm | ‘year’ |
| ōm | ‘water’ | kathōm | ‘liver’ |
| nūm | ‘piss’ | tŋūm | ‘under’ |
| ōm | ‘to bathe’ |

<p>| lān | ‘long (in space)’ | pǎθān | ‘five’ |</p>
<table>
<thead>
<tr>
<th>Thán</th>
<th>‘see nêy, ʔûn’</th>
<th>Θathán</th>
<th>‘old’</th>
</tr>
</thead>
<tbody>
<tr>
<td>ncèn</td>
<td>‘heavy’</td>
<td>ièn</td>
<td>‘low’</td>
</tr>
<tr>
<td>ʔón</td>
<td>‘he, she, it’</td>
<td>món</td>
<td>‘long (in time)’</td>
</tr>
<tr>
<td>phín</td>
<td>‘woman’</td>
<td>Φaphín</td>
<td>‘four’</td>
</tr>
<tr>
<td>khôn</td>
<td>‘child’</td>
<td>khôn~thîː</td>
<td>‘finger’</td>
</tr>
<tr>
<td>kaʔàŋ</td>
<td>‘wasp’</td>
<td>pàŋ</td>
<td>‘white’</td>
</tr>
<tr>
<td>tânh</td>
<td>‘to weave’</td>
<td>mén</td>
<td>‘star’</td>
</tr>
<tr>
<td>padéŋ</td>
<td>‘snake’</td>
<td>phín</td>
<td>‘to shoot’</td>
</tr>
<tr>
<td>thàŋ</td>
<td>‘dry’</td>
<td>ʔûn~thân:</td>
<td>‘grandfather’</td>
</tr>
<tr>
<td>‘ʔaʔûn’</td>
<td>‘father’</td>
<td>khûn</td>
<td>‘man; husband’</td>
</tr>
<tr>
<td>marûn</td>
<td>‘ant’</td>
<td>ntûn</td>
<td>‘mouth’</td>
</tr>
<tr>
<td>kaʔàŋ</td>
<td>‘bone’</td>
<td>kàŋ</td>
<td>‘house’</td>
</tr>
<tr>
<td>kàŋ~çáw:</td>
<td>‘heaven’</td>
<td>saplîŋ</td>
<td>‘shoulder’</td>
</tr>
<tr>
<td>kàŋkâŋ</td>
<td>‘iron’</td>
<td>kàŋ</td>
<td>‘flower’</td>
</tr>
<tr>
<td>màkâŋ</td>
<td>‘horse’</td>
<td>săŋ</td>
<td>‘bitter’</td>
</tr>
<tr>
<td>thàŋ</td>
<td>‘to kill’</td>
<td>pàvâŋ</td>
<td>‘tomorrow’</td>
</tr>
<tr>
<td>thàvàŋ</td>
<td>‘to ask’</td>
<td>cèŋ</td>
<td>‘foot’</td>
</tr>
<tr>
<td>lèŋ</td>
<td>‘high’</td>
<td>skèŋ</td>
<td>‘red’</td>
</tr>
<tr>
<td>ntièn</td>
<td>‘big’</td>
<td>thèŋ</td>
<td>‘drink’</td>
</tr>
<tr>
<td>phîn</td>
<td>‘many’</td>
<td>thêŋ~ʔîː:</td>
<td>‘bedbug’</td>
</tr>
<tr>
<td>khîŋ</td>
<td>‘tooth’</td>
<td>kîn~kîŋ:</td>
<td>‘knee’</td>
</tr>
<tr>
<td>màkîŋ</td>
<td>‘crab’</td>
<td>wîŋ</td>
<td>see θêk~</td>
</tr>
<tr>
<td>thîŋ</td>
<td>‘bamboo’</td>
<td>kîn</td>
<td>see khîŋ~</td>
</tr>
<tr>
<td>môŋ</td>
<td>‘to look’</td>
<td>?aʔûn</td>
<td>‘wasp’</td>
</tr>
<tr>
<td>xôi</td>
<td>Mekhong</td>
<td>yôŋ</td>
<td>‘good’</td>
</tr>
<tr>
<td>khûŋ</td>
<td>‘wet field’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>khàp</td>
<td>‘jaw’</td>
<td>pàp</td>
<td>‘to speak’</td>
</tr>
<tr>
<td>θaʔèp</td>
<td>‘rainbow’</td>
<td>lêp</td>
<td>‘blind’</td>
</tr>
<tr>
<td>cîp</td>
<td>‘to run’</td>
<td>?up</td>
<td>‘(cooked) rice’</td>
</tr>
<tr>
<td>ʔmphûp</td>
<td>‘lung’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔiık</td>
<td>‘fowl’</td>
<td>khîk</td>
<td>‘moon’</td>
</tr>
<tr>
<td>phaʔîk</td>
<td>‘bee’</td>
<td>kîk</td>
<td>‘finished’</td>
</tr>
<tr>
<td>phîk</td>
<td>‘to fly’</td>
<td>kîrk</td>
<td>‘to sew’</td>
</tr>
<tr>
<td>phaʔát</td>
<td>‘to swell’</td>
<td>khát</td>
<td>‘sick’</td>
</tr>
<tr>
<td>ʔanàt</td>
<td>‘gun’</td>
<td>nîât</td>
<td>‘comb; to comb’</td>
</tr>
<tr>
<td>nîât</td>
<td>‘to laugh’</td>
<td>thasît</td>
<td>‘lightning’</td>
</tr>
<tr>
<td>pêt</td>
<td>‘to spit’</td>
<td>khasêt</td>
<td>‘charcoal’</td>
</tr>
<tr>
<td>?axèt</td>
<td>‘bear’</td>
<td>ʔêt</td>
<td>‘to sleep’</td>
</tr>
<tr>
<td>ʔôt</td>
<td>‘to be at’</td>
<td>phôt</td>
<td>‘sambar deer’</td>
</tr>
</tbody>
</table>
Hu – A language with unorthodox tonogenesis

| θʔt | 'barking deer' |
| thūt | 'breast' |
| θaʔw | 'sour' |
| ?apaʔw | 'fish' |
| vəw | 'wide' |
| ?anəw | 'cat' |
| ?əy | 'we (dual)' |
| ?aləy | 'squirrel' |
| sanəy | 'eye' |
| nəy ~ thàn: 'grandmother' |
| khəy | 'to have' |
| kəʔy | 'three' |

References


ON AUSTRONESEAN LEXICON
IN VIETNAMESE

Kenneth Gregerson

The sources of Vietnamese lexicon have been much discussed as regards Chinese, Tai, and Mon-Khmer vocabulary (Maspero 1912; Schmidt 1926; Haudricourt 1954), and one may now presume that Vietnamese is rather safely established as an Austroasiatic language. The Austronesian (AN) subset of Vietnamese lexicon, however, while clearly not having gone unnoticed (qv. discussion in Benedict 1976), continues to invite a good deal of sorting out. The present paper mentions some well-known Austronesian forms as well as raising questions about whether certain other Vietnamese items are ultimately of Austronesian (or Austro-Thai?) origin. The effort here is to contribute towards a more deliberate Vietnamese perspective on Austronesian lexical associations with Austroasiatic.

1. Function forms
Vietnamese possesses several very regularly used grammatical or closed-set function words that appear to have Austronesian counterparts.

1.1. Aspectuals
The following forms function as regular Vietnamese (VN) aspectuals:

1) a. dā ‘already’
   Tōi dā mua xe rōi. I already buy auto already
   ‘I have already bought a car.’

b. sāp ‘about to’
   Tōi sāp mua xe. I about-to buy a car
   ‘I am about to buy a car.’

c. dang ‘in process’
   Ông ây dang làm viéc. grandfather that in-process do work
   ‘He is working.’

1. A version of this paper was presented at the Eighteenth International Conference on Sino-Tibetan Languages and Linguistics (Bangkok) in 1985. I gratefully acknowledge helpful comments from Paul Benedict, James Matisoff, Nguyễn Đình Hoaised, and Bill Gage, though certain things still survive that are perhaps not to their liking.

2. The ultimate Austronesian origin of Vietnamese is a notion that has existed for some time (cf. e.g. discussion in E. Sauvignet 1922; K. Wulff 1942; as well as by Binh Nguyễn Lộc, c.1960 in his Ngữ gõc MAlai, so Prof. Hồ informs me).

3. Some, though not all, Vietnamese illustrative forms are from Nguyễn Đình Hồased, 1971.
d. māi "continuative"
   Ho đì bō māi đên Thú-Dúc.
   They go by-foot continue to Thú-Dúc
   'They walked all the way to Thú-Dúc.'

These forms may be compared with the following in Bahasa Indonesia (BI):4

(2)a. sudah 'already'
   Ia sudah pergi.
   he already go
   'He has already gone.'

b. siap 'ready'
   Mereka siap untuk pergi.
   they ready for go
   'They are ready to go.'

c. sedang 'while, in process'
   Ia sedang membaca ketika saya datang.
   he in-process read when I arrive
   'He was reading when I arrived.'

d. masih 'still, yet'
   ia masih tidur.
   he still sleep
   'He is still sleeping.'

As exemplified above in 1a and 2a, VN dā 'already': BI sudah 'already' both occur in preverbal position to signal Perfective aspect. In VN it will be observed that a clause-final rōi reinforces the same completive meaning redundantly. This form is perhaps the 'true' Austroasiatic marker (cf. Rengao (Rg.) hadroi 'before') into whose territory dā has intruded.

The tone on dā accords well with diachronic expectations clarified originally by Haudricourt (1954), which may be generally summarised as follows:

(3)

<table>
<thead>
<tr>
<th>Original Finals</th>
<th>open syllable</th>
<th>final stop</th>
<th>final spirant</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless</td>
<td>C Yük(N)</td>
<td>CV</td>
<td>CV H high tones</td>
</tr>
<tr>
<td>original initials</td>
<td>CV</td>
<td>CV C</td>
<td>CV H low tones</td>
</tr>
<tr>
<td>voiced</td>
<td>CV</td>
<td>CV C</td>
<td>CV H low tones</td>
</tr>
</tbody>
</table>

4. A number of the sample Indonesian sentences and lexicon are from Echols and Shadily (1974).
Thus, depending on the original voicing status of the initial consonant (C-) and the closure status of the final (-V(N), -VC, or VH), a particular contrastive tone has developed in Vietnamese, typically, with a concomitant loss of most of the original conditioning features. The High Tones associated with old voiceless or imploded consonants are V (symbolised in our exposition with \( \overline{V} \) for clarity though it (the macron \( \overline{\text{\textnormal{\textbar}}} \)) is not written in Vietnamese orthography), \( \overline{V} \), and \( \overline{\text{\textnormal{\textbar}}} \). The Low Tones that originally occurred with voiced initial consonants are \( \dot{V} \), \( \acute{V} \), and \( \ddot{V} \).

At this point, however, when dealing with disyllabic forms, such as BI *sudah*, it is crucial to unravel the 'pecking order' that dictates which initial consonant counts as to voicing status in the selection of the High vs. Low tone set in (3) above. That is, in *sudah* does the s- or the d- take precedence? Clearly, the voiced initial d- had prevailed, and the \( \dot{V} \) tone has been appropriately selected, and *sudah* > *dåa*.

To clarify this consonant precedence, it is perhaps worth the digression to discuss the discoveries of Friberg and Hor (1977) as to initial consonant 'dominance' with respect to the register (+ATR* vs. -ATR) selection of stressed (main) syllables in Western Cham (Austronesian of southern Vietnam and Cambodia). Distinguishing register A (-ATR) consonants from register B (+ATR), Friberg and Hor (1977: 36) summarise how precedence is established and register effects determined on the phonation of vowel quality of the succeeding syllable:

\[
\begin{array}{cccccc}
A (-\text{ATR}) & B (+\text{ATR}) \\
1. & p & t & c & k & b & d & j & g \\
& ph & th & ch & kh & bh & dh & jh & gh \\
2. & s & h & m & n & ñ & ñ & w & l & y & r
\end{array}
\]

Register A and B in Western Cham words were accounted for by the following 'dominance' rules:

\((5)\)a. \( A + A = A \)
\( b. \ B + B = B \)
\( c. \ A1 + B1 = \text{second element} \)
\( d. \ B1 + A2 = B \)
\( e. \ A + B2 = A \)

These rules (following Purtle (1969) for Khmer), although neither ordered with respect to each other nor within the left-hand members, indicate register of tonic syllable. Quadrant A1 and B1 are equally strong; they both dominate A2 and B2. A2 also dominates B2. This analysis is based entirely on the consonant 'strength'; the consonant is seen to 'dominate' or determine the register characteristics of the following vowel. And, based on the five 'dominance' rules noted above, certain atonic syllable initial

\* = Advanced Tongue Register (Ed.).
consonants exert their strength over an intervening tonic syllable initial to determine the ‘registerness’ of the tonic syllable vowel.

In the following examples, consonant dominance is seen in the combining of various syllables to form words (a grave accent /`/ is added to indicate second (+ ATR) register clearly):

(6) A1 + B1 /ka/ + /baw/ /kabaw/ ‘buffalo’
    B1 + A1 no examples (historically B1 has become A1).
    B1 + A2 /bal/ + /haw/ /bahaw/ ‘new’
    A2 + B1 /ha/ + /dom/ /hadom/ ‘how much’
    A1 + B2 /ka/ + /ro/ /kar o/ ‘strong’
    B2 + A1 /la/ + /kaw/ /lakaw/ ‘to step over’
    A2 + B2 /ha/ + /nin/ /hanin/ ‘bow’
    B2 + A2 /la/ + /say/ /lasay/ ‘cooked rice’

Returning now to sudah > ḏā, one may usefully compare this process with Friberg and Hor’s A2 + B1, in which /ha/ + /dom/ > /hadom/, the word initial spirant yielding to the main syllable initial stop as the prosody determining element. Then, of course, Vietnamese ultimately reduces the form to a monosyllable.

Resuming our discussion of the other aspectuals, it will be observed that the pair VN sâp ‘about to’: BI siap ‘ready’ are also preverbal forms. The tone of sâp results straightforwardly from its syllable type as CVC (cf. 1.1.(3) above).

For VN dang ‘in process’: BI sedang ‘in process’ one must, as with ḏā, assume a simplification which ultimately drops the first syllable. As to tone assignment, one expects either level (unmarked) or ` with a CVN syllable (no. (3) above), but with Friberg and Hor’s rules an initial s-should have yielded to d- and resulted in *dâng. Perhaps the situation is more complex historically and Friberg and Hor-type dominance is further conditioned by other factors. A prime candidate is perhaps an original stress difference in Austronesian. Specifically, for example, note:

(7) BI ‘sudah ‘already’
    BI se’dang ‘in process’

in which VN ḏā derives from an unstressed syllable while dang corresponds to a stressed one.

The pair VN māi ‘continuative’: BI masih ‘still, yet’ do not function quite like the other aspectuals above, for māi occurs as a post-verbal while masih is a pre-verbal. Phonetically, one can assume a reduction masih > maih (i.e. CVH), after which, given the voicing of the initial, the resultant māi is completely expected (qv. (3) above).

1.2 Desiderative
The following sentences are instances of desiderative modality in VN or BI:
(8)a. VN: *em nay muon ve nhay.*  
younger sibling this want return home  
'He/she (young one) wants to go home.'

b. BI: *Dia mau datang sore ini.*  
3-sg. want return home this  
'She wants to come this evening.'

That is, VN *muon* 'want' and BI *mau* 'want' both function as conventional preverbal desiderative forms. The question is 'Do they have any historical connection?' The similarity of form is strengthened if one assumes that *muon* derives from the nominalised (*ke* ... *an*) form *ke-mau-an* 'a wish'.

If we again invoke consonant dominance (qv. 1.1 (5) and (6) above), the initial *k*- would determine one of the three High tones (*V, V, V*), but why *V* (*muon*), when CVN would seem to predict *V* (*muon*?). Again, the hypothesis of a nominalisation source may provide an answer, for suffixes in Indonesian (cf. also Philippine languages) regularly insert a (?) between vowel sequences, thus (BI):

*ke-mau-?an* 'wish (n.)'

which would provide the explanatory feature of stop in the final (-?N) to produce an expected *muon* 'want'.

1.3 Equative

Consider the following sentences in VN and BI:

(9)a. VN: *Anh toi la gio sa.*  
older brother my 'is' teacher  
'my older brother is a teacher.'

b. BI: *Bahasa Indonesia ialah bahasa kebangsaan*  
language Indonesia 'is' language national  
'Indonesian is the national language.'

Structurally, VN *lã* and BI *ialah* (lit. *i 3rd sg.* + *lah* 'emphatic') operate in remarkably similar ways. Assuming the reduction to one syllable *lah*, the problem for tone would be that CVH predicts a form *lã* rather than *lã*. The indication then, is that the Austronesian form, if the connection is authentic, must itself have been reduced to *la* at the time it was given a tonal interpretation in Vietnamese. This account of *lã* is problematical, however, for Nguyễn Đình Hoà has pointed out (pers. commun.) that in the old Chữ Nôm data this copular *lã* turns up as *làm* 'to do', and his suggestion is that that is, indeed, its origin. Since, on the other hand, it is acceptable even in Modern Vietnamese to use *làm* as well as *lã* in a copular sense, it is not clear to me that they could not be independent forms.
1.4 Pronouns
There appear in Vietnamese a number of pro-forms of a locative, personal or interrogative nature that bear a good deal of resemblance to Austronesian forms, among which are the following.6

(10) VN kia; kia ‘there’: UAN *ija ‘he, she, it’, BI ia, Chmr. gwid’a, SAt. hia

VN nay, này: PAN *iniH$_2$, BI ini ‘this’ sini ‘here’, Rade nei (cf. similar forms in Tai).

VN nó [arrogant] ‘3-sg.’, chủng nó ‘they’; PAN *na ‘3-sg.’, Agta na ‘3-sg.’, BI sana ‘there’.

VN ta [arrogant] I, chủng ta ‘we incl.’: PAN *(k)ita, *ta ‘we incl.’, BI kita.

VN mà ‘which’ (rel. pron.), mô ‘what, where?’: BI mana ‘where, which’ (interrog. pron.).

1.5 Adverbs


VN rải ‘very’: UAN *bọvat ‘heavy’, BI berat ‘heavy’ (cf. VN hoi lit. ‘vapour’ = ‘rather’).

VN lâm ‘very’: Haroi hlam ‘very’, cf. BI selama ‘as long as’, selama lamanya ‘at the most’.

VN xa ‘far’: PMP *za[h]ouq, BI jauh, Chmr. t’ag’o.

2. Content forms
Vietnamese has, in addition to the more ‘grammatical’ forms above, quite a large number of general lexical items that also bear enough resemblance to Austronesian forms to have been noted by a number of investigators.

5. Austronesian citations are in general from Dahl (1977). Rade forms are from Egerod (1978), other Chamic citations are from Burnham (1976). Waic references are from Diffloth (1980).

6. Abbreviations for languages cited without full text references are:

<table>
<thead>
<tr>
<th>Am.</th>
<th>Mlg.</th>
<th>Rg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ami</td>
<td>Malagasy</td>
<td>Rengao</td>
</tr>
<tr>
<td>Bs.</td>
<td>MVN Middle (17th-century) Vietnamese</td>
<td>Ru. Rukai</td>
</tr>
<tr>
<td>BsE in Encarnacion</td>
<td>NgD Ngaju Dayak</td>
<td>SA. Squilq Atayal</td>
</tr>
<tr>
<td>Bu.</td>
<td>PAA Proto-Austro-Asiatic</td>
<td>Sir. Siraya</td>
</tr>
<tr>
<td>Chmr. Chamorro</td>
<td>PAN Proto-Austronesian,</td>
<td>SL. Southern Li</td>
</tr>
<tr>
<td>CLi. Central Li</td>
<td>Dahl’s constructions</td>
<td>TB Toba-Batak</td>
</tr>
<tr>
<td>Jav. Javanese</td>
<td>PMP Proto-Malayo-Polynesian</td>
<td>Tg. Tagalog</td>
</tr>
<tr>
<td>Kv. Kuvalan</td>
<td>UAN Uraustronesisch,</td>
<td>Th. Thao</td>
</tr>
<tr>
<td>Lq. Laqua</td>
<td>Pu. Puyama</td>
<td>Dempwolff’s con-</td>
</tr>
<tr>
<td>Mak. Makassarese</td>
<td>Pw. Paiwan</td>
<td>structions</td>
</tr>
<tr>
<td>Md. Madurese</td>
<td>Pz. Pazeh</td>
<td></td>
</tr>
</tbody>
</table>

86
2.1. Body parts

VN luɾi ‘tongue’: UAN *dilah ‘tongue’, Jav. dilah, BI lidah, Pw. liɗaliɗ, Rade lah, Haroi caliah, Cham talah ‘id’.
VN trai ‘left hand’: PAN *uiyi ‘left (hand)’, BI kiri, Mak. ka-iri, Pu. tama-wiri’.
VN tai ‘ear’: PAN *t שהיו, BI telinga ‘ear’, cf. CLI. thāi.
VN dâu ‘head’: PMP *qulu ‘head’, Tg. ulo, BI hulu ‘upper end, head’, cf. also Lq. ru, SLi. dau ‘head’.

2.2. Humans and body functions

VN tō, trō ‘show’: UAN *tu(n)duh, Jav. tuduh, BI tunjuk ‘point’, Tg. turo ‘instruction’, Bs. tolo ‘finger’, Am. to ro’ ‘point’.
VN ān ‘eat’: PMP *ka’an ‘eat’, Tg. kaain, SAT. qan-iq. cf. Rg. kaq ‘eat (meat)’.
VN ngó ‘bend head down’: BI angguk ‘nod’ cf. BI anggut ‘nod, (ship) pitch’.
VN ngcii ‘scratch’: UAN *gatal ‘itch’, Jav. gatal, BI garit, Md. ghatal, Pw. gatsal, cf. also BI kais ‘scrape for food’.
VN nghe ‘hear’: UAN *daŋa ‘hear’, BI dengar, Tg. dinig, Pu. marrgai, Rade kna ‘ear’ (< *taliŋa).
VN đa (MVN đa) ‘belly’: UAN u(n)tah ‘vomit’ BI muntah; SAT. m-utaq. Note that, here, VN retains the Austronesian verbal prefix m-.
VN gāi ‘scratch’: UAN *gatal ‘itch’, Jav. gatal, BI garit, Md. ghatal, Pw. gaisal, cf. also BI kais ‘scrape for food’.
VN nghe ‘hear’: UAN *daŋa ‘hear’, BI dengar, Tg. dinig, Pu. marrgai, Rade kna ‘ear’ (< *taliŋa).
VN đa (MVN đa) ‘belly’: UAN u(n)tah ‘vomit’ BI muntah; SAT. m-utaq. Note that, here, VN retains the Austronesian verbal prefix m-.
VN đa (MVN đa) ‘belly’: UAN u(n)tah ‘vomit’ BI muntah; SAT. m-utaq. Note that, here, VN retains the Austronesian verbal prefix m-.
VN đa (MVN đa) ‘belly’: UAN u(n)tah ‘vomit’ BI muntah; SAT. m-utaq. Note that, here, VN retains the Austronesian verbal prefix m-.
VN thwa ‘respect form’: UAN *tuha/tuva ‘old’ Jav. tuwa.

2.3. Flora, fauna

VN cây ‘tree’: UAN *kaju, PMP *kaSju ‘tree, wood’, BI kayu, Tg. kaahoy, Am. kasui.

VN bông ‘flower’: BI bung ‘flower’, Rade mnya. cf. PMunda *ba[gl]a, Mon pkao, Aslian *bəkaw, PAA *baka[l].

VN mạ ‘rice seedling’: PMP *qumaH ‘field’, BI huma ‘field for dry rice’, Bu. humaq, Sat. qumah ‘work in the field’.

VN trảng (MVN lh'eng): *UAN *taluy ‘egg’, Tg. ‘itlog, BI telur, Am. lita'uy.


VN cỏng ‘claw, pincer’: UAN *bayan ‘molar’, Tg. bagan ‘molar’ Jav. wan ‘jaw’, Rade kaay ‘chin, jaw’.

2.4. World, weather


VN đất ‘earth’: UAN *datay ‘flat’, BI rata ‘flat, level’, Tg. lataq ‘carpet’, NgD datah ‘step, rung’.


VN trăng/giảng (MVN blang) ‘moon’: UAN *bulan ‘moon’, BI bulan, Rade mlaan, Jarai blan, Pu. vo'tan, Kv. buu ran.


VN đem ‘night’: UAN *đam đâm ‘keep quiet’, BI di'am ‘quiet’, Pw. dəm dzəm ‘last night before full moon’.


VN đường/dàng ‘road’: PMP *Zalan ‘road, path’, BI jalan, Tg. daan, Bu. daan, Sir. darang, Rade elaan, Haroi calian ‘id’.
2.5 Miscellaneous verbs

VN rải, trải ‘to sow’: BI beras ‘rice’, Rade rahi ‘sow rice in wet field’.
VN trả, giải (MVN blasph) ‘pay back’: UAN *bolah ‘split’, BI belah
‗split, part‘, Pw. velaq ‘split’. Cf. also VN búa ‘split open’ (except ~
tone is expected).

VN. két ‘fasten together’: UAN *dakat ‘to stick, BI dekat ‘near‘,
Tg. dikit ‘joined’. Pw. dɔ̥kat ‘to stick’.

VN tơi ‘arrive’: UAN *ha(n)td ‘deliver, convey’, BI antar
‗introduce‘ Jav., Md. ator, Pw. sādz ‘send’.

VN dổi ‘deceive, lie’: UAN *puta ‘rotate’, BI putar ‘turn, be
dishonest’. cf. Rg. podår ‘spin a top, deceive’.

VN nau ‘cook’: UAN *tunu ‘roast’, Jav. tuna ‘burn, Md. tono(h)
‘roast’, Pw. ma-tunu ‘hot’, Rade m?dau, Roglai pa?dau ‘warm’.

VN mất ‘lose’: UAN *mata ‘deliver, BI mati ‘die, dead’.

VN kiệm ‘seek’: UAN *ki[li]m ‘send’, BI kirim ‘id.’, Am., Bu.,
kilim, Pw. kim ‘search for’.

2.6. Descriptives

VN sai ‘wrong’: PMP *talaq, Tg. salâ ‘mistaken’, Pw. pa-
talaq ‘envy, jealous’.

VN sác ‘sharp’: PMP *hat’aq ‘whet’, Chmr. gwasa’, Tg. hása, Pw.
tataq ‘id’.

VN bê ‘big’: UAN *bɔvat ‘heavy’, BI berat ‘id.’, Mlg. be ‘big, great,
many’.

2.7. Miscellaneous

VN cuşı ‘end, least’: UAN *liku ‘back, behind’, BI ekor ‘tail’, Tg.
likod, NgD ba-rikor, Pw. likudz ‘behind’.

VN sung ‘gun’: UAN laieŋ, Tg. lusoy, Rade suŋ, Rg. iisuk ‘mortar’,
Pz. ludiŋ, SAT. luhun.

VN ván ‘plank, board’: UAN *papan, BI papan ‘board’, cf. Li. pen
‘classifier for people’, Thai *peen ‘plank’.

VN ná ‘bow’: UAN *panah, PMP *pana ‘bow, arrow, shoot’, BI
panah ‘bow and arrow’, Tg. paana, Rade hna, Am. pana’h. Cf. VN
bàn ‘shoot’, and BI panar ‘stunned, dull’, but also BI senapan
‘weapon’.

89
3. Some Phonetic Patterns

The Vietnamese forms cited above in sections 1 and 3, while presenting a number of 'irregularities', do, on the other hand, exhibit a number of likely phonetic associations with Austronesian forms in general, now summarised in 3.1-10.

3.1. VH high tones: AN voiceless initials

VN kêt ‘fasten together’: UAN *dakat ‘to stick’ BI dekat ‘near’, Tg. dikit ‘joined’, Pw. dɔkats ‘to stick’.

VN trúng (MVN thụng) ‘egg’: BI telur, Tg. ilog

VN ta [arrogant] I, chúng, ta ‘we incl.’: PAN *(k)ita, *ta, BI kita ‘we inclusive’.

VN cây ‘tree’: BI kayu ‘tree’.

VN tô, tro ‘show’: Jav. tuduh ‘point’, Tg. turo ‘instruction’.

3.2. VN low tones: AN voiced initials

VN bà ‘grandmother’: UAN *baji ‘wife, woman’, Tg. ba-bai’i, Am. va-vahi, BsE bayi ‘grandmother’ (cf. VN ṣọ ‘wife’).


VN mg ‘rice seedling’: BI huma ‘field for dry rice’, Bu. humaq, SAt. qumah ‘work in the field’.

VN gúc ‘bend head down’: BI angguk ‘nod’.

VA mèra ‘vomit’: UAN *u(n)tah, BI muntah, SAt. m-utaq ‘vomit’.

3.3. VN ? or ~ tone: AN -s-

VN mái ‘continuative’: BI masih ‘still, yet’.


3.4. VN ? or ~ tone: AN -s/h

VN ìa ‘defecate’: UAN *ia(h), Jav. p-ih ‘water conduit’, Bu. ‘isah ‘wine’.

VN rài ‘to sow’: BI heras ‘rice’, Rade rah ‘sow rice in wet field’.

VN lýi ‘tongue’: Jav. dilah, BI lidah, Pw. lidsidi’, Rade lah.
3.5. VN' or . tone: AN final stop

VN đất 'earth': UAN *data 'flat', BI rata 'flat, level', Tg. lataq 'carpet', NgD datah 'step, rung'.
VN sắc 'sharp': PMP *hat'aq 'whet', Chmr. gwas'a, Tg. haasa', Pw. t-ataq 'id'.
VN bùi 'dust': UAN *dabuk 'ashes, dust, grey', BI abu 'dust', Jav. dawu' 'grey'. cf. PMP *abuh, Rade ?bru'ih, Rg. həpəuh 'to dust', Waic. *pes.
VN gúc 'bend head down': BI angguk 'nod' (cf. 3.2.).

3.6. VN level (unmarked) or `tone: AN open syllable/final nasal

VN tai 'ear': BI telinga. cf. CLi. that 'ear'.
VN ta '[arrogant] I', chưng ta 'we incl.': PAN *(k)ita, *ta, BI kita 'we incl.'.
VN ăn 'eat': PMP *ka'an 'eat', Tg. kaa'in, BI makan, SAT. qan-iq 'eat'.
VN mà 'which (rel. pn.)': BI mana 'where, which (interrog. pn.)', cf. also VN mò 'what, where'.
VN đường, dăng 'road': PMP *Zalan 'road, path', BI jalan, Tg. daan, Sir. darang.

3.7. VN ch:-: AN t/pl-

VN chí 'thread': PAN *taliS' 'rope, cord', BI tali, Tg. taali?, Pw. tsalis, Pz. sariss 'cord'.
VN chac 'a collection of ten': UAN/PMP *puluh/puluq Tg. pulo', pu'o', Ru. porok, Pw. puluq 'ten'.

3.8. VN u or o: AN -l/r-

VN llevi 'tongue': UAN *dilah 'tongue', Jav. dilah, BI lidah, Pw. lidali, Rade lah.
VN trưng (MVN trúng): UAN *təluy 'egg', Tg. 'itlog, BI telur, Am. lita'uy 'egg'.
VN dvang, đàng 'road': PMP *Zalan 'road, path', BI jalan, Tg. daan, Sir. darang, Rade elaan.
VN bẹ 'big': UAN *bəvət, BI berat 'heavy', Mlg. be, 'big, great, many'.

91
3.9. **VN -i: AN -l/r**

- VN cuôi 'end, last': UAN *likuḍ 'behind, back', BI ekor, Tg. likod, NgD ba-rikor, Pw. likudz 'behind'.
- VN sai 'wrong': PMP *t'alaq, BI salah, Tg. sala', 'mistaken', Pw. pa-talaq 'envy, jealous'.
- VN tɔi 'arrive': UAN *ha(n)tő 'deliver, convey', BI antar 'introduce', Jav., Md. ațar, Pw. satədž 'send'.
- VN dɔi 'deceive, lie': UAN *putőd 'rotate', BI putar 'turn, be dishonest'.
- VN tai 'ear': PAN *taliLa, BI telinga 'ear'. cf. CLi. thai.

3.10. **VN tr-/gi-/bl-:AN b(v)l-**

- VN trāng/giāng (MVN blāng) 'moon': UAN *bulan 'moon', BI bulan, Rade mlaan, Pu. volan, Kv. buu ran.
- VN trā/giā (blā) 'pay back': UAN *bəlah 'split', BI belah 'split, part', Pw. volaŋ 'split'. Cf. VN bīə 'split open'.
- VN trɔn 'be round': UAN *bəluv. cf. Li. (p)luon, Thai *duan, 'don (Benedict 1966: 246).

4. **Concluding remarks**

This brief consideration of possible lexical affinities between Vietnamese and Austronesian makes no claim to far-reaching conclusions. At the same time, some of the VN forms observed above would, if valid, seem to call for at least two sources or periods of Austronesian (or Austro-Thai?) contact in order to explain their contemporary phonological constitution. These I will, for present purposes, distinguish simply as Immediate vs. Remote sources. The following sets of vocabulary are illustrative:

**Immediate**

- VN dā ‘already’: BI sudah 'already'.
- VN sāp 'about to': BI siap 'ready'.
- VN dang ‘in process’: BI sedang ‘in process’.
- VN lā ‘is’: BI ia lah ‘is’.
- VN thua ‘respect’: BI, Jav. tuwa ‘old’.
- VN cāy ‘tree’: BI kayu ‘tree’.
- VN rāi ‘to sow’: BI heras, Rade rah ‘sow rice in wet field’.

**Remote**

- VN gāi ‘scratch’: UAN *gatsol ‘itch’, BI garit, Pw. gatsol. cf. BI kais ‘scrape for food’.
On Austronesian lexicon in Vietnamese

VN mạ 'rice seedling': PMP *qumaH 'field', BI huma 'field for dry rice', Bu humaq, SAT. qumah 'work in the field'.

VN kiểm 'seek': UAN *kil[l]im 'send', BI kirim 'id', Bu. kilim, PW. kim 'search for'.

VN sắc 'sharp': PMP *hat'aq 'whet', Chmr. qwasa', Tg. hâasa', PW. t-ataq 'id'.

VN kết 'fasten together': UAN *dakat 'to stick', BI dekat 'near', Tg. dikit 'joined', PW. d'okots 'to stick'.

VN đường dăng 'road': PMP *Zalan 'road, path', BI jalan, Tg. daan, Sir. darang, Rade elaan.

Thus, some of the forms cited in this paper may well reflect contact between Austronesians or Austro-Thai and Austroasiatic, rather than with Việt-Mường proper. Even so, it is suggested that, for example, syllable reduction and tonal effects in the latter contribute significant perspective to historical processes even in early linguistic relationships.

REFERENCES


Diffoth, G. 1980. The Wà languages. (= Ling. Tibeto-Burman Area 5(2)).


KENNETH GREGERSON


1. Introduction
The phonological system of Middle (i.e. seventeenth-century) Vietnamese has been treated in Gregerson (1969), Haudricourt (1974), and others. This paper takes a look at the Vietnamese lexicon of the same period, as shown in the trilingual (Vietnamese-Portuguese-Latin) dictionary and in the Catechism authored by the Jesuit scholar-missionary Alexandre de Rhodes (1593-1660). Only 'full words' (thếc-từ) or content words in their archaic forms will be discussed, leaving out function words, known by traditional Sino-Vietnamese grammarians as 'empty words' (hu-từ).

The Dictionarium Annamiticum [sic] Lusitanum et Latinum² (Rhodes 1651a, hereafter Dict.) consists of 900 numbered columns, two to each page, listing Vietnamese entry words in alphabetical order according to the spelling of the time, each headword and illustrative example followed by its Portuguese equivalent, then by its Latin equivalent. Between the 'b' and 'c' sections, there is a section (cols. 65-74) devoted to words whose initial consonant is transcribed with the symbol ꜜ; this lenis obstruent was described by de Rhodes (1651b: 3) as 'almost like Greek β as in ἐπαν “enter”, ḍãi cá “fin [offish]”’. This feature is explained within a 31-page statement on Vietnamese grammar bound in the back of the dictionary, following an Appendix, which contains five unnumbered errata pages, and a 171-page index of Latin words: this Linguae Annamitae seu Tunchinensis Brevis Declarato (hereafter BD) is composed of eight headings.

On the other hand, the Catechismus pro iis qui volunt suscipere Baptismum in octo dies divisus, Phép giảng tám ngày cho kẻ muốn chịu phép rửa tội mà ḍãao đạo thành Đức Chúa Blãt (Rhodes 1651c; hereafter Cat.), constitutes 'the first work appearing in romanized Vietnamese' (tác-phẩm quóc-ngữ đầu tiên), as the work is referred to in the title Giáo-sĩ Đác-lô và tác-phẩm quóc-ngữ đầu tiên, (eds.) Nguyễn Khắc Xuyên and Phạm Đình Khiêm (1961), a book issued on the 300th anniversary of the death of the missionary – 'the evangelical apostle who codified the quóc-ngữ script,'

1. Maspero (1912) distinguished five periods in the history of the Vietnamese language: proto-annamite (before and during Chinese rule), pré-annamite (during independence), annamite ancien (fifteenth century), annamite moyen (seventeenth century), and annamite moderne.

2. For a general description of this trilingual dictionary, see Nguyễn Đình-Hoà 1985b.
and the reverend benefactor of the Vietnamese church and the Vietnamese nation.’

Since the compilation Diet. systematically listed the seventeenth-century lexicon, I have first tried to cull out archaic forms used in the Catechismus, then to check their listings in the trilingual dictionary. I also present other interesting items found in Diet., but not in Cat. In this preliminary gleaning process, of necessity not exhaustive, I shall successively discuss nouns, classifiers, adjectives, adverbs, preverbs, verbs, and postverbs.

2. Nouns

The Diet. lists several ‘taboo words’ denoting sexual organs, e.g. ben, dau, doc, doi, lon, ke (female), boi, cac, dai, lo (male). The entry dech is glossed as ‘semen humanum’: cf. ‘secretion from female genitalia’ and ‘female genitalia’ (Van Tahn 1977, under dac).

(anh)nain toi (Diet. 7) ‘to eat the bitter grass [called nan (Diet. 504)] ‘to show remorse for one’s sin’. Cat. 177h.

bai, con bai di bai (Diet. 19) ‘prostitute, lustful woman’. bym, con bym (53) ‘prostitute’, cf. modern VN di bym. Both Le Van Duc (1970) and Van Tahn (1977) have the compound bym bai ‘deceitful person’ (cf. bai, binh bai (also Diet. 19, but separate entry) ‘deceitful person’).

bau (30) ‘companion, partner’. Cat. 234. Le Van Duc et al. (1970) has both bau ‘you [to wife or girl friend]’ and bau-bau ‘friends’ whereas Van Tahn (1977) has only bau ban ‘same as be ban, friends’.

cai (79) ‘head, commander’ Cat. 71t, u, cf. modern VN ‘foreman’.
cang la (85), same as giang la (277) ‘wicker basket with handle’, cross-listed under la (389). This compound can be found in ‘Thap-gi6i c6-h6n qu6c-ngfr viin’ (‘Ten commandments to lonely souls’), in Thien-nam Du-ha T6p, a collection of poems by Emperor Le Thanh-tong and his co-members in Tao-dan Circle (fifteenth century):

Song viet lien tay; cang la, non anh.
H6m mai hop mat: noi co, vuon lau.

(Cao Hưu Lãng 1983: 12)

cat nh6n (89) ‘incense’. Cat. 165b-c.
cat (90) ‘back’. Cat. 185d. Diet. also gives sap cat cung ai and blai cat (cf. modern VN chung lang dau cat and trai cat).

che (101) ‘pretty large earthenware vessel’. Cat. 181f-g.

cham cu (95) ‘punctuation’. This expression is found in Poem 3, line 4, by Nguyen Trai (1380-1442): Phien sach ngay xu6n ngoi cham cu, although the word cu (136) for ‘sentence’ is cau in modern VN.

dinh lieu (415, under lieu) ‘huge bamboo torch’.

doi to (228) ‘chain, shackle’. Cat. 10f. Also spelled doi to, with toi itself listed separately (822) ‘chain’.

c6m (130) tang (804) ‘stock made of bamboo or wood to confine a

*Unless necessary for clarity, Diet. is dropped before page number to the entry (Ed.).
prisoner’s neck or ankles’. Cat. 10f. Cf. modern VN cúm and trón, respectively, with the second member of the compound occurring in the expression gông dòng trón mang to denote the status of a prisoner restrained by a cangue (gông, Dict. 299) around the neck and by stocks around the ankles.

dào (206), i.e. dòng ‘spear’. Cat. 232q, u.

hè (318) ‘mourning rites lasting three years’. Cf. modern hỏi hè ‘festivals, merry-making occasions’.

khôô (374) ‘period, time’ < Chinese *k’uâ 計, which explains why the variant forms thôô (776) and thôô (784), nowadays spelled thuôí/thuôí, are transcribed in the nôm script by the character 計. Cat. 150i, 176d, 195b, 196s-t, 196b-c.

khu (376) ‘rear end, buttocks’.

là dâ (191 under dâ and 390 under là) ‘stone slab’ Cat. 280o,p,r. This archaic form is found for instance in Poem 21, line 1 by Nguyễn Trãi: Dâu người dì là dâm.3

lâm (396) ‘mud’ Cat. 195n, 196x. Cf. châm lâm tay bùn.


nê in tô nê ‘ancestors’ Cat. 81x. Cf. tô tôû (817) and both compounds appearing as subentries under tôû, i.e. tôôí (827).

pheo (598) ‘bamboo’: the second member of the synonym compound tre pheo retains this meaning in Mường.

răc (632) ‘prison, jail’. Cat. tă răc 10z, 10e. Dict. also gives the phrase răc di nguc ‘hell’.

táp (725) ‘tempest, hurricane’, nowadays occurs only in such compounds as băo táp.

tang (721) ‘mourning’ provides cultural meaning of the expression đê tang đê tôc [tăoc] (the bereaved man letting his hair grow on top and in front and the bereaved woman cutting part of her hair), which confirms that tang tôc is another synonym compound and not a reduplication.

thê (754) ‘world’. Cat. 5b, 6i: thê nay ‘this world’.

thôi (774) ‘short moment’. Cat. 241m.

thôâ (773) ‘brothel; prostitute’, with con thôâ meaning ‘prostitute, harlot’ (cf. the adjective đê thôâ in modern VN).

tôí (822) ‘servant, slave’ with the compound tôí tâ ‘servants’ Cat. 6t,x,y (cf. modern VN tôí tôí).

3. Determiners, Classifiers and Demonstratives

Two items are used to denote ‘all, the whole’: că and thay thay. The former occurs with và: că và nhâ ‘the whole house’, că và thien-hâ ‘the whole universe’ (77), cho că và loài người ta ‘for our whole mankind’ (Cat. 88f), că và nước ‘the whole country’ (Cat. 21c). Dict. also lists că hoâ as the equivalent of că và under the entry hoâ (329). The latter word can be found in Poems 80, 90, 93 and 247 by Nguyễn Trãi, and it is tempting to posit the following development: hoâ/huâ/ > uâ/wâ/ > và/vâ/. One can
also note **và hai** ‘both’ and the use of **và** as correlative conjunction in **và bánh và cá** ‘both bread and fish’ (Cat. 184n). The determiner **thay thay** is found with **hết** (734) although under **hết** (322) the spelling **hêt thay thài** is presented.

In addition to pluralisers **các** (78), **chúng** (121-2), **những** (559), Dict. lists **mỏ** (475), as in **mỏ tôi** (475),\(^4\) also discussed in **BD** 12, and **phô** (601), as in **phô ông phô bà** ‘ladies and gentlemen’, **phô thây** ‘masters’ and **những phô ông** ‘gentlemen’ (BD 12, 19). This last example is also listed under **những** that is glossed ‘only’, so **những phô ông** means ‘only gentlemen’ with **chảng những** meaning ‘not only’ whereas **những**, means ‘all’, and also ‘nothing but’ as in **những mong** ‘only hoping’, **dây những mủi** ‘full of mosquitoes’, **những mủi là mủi** ‘nothing but mosquitoes’ in modern VN.

Concerning **những**, the sense ‘as many as’ (e.g. **những tân đa con** ‘as many as eight kids’) seems to have been non-existent in middle Vietnamese, and a dictionary published in the late nineteenth century, (Huinh-Tĩnh Cua 1896: 143) does not list this sense, either, under the entry **những**.

On the other hand, for the second-person plural pronoun, the two arrogant expressions **m hoá** and **m may** are noted together with **chúng may**, **chúng bay** and **bay** ‘you guys’ (BD 12).

The pluraliser **phô** (Cat. 197 and passim) had been used in the fifteenth century by Nguyễn Trãi, transcribed by the nôm character 鳄 (Nguyễn Đình-Hoa 1985a: 471); it was still listed in Theurel (1877: ix, 363), and in Huinh-Tĩnh Cua (1896: 200) with the character 鳄. Under the rubric ‘classifiers’, it is worth mentioning that **cái**, nowadays used for inanimate, non-living things, appears in the Dict. with such nouns as **cóc** ‘toad’, **dẹn** ‘spider’, **ếch** ‘frog’, **kiến** ‘ant’ (128, 167, 249, 380, respectively), etc.\(^5\)

Beside the demonstratives **này**, **no**, **áy**, **kia**, **kia**, the Dict. also has **tē** as a synonym of **kia** ‘that... over there’: **ông tē** (728; BD 20). The latter is still used in Central Vietnamese.

**nào** (507) is shown as occurring following the noun it modifies: **sao nào?**, **thế nào?**, **cách nào?**, **dưới/nhườm nào?** (BD 21) to denote ‘which...?’ and as occurring in such idiomatic expressions as **nào có ai biết?** ‘who knows?’, **nào có tôi gì?** ‘which sin? which offence? which crime?’ with the connotation of negativity (508; BD 22). But there is also another usage, in which **nào** precedes a noun: **nào sách?** ‘where’s the book?’, **nào thây?** ‘where’s the teacher?’ (508). This word order can be found in some of Nguyễn Trãi’s poems in the fifteenth century: **nào noi** ‘which place?’: (poem 47), **nào hoa** ‘which flower(s)?’ (poems 59 and 224), **nào cua** ‘which thing?’, **nào thô** ‘which time?’ (poem 164).

\(^4\) Theurel (1877: 279) has the core meaning of **mô** as ‘10,000’ with its derived meaning as ‘parum, quoddam numerale’ to designate ‘a bunch’ [of vegetables **mô rau**] or ‘a handful’ [of rice **mô gạo**]. Huinh-Tĩnh Cua (1896: 41) gives the nôm character 鳄 for **mô**, and provides two illustrations **mô tôi** and **mô qua** for ‘we, us (exclusive)’.

\(^5\) According to Theurel (1877: pp. vii-viii) **rắn** ‘snake’, **kiến** ‘ant’, and **sâu** ‘insect, bug’ could take either **con** or **cái** as classifier.
4. Adjectives

There is a striking number of synonym compounds among the adjectives, in which one member somehow has lost its original meaning so that in modern VN the compound sounds like a reduplication. đót, for example, meant ‘ignorant, feeble-minded’ (178), occurring in the compound cái-dót ‘imprudent, unwise’ (155). hối (334) was a synonym of hêt (321), and the compound meaning ‘narrow [of mind]’ is not an alliterative reduplication. However, the adjective chịu bôi ‘to indulge in promiscuity’ is merely a rhyming reduplication (52), in which bôi is not the homophonous entry bôi (also 52) meaning ‘a lot’ which Lê Trung Hoà (1982: 30) interprets as a modifier of the verb chịu ‘to play (around)’.

kham (360) ‘sufficient’ with the example chăng kham ‘not enough’ (cf. no below).

khô (374) ‘difficult’ [and also ‘wise’]: this word was used by Nguyễn Trãi in poems 1, 6, 7 and 65 (Nguyễn Đình-Hoà 1985a: 468), and also discussed in Theurel (1877: p. xxi) as a synonym of khổ.

không (spelled khoû, Ddict. 375) is glossed ‘vacuus, -a, -um’ as in examples an com không ‘to eat plain rice [without meat or vegetables]’, làm (việc) không ‘to work without pay’, nhà không ‘empty house’, and, in the following entry, không góa ‘widow’: cf. những xã không (Cat. 15x), những không doi (Cat. 237p.), next to the compound hû-không (Cat. 54h) which is glossed ‘vacuum’, i.e. ‘nothing, nothingness, void, vacuum’ in such collocations as hoa hû-không, lai ra hu-không (Ddict. 341). (On the meaning of không as negative particle, see the section on verbs below).

lã dào (400) or lão nhâo ‘disordered, chaotic’ (Cat. 19y, 87u, etc.)

lốp (425) is given as a synonym of kip (384) ‘opportuné’ (để cho kip): with such illustrations as đi (cho) lốp ‘to go in time’, chăng lốp ‘not in time’. Cat. does not use lốp (it has only kip), and Theurel (1877), which lists a number of illustrative examples, ~ thi, ~ sinh, ~ vi ~, ~ chán, ~ vua ~, ~ chua ~, dã ~, ~ an nàn ~, ~ trán ~, ~ for kip (218), does not have the entry lốp.

mìa (487) ‘more than sufficient, superabundant’ (synonymous with dư (179)) nowadays occurs only in the compound thúa mìa (487) and in the expression bô mìa ‘to leave one’s food unused’: cf. Theurel (1877; add. 44) bô mìa, mìa mìa ‘leave work unfinished’.

nát (509) has the sense ‘rotten, spoiled, decadent’ used in the moral sense in Cat. 70 (also hu nát).

nghi (526) ‘easy’ (chăng nghi chịu lợi nó cảm dô, nghi đối dàn bô được in Cat. 83 n, r; nghi chịu được tha vây in Cat. 90o); càng nghi làm việc ấy in Cat. 118y; nghi làm lên in Cat. 251o) has been found in Chi-nam Ngọc-âm Giải-nghi, a dictionary of Chinese and Nôm characters, compiled some time after the sixteenth century (Trần Xuân Ngọc-Lan 1982; Lê Văn Quân 1981). In the preface, its author, a monk by the name of Pháp-tính, said:

Bây giờ Nôm dây chỉ dòng,
Cho người mồi hoc nghi xem, nghi nhuần.

‘The nôm script is taught in here as single characters,
‘So as to make it easy for beginners to read and to learn them’.
như nhởm (556) ‘deceitful, false’, synonymous with đời trava (dội blâ), as in ông Adam chẳng có phải lời như nhởm đâu (Cat. 88-I).

no (562) ‘full; sufficient’ has such run-on entries as no moi noi, no moi sù, chẳng no, dâ no mût, and also no đó, no khâm [see khâm above]. Cf. đến khi ở thượng dâ no (Cat. 81x), ta phải kính ông moi noi (Cat. 51s), the latter structure no moi... being the equivalent of English ‘every..., all...’ with the meaning of this lexeme restricted in modern VN to the notion of ‘fullness (from eating), satiety’.

nót (571-2) ‘premature’, synonymous with sinh non, also has yếu nót ‘feeble in body and mind’; the form non nót ‘tender, immature’ in modern VN, made up as a synonym compound, is often considered an alliterative reduplication.

(rần may) rần mặt under rần (636) ‘obstinate, not amenable to control’ (ở rần mặt cùng nó, in Cat. 87r: cf. Theuret 1877: 383). The equivalent in modern VN is rần đâu or cõng cõ, and involves the head or neck, not the face.

tày (716) ‘equal’ is cited as appearing freely (like its peer bằng) followed by a complement: tày người nén hai mươi tuổi, ‘equal in size to a 20-year-old person’. Cf. modern VN idiomatic expressions tày đỉnh and tày trì as modifiers of tội ‘offence’, or việc ‘offence, fine, calamity’.

vậy (857) ‘slanting, crooked’ used with the noun đạo ‘way, path, religion’ just as in đạo tà ‘false or erroneous precept’ (712); Cat. 29m, 30p, where đạo vây refers to Confucianism, Taoism, Buddhism and other local beliefs of the time.

5. Adverbs

Only a few adverbs deserve mention:

gìa giét (271) is a modifier like lâm, thay or ráp ‘very’: BD 14 gives tiết lâm, lâm thay, xấu ráp, and mrowth, i.e. lớn, gia giét as examples, where such an intensifier follows the modifier. Cf. lớn gia giét hồn cá (Cat. 62y), đập gia giét dâu may (Cat. 92e), lo gia giét làm vậy cho cha mẹ (Cat. 120n).

Lê Văn Đức (1970: 537) also gives the meaning ‘severely, sternly’ illustrated by the expression phê-bình gia giét.

mốt (482) is listed and glossed only as a numeral. However, its use in the sense of ‘only’ is found in more than one source: môt lo sự thể gian mà thôi (Cat. 95l), chàng thây lôi, môt thây hai mà thôi ‘non lucrum sed tantum detrimentum video’ (Theuret 1877: 284, under môt); as well as môt nói sự mình mà thôi (in Sách số sang chẽp các việc, a diary by an eighteenth-century priest, Philippe Binh (1968: 152)).

rắp (641) ‘very’ as in tốt ráp ‘very good’ (and in xấu ráp, BD 14). Cf. năng ráp (Cat. 221s), đất ráp (Cat. 217m).

sê sê (684), synonym of dân dân (268), as in tó ra sê sê (256p), sê sê cõ chrious (268x).

vồi, often considered an adverb modifying xa ‘far’ (872-3) is glossed as a substantive meaning ‘high seas’ so that xa vồi really means ‘far off the coast, on the high seas’.
6. Proverbs

chàng is used more often than không (375) as the negative preverb, whereas đòng is synonymous with cùng ‘all, together’ as in đòng ở một ngăn, đòng nhất thể (236).

khả ‘can; appropriately, fittingly’ (359) < khá ~ (Cat. 11y-z has Dây bèn khá bước sang trái / Lê thất khả học được lòng người ta, which is a translation of the Sino-Vietnamese couplet Kiến-thẳng khả kề ngầu-giác / Lấy-ngử nâng phục nhân-tâm).

khừng (377) ‘agree to, consent to, accept to’; chăng không wàng phép cha mẹ (Cat. 181).

mưa (487) is a prohibitive form ‘don’t, should not’ found in fifteenth-century utterances (Nguyễn Đình-Hòa 1985a: 469); the entry mưa nói đồi ‘don’t lie, should not tell lies’ is repeated in BD 24 as mưa hề nói đồi ‘should never tell lies’.

rính (649) ‘be about to, approach [a condition]’ with three examples rính chết, rính ngã, rính dè, the first of which is found in Cat. 67, 69 with the meaning ‘dying, about to die’, also cited in Theuret (1877: 390) ‘morti vicinus’. Cat. 98b also has rính lút about the Deluge.

7. Verbs

The subgroup comprising bì, chịu, duốc and phải deserves priority treatment simply because these suggest the passive voice of Western languages.

The first one, bì (34), not to be confused with the preceding entry, which means ‘bag’ (‘mantica’), is given only one illustration: bì phũ [phong] ba ‘run into a storm’ although in modern Vietnamese (Clark 1974, Nguyễn Đình-Hòa 1972) its [-pleasant] or [+ adversative] feature tempts many an analyst to translate it as ‘be... , get... ’ in English.

The verb chịu, with the meaning ‘undergo, suffer, experience’ must have been used more frequently than bì. Dict. (109) lists chịu tội, chịu chét, chịu lội, chịu lụp, chịu nờ, chịu miếng, chịu thai.

Phải, on the other hand, which denotes involvement in an unpleasant or unfortunate situation, is illustrated by no less than twelve examples, in which the ‘predicament’ can be as serious as thunder and lightning (phải sâm sét), paralysis (phải liệt), or as inadvertent as eating meat by mistake (mắm phải thịt), or as committal as falling in love (phái lòng). (590)

The opposite of these three terms, the entry duốc (243) is often glossed as ‘to gain, get, acquire, earn, find, obtain’ with the feature [+ pleasant]. At any rate, all four lexemes fully qualify as regular verbs.

bố vạ (49) ‘to slander, calumniate’; Cat. 65, 190.

chắc (93) ‘to purchase’, found earlier in Nguyễn Trãi’s poems.

dài (154) ‘to fear’; Cat. 59, 91, 221x. Cf. modern VN expression khốn cho người ta dài, dài cho người ta thương.

dâu (160) ‘to love’ with the examples yêu dấu con and thuộc dấu, the latter still found in bò yêu thuộc dấu ‘love charm, philtre, love potion’.

dê (161) ‘to despise’ as in dê người ‘to be haughty’.
dá (191) ‘to touch’, with co-verb dến or phải preceding the object: téy
dá dën cái sâng ‘His hand touching the coffin’ (Cat. 187), Tây tói dá dën
duộc chân áo dúc Chúa iesu ‘As my hand touched the hem of Jesus’ robe
(Cat. 185f).

dá (191) ‘to heal’: some examples in Cat. are ké dâu nunng thì cho dá
‘curing persons who were seriously ill’ (179r), song châng dá cho ‘but they
could not cure her’ (185e), thì tói dá ‘then I was cured’ (185g), chúa dá tâ
ta ‘to cure our ailments’ (186m). Cf. modern saying Thuóc dâng dá tâta /Sự
thất mát lòng.

dem [dàng] (195) ‘to guide, lead, show the way’ as in [cân sao] sô cho mà
dem dâng ‘[the star] guiding them and showing them the way’ (Cat. 164u),
[ông thành Joong Baptistä] khi dem dâng cho dúc Chúa iesu ‘showing the
way to Jesus’ (Cat. 177k).
de (211) ‘to threaten’: Cat. 86h has de chét, and Theurel (1877: 114) lists
three synonymous compounds de nét, de loi, de phat.
dài (228) ‘to imitate, copy’ with examples theo dài, lâm dài listed as
synonyms of bôt chuíc: Theurel (1877: 125) has both theo dài and học
dài.
dóm (231) ‘to add, augment’, about which Dict. has next to them dóm
the expression nói dóm đât ‘to exaggerate’; cf. thiệu một chúa, vì vậy
thì phải thêm dóm chúa ấy ‘since the word chúa is missing, we have to add
that word on’ (Cat. 16z).
giá (272) ‘to show gratitude to, thank’, as in giâ ôn; cf. có giâ nhâ hay là
giá chúa nhâ (Cat. 14c), vì hàng có ai giâ nhâ mà chông giâ chúa nhâ (Cat.
14e).
gúom (185) ‘to wait, hold it’; Gúom [spelled duóng] dâ! (Cat. 306p).
le (407) ‘to be afraid, fear’, a lexeme that occurred freely in the fifteenth
century (Nguyen Đình-Hoa 1985a: 468) but later was – and is in modern
parlance – used only in compounds: for instance e lê, sô lê recorded in
Theurel (1877: 233).
mâng (452) ‘to hear, perceive through the sense of hearing’, therefore
the equivalent of nghe thây, as opposed to nghe ‘to listen’. Theurel (1877:
268) lists mâng tin, mâng tai, mâng nghe, mâng tiêng, and a recent article
by Nguyen Bat-Tuy (1976) points out, using cognates in several Mon-
Khme languages, that the common reading of the nôm character 「」
as măng is wrong.
nên (513) ‘to become [so many years old]’ as in nên muôí tuôí, or nên
môt, nên hai (Theurel 1877: 297) – colloctions in which modern VN
would substitute lêm ‘to go up to, reach’, Cf. ông Noe nên sau trâm tuôí
(Cat. 99p), bà Sara dâ nên chín muôí tuôí (Cat. 125a).
phen lê (405) ‘to envy, begrudge’; cf. phen bi, phen lê in Theurel (1877:
360), or phen bi with the meaning ‘to compare’ in modern VN.
tây (716) ‘to be partial to, favour’, for which the example in Dict. (dúc
Chúa blôí) cháng tây ai is also cited in Cat. 56k: phân xet chãng tây ai ‘in
his judgement, was not partial to anyone’.
8. Postverbs

doan (226) 'to complete, finish'; phạm tôi doan (Cat. 89u), khi đề con doan (Cat. 161e), khi đề doan (Cat. 162s, t, u) 'after She gave birth [to Jesus]. The example thời doan 'after stopping' cited in Dict. shows that the usage has not changed, since modern VN has nói doan 'so saying...' comparable to đã chép the doan in BD 28.

liên (413), modern VN liên, as in học liên 'to study continuously'; cf. giữ liên (Cat. 40g, h), which could be likened to luôn.

rồi (653) 'to finish, complete' > 'already': the early meaning persists in Saigonese (Rồi chúa? – Rồi. ‘Did you finish yet? – I did’), the equivalent in the Hanoi dialect being Làm xong chúa? – Làm xong rồi. Dict. gives the basic meaning of ‘leisure’ ['otium'], encountered in vô công rồi nghỉ (Theurel 1877: 392) or in the modern expression ăn không ngồi rồi ‘to sit idle, live in idleness’.

REFERENCES


103
NGUYỄN ĐỊNH-HOÀ


Nguyễn Khắc Xuyên 1960. Lạng-cô-khảo về cuốn tư-dĩển Việt-Bồ-La (1651). Bách-khoa 77 (3-8); 78 (3-7).


1651b. Linguae annamiticae seu tunchinensis brevis declaratio. – ope Sacrae Congregationis de Propaganda Fide. Roma. [Appendix to the Dictionarium.]


THE PHONOLOGY OF KOMPONG THOM CHAM

Robert K. Headley

0. Introduction
Cham is a member of the Chamic Group of Austronesian. Other Chamic languages include Chru, Haroi, Jarai, Radé, and Roglai. It is classified into two major dialects: Eastern Cham (spoken by about 30,000 people in the Socialist Republic of Vietnam in the vicinity of Phan Rang) and Western Cham (spoken by about 150,000 people in Vietnam in the vicinity of Châu Đốc and in Tây Ninh Province, and in Cambodia along the Mekong River and its tributaries). This study describes the synchronic and diachronic phonology of the Western Cham dialect spoken in Kompong Thom Province, Cambodia, 1 (KTC). This dialect differs somewhat from the Western Cham dialect of Châu Đốc (CHD) described by Friberg and Hor (1977); it differs substantially from the Eastern Cham of Vietnam (CVN).

The Western Chams apparently represent the descendants of immigrants who left Vietnam after the collapse of the Cham Kingdom in the sixteenth century. The majority of Western Chams are Muslims and use an adoption of the Arabic script—via Malaysia—to write Cham. Recently, a Latin script has been devised to write Eastern and Western Cham.

1. The Word
The native Cham word—which may differ significantly in pattern from borrowed words—has the following syllabic patterns:

MONOSYLLABLES: (C₃V₂)C₁(C₂)V₁F

<table>
<thead>
<tr>
<th>CV</th>
<th>/la/</th>
<th>['la:]</th>
<th>‘leaf’</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVF</td>
<td>/saw/</td>
<td>['saw]</td>
<td>‘dog’</td>
</tr>
<tr>
<td>CVCV</td>
<td>/taho/</td>
<td>['tho:]</td>
<td>‘large jar’</td>
</tr>
<tr>
<td></td>
<td>/pato/</td>
<td>[pə'to:]</td>
<td>‘teach’</td>
</tr>
<tr>
<td>CVCVF</td>
<td>/kanɔŋ/</td>
<td>[kɔ'nɔŋ]</td>
<td>‘be angry’</td>
</tr>
</tbody>
</table>

---

1. This study is based on a corpus of data obtained from a male speaker in his late twenties from Phum Baray in southern Kompong Thom province. The Proto-Chamic reconstructions are mainly from Lee (1965) with some minor modifications, such as *ca'buay ‘mouth’ for Lee’s *ca'buai, and a few new reconstructions by the author. Additional Western Cham data are from Friberg et al. (1977), and Eastern Cham data are from Blood (1967) and Moussay (1971). Forms between slants are phonemic while forms in italic are generally in the popular orthography of the language.
POLYSYLLABLES: “C₄V₂C₃V₂F₂'C₁V₁F₁

CVCVCV /marasa/ ['mayəsə:] ‘maybe’
CVCVCVF /panatən/ ['paŋətən] ‘animal’
CVCVFCVF /patanrawu?/ ['paŋən'rawu?] ‘make heavy’

In the patterns above: C = any consonant, V = a vocalic nucleus, ’ marks primary stress of the following syllable, ‘ marks secondary stress of a following syllable, and F = any final consonant. Borrowed words from Khmer, Vietnamese, Arabic, Malay, Sanskrit, or European sources may show varied syllabic patterns. In some cases, there has been a reshaping of a foreign word into a Cham syllabic pattern. More recent borrowings are less likely to be reshaped than older ones.

/šərga/ [šərga] ‘heaven’ (Skt. svarga)
/fim/ [fi:m] ‘film’ (ultimately from French or English ‘film’)

2. Register

Western Cham is a two-register language. The effects of register are seen most obviously on the vocalic nucleus. The High Register (HR) vowels are higher (in terms of tongue height), rather ‘breathy’ in voice quality, and associated with low pitch. Low Register (LR) vowels are lower in tongue height, often with lower on-glides, rather ‘clear’ in voice quality, and associated with higher pitch. Friberg and Hor (1977: 18-19), following Gregerson (1976), suggest that the physiological basis for register is the advancement or retraction of the tongue root. They describe First Register (with the tongue root retracted, called Low Register in this paper) vowels as generally lower and tense and Second Register (with tongue root advanced, called High Register in this paper) vowels as generally higher and lax with associated lower pitch and slightly breathy quality.

The High Register developed in syllables which originally, in an earlier stage of Cham,² began with voiced consonants. In the dialect of Western Cham described by Friberg and Hor (1977), the High versus Low Register opposition was noted only following stops. Low register vowel nuclei followed originally voiceless stops and high register vowel nuclei followed originally voiced stops. The remaining consonants seemed to be associated with the high register. In the Kompong Thom dialect, original clusters of /h/+ /w, l, y, r, m, n, η/³ have lost the /h/⁴ and have vowel nuclei associated with the low register. In some cases, the original Proto-Chamic (PC) cluster had initial /s/ which must have become /h/ before the Cham Empire broke up, e.g.:

*humə ‘field’ > /mi/ [mə] cf. CHD hamu and CVN hmu or hamu
*hanə ‘asthma’ > /ni?/ [nə] (cf. CHD hanuc and CVN hanuc)

2. Since the traditional Cham script distinguishes between voiced and voiceless stops and it was adopted prior to the eighth century A.D., it is believed that the voiced-voiceless distinction was present in Cham at that time.
3. /h/ + /b/ has not been found.
4. In careful speech, the /h/ may be pronounced, but it has been lost in the everyday spoken language.
The phonology of Kompong Thom Cham

*sana ‘roast’ > /na/ [na:] (cf. CHD hana and CVN hna)
*hure? ‘vine’ > /reʔ/ [yeʔ] (cf. CHD hrēk and CVN harē')
*surʔ ‘writing’ > /rāʔ/ [yaʔ] (cf. CHD hrak and CVN hraʔ or sraʔ)
*sula ‘leaf’ > /la/ [la:] (cf. CHD hla and CVN hala)
*haway or huway ‘rattan’ > /wāʔ/ [waj] (cf. CHD haway or CVN hawēy)

The contrast between the reflexes of *w, *l, *y, *r, *m, *ŋ and *h(V) + *w, *l, *y, *r, *m, *ŋ, *ŋ can be seen in pairs like the following.

/ni/ [naiː:] ‘bee’ < *hunī
/ni/ [niː:] ‘this’ < *ʔuni

/miʔ/ [miːt] ‘hear’ < *hmāt
/miʔ/ [miːt] ‘always’ < *(mit)†

/wāʔ/ [yaʔ] ‘eat’ < *hua?
/wāʔ/ [ŋaʔ] ‘pull down’ < *(wāʔ)

/yāw/ [jaʔ] ‘like’ < *(hyəʔ)
/yāw/ [jaʔ] ‘yoke’ < *yəʔ

/rāʔ/ [yaʔ] ‘writing’ < *surā?
/rāʔ/ [yāʔ] ‘vein’ < *ʔurāt

3. Consonants

The following consonants occur in KTC:

<table>
<thead>
<tr>
<th>BILABIAL</th>
<th>DENTAL</th>
<th>PALATAL</th>
<th>VELAR</th>
<th>GLOTTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>STOPS</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>NASALS</td>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td>ŋ</td>
</tr>
<tr>
<td>SONANTS</td>
<td>w</td>
<td>l</td>
<td>y</td>
<td>r</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

/ʃ, f, ʂ, z/ also occur in borrowed words.

The consonant phonemes of KTC are described in detail below. /p t c k/ are voiceless, unaspirated bilabial, dental, palatal, and velar stops respectively. They are unreleased in final position. When followed by /l/, they are lenis.

/cip/ ‘Thursday’ [cip] /tāʔ/ ‘chop trees’ [taʔ]  
/taraʔ/ ‘sow rice’ [tājʔ] /tahāw/ ‘know’ [thâj]  
/matāh/ ‘raw’ [mɔtah] /ŋāt/ ‘be careful’ [ŋat]  
/cim/ ‘bird’ [ĉim] /cakāw/ ‘cut with scissors’ [ĉaʔkaŋ]  
/tacūʔ/ ‘spit’ [tɔchuʔ] /kēʔ/ ‘to bite’ [kē]  
/kapal/ ‘thick’ [kɔpɔ:l] /take/ ‘horn, antler’ [tɔ'ke:]  

†Asterisked entries in parentheses ( ) are tentative reconstructions (Ed.).
Modern KTC initial /p t c k/ are reflexes of Proto-Chamic *p, *t, *c, *k and *b, *d, *j, *g. The four voiced PC stops conditioned following high register vowel nuclei.

*patːiːw ‘stone’ > /pataw/
*tubaː ‘sugarcane’ > /tapːaw/
*pagaː ‘tomorrow’ > /pakːə/
*ʔadaː ‘younger sibling’ > /tːiːy/
*cimaː ‘bird’ > /cim/
*jukaː ‘black’ > /kuʔ?

tupaː ‘straight’ > /tapːa?
*duːrə ‘thorn’ > /truːay/
bala ‘tusk’ > /plːə/
tiːn ‘abdomen’ > /tean/
keʔ ‘bite’ > /kɛʔ/
groh ‘to bark’ > /krəʔ/

dleː ‘tired’ > /klɛːh/

The clusters *dl- and *tl become /kl-/

Voiceless stops are retained in words borrowed from Khmer.

Khm. /phaːŋ/ ‘inch’ > /phəŋ/
Khm. /ciːl/ ‘gouge’ > /cakial/
Khm. /kmaːc/ ‘ghost’ > /kamot/
Khm. /tnaːt/ ‘sugar palm’ > /tanot/
Khm. /tːiː/ ‘only’ > /tːiː/.

Some words which contained voiced stops in Middle Khmer were apparently borrowed before these stops were devoiced in Khmer and then subsequently devoiced in KTC. The vowel nuclei in KTC that follow originally voiced Khmer stops are those associated with the high register.

Khm. /peːl/ (beʔ) ‘time’ > /pəl/
Khm. /koːm/ (gəm) ‘lamp’ > /kəm/
Khm. /tːap/ (daʔpa) ‘army’ > /tːap/
Khm. /tup/ (daʔpa) ‘stop’ > /tːap/

There are several problems with the reflexes of the final PC stops. Generally, KTC has /ʔ/ as the reflex for *-p *-t *-c *-k and *-ʔ. Final *-p
The phonology of Kompong Thom Cham

becomes /-u?/; final *-t becomes /-ʔ/; final *-c becomes /-ʔ/ or /-ʔ/ after front vowels, and final *-k and *-ʔ become /-ʔ/. Examples of the normal reflexes are given below.

*găp ‘each other’ > /kâuʔ/  *lăp ‘fold’ > /lăuʔ/
*ʔasăp ‘smoke’ > /săuʔ/  *hadip ‘alive’ > /tiuʔ/
*chiăp ‘wing’ > /ceauʔ/  *ʔdiăp ‘sticky rice’ > /diauʔ/

There are numerous examples of final /-p/ in all phonetic environments in KTC. Many of these words are clearly borrowings, but others are not identifiable as such.

/kũmsup/ ‘dawn’ cf. CHD masup and N. Roglai mōsup.
/pakăp/ ‘compare’ cf. Chru pogăp.
/tănăp/ ‘low’ cf. Old Khm. dănŭp ‘lowland’ < dăpa ‘low’
/ăp/ ‘accompany’ cf. Khm. /aep/ (‘ēpa) ‘take shelter’
/răp/ ‘body’ cf. Skt. rūpa.

*haăt ‘what’ > /kēʔ/  *tuăt ‘knee’ > /taʔūʔ/
‘jhit ‘sew’ > /chiʔ/  *lańt ‘sky’ > /lańjiʔ/

Lee (1965) has two reconstructions of PC words with final *-t that have kept /-t/ in Cham.

*capăt ‘squeeze’ > /capēt/  *kawăt ‘wire’ > /kawēt/

There are other words with final /-t/ for which reconstructions are not available, but which have cognates in the other Cham dialects.

/klūt/ ‘stuck, jammed’ cf. CHD glūt and CVN klūt (< glut)
/cahēt/ ‘jump rope’ cf. CHD chhēt
/thīt/ ‘cured, recovered’ cf. CHD and CVN dhīt
/mit/ ‘hear’ cf. CHD and CVN hamūt
/nāt/ ‘be careful’ cf. CHD ngāt
/padăt/ ‘compare’ cf. CHD padăt
/pawēt/ ‘cut in half’ cf. CHD pawēt
/sīt/ ‘small’ cf. CHD sīt

5. The Chamic languages are in the process of losing final /-p/. This process has almost reached completion in Cham, Chru, Roglai, and Haroi. Compare the following chart:

<table>
<thead>
<tr>
<th></th>
<th>fold</th>
<th>smoke</th>
<th>tie</th>
<th>hide</th>
<th>full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cham</td>
<td>lău?</td>
<td>său?</td>
<td>cău?</td>
<td>tă?</td>
<td>pău?</td>
</tr>
<tr>
<td>Chru</td>
<td>lōu?</td>
<td>asău?</td>
<td>că?</td>
<td>(pŏ)doă?</td>
<td>—</td>
</tr>
<tr>
<td>South Roglai</td>
<td>lou’</td>
<td>sao</td>
<td>că’</td>
<td>dou</td>
<td>—</td>
</tr>
<tr>
<td>North Roglai</td>
<td>—</td>
<td>asă?</td>
<td>ica’</td>
<td>da?</td>
<td>bă’</td>
</tr>
<tr>
<td>Haroi</td>
<td>—</td>
<td>asău?</td>
<td>cău?</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Jārai</td>
<td>lăp</td>
<td>asăp</td>
<td>că?</td>
<td>dăp</td>
<td>bō</td>
</tr>
<tr>
<td>Rade</td>
<td>lăp</td>
<td>săp</td>
<td>căp</td>
<td>dăp</td>
<td>bō</td>
</tr>
</tbody>
</table>

109
There are also a fair number of words with final /-t/ in KTC for which no cognates have been found: /pāṭ/ 'of', /prēṭ/ 'run fast', /cawēt/ 'erratically', /cūt/ 'small bamboo', /mīt/ 'run', /padīt/ 'make up a story', /padēt/ 'make a fingerprint', /prīt/ 'to set limits', /srēt/ 'collapse', /wēt/ 'shake the head', /yīt/ 'keep on doing'. Of interest is the fact that, in all of these words, /-t/ is preceded by a short vowel and in most of them the vowel is a front vowel. Borrowed words with final /-t/ are numerous: /ahāt/ 'Sunday' (Malay ahat, ahad), /pēt/ 'hospital' (Khm. pseṭ), /bōt/ 'turn' (Khm. bōt, /caṛīt/ 'cricket' (Khm. caŋrōt), /makēt/ 'female' (Malay mēgat?). There may be some modifications of the vowel: /pēt/ 'surround' (Khm. pōat/ba' ta), /tīt/ 'kick with the toe' (Khm. toat/da' ta). Final /-c/ in Khmer loans becomes /-t/ in KTC: /khut/ 'broken' (Khm. khooc), /kamot/ 'ghost' (Khm. kmaoc).

*būc 'pull up' > /pūi?/  *duac 'run' > /duai?/
*pruec 'intestine' > /proi?/  *sāc 'shake out' > /sai?/
*hūe 'whistle' > /huī?/  *huac 'afraid' > /huai?/
*?amrec 'pepper' > /mrē?/
*brūk 'rotten' > /prū?/  *pirāk 'silver' > /prea?/
*bruā? 'work' > /prū?/  *maṇāk 'oil' > /maṇī?/
*katūk 'flatus ventrus' > /katē?/  *pītū' 'star' > /patū?/
*pā? 'four' > /pa/?  *ṇō? 'upgrade' > /ṇē?/
*hua? 'eat rice' > /hoa?/  *tasī? 'sea' > /tasi?/

Usually, following a nasalised vowel in PC and in borrowed words, /-k/ and /-ʔ/ become /-k/:

*ńūʔ 'dive' > /ńūk/  *tulōk 'disk shape' > /talōk/
*maṣāk 'brave' > /maṣāk/

but, in some cases, no evidence of a nasalised vowel has been found:

*kutok 'grasshopper' > /katok/  *jāk 'clever' > /cāk/.

KTC also retains final stops intact in borrowings from Khmer.

Khm. /sroop/ 'absorb' > /srop/
Khm. /aep/ 'move' > /ep/
Khm. /book/ 'hump' > /bok/
Khm. /lout/ 'jump' > /lōt/

/b d j/ are voiced, fortis, optionally preglottalised, bilabial, dental, and palatal stops. The preglottalisation is especially noticeable in /j/ which appears to vary between [ʔj] and [ʔi]. These stops do not occur in final position.

6. There may have been metathesis here: *pirāk (I am unsure of the length here, it may have been *pirāk which would explain the Cham form better) > *priak > Cham /prea?/. 110
The phonology of Kompong Thom Cham

/beah/ ‘until’ [/beah/]
/bînj/ ‘eat’ [/bînj/]
/kadôŋ/ ‘bell’ [/kâ’dôŋ/]
/pajuan/ ‘move’ [/pâ’juan/]
/dêh/ ‘sputter’ [/dêh/]
/labâh/ ‘miserable’ [/lâ’bâh/]
/jâu/ ‘must’ [/jâu/]

The KTC preglottalised series reflects preglottalised consonant in PC:

*?barJ ‘eat’ > /bînj
*ca?buay ‘mouth’ > /cabuay/
*?duac ‘run’ > /duaiʔ/
*?jau? ‘must’ > /jauʔ/
*labiih/ ‘miserable’ [/lâ’bâh/]
*?ja11? ‘must’ > /ja11/

Preglottalised stops are retained in words borrowed from Khmer:

Khm. /book/ ‘hump’ > /bok/
Khm. /samba;}m/ ‘great’ > /sab;}m/
Khm. /dael/ ‘which’ > /del/

/m n n / are bilabial, dental, palatal, and velar nasals respectively. All but /n/ have been found in final position.

/mût/ ‘hear’ [/mi:t/]
/môh/ ‘place’ [/mo’h/]
/pâm/ ‘kind of pancake’ [/pam/]
/ni/ ‘bee’ [/nî:/]
/kôn/ ‘group’ [/kôn/]
/npj/ ‘noun’ [/nâŋ/]
/mañum/ ‘drink’ [/mâ’ñum/]

The PC nasals *m, *n, *ŋ, and *ŋ are retained unchanged in KTC.

*mata ‘eye’ > /mata/
*tamä ‘enter’ > /tami/
*mañâ ‘oil’ > /mañâ/ *
*?inä ‘main’ > /ni/
*jaŋ? ‘do’ > /jaŋ?/
*sâŋ ‘house’ > /saŋ/

The nasals are also retained in words borrowed from other languages.

Khm. /mei/ ‘chief’ > /mê/ Khm. /maak/ ‘type’ > /mâk/
Khm. /niên/ ‘addicted’ > /niân/ Khm. /aen/ ‘self’ > /eŋ/
Mal. hîna (ult. from Skt. hîna) ‘inferior’ > /hina/ ‘difficult’
Mal. saman (ult. from Ar. zaman) ‘era’ > /samân/ or /zamân/.

/w/ is a bilabial semi-vowel. It occurs in C¹ and F positions, and, in rare occasions, in C² position; and it is associated with high register nuclei. Original *hVw- clusters have become /w/ following by low register nuclei.

/wâʔ/ ‘pull down’ [/uâʔ/]
/lawây/ ‘stir’ [/lâ’yâ/]
In at least one case, the diphthong /oa/ which is a reflex of PC *ua and which seems regularly to become /wa/ in Eastern Cham, formed a cluster with /w/ as C\(^2\). This forms alternates with the expected /oa/ form.

/hoa?/ and /wa?/ (< *hwä? < PC *hua?) 'eat rice'

/w/ is a reflex of PC *w. In borrowings from Khmer, /v/ is reflected by KTC /w/ in all positions.

*wař 'stable' > /wär/
*lawa or jawa 'soul' > /niawä/ *niąw 'left (side)' > /iw/
*haway (or) huway 'rattan' > /wäy/
Khm. /väl/ 'field' > /wał/ Vn. čiēu 'mat' > /ciaw/
Khm. /sievphiv/ 'book' > /saphäw/ or /saphäw/

/l/ is a lateral sonant. It occurs in C\(_1\), C\(_2\), C\(_3\), and F positions.

/laʔu/ 'coconut' [ləʔu:] /bläy/ 'buy' [ʔblaí]
/lanuŋ/ 'eel' [lənunŋ] /täl/ 'until' [tal]

/l/ is a reflex of PC *l; the clusters *h(V)l- and *s(V)l- are reflected in KTC by /l/ followed by a low register nucleus. /l/ in borrowed words is retained.

*ilu 'much, many' > /lò/ *leʔ 'fall' > /lè/
*glo 'brain' > /klô/ *tulän 'bone' > /talän/
*lupän 'centipede' > /lapan/ *jäl 'net' > /cål/
*kapäl 'thick' > /kapal/ *täl 'until' > /tal/
*hVläʔ 'rice dust' > /ləʔ/ *hVläaʔ 'sharp' > /luə?/
*hulun 'slave' > /lin/ *sula 'leaf' > /la/
Khmm. /lout/ 'jump' > /lot/ Khm. /peil/ 'time' > /pēl/
Khmm. poolih> (ult. from French, police) > /plih/ 'police'
French chaland 'ferry' > /chalaɲ/
Skt. dhāli 'dust' /thūl/
Khmm. ḍobua (spoken: ḍobuah) 'wounded' > /labuah/

/r/ has two allophones. In all positions but F it is a voiced velar fricative [ɣ] which sometimes seems to approach a voiced stop [g]. In C\(_2\) position preceding high central vowels, /r/ shows a tendency to be lost; in F position, it is a voiced, high-mid, unrounded semi-vowel [u], while in C\(_3\) position, there is some alternation between /r/ and /l/.

/rüh/ 'shake out' [yuh] /prëy/ 'give' [pəyâi]
/krim/ 'bamboo' [kši:m] or [ki:m] /pär/ 'to fly' [pauʃ]
/ránəŋ/ or /láníŋ/ 'wide' [yəniŋ] or [lãnəŋ]

/r/ in all positions is a reflex of PC *r, which in F position following a high back rounded vowel or a diphthong is lost in KTC. /r/ in borrowed words has several treatments. In most borrowings from Sanskrit and Khmer, it has the same allophonic distribution as it has in native Cham words, but in some Sanskrit and Arabic loans, /r/ is a flap [ɾ].
*ruah ‘choose’ > /roah/
*krām ‘bamboo’ > /krim/
*cur ‘lime’ > /cu/
*kuar ‘gather up’ > /kua/
Skt rūpa ‘body’ /rup/
Skt nagara ‘fortress’ > /lakār/ ‘country’
Skt. naraka ‘hell’ > /nārka/ Skt svarga ‘heaven’ > /śārga/
Malay murtabat or martabat (ult. from Arabic) ‘grade, step’ > /muftabat/ ‘knowledge’
Khm. /riep/ ‘level’ > /rāp/ Khm. /sroop/ ‘absorb’ > /srop/

/y/ is a palatal semi-vowel. In C₁ position preceding a vowel it may have considerable friction, approaching [j].

/yaw/ ‘like’ [iau] /yaw/ ‘yoke’ [iu]
/hatay/ ‘after’ [hɔ’ti] /kayaw/ ‘tree’ [kɔ’ja]

/y/ in all positions reflects PC *y. The cluster *hVy- gives /y/ followed by low register vowel nuclei.

*yaw ‘yoke’ > /yaw/ *hayaw ‘like’ > /yaw/
*hayuk ‘harvest’ > /yoa?/ *tuay ‘stranger’ > /tuay/

/y/ in borrowed words remains /y/.

Khm. /claay/ ‘answer’ > /chlay/
Mal. dunia (ult. from Arabic) ‘world’ > /tûnya/

/s/ is a voiceless dental spirant. It occurs in C₁, C₃, and C₂ positions. It occurs in F position only in some recently borrowed words.


/s/ is a reflex of PC *s. Final s in early borrowings is usually reflected by -ih/; final PC *s has become -h/.

*suay ‘slow’ > /suay/ *susr ‘seed’ > /sår/
*srūh ‘nest’ > /srūh/ *brās ‘husked rice’ > /prāh/
*tikus (or) *tukus ‘rat’ /takūh/ *kakās ‘scale’ > /kakāh/
Khm. /saak/ ‘attempt’ > /sak/ Skt. doṣa ‘fault’ > /tūih/

Skt. mānuṣa ‘man’ > /manuih/
Khm. /cbah/ < (cbās) ‘clear’ > /cabaih/ Skt. aṣoka ‘fault’ > /tūih/
Khm. /cumnuah/ < (jamnuas) ‘substitute’ > /camnuih/

PC *s in C₃ position may be lost before stops or retained before other consonants. In the latter case, *s sometimes becomes *h which in turn is lost after causing a change in the register of the following syllable.
*sagor ‘drum’ > *kâr/  *sara ‘salt’ > *sra/
*sidor (or) *sideshow ‘remember’ > *târ/
*sidOm (or) *sudOm ‘ant’ > *tôn/ 
*samû (or) *sumû ‘flat’ > /samû/ 
*surâ? ‘writing’ > *harâ? > /râ?/ 
*sula ‘leaf’ > *hala > /la/  *sana ‘roast’ > *hana > /na/

/h/ is a voiceless glottal spirant. It occurs in all positions.

/hoa/ ‘pull tight’ [hoô]  /pahi/? ‘bitter’ [phi:?]
/hawã/ ‘dissatisfied’ [hô’ya:]  /plâh/ ‘split’ [plâh]

/h/ in all positions reflects PC *h. Final PC *h and *s have merged, and are reflected by /-h/. In borrowed words, /h/ remained /h/. Arabic ḥ merged with Cham /h/.

*hâ ‘yes’ > /hi/  *hia ‘cry’ > /hea/ 
*hûc ‘whistle’ > /hui/?  *bahraw ‘new’ > /phaw/ 
*hadûm ‘how many’ > /hatôm/  *?abih ‘all’ > /pîh/ 
Mal. hakêm ‘judge’ (ult. from Ar. hakêm) < /hakêm/ 
Khm. /cneah/ (< jhnah) ‘win’ > /canê

/f/ s/ z/ appear to be limited to borrowed words especially those from Arabic through Malay. /f/ is a voiceless bilabial spirant; /s/ is a voiceless palatal spirant, and /z/ is a voiced dental spirant. In casual speech, /s/ and /z/ usually become /s/.

/fô/ /sô/ /zô/ ‘heaven’ < Skt. svarga  /fâl/ ‘memorise’ < Ar. fal
/işrât/ (or) /isrât/ ‘show’ < Mal. isharat (ult. from Ar.) ‘signal’
/îzîn/ ‘permission’ < Mal. idzin (or) izîn (ult. from Ar.)

There seem to be two kinds of initial consonant clusters in KTC. One may be represented as /C1C2/ and consists of two consonants in close juncture without any intervening vowel. C2 in these clusters is limited to /l r h/. The following examples have been noted in this corpus:

<table>
<thead>
<tr>
<th>SECOND MEMBER</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>l</td>
<td>r</td>
</tr>
<tr>
<td>p</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>c</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>k</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>n</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

The second type of cluster may be represented as /C1aC2/. In clusters of this type, there is an intervening short vowel. Some of these show considerable variation. One speaker, in slow, careful speech may pronounce a cluster with the short vowel, e.g., [tôhun] ‘year’, while the
same speaker in more casual, rapid speech may pronounce the same cluster without any perceptible intervening vowel, e.g. [thun]. The C₂ in these clusters can be /p t c k ? j m n ñ w l y r s h/. As described in the second type above, there has been some simplification of PC clusters of *hVC- and *sVC-. In at least one case, the historical treatment of clusters differs. PC *tll- and *dll- > CHD /kl-/ but *tll- > KTC /tal-/, *tlaw ‘three’ > /klaw/, *tulaj ‘bone’ /tala/. Cham also has consonant clusters in polysyllabic words. In these, the common pattern is CVCCVF, as in:

/cakla/ ‘lightning’ [tɕʰ'klaː]
/paplay/ ‘to sell’ [pʰ'plɒj]
/karum/ ‘trap’ [kʰn'yam]
/masrüh/ ‘war’ [mwares'yuh]
/taple/ ‘scrapped off’ [tɒ'pleː]

Another pattern has a short vowel between the two middle consonants. The third syllable receives the main stress while the first receives secondary stress.

/marasa/ ‘maybe’ [r'maɣ'saː]
/masatah/ ‘from a distance’ [r'masə'tah]
/palahu?/ ‘to destroy’ [palaθiuyʔ]

Finally, there is the pattern CVCVFCVF as in:

/patanrauʔ/ ‘make heavy’ [r'patən'yaʊʔ]

4. Vowels
KTC has the following vowels:

<table>
<thead>
<tr>
<th></th>
<th>FRONT</th>
<th>CENTRAL</th>
<th>BACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH</td>
<td>i i'</td>
<td>i i'</td>
<td>u u'</td>
</tr>
<tr>
<td>MID-HIGH</td>
<td>e</td>
<td>e</td>
<td>o o'</td>
</tr>
<tr>
<td>MID</td>
<td>o o'</td>
<td>o o'</td>
<td>o o'</td>
</tr>
<tr>
<td>MID-LOW</td>
<td>e ĝ</td>
<td>a ĝ</td>
<td>o o'</td>
</tr>
<tr>
<td>LOW</td>
<td>a a'</td>
<td>a a'</td>
<td>o o'</td>
</tr>
<tr>
<td>GLIDES:</td>
<td>/ia, iu, ea, au, ua, oə/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All vowels that precede /-h/ are short, but they are still marked with the breve (˘) in this article. First or low register vowels are generally lower than their second or high register counterparts and may have on-glides. The vowels are now described in detail:

/i i̯/ are high front vowels. First register /i i̯/ is [ɨ i̯]:

/ni/ ‘this’ [ni:i̯:]       /ni/ ‘bee’ [n³i:i̯:]
KTC /i/ and /ĩ/ are reflexes of PC *ᵣi, *ᵣɨ, *ᵣɨ, and *ᵣɨ.

*cīm ‘bird’ > /cīm/  *?abīh ‘all’ > /pīh/
*jhīt ‘sew’ > /chīʔ/  *gīŋ ‘stove’ > /kīŋ/
*ʔaŋīn ‘wind’ > /ʔnīŋ/  *laŋīp; ‘sky’ > /laŋīʔ/
*phil ‘bitter’ > /phiʔ/

KTC /i/ and /ĩ/ may reflect /i/ and /ːi/ in borrowed words.

Pali bhikkhu > Khm. /phiʔkhoʔ/ > /phiʔ/ ‘Buddhist monk’
French police > Khm. /pooli̠h/ > /plīh/ ‘police’
Khm. /kan/ > /kīn/ ‘to mill rice’
Skt. sākṣin > Malay sakṣi > /sāksi/ ‘witness’

/e/ is a long mid-high front vowel; no short /e/ has been found in the present material except preceding /-h/.

/mè/ ‘chief’ [mè:]  /pēl/ ‘time’ [pē:l]
/ʔen/ ‘self’ [ʔen]  /sēh/ ‘student’ [sēh]
/take/ ‘horn, antler’ [tə’kə:]  

KTC /e/ is a reflex of PC *ᵣi.

*paggi ‘tomorrow’ > /pake/  *tuki ‘horn, antler’ > /take/

It also occurs in words borrowed from Khmer in which it is a reflex of /ei/ and /ae/.

Khm. /peil/ ‘time’ > /pēl/  Khm. /mei/ ‘chief’ > /mē/
Khm. /ʔaen/ ‘self’ > /ʔen/  Khm. /dael/ ‘have ever’ > /dēl/  

/e/ /e/ are mid-low front vowels; first register /e/ is [æ] in open syllables.

/ʔeʔ/ ‘fall into’ [ʔeʔ]  /cēh/ ‘insert’ [cēh]
/sare/ ‘debt’ [sə’yaæ]

PC *e and *e and *e are reflected by KTC /e/ and /e/.

*kēʔ ‘bite’ > /kēʔ/  *ʔeh ‘excrement’ > /ʔeh/
*bube ‘goat’ > /papē/  *srē ‘debt’ > /sare/

/i/ /i/ are high central unrounded vowels; first register /i/ is [ᵣiː] in open syllables.

/bīg/ ‘eat’ [ʔbiŋ]  /mī/ ‘field’ [mᵣiː:]
/mīʔ/ ‘get’ [mᵣi]  /nī/ ‘chief, main’ [nᵣiː:]

KTC /i/ and /ĩ/ are reflexes of PC *ã and *ã and *ã and *ã.

*tanã? ‘cook’ > /tanãʔ/  *lanãŋ ‘worm’ > /lanãŋ/
*limã ‘five’ > /lami/  *lumãn ‘elephant’ > /lamin/
The phonology of Kompong Thom Cham

*lumiǎ? ‘prepare’ > /lamiːʔ/ or /ramiːʔ/
*ḥā ‘yes’ > /bī/  *ʔbāŋ ‘eat’ > /bīŋ/
*čǐʔ ‘mountain’ > /cĩʔ/  *dāŋ ‘stand’ > /tĩŋ/

KTC /i/ may reflect /a/ in ancient borrowings.

Khm. /cneah/ (jhnah) > /canǐn/ ‘win’
The word /šŋąuʔ?/ ‘verb?’ is probably a loan from Malay.

/ə ǝ/ are mid-central unrounded vowels. The perceived difference between first register /i/ and second register /â/ is often very slight.

/ya/ final emphatic particle [jə]  /pɔn/ ‘blown away’ [pɔn:]  
/pɔʔ?/ ‘small dam’ [pɔʔ?]  /pɔʔ?/ ‘(banana) sheath’ [pɔʔ?]
/caln/ ‘very tall’ [cɔlɛn]

KTC /a/ and /â/ are reflexes of PC *ǝ and *ə. They also derive from Khmer /a/ and /aa/ in borrowed words.

*pɔn ‘to nail, pound’ > /pɔn/  *wɔʔ ‘instead’ > /wɔʔ/  
*pɔʔ ‘open’ > /pɔʔ/  *(...)*tą ‘tremble’ > /cataʔ/  
Khm. /clɔy/ ‘answer’ > /clɔy/  
Khm. /sɔmbaɔm/ ‘important’ > /sabɔm/  
Khm. /pɔhoɔn/ or /pɔhoɔn/ ‘arrogant, insolent’ > /phɔn/

There is one instance of KTC /a/ from Khmer /aa/.

Khm. /tae/ ‘only’ > /tɔə/ (cf. p.108 above).

/a ą/ are low central unrounded vowels. /a/ in unstressed presyllables represents a very short mid-central vowel [ǝ]; when it occurs before /m/ in presyllables, it is backed and rounded to [ɓ].

/saŋ/ ‘house’ > [sa:ŋ]  /nān/ ‘that’ [nàn]  
/tāŋ/ ‘younger sibling’ [tāŋ]  /kapaʔ?/ ‘walk’ [kɔpæʔ?]  

KTC /a/ and /ą/ are reflexes of PC *a and *ą. PC *ő preceding *r and *l is also reflected by /ą/.

*blâh ‘split’ > /plâh/  *hān ‘spicy’ > /hâŋ/  
*dalâm ‘deep’ > /talâm/  *katâl ‘itchy’ > /katâl/  
*pha ‘thigh’ > /pha/  *tulâŋ ‘bone’ > /talâŋ/  
*pâʔ ‘four’ > /paʔ/  *hujâŋ ‘rain’ > /câŋ/  
*ułâ ‘snake’ > /là/  *gɔr ‘handle’ > /kâr/  
*tâl ‘until’ > /tâl/  

KTC /a/ may reflect /a/ in borrowed words as well as /ie/ in some borrowings from Khmer.
Pali *akkhara* > /akhār/ ‘letter’
Pali *kappāsa* > Khm. /kapbaah/ > /kapāh/ ‘cotton’
Malay *akal* (ult. from Arabic) > /akāl/ ‘intelligence’
Skt. *rājakāra* > Khm. /rieccaaa/ > /racaka/ ‘government’
Khm. /riep/ (rāpā) > /rāp/ ‘level’

/u/ are high back rounded vowels; first register /u/ and /ū/ have lower on-glides [ʰu] or [u].

/hu/ ‘have’ [ʰu] /kūnj/ ‘who’ [kun]

/ūu/ ‘he, she, it’ [nū:] /cāruh/ ‘fall down’ [cōvūh]

KTC /u/ and /ū/ are reflexes of PC *ū* and *ū and PC *ū* and *ū.

*tūh* ‘pour’ > /tūh/ *pitū* ‘star’ > /patū/?
*tupūn* ‘flour’ > /tapūn/ *thūn* ‘year’ > /thūn/
*manū*? ‘chicken’ > /manū/? *srūh* ‘nest’ > /srūh/

KTC /u/ and /ū/ reflect /u/ in borrowed words.

Skt. *dhulī* > Khm. /thuuli/ > /thul/ ‘dust’
Malay *dunia* (ult. from Arabic) > /thunya/ ‘world’
Khm. /puun/ (būna) > /pūn/ ‘to mound up’
Skt. *guru* > Khm. /kruu/ > /kru/ ‘sorcerer’
Malay *kubur* (ult. from Arabic) > /kupur/ ‘tomb’

/o/ are mid-high back rounded vowels.

/o/ final negative particle [ʔo:] /pahōŋ/ ‘group’ [phoŋ]

/pōh/ ‘fruit’ [pōh] /lō/ ‘very’ [lōː]

KTC /o/ and /ö/ are reflexes of PC *ū*, *ū and *ū. In loanwords, KTC /o/ may also reflect Khmer /oa/, /ou/, and /ao/.

*/buh* ‘see’ > /bōh/ *hūŋ ‘papaya’ > /lahōŋ/
*/hadūm* ‘how long’ > /hatōm/ *lū ‘many’ > /lō/
*/tāgū*? ‘arise’ > /takōʔ/ *plūm ‘leech’ > /plom/
*/tū*? ‘receive’ > /tōʔ/

Khm. /tanot/ ‘sugar palm’ > /tanot/ Khm. /sroop/ ‘absorb’ > /srop/
Khm. /yong/ ‘two’ > /yōŋ Khm. /lout/ ‘jump’ > /lōt/

/o/ are mid-low back rounded vowels.

/lamō ‘cow’ [lāmō:] /tōʔ? ‘to be located at’ [tōʔ?]
/tōp/ ‘hit with the flat of the fist’ [tōp]
The phonology of Kompong Thom Cham

/kəh/ ‘knock’ [kəh]

KTC /o/ and /ɔ/ are reflexes of PC *o, *ö, *ö, and *ö.

*ʔakö ‘head’ > /kəʊ/  *gröh ‘to bark’ > /krəh/
*patö ‘teach’ > /pətə/  *rön ‘nourish’ > /rən/  
*šök ‘to peel’ > /ləʊ/  *néö ‘on’ /nəʊ/  
*höög ‘wasp’ > /hoʊ/  

Lee’s (1965) *O sporadically becomes KTC /ɔ/.

*ʔbrOm ‘arrow’ > /brəm/  
*sudOm (or) *sidOm ‘ant’ > /təm/ 

Khmer /oa/ and /ɔa/ become /ɔ/ in KTC; /oa/ becomes /e/ before -t/. 
Khm. /toap/ ‘army’ > /təp/  Khm. /croök/ ‘pickle’ > /croök/ 

There is abundant evidence in KTC for a split in the PC high and low vowels: *i > /i/ and /e/, *u > /u/ and /o/, and *a > /a/ and /i/. The conditioning factor for the split of *a is apparently nasalisation: *a > /a/, while *ã > /i/. The conditioning factor for the split of the high vowels has not been discovered. Lee (1977: 89-92) relates similar vowel splits in Haroi to: (1) the voicing feature of a preceding obstruent, and (2) the height feature of a preceding vocoid. Both reflexes of *i and *u have been found in apparently identical environments in KTC. It is always possible that the /e/ and /o/ reflexes, which are statistically much rarer than the /i/ and /u/ reflexes, are limited to borrowings from other dialects or from related languages, but so far, in most cases, no sources have been found.

The following glides occur:

![Diagram of glides](_diagram.png)

/sia/ ‘near’ [sia]  /hapian/ ‘when’ [hɐ’piən]
/laihu?/ ‘lose’ [lɐ’hiə?]  /seam/ ‘good’ [seam]
/cəən/ ‘craftsman’ [cəən]  /yuan/ ‘Vietnamese’ [jʊən]
/kua/ ‘embrace’ [kua]  /rəah/ ‘seek’ [rəah]
/təə/ ‘two’ [təə]  /səʔ?/ ‘smoke’ [səʔ?] 

119
KTC /ia/ is a reflex of PC *iē, *ia, and *io. It also occurs in borrowings from Khmer and Vietnamese.

*biēr ‘dwarf’ > /pia/ *palia ‘hail’ > /plia/
*pioh ‘keep’ > /mapiah/  
Kh. /ckiel/ ‘scrape’ > /cakial/ Vn. chīêu ‘sleeping mat’ > /ciaw/  
Khm. /nien/ ‘addicted’ > /niian/

KTC /iu/ occurs only before /-ʔ/ where it is a reflex of PC *iā?.

*lumiā ‘put away’ > /lamiwʔ/ *lahiā ‘lose’ > /laiwʔ/  
*maʔiā ‘urinate’ > /maʔiuʔ/

KTC /ea/ is a reflex of PC *iā/

*hiā ‘cry’ > /hea/ *tiān ‘abdomen’ > /tean/  
*liah ‘lick’ > /leah/ *biāk ‘true’ > /pēʔ/  
*paʔiā ‘hot’ /padeaʔ/

KTC /ua/ is a reflex of PC *ua and *uā. It also occurs in loanwords from Khmer.

*ruay ‘fly’ > /ruay/† *luan ‘to swallow’ > /luan/  
*suay ‘slow’ > /suay/ *kamuān ‘nephew’ > /kamuan/  
Khm. /phuay/ ‘blanket’ > /phuay/ Khm. /sua/ ‘ask’ > /sua/  
Khm. /tōtal/ ‘accept’ > /tātal/

KTC /oa/ is a reflex of PC *ua and *uā.

*dua ‘two’ > /tōa/ *cuah ‘sand’ > /coah/  
*dua ‘carry on the head’ > /dōa/ *juaʔ ‘step on’ > /cōʔaʔ/  
*caʔuā ‘basket’ > /caʔoа/

KTC /aʊʔ/ is a reflex of PC *-āp and *āʔ.7 (See also the treatment of PC final *-p above p. 109).

*cāp ‘tie’ > /cǎuʔ/ *anāp ‘in front’ > /tanǎuʔ/  
*troāu ‘heavy’ > /trǎuʔ/ *hǎuʔ ‘sweat’ > /hǎuʔ/

Presyllable Vowels

Lee (1965), mainly on the basis of evidence from Roglai, reconstructed three presyllable vowels: *i, *u, and *a. In KTC these have all merged to /a/, which is [ə]. When the presyllable began with *ʔ-, *h-, and sometimes *s-, the presyllable has been lost in KTC.

† cf. VN ruā ‘housefly, fly’ (Ed.).
7. Although I suspect that *āʔ derives ultimately from *āp.
The phonology of Kompong Thom Cham

*ʔinā 'female animal' > /ni/ *bilān 'season' > /plān/
*dilah 'tongue' > /talāh/ *ʔuni 'this' > /ni/
*sula 'leaf' > /la/ *limā? 'five' > /lami/
*kubaw 'water buffalo' > /kapāw/ *ʔadh~y 'forehead' > /thāy/
*hawāy 'rattan' > /wāy/ *patōw 'stone' > /patāw/

There may also be some evidence of the presyllable vowel in cases of metathesis.

*bunja 'flower' > /paŋJu/ *pirāk 'silver' > *pria? > /prea?/
*duray 'thorn' > /trūay/ *manī 'bathe' > *mnāy > /manāy/

REFERENCES

ASPECTS OF INTER-CLAUSAL RELATIONS IN KHMU

Suwilai Premsrirat

0. Introduction
This paper attempts to investigate certain mechanisms involved in the manifestation of inter-clausal relations in Khmu,\(^1\) and will discuss the Khmu that is spoken in Huoyen village, Chiengkhong district, Chiangrai province, Thailand. This variety is also found widely spoken in the Park Bang area of northern Laos.

Khmu is a Mon-Khmer language spoken by a group of hill tribesmen found primarily in northern Laos, northern Thailand—especially the area along the Thai–Lao border—northern Vietnam and southern China. They are believed to have previously occupied a larger area than at present, their arrival to this area predating that of Thai speakers and other more recent hill tribes such as Miao, Yao, Lisu, Karen, and others.

A. General syntactic structure\(^2\)
Words in Khmu are mainly monosyllabic, though disyllabic words are still numerous. It is possible that Khmu once had a very complex system of affixation even though nowadays many affixes are found only in fossilised form and have often lost their meanings.

In text materials, although most of the items can obviously be defined as belonging to the natural word classes such as noun, verb, adjective, adverb, preposition, demonstrative, etc., there are a number of ‘syntactic’ particles that do not seem to have meaning by themselves but do have certain functional meanings. They have a possible word shape, Cv(C), but they are never stressed. A number of them occur before the main verb. Many seem to be semantically related to the verb morpheme while in other cases they are related to the whole clause or even to a higher level. Without a knowledge of functional meanings of the particles (PT) it is almost impossible to understand the subtlety of the expressions and thus to follow a conversation, story, or talk thoroughly, even though the meanings of all the major words are clearly known. In order to give a clearer view of this curious syntactic feature which may lead to the understanding of inter-clausal syntax in Khmu, an overview will be given here, including a discussion of the verb phrase.

1. I am grateful to my Khmu friends who patiently gave me the data discussed here and for David Thomas’ comments and suggestions.
2. Further information on Khmu structure (phonology and syntax) can be found in Suwilai Premsrirat (1987c). Detailed information on phonology (different dialects) and morphology will also be found in Smalley (1961) and Svantesson (1983).
The normal word order in Khmu is SVO. In case of emphasis the OSV order is very common. Examples of normal clause patterns are:

Intransitive clause - կը:n ճամ³
child  cry
‘The child is crying.’

Transitive clause - տո: pa? մահ
they  eat  rice
‘They are having a meal.’

Bi-transitive clause - ո՞ տան մահ ջոյ տե? կմն:լ
I  give  parents  take  money
‘I gave my parents some money.’

Semi-transitive clause - ճա? ճատ տա կա:յ
grandmother  stay  at  home
‘My grandmother is at home.’

Equative clause - ո՞ մահ քբռ:ու:
I  am  teacher
‘I am a teacher.’

Existive clause - տա կա:յ ահ հիտոր, ահ սջ?, ահ սինջ
at  home  have  hen,  have  dog,  have  pig
‘There are hens, dogs and pigs at home.’

Quotative clause - ջոյ պրա:կ կավ ժահ լո? լո?
father  say  before  say  go  well  well
leaving
‘My father wished them a safe trip.’

Modifiers (Mod.) in a noun phrase normally follow the head noun (HN).

HN + Mod.

սջ? հիտոն պա:ր (տո:) կամ նամ կոնա:ջ
dog  black  two (class.)*  which  big  that
‘those two big black dogs’

However, it is noticeable that the reversed orders of clause pattern, (O)VS, and of noun phrase, Mod.+HN, have also been found in text materials.

(1) լամ մահ լամ մա:ր
delicious  rice  delicious  salt
‘Have a good life.’

3. * refers to the clear tense register of this Khmu variety, whereas the breathy lax register is unmarked.
* = classifier (Ed.).
Aspects of inter-clausal relations in Khmu

(2) *puh kmá?*

fall rain (n.)

'It is raining.'

The Khmu verb phrase normally consists of one or more main verbs (MV) and their modifiers (auxiliaries and preverbal adverbs) occurring mainly in the pre-verbal positions. The pre-verbal modifiers (preV) can be classified into 5 categories according to their positional occurrences away from the main verb.

preV 5  preV 4  preV 3  preV 2  preV 1  MV

PreV 1 indicates state of existence like /pion/ 'able', /jat/ 'to have been happening, still', /hó:c/ 'already', /cú'/ 'want to', /?k/ 'not want to', /ko:j/ 'slowly, gradually', /lak/ 'completely', /ph5:n/ 'never,' etc.

PreV 2 indicates negative meaning like /pe/ 'not', /pl5?/ 'not any more'.

PreV 3 indicates intention, determination and future action like /ce/ 'will'.

PreV 4 indicates tense and aspect like /há/ 'already (past action)', /nó:η/ 'still, will'.

Some of the pre-verbals of category 1 and category 2 like /pion/, /hó:c/, /lak/, and /pl5?/ may sometimes be shifted to the post-verbal position.

Examples of verb phrases containing preV 1, 2, 3, and 4 functioning in the clause are

\[
\text{na: p}^{h5:n} \text{ ?iak pu:c} \\
\text{ she never drink wine} \\
\text{ 'She has never drunk wine.'}
\]

(3) *sna: há r2:t*  

they (two) already arrive  

'They have come.'

(4) *?o? pë pion jh ta jü?*  

I not able go to forest  

'I did not have a chance to go to the forest.'

(5) *?o? n5:η ce jh m3:t kmú:l ká:l*  

I still will go take money before  

'I still have to go to take the money first.'

(6) *nó: há ce pl5? pion jh ja:η ki:*  

they already will not any more able do like this  

'They will not be able to act like this any more.'
PreV 5 indicates modality such as the actor's volition, obligation, necessity or ability in doing something. Since it involves a relationship between a causer which may be stated or unstated and the actor and the action it is more like a clause rank notion than a phrase one. In Khmu it contributes a good deal to the inter-clausal relationship. This category includes /tan/, /le/, /cen/, mák/, /kaj/ and /th/.4 These will be dealt with in detail in the next section. Examples of Khmu utterances with various pre-verbal particles including the category 5 pre-verbals are

(1) prap loh las ?n? mák pe pían j?h
   5 2 1 MV
dress body quickly we afraid not able go
   'Dress quickly otherwise we will not be able to go.'

(2) náj ?o? law ce pe j?h ?o? tan ha ce
   3 2 MV 5 4 3
a moment ago I say not go I certainly already will
   2 MV MV 5 MV
not go you go I therefore go
   'I said that I would not go so I should not go. But you will go so I will too.'

(3) po j?h kâ:l ?n? tan ce n5:y j?h
   5 3 4 MV
you go before I certainly will still go
   'You go first and I will follow you.'

It should be noted here that the preV 3 /ce/ which indicates the future and the speaker's intention and determination can also function at the inter-clausal level (as will be seen in the next section) whereas the preV 5 /tan/ which also indicates the speaker's determination or obligation (though in a higher degree) can be manifested at the clause level.

B. Inter-clausal Relations
In the context of speech, a semantic sentence represents a minimum speech act consisting of one or more predications (semantic clauses). A structural sentence is therefore composed of at least one main independent clause with or without subordinate clauses and occurs with a distinct illocutionary force (mood) which may be a statement, question or command. In Khmu, two clauses or series of clauses whose meanings are related are structurally joined by various devices. The formal linguistic devices used may be conjunctions, pre-verbal particles, and reversed word

4. Since these pre-verbal particles do not have any content meaning, the explanation here is presented without giving the meaning of each item. However, in the following examples the approximate meaning of each item has been attempted.
5. The category 4 pre-verbs can appear after the category 3 in certain cases.

It is noticeable that consciously and off context the co-occurrence between the category 4 pre-verbs /há/ and /n5:y/ and the category 3 pre-verbal /ce/ is considered as 'possible' but 'not quite right'. The /ce/ seems to be the main item.
Aspects of inter-clausal relations in Khmu

order in the subordinate clauses. Apart from this, there are structurally independent clauses that are just juxtaposed (strung together within a common intonation contour but without any formal conjunction markers). I shall discuss inter-clausal relationships in terms of temporal, logical, and compounding sequences.

1. **Temporal sequence relations**: these have an initial action followed by one or more succeeding actions. They are normally joined by the medial conjunctions /hó:c/ 'after that, already, then' /jna:m/ 'at the time, when, time', /ká:l/ 'before, in front of' and /kʰni?/ 'after, at the back'. However, in certain contexts the succession of actions may appear without any conjunction. The structural manifestation of this kind of relation is:

   SVO (conj.) SVO

   (1) peʔ mah hó:c ?əh wiak
eat rice then do work
   ‘I will eat rice first then I will go to work.’

   (2) ?oʔ ?əh rnaː hó:c kə:j sih
I do rice field then come sleep
‘When I finish working on the rice field I will come to sleep.’

   (3) kaː jəh naːm ?oʔ nòːn sih
he go when I still sleep
‘He went while I was sleeping’

   (4) ?oʔ muːm ?əm ká:l pəʔ mah
I bath water before eat rice
‘I take a bath before having a meal.’

   (5) ?oʔ rəh tə sih kʰniʔ maʔ jəh ta rəʔ
I get out from sleep after mother go to field
‘I got out of bed after my mother had gone to the field.’

Sometimes the meaning of words chosen indicates the temporal sequence.

   (6) jeʔ jəh ká:l ?oʔ ce jəh nam ta kʰniʔ
you go before I will go follow at the back
‘You go first. I will follow you.’

In some contexts a series of verbs following one another indicates the sequence of actions that happen one after the other. They may or may not have the conjunction /hó:c/ in between. The /hó:c/ normally occurs before the last verb of the utterance.

   (7) paːt pluʔ ə əɾéh (hóːc) kəh
cut off leg remove flesh then chop
‘Sever and bone the leg and then chop it up.’
8. *pok* 0 *pliut* 0 *pen* to:n 0 *tran* 0 wah *pe:*r
cut off cut short to be piece cut square split small
*pe:*r (*hô:c*) *hô:c*
small then cut smooth
‘Cut down (a tree) cut it into (short) pieces, smooth off the rough edge, split it into small pieces, then smooth them.’

The pre-verbal particle category 5 /tʰa/ or its variants /ta- da-sa/ has also been found optionally co-occurring with the pre-verbal category 3 /ce/ and the conjunction /hô:c/ in the following clause to indicate that the event happens after another. Its structure can be manifested as

```
SVO (*hô:c*) S (tʰa) ce VO
```

1. *(tʰa)* kma? hâ *phri?* hô:c 0 *pe:*r (tʰa) ce *joh* pe? mah
   rain already stop after that I then will go do work
   ‘When the rain has already stopped I will then go to work.’

2. *nô:m* mah *hô:c* 0 *pe:*r (tʰa) ce *hian* to:
   I still have to eat rice after that I then will learn continue
   ‘I have to have a meal first then I will continue learning.’

3. *(tʰa)* *mu:m* ?om hô:c 0 *pe:*r (tʰa) ce *sih*
   I bath water after that I then will go sleep
   ‘I will go to sleep after I have taken a bath.’

2. Logical sequence relations: these will be discussed mainly in terms of various ‘conditional’ sequences that are different in the degrees of determination, possibility, certainty, expectation and counter-expectation, including various shades of overtone. This type of sequence normally consists of a main clause stating ‘condition’, ‘cause’ or ‘environment’ and one or more dependent clauses stating ‘consequence’. Notions like ‘causal’, ‘resultant’ and ‘co-varying’ will also be included in the discussion of this type of relationship.

2.1. Conditional relations: six different pre-verbs /ce/, /le/, /tañ/, /cen/, /mâk/, and /kaj/ are found attached to the consequence clause distinguishing six different types of conditional relations. The first particle has been discussed as pre-verbal category 3, whereas the last five belong to category 5. They can be manifested structurally as

```
SVO S preV VO
```

<table>
<thead>
<tr>
<th>ce</th>
<th>le</th>
<th>tañ</th>
<th>cen</th>
<th>mâk</th>
<th>kaj</th>
</tr>
</thead>
</table>

128
2.1.1. /ce/ normally occurs within a clause to indicate futurity and the actor’s determination to do something. Occurring in the consequence clause, it indicates that the action it is attached to will normally happen as the possible consequence of the preceding event. It is the most common type of conditional relationship. There is more than one way of manifesting this type of conditional relationship. Apart from the fact that the verb in the consequence clause is introduced by /ce/, the conditional clause may be introduced by the Lao loan conditional conjunction /kʰán/ or the Thai loan /tha:/, both meaning ‘if’. The structural manifestation of this relation can be summarised as:

(conj.) SVO S preV VO

1. (kʰán) ?o? pup rō:j ?o? ce tar
   If I meet spirit I will run
   ‘If I see the spirit I will run.’

2. (kʰán) kʰa:t ḱo:t ḱo? ce pian woh
   If rain arrive paddy will get plenty
   ‘If it rains we will get plenty of paddy.’

   If I not do work the second day I will go down after tomorrow
   ‘If I do not work on the second day after tomorrow I will go to see you.’

4. (tʰa:) ?o? ah snā:t ?o? ce piŋ si:m
   If I have gun I will shoot bird
   ‘If I have a gun I will shoot the birds.’

5. (tʰa:) pa jōh ḱo? ce jōh dē
   If you go we will go together
   ‘If you go so will we.’

Some speakers put the word /ʔuan/ ‘give’ before the first clause of a sentence with /ce/.

6. ʔuan mar tōk nx: nx: ce há:n
   give snake bite then they will die
   ‘If they are bitten by a snake they will die.’

7. ʔuan jōŋ ḱo:t ḱo? ce jōh
   give father arrive I will go
   ‘If my father comes I will go.’

8. ʔuan ḱo? ah pnīr ḱo? ce tō:r miŋ si:m
   give I have wing I will fly like bird
   ‘If I had a wing I would fly like a bird.’
2.1.2. /le/ indicates that the action it is attached to has naturally happened or has to happen as the consequence of the preceding action. It is information which is being reported to somebody.

(1) ʔoʔ mu:m ʔom hə:c ʔoʔ le ʔoʔ le joh sih
I bath water already I therefore go sleep
‘I go to sleep after taking a bath.’

(2) maʔ céh ʔoʔ ʔoʔ le ja:m
mother scold me I therefore cry
‘My mother scolded me so I cried.’

(3) hic ʔiʔ ʔiʔ ʔun ʔak hə:c le ku:m
crush paddy cause small piece already therefore winnow
‘Crush the paddy until it becomes small pieces then winnow it.’

(4) joŋ kaʔ ʔiʔ ʔak pu:c ka: le te: pa:j
father he not want drink wine he therefore make himself drunk
‘My father did not want to drink wine so he pretended to be drunk.’

(5) ʔaʔ jəʔ kləm pion le ʔoʔ pə:j ʔoʔ ʔəm ʔəm
init. part. you carry able therefore take lead go fin. part.
‘If you can carry it then take it.’

2.1.3. /tan/ indicates that the action it is attached to certainly and obligatorily happens as determined by the event or circumstances which may be expressed in a clause, or even by a non-verbal behaviour or situation. It implies the speaker’s expectation and determination. It may also co-occur with /ce/.

(1) naʔ jəʔ pion ʔiʔ tan jəʔ pion
they live able we certainly live able
‘If they can stay so can we.’

(2) nk:i:n ʔoʔ joh ta júʔ mi: skı: ʔoʔ tan ce joh me?
yesterday I go to forest today I certainly will go new
‘Yesterday I went to the forest, today I will go again.’

(3) ma wát tan wát
init. part. fight certainly fight
‘Let’s fight.’

(4) ten ta kə: tan pion
sit at here certainly able
‘Have a seat here.’

(5) ʔoʔ cúʔ lu:j ʔoʔ tan joh bo:k pa:
I have a pain stomach I certainly go tell you
‘If I have a stomach-ache I will tell you.’

6. fin. part. is abbreviated from ‘final particle’. The other related abbreviation is init. part. (= initial particle).
2.1.4. /cên/ indicates that the action it is attached to is unavoidable and would logically happen as a result of a specific action or state. It normally has an overtone of ‘sarcasm’, ‘unconcern’, or sometimes ‘annoyance’. In some cases it indicates that the action it is attached to is an unreal situation. It normally occurs with the final particle /(ʔu) ?eh/.

you know the right thing you definitely not do like this fin. part.
‘If you know what is right and wrong you will definitely not do like that.’

(2) ka:j cen ka:j pa: na jat ?zh mûh ta ki:
go back definitely go back you part. live do what at here
‘If you want to go (home) then go. Why continue living here?’

(3) hûn cen hûn ?zh pà: na pri?l ?zh mûh
die definitely die fin. part. you part. alive do what
‘If you want to die then die. Why stay alive?’

bird it have wing it fly definitely able fin. part.
(je? pê ?zh pnîr je? ti:r ce pûn mûh)
you not have wing you fly will able what
‘Birds have wings so they can fly. (You don’t have wings, how can you fly?)’

you leave like this it definitely fin. part. lose
‘You left it here like this so (of course) it got lost!’

Sorn, you definitely sleep get up early fin. part. go do work
rûa ce tan
then will be in time
Sorn, if you had got up early you would have been able to go to work in time.’

(7) na: cen tar las ?u ?zh na: nû:ŋ nîm
she definitely run quickly fin. part. she still young
‘She is still young, of course she can run quickly.’

2.1.5. /mák/ indicates that the action to which it is attached is undesirable, and will (probably) happen as a result of another action. It has the overtone of ‘warning’ or ‘concern’. It may occur with the conjunction /hi?/ ‘in a moment, otherwise’.

(1) hû ta: lor re:ŋ (hi?) ka: mák mec
already not talk loud otherwise he probably hear
‘Don’t talk loudly or he will hear what you say.’
(2) ta: lor ja:ŋ koni no: mâk mo:r je?
not talk like this they probably fine you
‘Don’t talk like this or they will fine you.’

(3) krî rîh ri? ka:j las mâk puh
thunder loud noise we go back quickly probably be soaked in
krî mâk pê tan
rain go back probably not be in time
‘It is thundering so we have to go quickly, otherwise it will rain
and we will not be able to go back.’

(4) ta: joh sê:t hi? mâk la:c
not go alone otherwise probably get lost
‘Don’t go alone or you will get lost.’

(5) ta: hô:j we:l hi? mâk òoh
not play knife otherwise probably cut oneself
‘Don’t play with the knife, otherwise you will cut yourself.’

(6) ta: joh krip kâ te: jô no: mâk pîng po:
not go seize fish other people because they probably shoot you
mâk hân le:ŋ
probably die free
‘Don’t go to catch others’ fish or they will shoot you and you will
die for nothing.’

2.1.6. /kaj/. In concessive or adversative sequence, the main independent
clause and the subordinate clauses that follow are contrasted in content.
The subordinate clause normally posits an exception or ‘concession’ to
that which was or will be stated in the main clause. In Khmu this
relationship is expressed by the occurrence of the pre-verbal particle /kaj/
in the subordinate clause. /kaj/ indicates that the event it is attached to
happens in a contra-normal sequence, as an unexpected consequence of
the preceding event, or even as a kind of conflict with the condition.

(1) ñe: jat ta kăng je? kaj òah kmû:l woh òó kâm
init. part. live at village you become have money a lot I who
joh ca:ŋ kaj pê òah
go be employed become not have
‘How come you who stay at home are rich but I who go to work
outside am not?’

(2) bô:k ñûan pa: joh òô wîsk hò:c pa: kaj jat òô môh
tell give you to do work already you become stay do what
‘I told you to go to work. Why are you still here?’

(3) no: jat pîô: ì: kaj jat pê pîô:
they live able we become live not able
‘They can stay; how come we can’t?’
Aspects of inter-clausal relations in Khmu

(4) mi: skì: ?ò? pè ?ah wìák ta sù? ?ò? kaj ?òh wìák ta kì: today I not have work there I become do work here ‘I come to work here because there was not any work over there.’


(6) kà: ce sròn kò:n te: ku:t rò:ngia kà: kaj law he will send child his enter school he become say pè ?ah ma? pò:s hè:m te: not have who look after younger sibling his ‘Instead of sending his child to school he said that there was no one else to look after his younger sibling.’

2.2. For ‘causal’ sequences the conjunction /jò?/ ‘because’ is used between the consequence and its conditional cause. The structural manifestation of this type of relationship can be written as:

SVO conj. SVO

(1) kà: hà:n jò? mar tòk he die with snake tòk ‘He died of snake bite’ (lit. ‘The snake bit him to death’)

(2) kò:mù? hììn jò? pò:h wà:r Khmu black because to be in the sun hot ‘The Khmu are dark because they work in the sun.’

(3) na: tòn ?ìk je jò? jò? hà ?òh tmkò? she certainly not want you because you already have wife ‘She must not want you because you already have a wife.’

(4) ?ò? jò: tò jì: jò? te: ce ?ò:h hò? I go to forest because myself will do firewood ‘I will go to the forest to get the firewood.’

2.3. Co-varying sequence or a sliding-scale conditional sequence consists of two clauses that are grammatically balanced but semantically opposing, with a free variable and a conditioned variable. It indicates the increasing degree in the event of the second clause. The conjunctions normally used are the Thai loan /jì:ò/... jì:ò/ and the Lao loan /hè:ò/... hè:ò/. The structural manifestation of this type of relationship can be written as:

conj. SVO conj. SVO
2.4. In certain conditional constructions the subordinate clause which states the result has the normal clause order reversed.

SVO OVS

(1) ʔo? srʔeːŋ maʔ jon lian ?oʔm mat
    I miss father mother come out water eyes
    ‘I miss my parents so much that the tears fall.’

This characteristic has been extended to other sequences such as certain contrastive compounding (which is treated in the next section) and also to an immediate answer to the question in a question-answer sequence.

(2) maʔ ce ra: snʔeh ra?: ?i?
    who will clean dishes clean we
    ‘Who will do the dishes? We will.’

(3) maʔ ruŋ puːc kŋ ki: lam lak puːc ruŋ ?o?
    who cook make wine this delicious a lot wine make I
    ‘Who made this delicious wine? I did.’

(4) maʔ pęŋ sʔʔeːŋ sʔʔeːŋ pęŋ ?o?
    who cut down tree tree cut down I
    ‘Who cut the tree down? I did.’

However, it must be noted that this kind of syntactic reversal occurred only in certain contexts. The effect is probably topicalisation.

3. Compounding sequence relations: these refer to the co-ordination of two independent clauses acting as a single unit in the higher grammatical level or sentence level. I shall discuss this type of relation in Khmu in terms of additive, contrastive and alternative sequences. There is no formal conjunction between clauses. The clauses are just juxtaposed. However, they normally occur within a common intonation contour. The structural manifestation of this relation is

SVO SVO

However, in contrastive sequence the subordinate clause may have a reversed order of the elements.

SVO OVS / OVS SVO
Aspects of inter-clausal relations in Khmu

3.1. Additive compounding consists of clauses with only either the subject, predicate or the object in contrast.

(1) na: joh dé: ja:m dé: she go also cry also
   ‘She was crying while going.’

   at home I do field do trap feed dog feed pig
   ‘At home I worked on the rice field making traps, feeding dogs and pigs.’

(3) nak nō? sēh nak kla:n sēh
   heavy paddy put in heavy rock put in
   ‘It is heavy because both paddy rice and rocks have been put in there.’

(4) na: ku? no: woh na: pé lōn no:
   she love them a lot she not forget them
   ‘She loves them a lot and she will not forget them.’

(5) co:n la:c kʰŋ tō y ia:n nax ha:n
   thief steal thing kill them die
   ‘The thieves took the things and killed them.’

3.2. Contrastive compounding consists of two contrasting clauses in subject and predicate (verb or object).

(1) je? nām ?o? ne?
   you big I small
   ‘You are big (but) I am small.’

(2) ?o? tuk pa: rma:n
   I poor you rich
   ‘I am poor (but) you are rich’.

(3) ?o? hē:t ka: pē jās
   I call he not answer
   ‘I called (him) (but) he did not answer.’

(4) ?o? ce joh ta mok na ce joh ta ?ōm
   I will go to mountain she will go to water
   ‘I will go to the mountain (but) she will go to the river.’

(5) me:w pē pēn sō:n kʰmū pēn sō:n
   Meo not cut down tree Khmu cut down tree
   ‘The Meo do not cut the tree down (but) the Khmu do.’

   I not know how to sing he know how to
   ‘I do not know how to sing Khmu traditional song (but) he does.’
The order of the subordinate clause in the last two examples is naturally reversed in the text.

(5)a $\text{me:}:\text{w } \text{pē} \text{ pēν} \quad s^{?}:\nu \quad s^{?}:\nu \text{ pēν} \quad k^{\text{h}}\text{mu?}$
Meo not cut down tree tree cut down Khmu
‘The Meo do not cut the tree down (but) the Khmu do.’

or (5)b $s^{?}:\nu \text{ pēν} \quad k^{\text{h}}\text{mu?} \quad \text{me:}:\text{w } \text{pē} \text{ pēν}$
Tree cut down Khmu Meo not cut
‘The Khmu cut down the tree (but) the Meo do not.’

(6)a $\text{?o? } \text{pē } \text{c}^{\text{h}}\text{ā}:\text{j} \quad \text{tō:} \text{m } \text{c}^{\text{h}}\text{ā}:\text{j} \quad \text{ka:}$
I not know how to sing know how to he
‘I do not know how to sing Khmu traditional song (but) he does’.

3.3. Alternative compounding consists of two clauses with alternative predicate or alternative participants.

(1) $\text{pa: } \text{ce } \text{j}\text{sh} \quad (\text{pa:} ) \quad \text{ce } \text{pē } \text{j}\text{sh}$
you will go you will not go
‘You will go or not.’

(2) $\text{pa: } \text{ce } \text{j}\text{sh} \quad \text{pa: } \text{ce } \text{jat} \quad \text{tan} \quad \text{pian } \text{pā?} \quad \text{pā:}\text{r}$
you will go you will stay certainly able both two
‘You will go or you will stay.’

(3) $\text{pa: } \text{lor } \text{‘mo:j } \text{cum’} \quad \text{tan} \quad \text{la?} \quad \text{‘mo:j } \text{kū:’} \quad \text{tan} \quad \text{la?}$
you say ‘mo:j cum’ certainly right ‘mo:j ku:’ certainly right
‘You can say either ‘mo:j cum’ or ‘mo:j ku:’.’

(4) $\text{lor } \text{ro?} \quad \text{lor } \text{hō:j}$
speak real speak play
‘Are you just joking or not?’

(5) $\text{pa: } \text{‘a} \text{h } \text{kmā:l } \text{woh} \quad \text{pa: } \text{ce } \text{we:t } \text{lōt} \quad \text{pa: } \text{ce } \text{ro?} \quad \text{ka:} \text{ν}$
you have money a lot you will buy car you will build house
‘If you have a lot of money will you buy a car or will you build a house?’

The correspondences between various types of semantic relations and syntactic devices is summarised in Fig. 1 on the following page.
Aspects of inter-clausal relations in Khmu

**semantic relations**

1. temporal
   - 2.1 possible cons.
   - 2.1.1 initial preV conj.
   - 2.1.2 reported cons.
   - 2.1.3 expected cons.
   - 2.1.4 unavoidable cons.
   - 2.1.5 undesirable cons.
   - 2.1.6 unexpected cons.

2. logical
   - 2.2 causal
   - 2.3 resultant
   - 2.4 co-varying
     - 3.1 additive
     - 3.2 contrastive
     - 3.3 alternative

3. compounding

**syntactic devices**

1. conjunctions
   - med. conj.
   - init. preV conj.
   - init. med. conj.

2. pre-verbal particles

3. reversed word order

4. juxtaposition

**Fig. 1**

**Conclusion**

Inter-clausal relations in Khmu have been discussed both semantically and structurally. Three main types of semantic relations, temporal, logical and compounding sequences, including various sub-types, have been discussed. The syntactic devices used in manifesting these semantic relations are conjunctions, pre-verbal particles, reversed word order and juxtapositions or a mixture of these. It is obvious that in Khmu conjunctions are not in predominant use, only a few examples having been found. Some are Thai or Lao loanwords, for instance, those initial, or initial and medial, conjunctions discussed in 2.1.1 and 2.4. The common Khmu medial conjunctions are /hō:c/ 'then, after that', /jɔŋ/ 'because', and /hiʔ/ 'otherwise'. It is noticeable that their occurrences are not obligatory. They may be left out, or some other way of expressing the relationship may be taken instead. They also occur normally in positions of other word classes such as /jɔŋ hō:c/ 'to have already gone', /ckō hiʔ/ 'this coming evening', /pɔŋ jɔŋ tɨʔ/ 'to eat with hand', /jɔŋ jɔŋ jɔŋʔ/ 'to go with others'.

The reversed word order has been found used mainly in special or emphasised cases. The use of pre-verbal particles and juxtaposition is quite clear and obligatory. Pre-verbal particles are mainly used for subordinative constructions manifesting temporal and various logical (conditional) sequences, whereas juxtaposition is mainly used for co-ordinative constructions manifesting various compounding sequences. The systematic use of different pre-verbal particles for distinguishing different conditional relations in Khmu is unique. It also shows certain things about how the speakers of this language view and systematically categorise their 'conditional (cause and consequence)' experiences.

The young Khmu consider the pre-verbal particles as the conservative
forms of Khmu. Those who work in urban areas among the Thais tend to lose them gradually and unconsciously adopt the Thai pre-verbal particles /k5:/ ‘then’ /t5η/ ‘must’ and final particles /sì/ and /lɔːj/, even though they are not exactly equivalent, whereas the people at home, especially the old and the women, still make quite a lot of use of the old forms. They sometimes adopt Thai conjunctions, /tʰâː/ ‘if’, /pʰr5? (wâː)/ ‘because’, /lèʔ/ ‘and’, /têː/ ‘but’, and /rīː/ ‘or’. However, they said that they still found it difficult and unnatural to express the contrastive, alternative or additive compoundings using the conjunctions /têː/, /rīː/ and /lèʔ/, respectively. Looking at the text materials of another Khmu variety which is spoken in the Nam Pan, Huay Moy, and Nam Lu villages in Chiangklang district, Nan province, Thailand, it is noticeable that the syntactic constructions in that area have been simplified quite a lot and become more like those of Thai. However, unconsciously, in certain contexts, most of the pre-verbal particles mentioned above are still retained, though the pronunciation may be slightly different. The exception is /cen/ which is generally replaced by /ce/, the equivalent of the Thai /câʔ/ which indicates the speaker’s intention and future action. This pre-verbal seems to be very common and can sometimes replace other pre-verbals such as /mâk/ and /râː/.

REFERENCES


Aspects of inter-clausal relations in Khmu


AN INSTRUMENTAL STUDY OF CHONG REGISTERS

Theraphan L. Thongkum

0. **Introduction**

The Chong language belongs to the Pariic branch of the Mon-Khmer language family (Thomas & Headley 1970; Diffloth 1974; Huffman 1976, 1985; Headley 1977, 1978, 1985). Both amateur and professional linguists who have worked on Chong (eg. Baradat 1941; Martin 1974; Huffman 1985; Gainey (personal communication); Suphanphaiboon 1982) seem to recognize the 'glottal feature' or 'glottalization' which occurs in some Chong words. Some of them (Huffman, Gainey, and Suphanphaiboon) hear phonation types—normal voice vs. breathy voice, etc.—and pitches. On the basis of linguistic descriptions, there is no doubt that Chong is a register (R) language.

In 1983, Gérard Diffloth and I made several linguistic field trips to Chong communities in Makham District, Chanthaburi Province; we also visited Chong villages in Pong Nam Ron district, Chanthaburi Province, in Bo Rai District, Trat Province. During 1983-85, two female Chong informants from Krathing Village, Phluang Sub-district, Makham District, were brought to Bangkok many times for the purpose of checking language data and making high quality recordings and instrumental studies.

In spite of my training as a phonetician and my knowledge of Mon-Khmer languages such as Bru, Nyah Kur (Chao Bon), Mon, Kui (Suai), and Mia Bri, I still think that Chong is very exotic. In my opinion, the cause of the complexity lies in the process of Chong becoming a tone language. In fact, some dialects of Chong, such as the one spoken in Chamkhlo’ Village, Takhianthong Sub-district, Makham District, have already become tonal: presyllables are dropped; phonation types are less prominent and in some cases disappear; and pitch differences can be heard clearly, especially in slow speech. Our Chong informants also describe their language as having high, higher, mid and low tones.

The instrumental analysis presented in this paper is based on the speech of four Chong informants from Krathing Village. For the sake of convenience, they will be addressed as MA (first male speaker), MB (second male speaker), FA (first female speaker), and FB (second female speaker).
speaker). The Krathing dialect was chosen because its register phenomena suit my major interest—the acoustical measurements of the register complexes which involve several phonetic parameters. Moreover, the place where it is spoken is easy to reach, and the villagers are also very co-operative.

The phonetic instruments used in the study are as follows:
- Kay Sono-Graph 6061-B;
- Fundamental Frequency Meter, type FFM 650 (F-J);
- Intensity Meter, type IM 360 (F-J);
- Electro-glottograph, type EG 830 (F-J);
- Electro-aerometer, type EA 510/4 (F-J);
- Mingograf 34 T (Siemens AB).

1. Definitions of 'register'
The term 'register' has been used in many different ways. As a result, many definitions can be found in the literature depending upon who uses it—music and voice specialists, phoneticians, linguists, or language teachers.

1.1. Voice register
Music and voice specialists describe the rate of vocal-fold vibration in terms of registers. Garcia (1855) recognises three voice registers or ranges of pitch: the chest register, the mixed or middle register, and the head register. More often, only the chest and head registers are used. Others have tried to clarify the problems of voice-register terminology: 'The terminology with regard to voice pitch level, i.e. "registers", suffers from the existence of an abundance of terms and an ambiguity of their use.' (Mörner, Fransson & Fant 1963: 18). They therefore define a register by means of its range on the musical scale, suggesting five basic registers, namely: deepest range, deep level, mid level, high level, and highest level.

The approximate ranges and boundary limits of these registers are illustrated, and some synonyms are listed, for example:

<table>
<thead>
<tr>
<th>Deepest range</th>
<th>Deep level</th>
<th>Mid level</th>
<th>High level</th>
<th>Highest range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayon profond</td>
<td>Chest register</td>
<td>Falsetto I</td>
<td>Falsetto II</td>
<td>Pipe register</td>
</tr>
<tr>
<td>Chest voice</td>
<td>Medium</td>
<td>Falsetto voice</td>
<td>Flute</td>
<td></td>
</tr>
<tr>
<td>Long-reed</td>
<td>Mid voice</td>
<td>Short-reed</td>
<td>Whistle</td>
<td></td>
</tr>
<tr>
<td>Site grave</td>
<td>Site moyen</td>
<td>Site aigu</td>
<td>Rayon élevè</td>
<td></td>
</tr>
</tbody>
</table>

A particular mode of vocal-fold vibration is usually confined within a pitch range. Zemlin says that when an individual reaches the upper limits of his normal pitch range, the mode of vocal-fold vibration may be modified. He states:

This modification of the mode of vocal-fold vibration may be regarded as an operational definition of voice register. Thus, as a person transcends the limits of a particular vocal register, the voice may undergo an abrupt modification of quality. This vocal quality is often the primary characteristic of voice register. (Zemlin 1968: 193)
Also, according to Zemlin (op.cit.: 206-9), besides normal or acceptable vocal quality, there are three types of unacceptable vocal quality: breathiness (incomplete blockage during the closed phase results in a continuous flow of air during the entire vibratory cycle), harshness (irregular vocal-fold vibration), and hoarseness (combination of the features harshness and breathiness).

1.2. Register vs. Contour
When Pike discusses the types of tone languages, he defines registers as contrastive level phonemes. A language can have two, three, or four registers. The labels for two-, three-, and four-register systems are as follows:

<table>
<thead>
<tr>
<th>Language A</th>
<th>Language B</th>
<th>Language C</th>
</tr>
</thead>
<tbody>
<tr>
<td>high —</td>
<td>high —</td>
<td>high —</td>
</tr>
<tr>
<td></td>
<td>mid —</td>
<td>norm —</td>
</tr>
<tr>
<td>low —</td>
<td>low —</td>
<td>low —</td>
</tr>
</tbody>
</table>

Thus, a register tone language is a tonal language that has a register-tone system, and a contour tone language is the one in which gliding tonemes are basic to the system (Pike 1948: 5-9).

1.3. Designative register
Register can also be regarded as part of tone of voice. In some languages, changes of register may be used to express different emotional states and attitudes of the speaker. The same register might not carry the same affective indices in different cultures (Abercrombie 1967: 101). This is a paralinguistic use of register.

1.4. Register vs. Tone
Register as used by Henderson (1952) is a phonological concept. It is a cover term not only for laryngeal activity but also for a cluster of activities in the vocal tract. She states:

The Cambodian ‘registers’ differ from tones in that pitch is not the primary relevant feature. The pitch ranges of the two registers may sometimes overlap, though what I shall call the Second Register tends to be accompanied by lower pitch than the First Register. The characteristics of the first register are a ‘normal’ or ‘head’ voice quality, usually accompanied by a relatively high pitch. The characteristics of the second register are a deep rather breathy or ‘sepulchral’ voice, pronounced with lowering of the larynx, and frequently accompanied by a certain dilation of the nostrils. Pitch is usually lower than that of the first register in similar contexts. The register of a syllable is closely bound up with the vowel nucleus of that syllable, the two being mutually interdependent in a way that will be shown hereafter.
In sentences the word registers are modified according to intonation and by emotional factors. Register may be used, as in many other languages, to express emotion, and when this happens the emotional register may overlie the lexical register, much as in many tone-languages intonation may overlie lexical tone. (Henderson 1952: 151-2).

This new concept of register which was introduced into the field of South-East Asian linguistics by Henderson was adopted by Shorto (1966) and also by linguists of later generations, including myself. Abercrombie established the term. In his book *Elements of general phonetics*, besides mentioning Henderson's work on Cambodian, he also points out (1967: 101-20) that Gujarati, Danish, some dialects of Scots Gaelic, and various West African languages make use of register contrast. He finally concludes: 'It is to be expected that future research will disclose many more examples of the linguistic use of register.' (op. cit., 102). In this sense, a register language may be defined as a language that has a lexically contrastive register complex (a combination of vowel quality, pitch, phonation type, etc.), whereas a tone language has lexically contrastive pitch.3

2. Brief sketch of Chong phonology

Consonant system

- **Initial consonants**
  - p t c k ? ph th ch kh b d m n ə s h w r l j
- **Consonant clusters**
  - pr tr kr phr khr pl kl phl khl ml mr kw
- **Final consonants**
  - p t c k ? h m n ə j

Vowel system

- **Monophthongs**
  - i e ə w y a u o o
- **Diphthongs**
  - iə iə wə un yə a ui oo oo o

Register system

- **Static registers**
  - R₁ (clear voice, higher pitch, more open or on-gliding vowel)
  - R₃ (breathy voice, lower pitch, raised vowel)
- **Dynamic registers**
  - R₂ (clear-creaky voice, high-falling pitch, more open vowel)
  - R₄ (breath-creaky voice, low-falling pitch, raised vowel)

The co-occurences of registers with initial consonants, final consonants and vowels are given in Charts 1, 2 and 3 below.

---

3. Diffloth does not like this definition. He thinks that a register language should be defined as 'a language that has contrastive phonation type' (personal communication). Certainly, this definition is more specific. In my view it is also problematic. Both pitch and phonation type can be heard clearly in all register languages that I have come across. Without doing perception testing, I do not think that we can make a definite claim. Native speakers of register languages might hear both or more phonetic features at the same time.
| p | t | c | k | ph | th | ch | kh | b | d | m | n | s | h | w | r | l | j | pr | kr | phr | khr | pl | kl | phl | khl | ml | mr | kw |
| R1 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| R2 | + | + | + | + | + | + | + | - | - | + | + | + | + | + | - | - | + | + | + | + | - | - | - | - | - | - | - |
| R3 | + | + | + | + | - | - | - | - | - | - | + | + | + | - | - | + | + | + | - | - | + | - | - | + | + | + | + |
| R4 | + | + | + | + | - | - | - | - | - | - | + | + | + | - | - | + | + | + | - | - | - | - | - | + | - | + | + |

Chart 1: Co-occurrence of registers with initial consonant.

<table>
<thead>
<tr>
<th>p</th>
<th>t</th>
<th>c</th>
<th>k</th>
<th>h</th>
<th>m</th>
<th>n</th>
<th>s</th>
<th>w</th>
<th>j</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>R2</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>R3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>R4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Chart 2: Co-occurrence of registers with final consonants.

| i | e | e | w | y | a | u | o | i | e | e | w | y | a | u | o | i | e | e | w | y | a | u | o | i | e | e | w | y |
| R1 | + | + | + | + | - | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| R2 | + | - | + | + | + | + | - | + | + | + | + | + | + | + | + | + | - | - | - | - | - |
| R3 | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| R4 | + | + | + | + | + | - | - | + | + | + | + | + | + | + | + | + | + | + | - | - | - | - |

Chart 3: Co-occurrence of registers with vowels.
3. Acoustic measurements
In this section, an acoustic analysis of Chong vowels in respect of the formant frequency, power spectra, fundamental frequency, duration, and overall intensity will be given. The results of the measurements will indicate the characteristics of the register complex in Chong. Wide band spectrograms were used for measuring formant frequencies and power spectra, and narrow band spectrograms for fundamental frequencies.

3.1. Formant frequency
The frequencies of $F_1$ and $F_2$ of the following 17 clear vowels, 15 clear-creaky vowels, 18 breathy vowels, and 15 breathy-creaky vowels were measured:

<table>
<thead>
<tr>
<th>Clear ($R_1$)</th>
<th>Clear-creaky ($R_2$)</th>
<th>Breathy ($R_3$)</th>
<th>Breathy-creaky ($R_4$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>e</td>
<td>–</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
<td>e</td>
<td>–</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>o</td>
<td>–</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>o</td>
<td>o</td>
<td>–</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>ii</td>
<td>ii</td>
</tr>
<tr>
<td>ee</td>
<td>ee</td>
<td>ee</td>
<td>ee</td>
</tr>
<tr>
<td>ee</td>
<td>ee</td>
<td>ee</td>
<td>ee</td>
</tr>
<tr>
<td>uu</td>
<td>uu</td>
<td>uu</td>
<td>–</td>
</tr>
<tr>
<td>yy</td>
<td>yy</td>
<td>yy</td>
<td>yy</td>
</tr>
<tr>
<td>aa</td>
<td>aa</td>
<td>aa</td>
<td>aa</td>
</tr>
<tr>
<td>uu</td>
<td>uu</td>
<td>uu</td>
<td>uu</td>
</tr>
<tr>
<td>oo</td>
<td>oo</td>
<td>oo</td>
<td>oo</td>
</tr>
<tr>
<td>oo</td>
<td>oo</td>
<td>oo</td>
<td>oo</td>
</tr>
</tbody>
</table>

Five test words said in isolation by the two male speakers (MA, MB) were used for each vowel: there were altogether 650 ($65 \times 5 \times 2$) test tokens. The mean values of $F_1$ and $F_2$ were plotted separately for each speaker on vowel charts (see Figures 1, 2, 3 and 4).

In general, the results of the measurements confirm what can be perceived auditorily: breathy and breathy-creaky vowels are higher than clear and clear-creaky vowels. The two male speakers do not seem to diphthongise their clear vowels. This finding supports Gregerson’s (1976) hypothesis that in most Mon-Khmer languages first register (clear) vowels which are produced with retracted tongue-root are always more open than second register (breathy) vowels which are produced with advanced tongue-root. The four vowel formant charts exhibit obvious patterns: clear voice and semi-clear voice vowels of the first and second registers are more open than the breathy voice and semi-breathy voice vowels of the
An instrumental study of Chong registers

Fig. 1: Formant frequencies ($F_1$ and $F_2$) of 1st, 2nd, 3rd and 4th register short vowels (speaker MA)
+ clear vowel ($R_1$)
• clear-creaky vowel ($R_2$)
▲ breathy vowel ($R_3$)
■ breathy-creaky vowel ($R_4$)

Fig. 2: Formant frequencies ($F_1$ and $F_2$) of 1st, 2nd, 3rd and 4th register long vowels (speaker MA)
+ clear vowel ($R_1$)
• clear-creaky vowel ($R_2$)
▲ breathy vowel ($R_3$)
■ breathy-creaky vowel ($R_4$)
Fig. 3: Formant frequencies ($F_1$ and $F_2$) of 1st, 2nd, 3rd and 4th register short vowels (speaker MB)

+ clear vowel ($R_1$)  
△ breathy vowel ($R_3$)  
● clear-creaky vowel ($R_2$)  
■ breathy-creaky vowel ($R_4$)

Fig. 4: Formant frequencies ($F_1$ and $F_2$) of 1st, 2nd, 3rd and 4th register short vowels (speaker MB)

+ clear vowel ($R_1$)  
△ breathy vowel ($R_3$)  
● clear-creaky vowel ($R_2$)  
■ breathy-creaky vowel ($R_4$)
third and fourth registers. However, it is still too early to accept Gregerson’s claim. Although it works well for Chong, unfortunately it cannot explain the register phenomena in Nyah Kur and Kui (Thongkum 1982, 1985).

3.2. Power spectra
Kirk and others point out that the power spectra enable phoneticians to quantify the relative amount of energy in different harmonics. For their study of phonation types in Jalapa Mazatec, they measured the difference in dB between the intensity of the fundamental and the intensity of the first formant. They conclude, ‘There is considerable variation from speaker to speaker in the three phonation types; but for each speaker on this measure the value for creaky voice is less than that for modal voice, and the value for modal voice is less than that for breathy voice’ (Kirk et al. 1984: 109). Following their recommendations I did the same measurements for Chong. For the measurement of power spectra, each of the two male speakers said 325 test words. The results of the measurements may be summed up as in Table 1:

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Number of test token</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>40</td>
<td>-8.09</td>
<td>2.01</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>-8.35</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>-8.27</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>-8.61</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>-9.18</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>-9.28</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>-8.86</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>-8.95</td>
<td>1.58</td>
</tr>
</tbody>
</table>

It is unfortunate that the results do not meet my expectation; i.e. the measure does not seem to separate out successfully in that the value for modal voice is higher than that for creaky voice and less than that for breathy voice. Why? The reasons that I can think of are as follows:

(i) The measurements were done by hand because I did not have a Kay digital sound spectrograph, and so they could be less accurate.

(ii) As stated by Kirk (op. cit. 109), ‘This measure can be used for comparing phonation types only in cases in which the vowels being
compared have similar formant frequencies; as the relative intensity of each formant is a function of its frequency. It is true that in Chong the vowels of each register have specific quality as described in 4.1 above. Moreover, \( R_2 \) and \( R_4 \) vowels have dynamic or combined phonation types: clear followed by creaky phonation and breathy followed by creaky phonation; this can actually cause problems for the measure.

3.3. **Fundamental frequency (\( F_0 \))**

The word list used for \( F_0 \) measurement consisted of 71 minimal or analogous sets; for example,

Set 20

<table>
<thead>
<tr>
<th>Set 20</th>
<th>( puy^1 )</th>
<th>( pum^2 )</th>
<th>( puy^3 )</th>
<th>( puy^4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'to get pregnant'</td>
<td>'rash'</td>
<td>'the entrails of animals'</td>
<td>'water melon'</td>
</tr>
</tbody>
</table>

284 meaningful items were said by each of the two female speakers (FA, FB). Narrow band spectrograms were made and measured at ten points starting from the onset to the end of vowel. For plotting the results of \( F_0 \) measurements, the 284 test words were divided into 20 sets based on register differences and different types of syllable:

<table>
<thead>
<tr>
<th>Clear</th>
<th>Clear-creaky</th>
<th>Breathy</th>
<th>Breathy-creaky</th>
<th>Number of Test Token</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVN</td>
<td>CVN</td>
<td>CVN</td>
<td>CVN</td>
<td>40 (10 \times 4)</td>
</tr>
<tr>
<td>CVS</td>
<td>CVS</td>
<td>CVS</td>
<td>CVS</td>
<td>16 (4 \times 4)</td>
</tr>
<tr>
<td>CVVN</td>
<td>CVVN</td>
<td>CVVN</td>
<td>CVVN</td>
<td>100 (25 \times 4)</td>
</tr>
<tr>
<td>CVVS</td>
<td>CVVS</td>
<td>CVVS</td>
<td>CVVS</td>
<td>88 (22 \times 4)</td>
</tr>
<tr>
<td>CVH</td>
<td>CVH</td>
<td>–</td>
<td>–</td>
<td>12 (6 \times 2)</td>
</tr>
<tr>
<td>CV^p</td>
<td>–</td>
<td>CV^p</td>
<td>–</td>
<td>8 (4 \times 2)</td>
</tr>
</tbody>
</table>

The mean values of \( F_0 \) (in Hz) are shown in Figures 5 and 6. From the data presented there, the following points can be made:

(i) In CVN, CVS, CVVN and CVVS types of syllable, clear-creaky (\( R_2 \)) vowels have the highest fundamental frequency. Only the clear part of \( R_2 \) and \( R_4 \) vowels could be measured because of the irregularities of the creaky part. However, it is possible to claim that \( R_2 \) and \( R_4 \) vowels have high rise-fall and high fall \( F_0 \) contours, respectively. The insertion of laryngealisation or creakiness seems to be the cause of falling \( F_0 \) contour. This may be an explanation of how falling tones are acquired in some tone languages. The narrow band spectrograms of sets 5, 16, 17, 18 and 32 (speaker FB) illustrate the absence of creaky phonation. Thus, clear-creaky (\( R_2 \)) vowels become clear vowels with higher fall \( F_0 \) contour, and breathy-creaky (\( R_4 \)) vowels become breathy-clear vowels with lower fall \( F_0 \) contour (see Figure 7 below).

\[4. N:\] represents nasals and semi-vowels (\( m\ n\ p\ \eta\ w\ j\ )); S: stops (\( p\ t\ c\ k\ )); H: glottal fricative (\( h\ )); and \( ?\): glottal stop (\( ?\ )), respectively.
An instrumental study of Chong registers

Fig. 5  Mean $F_0$ values (in Hz) of 1st, 2nd, 3rd and 4th register vowels in 6 syllable types (speaker FA)
- - - - clear vowel (R1)
----- clear-creaky vowel (R2)
-- -- breathy vowel (R3)
- - - - breathy-creaky vowel (R4)
Fig. 6  Mean $F_0$ values (in Hz) of 1st, 2nd, 3rd and 4th register vowels in 6 syllable types (speaker FB)

- clear vowel (R₁)
- breathy vowel (R₃)
- clear-creaky vowel (R₂)
- breathy-creaky vowel (R₄)
An instrumental study of Chong registers

![Diagram](image)

**Fig. 7:** Mean $F_0$ values (in Hz) of 2nd and 4th register vowels in CVVS syllable type when creaky phonation disappears and is replaced by falling $F_0$ contour (speaker FB)

(ii) Although breathy vowels can be perceived auditorily as having the lowest pitch, their $F_0$ does not start low at all. According to the tracings (from 0% up to 30% of the vowel duration) breathy vowels can have even higher $F_0$ than clear vowels. The point where the higher pitch of clear vowels and the lower pitch of breathy vowels can be differentiated is the $F_0$ from 30% up to the end of vowel duration.

(iii) There is a tendency for $R_1$, $R_2$, $R_3$ and $R_4$ vowels in CVN and CVS types of syllables to have higher $F_0$ than those in CVVN and CVVS syllable types.

(iv) Only clear and breathy vowels can occur in CVH and CV? types of syllable. In CVH syllable structure, both types of vowel have rising $F_0$ contour; however, breathy vowels do not always have lower $F_0$ than clear vowels. For speaker FA, the $F_0$ tracings of clear and breathy vowels do not exhibit much difference; i.e. both types of vowel have rising $F_0$ contour. In CV? syllable structure, clear vowels have rising $F_0$ contour with an abrupt fall at the end, whereas breathy vowels have falling $F_0$ contour.

3.4. **Duration**

The word-list used for $F_0$ measurement was used again for the measurement of vowel duration and overall intensity. The test words were said by two female speakers of Chong. The audio recordings used as input were made in the recording studio of the Linguistics Research Unit (LRU) at the Faculty of Arts, Chulalongkorn University, and the results
of the measurement can be found in Tables 3 and 4.

It is noticeable that breathy-creaky ($R_4$) vowels seem to be shorter than the other kinds of vowels no matter what the types of syllable. However, there is one exception; i.e. in CVN syllable type, $R_4$ vowels are somewhat longer.\(^5\)

I have earlier commented that differences in duration caused by differences in phonation types might not be important when the language in question possesses distinctive vowel length (Thongkum 1985: 12). The results of the measurements of Chong vowels give me more confidence. Although breathy or murmured vowels in Gujarati (Fischer-Jørgensen 1977) and Jalapa Mazatec (Kirk et al. 1984) have longer duration than clear vowels, we still cannot claim that it is a universal phonetic characteristic.

Regarding clear-creaky ($R_2$) and breathy-creaky ($R_4$) vowels, the creaky part of $R_4$ vowels is longer than that of $R_2$ vowels, although the overall duration of $R_4$ vowels is shorter than that of $R_2$ vowels. This quality can also be perceived auditorily. The proportion (in percentage) of the duration of the first and second parts of vowel pertaining to dynamic registers ($R_2$ and $R_4$) is given in Table 2 below.

Table 2: \textit{Mean duration (in percentage) of the clear and creaky parts of the 2nd register vowels and of the breathy and creaky parts of the 4th register vowels.}

<table>
<thead>
<tr>
<th>Register</th>
<th>SPEAKER FA</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CVN</td>
<td>CVS</td>
<td>CVVN</td>
<td>CVVS</td>
</tr>
<tr>
<td>$R_2$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
</tr>
<tr>
<td>$R_4$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Register</th>
<th>SPEAKER FB</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CVN</td>
<td>CVS</td>
<td>CVVN</td>
<td>CVVS</td>
</tr>
<tr>
<td>$R_2$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
</tr>
<tr>
<td>$R_4$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
</tr>
</tbody>
</table>

5. Since the segmentation was done by hand without any computer aids, defects may have been caused by inaccurate segmenting. The beginning part of laryngealised or creaky voice nasals and semi-vowels could have been segmented as part of clear-creaky voice and breathy-creaky voice vowels.
3.5. Intensity

Regarding overall intensity, clear vowels and clear-creaky vowels seem to have higher amplitude than breathy and breathy-creaky vowels. The loss of intensity is due to leaking glottis during breathy phonation (Fischer-Jørgensen 1977: 119), and the acoustic energy is lost by the damping effect of the general relaxation of the muscles of the whole vocal system in lax voice (Laver 1980: 135). The intensity curves of clear and breathy vowels look similar (i.e. look more bell-shaped), whereas those of clear-creaky and breathy-creaky vowels look different (i.e. more cone-shaped), which is due to a sudden drop of intensity when clear and breathy vowels become creaky caused by the abrupt closing of the vocal folds. The results of the measurements are given in Tables 3 and 4.

1-5 = overall vowel duration
1-3 = distance from the onset of vowel to the peak of intensity curve
3-5 = distance from the peak of intensity curve to the end of vowel
1-a = amplitude at the onset of vowel
2-b = amplitude at the half distance of the onset of vowel and the peak of intensity curve
3-c = peak of amplitude
4-d = amplitude at the half distance of the peak amplitude and the end of vowel
5-e = amplitude at the end of vowel

3.6. Manifestation of \(R_2\) and \(R_4\) vowels

In the recorded material, clear-creaky (\(R_2\)) and breathy-creaky (\(R_4\)) vowels behave in many different ways. The following are the most obvious manifestations of clear-creaky and breathy-creaky vowels:

(i) Complete change of the oscillatory pattern in the middle or at \(\frac{3}{4}\) of vowel duration;
(ii) Jagged oscillations into the following consonant;
(iii) Brief change of waveforms in the middle of vowel with a decrease in intensity and fundamental frequency;
(iv) Slight change of waveforms which indicates laryngealisation or creakiness throughout the vowel, not found very often.

---

6. Danish, it is maintained, is a register language: 'In Danish two such words as hun "she" and hund "dog" are pronounced alike except for a difference of register, the second having creaky voice.' (Abercrombie 1967: 101).

Peterson (1973) did an instrumental investigation of the Danish 'stød', which is regarded as a phonologically distinctive element. It is amazing to see that the results of my instrumental study of clear-creaky and breathy-creaky vowels in Chong more or less agree with these of Petersen who also refers to another instrumental investigation of the stød which was done by Smith (1944). Smith, who included electromyographic, oscillographic and kymographic registrations, thinks that it is a stress accent, concluding that it often appears to be three-phased:

(1) A ballistic contraction of the expiratory muscles;
(2) Cessation of this activity which causes a lack of balance in the reaction of the vocal folds;
(3) A new activity in the expiratory muscles.
Table 3: Mean amplitude, mean duration, and standard deviation of 1st, 2nd, 3rd and 4th register vowels in six types of syllable (speaker FA)

<table>
<thead>
<tr>
<th>Syllable type</th>
<th>Register</th>
<th>Amplitude (dB)</th>
<th>Duration (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1-a</td>
<td>2-b</td>
</tr>
<tr>
<td>R1</td>
<td></td>
<td>23.52</td>
<td>36.32</td>
</tr>
<tr>
<td>R2</td>
<td>(C)V(V(N)</td>
<td>17.20</td>
<td>38.48</td>
</tr>
<tr>
<td>R3</td>
<td>(C)V(S)</td>
<td>21.08</td>
<td>33.12</td>
</tr>
<tr>
<td>R4</td>
<td>(C)V(N)</td>
<td>19.29</td>
<td>33.17</td>
</tr>
<tr>
<td>R5</td>
<td>(C)V(H)</td>
<td>19.50</td>
<td>32.40</td>
</tr>
<tr>
<td>R6</td>
<td>(C)V(?)</td>
<td>19.15</td>
<td>32.40</td>
</tr>
</tbody>
</table>

Mean amplitude (dB) Syllable type
- R1: (5.34) (3.63) (3.59) (3.73) (4.67) (10.45) (10.02) (1.87)
- R2: (7.71) (3.55) (3.27) (3.89) (6.42) (6.19) (6.42) (2.11)
- R3: (6.35) (4.09) (3.43) (3.51) (3.52) (6.00) (5.62) (2.17)
- R4: (5.62) (4.76) (3.98) (5.91) (5.54) (4.45) (4.87) (2.81)

Mean duration (msec)
- R1: 28.50 34.75 38.50 39.00 17.25 453 090 143
- R2: (4.09) (4.60) (3.77) (5.09) (11.03) (2.27) (1.87) (0.83)
- R3: (13.75) (3.10) (8.50) (29.75) (11.25) (0.72) (103) (175)
- R4: (8.32) (2.23) (3.5) (2.05) (6.83) (2.58) (2.58) (0.5)
- R5: (5.85) (6.08) (4.84) (3.70) (4.12) (3.63) (1.58) (2.58)
- R6: (19.25) (26.50) (32.00) (27.00) (6.50) (0.40) (0.88) (128)
- R7: (6.83) (2.60) (4.74) (4.58) (8.17) (1.87) (1.79) (1.09)
- R8: (20.17) (37.17) (40.17) (34.00) (19.67) (0.80) (132) (212)
- R9: (9.5) (4.59) (5.75) (5.36) (7.34) (2.58) (1.34) (1.57)
- R10: (11.33) (35.83) (37.00) (29.83) (10.83) (1.02) (117) (229)
- R11: (7.09) (4.71) (5.42) (4.71) (8.13) (2.03) (2.43) (1.57)
- R12: (29.00) (39.75) (42.50) (40.50) (31.25) (0.98) (0.87) (185)
- R13: (5.20) (1.48) (2.69) (2.18) (5.36) (0.83) (0.83) (0.50)
- R14: (16.75) (35.25) (38.50) (35.50) (27.00) (0.82) (0.98) (180)
- R15: (10.52) (2.38) (2.06) (2.60) (4.74) (2.86) (3.70) (1.22)
### Table 4: Mean amplitude, mean duration, and standard deviation of 1st, 2nd, 3rd and 4th register vowels in six types of syllable (speaker FB)

<table>
<thead>
<tr>
<th>Syllable type</th>
<th>Register</th>
<th>Amplitude (dB)</th>
<th>Duration (msec)</th>
<th>Standard Deviation (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>1-a</td>
<td>11.00 (6.58)</td>
<td>141 (3.65)</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>2-b</td>
<td>24.92 (3.19)</td>
<td>150 (6.69)</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>3-c</td>
<td>27.36 (3.54)</td>
<td>105 (6.69)</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>4-d</td>
<td>25.00 (3.37)</td>
<td>(4.28)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>5-e</td>
<td>17.72 (3.73)</td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
<td></td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>3-5</td>
<td></td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td></td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td>R2</td>
<td>1-a</td>
<td>9.96 (6.31)</td>
<td>105 (6.69)</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>2-b</td>
<td>26.72 (6.28)</td>
<td>105 (6.69)</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>3-c</td>
<td>31.32 (5.24)</td>
<td>105 (6.69)</td>
<td>267</td>
</tr>
<tr>
<td></td>
<td>4-d</td>
<td>22.48 (7.76)</td>
<td>(4.28)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>5-e</td>
<td>14.92 (6.05)</td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
<td></td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>3-5</td>
<td></td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td></td>
<td>(4.19)</td>
<td>(2.78)</td>
</tr>
<tr>
<td>R3</td>
<td>1-a</td>
<td>10.28 (5.18)</td>
<td>96 (4.71)</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>2-b</td>
<td>21.28 (3.70)</td>
<td>96 (4.71)</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>3-c</td>
<td>24.40 (3.05)</td>
<td>96 (4.71)</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>4-d</td>
<td>21.44 (3.47)</td>
<td>(4.19)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>5-e</td>
<td>13.88 (5.17)</td>
<td>(4.19)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
<td></td>
<td>(4.19)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>3-5</td>
<td></td>
<td>(4.19)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td></td>
<td>(4.19)</td>
<td>(2.50)</td>
</tr>
<tr>
<td>R4</td>
<td>1-a</td>
<td>9.48 (5.55)</td>
<td>108 (4.73)</td>
<td>261</td>
</tr>
<tr>
<td></td>
<td>2-b</td>
<td>22.08 (3.78)</td>
<td>108 (4.73)</td>
<td>261</td>
</tr>
<tr>
<td></td>
<td>3-c</td>
<td>28.52 (3.77)</td>
<td>108 (4.73)</td>
<td>261</td>
</tr>
<tr>
<td></td>
<td>4-d</td>
<td>16.16 (5.61)</td>
<td>(4.73)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>5-e</td>
<td>12.68 (4.75)</td>
<td>(4.73)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>1-3</td>
<td></td>
<td>(4.73)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>3-5</td>
<td></td>
<td>(4.73)</td>
<td>(2.50)</td>
</tr>
<tr>
<td></td>
<td>1-5</td>
<td></td>
<td>(4.73)</td>
<td>(2.50)</td>
</tr>
</tbody>
</table>

- **(C)VV(N)**
- **(C)VV(S)**
- **(C)V(N)**
- **(C)V(S)**
- **(C)V(H)**
- **(C)V(?)**
4. Notes on laryngeal waveforms and airflow
For the investigation of laryngeal vibration and airflow, two female speakers of Chong were brought to the Phonetics Laboratory at Chulalongkorn University. No quantitative measurement was attempted. However, it can be seen clearly that the laryngograms of the four sets of Chong vowels look different. At the onset of clear and clear-creaky vowels, there is a relative rise which is produced by the rapid closing of the vocal folds for normal voicing, and on the other hand, there is a relative fall at the onset of breathy and breathy-creaky vowels which is caused by a more open glottis during breathy phonation. There is also a relatively sharp rise after the release of final consonant following clear-creaky and breathy-creaky vowels.

Regarding airflow, the most prominent characteristic of breathy vowels is strong airflow. During the creaky part of R₂ and R₄ vowels, there is a sudden drop of airflow caused by a rapid closing and tightening of the vocal folds.

6. Conclusion
The instrumental investigation of Chong vowels supports what can be perceived auditorily: that is, there are four sets of register complex, namely, R₁, R₂, R₃, and R₄. The phonetic correlates of the four registers may be summed up as follows:

*Register 1:* higher frequency of F₁ (more open vowel); rather level F₀ contour (level pitch); regular audio and laryngeal waveforms; higher amplitude; bell-shaped intensity curve; lower amount of airflow in comparison with R₃; normal voice phonation;

*Register 2:* higher frequency of F₁ (more open vowel); high rise-fall F₀ contour; regular followed by irregular audio and laryngeal waveforms; highest peak of amplitude; cone-shaped intensity curve; sudden decrease of intensity and fundamental frequency; sudden drop of airflow; normal voice followed by creaky voice phonation;

7. In normal voicing there are three distinct parts of the waveform. First, there is a relatively sharp rise which is produced by the rapid closing of the vocal folds which is so characteristic of their normal vibration and is associated with the interval of greatest acoustic excitation of the vocal tract. Second, there is the more gradual fall which is associated with the parting of the vocal folds as the sub-glottal pressure is increased; and third, there is the flatter base of the waveform which corresponds to the interval during which the glottis is open and the vocal folds are out of contact.

(Fourcin & Abberton 1972: 165-66)

8. Fischer-Jørgensen measured the airflow of clear and murmured or breathy vowels in Gujarati by means of aerometer. She concludes (1967: 153) that murmured vowels are characterised by a strong airflow which is due to the presence of a small opening in the rear part of the glottis, and/or by an increased activity of the expiratory muscles.

9. When a laryngealised or creaky voice sound is made, the arytenoids are held tightly together while a small length of the ligamental vocal cords is vibrating (Ladefoged 1971: 8).
An instrumental study of Chong registers

Register 3: lower frequency of $F_1$ (more closed vowel); gradual fall $F_0$ contour; regular audio and laryngeal waveforms; lowest amplitude; bell-shaped intensity curve; strong air flow; breathy voice phonation;

Register 4: lower frequency of $F_1$ (more closed vowel); high fall $F_0$ contour; regular followed by irregular audio and laryngeal waveform; rather low amplitude; cone-shaped intensity curve; sudden decrease of intensity and fundamental frequency; sudden drop of airflow: breathy voice followed by creaky voice phonation.

REFERENCES


159
Lee, T.


Martin, Marie A.


Mörner, M., Fransson, F. & Fant, G.


Petersen, Pia R.


Pike, K.L.


Shorto, H. L.


Suphanphaiboon, Surekha


Thomas, D., & Headley, R. K.


Thongkum, Theraphan L.


Zemlin, W.R.

KEEPI NG THINGS U P FRONT: 
ASPECTS OF INFORMATION PROCESSING 
IN MAL DISCOURSE STRUCTURE

David Filbeck

0. Introduction
Mal is a Mon-Khmer language spoken in Pua and Chiang Klang Districts of Nan Province, Thailand. It is also one of the many dialects spoken by the Tin tribal people who are located in this part of Thailand and across the border into Laos (Filbeck 1978). The number of Mal speakers probably does not exceed 6-7000 people but an accurate census is not available since the Thai government includes all Tin people, regardless of dialect, in a single census.

The purpose of this paper is to describe discourse-level phenomena found in the Mal language. Related phenomena occur in other Tin dialects, but it appears that only in Mal are they so extensively used. Since Mal is the most conservative of the Tin dialects, retaining more of Proto-Tin than other dialects (Filbeck 1978), we may presume that these phenomena were part of Proto-Tin as well. However, there is internal evidence showing that at least some of these discourse features are developments unique to Mal. For example, the emergence of a rising tone in Mal (Filbeck 1972) has evidently been used to signal at least one discourse feature in Mal. (See discussion in 3. below.)

The particular viewpoint taken here in describing these aspects of Mal discourse structure is information processing within a discourse.\(^1\) As a descriptive approach to language, information processing differs from the more traditional linguistic approach. The latter views language as composed of different hierarchies or systems, e.g. phonology, morphology, semantics, and discourse, each of which is then described as a more or less autonomous part, and when all the hierarchies or systems have been described separately for a particular language, the task of the linguist is considered to be complete.

While this 'divide and conquer' approach has been productive, it has also had its costs. Logically such hierarchies or systems appear to exist, and in the way that linguists say they exist. However, in terms of process, not everything that linguists claim to exist—or to exist in the way that they claim—plays a part in the main function of language, viz. communicating or transmitting meaning. Not everything is needed or used to process, i.e.

---

1. At this stage, information processing should not be confused with information theory. The latter is basically a form of statistical theory dealing with the number of alternatives available in a code and the results of choosing one such alternative. On the other hand, as will be seen in various parts of the following discussion, information processing does use a number of insights drawn from the more statistically oriented information theory.
encode and decode/interpret, meaning by means of language. At least, not everything is used at all times. In short, processing in language follows its own 'rules' which, though drawing upon the traditional categories of linguistics, describe language from a different perspective and with different goals in mind.

Information processing—viewed here in perhaps a narrower sense than that which is found in the literature—is a specialised part of the encoding/decoding aspect of language. On the one hand, there is content that is encoded in a text, whether written or oral, for transmission and decoding. Yet, content is not just encoded and decoded; it is also 'processed' in various ways by means of linguistic structure—more precisely, the categories used in linguistic description. The linguistic structure of a language, therefore, plays a 'mediating role' in the encoding, transmitting, and decoding, of meaning. In this view of linguistic structure, processing may be considered a part of the total semantic structure of language, or, to state it another way, meaning as transmitted by language is composed of content plus processing.2

Content is processed in two ways. First, it is distributed or mapped linearly along a syntactic code. For example, content may be 'tightly packed' into a few sentences (as a technical exposition), or it may be distributed over many sentences (as in popular exposition). Second, content may be processed globally as theme, prominence, topic, focus, old and new information, etc., in the course of a text or discourse. In either way, what is processed is additional information 'about' the content that is similarly encoded in the code. Moreover, just as the speaker processes content for additional information while encoding content, so the hearer processes the incoming code for both content and information in the act of decoding. That is, the hearer (or reader of a text) comes to certain judgements about the content he or she has received on the basis of the code, thus providing (additional) information 'about' the content.

For Mal, this paper discusses information processing in this second way only. More specifically, Mal employs various linguistic means to signal and/or differentiate—i.e. process—various types of information flowing globally through discourse. As will be seen in the following discussion, such processing forms a crucial dimension of the Mal language. A description of such, then, is necessary for a full linguistic description of Mal. Moreover, as a by-product, it is hoped that this paper will make a contribution to describing the nature of discourse phenomena in Mon-Khmer in general.3

Data for this paper are taken from both recorded texts that I have collected and from the recently completed translation of St. Mark's

2. Or, in terms of information theory: content, plus information about content, constitutes the message of a text.
3. Some speak also of a pragmatic level of information processing (qv. Smith & Wilson 1979), where information is processed according to shared knowledge, assumptions, or beliefs, but not necessarily according to the characteristic of the linguistic code itself. This pragmatic aspect of information processing is also not discussed in this paper.
Gospel into Mal. Examples will also be taken from the first two chapters of the Book of Genesis, which is currently being translated into Mal. Illustrations from the latter two sources are especially revealing in that they show how productive discourse-level phenomena are in Mal. Greek, English and Thai obviously do not have the same type of formal means for dealing with such phenomena. Consequently, it was fascinating to observe the translator (translating from Thai into Mal in this case) determine from the context how best to structure the text formally according to Mal discourse structure.

1. A discourse characteristic
For this paper, a deductive approach, i.e. from the ‘top down’, to informational processing in Mal is assumed. That is, we begin by observing a general characteristic—a ‘flavour’—typical of Mal discourse, whether interpersonal communication or story-telling. When we have established what this discourse characteristic is, we will then illustrate how it is realised formally or linguistically in discourse.

In analysing the following data we can observe a general ‘theme’ operating throughout the discourse structure of Mal, namely, the theme of inclusivity vs. exclusivity, or ‘this/that-including-others’ vs. ‘this/that-excluding-others’. Of course, inclusivity dominates discourse activity in Mal; on the other hand, exclusivity can be forcefully and unambiguously marked with formal linguistic mechanisms. That it is so marked quite often gives Mal discourse a distinct character.

In this regard, then, we may classify exclusivity as the ‘marked’ process and inclusivity as the ‘unmarked’ process. Also, a text may be ‘more or less’ highly marked according to the number of marked processes it contains. Moreover, when considering the Mal language from a discourse perspective, we can observe that Mal is a comparatively ‘highly marked’, i.e. more complex, language with respect to the inclusive-exclusive processing of information. For, characteristically, this distinction is a real option in everyday speech, and it is an alternative that is chosen often enough in interpersonal communication as well as in other forms of discourse.

The communicative function of this inclusive/exclusive dichotomy in Mal is, as the title of this paper suggests to ‘keep things up front’, i.e. to keep track in a formal way of what is and is not important during the flow of discourse. It is a formal way of processing information, both for the speaker and hearer, in communication, whereby a piece of content is ‘kept up front’, or made prominent, in discourse. Conversely other pieces of content are ‘shuttled’ to the background but, should any of these need to be made prominent, it is possible to do so by ‘bringing it up front’, as it were, by formal linguistic means.

2. Pronouns: ‘we vs. they’
The dichotomy between inclusivity and exclusivity in Mal discourse is nowhere better observed than in the pronouns, which form a well-
DAVID FILBECK

developed system of exclusive–inclusive relations.

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Dual</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ən/ ‘I’</td>
<td></td>
<td>/iaʔ/ ‘we (excl.)’</td>
<td>/ii/ ‘we (excl.)’</td>
</tr>
<tr>
<td>/mah/ ‘you’</td>
<td></td>
<td>/paaʔ/ ‘you’</td>
<td>/pee/ ‘you’</td>
</tr>
<tr>
<td>/nam/ ‘he, she’</td>
<td></td>
<td>/paam/ ‘they’</td>
<td>/ah/ ‘they’</td>
</tr>
</tbody>
</table>

The unmarked pronouns include the singular pronouns (‘I, you, he, she’) and the inclusive plural. These are used the most often in conversation and in recorded texts, while the marked pronouns—the duals and exclusives—occur more rarely in those environments. Indeed, the marked pronouns in Mai occur not just in one degree of markedness but two: marked and highly marked.

The marked pronouns are the dual and/or exclusive plurals, e.g. /iaʔ/ ‘we (dual excl.)’, /paaʔ/ ‘you (dual)’, /paam/ ‘they (dual)’ and /ii/ ‘we (pl. excl.)’. These occur often enough in conversation and in texts to lend a flavour of exclusivity to Mai communication.

Only /aaʔ/ ‘we (dual incl.)’ is a highly marked pronoun. It is classified here as an inclusive dual pronoun while /iaʔ/ is classified as its exclusive counterpart. The reason for this is that /iaʔ/ is used in third-party conversation where the listener is not included, i.e. is ‘excluded’ from the dual grouping. For example, in a recent conversation two young boys requested leave to go somewhere. One of the boys, who was the spokesman, used the pronoun /iaʔ/ to include himself and the other boy while at the same time excluding me as the listener in the request. By contrast, the other pronoun, /aaʔ/ ‘we (dual incl.)’, is used to include each partner of the dyad in the conversation. Consequently it is rarely used, being limited to conversations between two people who are alone together.

In the Mai story ‘The Magic Trail’ (which we shall refer to again in this paper), a husband and wife set out alone to seek a fortune. In the version I have recorded, the narrator relates the story using a great number of direct quotes, especially when the husband speaks to the wife. Since the couple are poor and own no livestock, they are seeking in the forest for animals which can be taken back and raised as domesticated stock in the village. One of the first creatures they find is a wild boar, at which the husband exclaims to his wife:

*oo, aaʔ hany əc!
Oh, we-two are-rich already*

He then instructs her to take the boar back to the village and pen it up. Whenever a different forest animal is found, the husband makes the same exclamation in addressing his wife, each time using the pronoun /aaʔ/. The other dual pronoun, /iaʔ/, is not the correct pronoun here as there is no third person whom the husband is addressing. Presumably, on
Information processing in Mal discourse structure

returning to the village with his wife and his entourage of newly-found animals, he could announce to the villagers that /iaŋ hän ac/ ‘we-two are now rich’, since he would be excluding the villagers, in their role as listeners, from any participation in this new-found wealth.

A major role of pronouns in language is to keep track of participants, especially on the discourse level in a way that is not excessively redundant. Or, from the standpoint of information processing, pronouns serve to process information about participants that figure in a discourse. For example, to repeat a noun or a person’s name continually in a discourse or text is tediously redundant. Substituting pronouns for such items allows redundancy to be reduced while at the same time allowing the processing of essential information regarding the participants. Obviously, then, the more complex a pronominal system the better a language is able to keep track of, or to process, information about participants in a text or discourse.

Because of a well-developed system of duals, and inclusive-exclusive relations, pronouns in Mal are able to fulfil the above role in perhaps a more refined way than that which is possible in some other languages. This can be seen through the use of the dual pronouns. Examples are found in the Mal translation of Chapter 14 of Mark’s Gospel in which Jesus selects two of his disciples to go and prepare a place for the Passover meal:

And he sent two of his disciples, and said to them (dual), ‘Go into the city, and a man carrying a jar of water will meet you (dual); follow him, and wherever he enters, say to the householder, ‘The Teacher says, Where is my guest room, where I am to eat the passover with my disciples?’ And he will show you (dual incl.) a large upper room furnished and ready; there prepare for us (pl. incl.)’ And the disciples set out and went to the city, and found it as he had told them (dual); and they (dual) prepared the passover. (Mark 14:13-16, RSV).

These instructions were originally clearly addressed to the two disciples, but of course there are no dual pronouns in either Greek or English. Consequently, the information about the ‘duality’ of the participants (i.e. the two disciples) is processed in a more refined way in Mal; once it is established in the text, dual pronouns are used thereafter in order to keep track of them. Viewing the process another way, one might say that—just as in the morphology of certain languages—in Mal discourse structure a ‘principle of informational agreement’ is in operation. That is, once two participants have been established as a ‘dual unit’ in a textual discourse, subsequent anaphoric reference to them must be kept ‘in agreement informationally’ by employing the appropriate dual pronoun(s). Use of other plural pronoun(s) would not be in agreement and would transmit incongruous information. In Mal, then, it is necessary to use dual pronouns in order to process correctly certain types of information about participants who figure in a text or discourse.
Even in the Mal story about the magic trail, mentioned above, we can see how certain information regarding the participants is processed through the use of the pronoun /aaʔ/ 'we (dual, listener inclusive)’. At the time of recording, the story was told by a narrator to a group of people. By casting the husband’s exclamation (‘We are now rich!’) to his wife in direct quotes, the narrator was able to use the pronoun /aaʔ/ in his narrative, thus transmitting to his listeners that not only were there two participants but that they were also alone together in their adventures along the magic trail. Again we can see the ‘principle of informational agreement’, mentioned above, in operation here. In the story, the actual situation of a husband and wife starting out alone on the magic trail is established immediately. Consequently, any anaphoric reference to the two of them must necessarily involve the pronoun /aaʔ/; the information transmitted by any other pronoun would not be in agreement with the established context.

Exclusive pronouns in Mal also function to process information about participants in a text or discourse. An example of this may be seen in Mark 9: 28, which tells how nine of Jesus’ disciples were unable to cast out a spirit from a little boy. Jesus did so and, after they had entered the house, the disciples asked him:

\[
ay \ phiʔ \ naa \ ii \ \ yaap \ldots ay \ poon
\]

why we (excl.) cast un-able

‘Why could we (excl.) not cast it out?’

For Mal, the pronoun ‘we’ is here translated as /ii/ ‘we (pl. excl.)’, which both refers unambiguously to the nine disciples and excludes Jesus; that is, the information that these participants are mutually exclusive is processed in a straightforward manner.

3. Demonstratives: ‘keeping things up front’
Mal parallels other languages in that it has demonstratives, relative pronouns and nominalisers, i.e. words or particles that turn verbs and clauses into nominals. The Phrase Structures underlying these categories are as follows (see also Filbeck 1976):

(for demonstratives and rel. clauses)  
\[
\begin{array}{c}
NP \\
N - X - DEM
\end{array}
\]

\[
X = \{\text{Adj.} + \text{Adj.} + \text{No.} + \text{Class} \}
\]

(for nominalization)  
\[
\begin{array}{c}
NP \\
\text{NOM} - \text{S}
\end{array}
\]
However, this is where the parallel with other languages often ends, for the information processed by these grammatical categories and structures in Mal frequently turns out to be different. To begin with, Mal demonstratives also function like determiners (e.g. ‘the’ in English) to introduce old information as well as to specify (e.g. ‘this’ and ‘that’). Moreover, there are two sets of the above categories for demonstratives, relative pronouns and nominalisers, which include also a null element, as follows:

<table>
<thead>
<tr>
<th></th>
<th>Set 1</th>
<th>Set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstratives</td>
<td>/nee/ ‘this, here’</td>
<td>/ar(^{42})/ ‘this, here’</td>
</tr>
<tr>
<td></td>
<td>/een(^{42})/ ‘that, there’</td>
<td>/ee(^{42})/ ‘that’</td>
</tr>
<tr>
<td>Rel. pron.</td>
<td>/O/ ‘null’</td>
<td>/ee/ ‘who, which’</td>
</tr>
<tr>
<td>Nominaliser:</td>
<td>/ta-/ ‘that which is’</td>
<td>/ee/ ‘that which is’</td>
</tr>
</tbody>
</table>

These differ in that Set 1 may be considered the unmarked set with its items being used to signal old information (in the case of ‘this’ and ‘that’), or to process information internally on the sentence level (as in the case of the relative pronoun and nominaliser). Set 2, then, is the marked set whose items are used on the discourse level to process information on a more global scale across sentence boundaries. Also, in Set 2, /ee/ may be either a relative pronoun or nominaliser. In historical items, /ee\(^{42}\)/ ‘that (dem.)’ and /ee/ ‘who, which (rel. pron.)’ may have been the same word but, because of the development of the rising tone in Mal,\(^5\) it may have been split in order to process in an unambiguous manner the difference in information (demonstrative vs. relative pronoun/nominaliser) that the two words now convey.

The demonstratives may also be used as locatives, with the exception of /ee\(^{42}\)/ in Set 2. Yet even this word may have a locative sense, meaning ‘there’, under certain circumstances, so that it can correspond to the /ee\(^{42}\)/ ‘there’ of Set 1. However, this locative sense, due to the special or marked role of Set 2, is clearly secondary when used in discourse. That is, the special role of Set 2 takes precedence in the usage and interpretation of /ee\(^{42}\)/, as will be seen below.

The special role of Set 2 is to mark the focus of a discourse. More specifically, Mal employs several linguistic markers for singling out and keeping track of this focus, which may be a noun (in which case it is modified by a demonstrative), or a verb or even a clause (in which case the relative pronoun/nominaliser is used). If the focus changes during the discourse, these same linguistic markers are used to indicate that change. They stand in opposition to ‘non-markers’ which signal that other words are not foci. Stated another way, Mal employs certain morphemes to ‘mark’ certain pieces of content as more salient in discourse; other pieces

\(^4\) I am indebted to Hermann Janzen for this particular way of describing the phenomenon.

\(^5\) This single tone is signalled here and above by the raised numerals [\(^{42}\)] showing that the tone starts on a [\([4]\)] or low-level and rises to a [\([2]\)] or high-level pitch.
are then left 'unmarked' and thus are not as salient.\textsuperscript{6}

In terms of information processing, the function of the words contained in Set 2 is to 'keep things up front', i.e. to keep track in a formal way of what is, and what is not, the focus of a discourse. It is a way of processing information 'about' content in discourse, viz. to signal which piece of content is the focus; other pieces of content not so marked are not foci but rather form background information. But should any of these need to be made prominent, this can be achieved by 'bringing them up front', as it were, by means of these same linguistic devices.

We now turn our attention to a more detailed description of these two sets of demonstratives/relative pronouns/nominalisers. Most of the discussion will focus on Set 2 since it is the more interesting with regard to discourse structure in Mal.

3.1 /nee/ and /ar\textsuperscript{42}/: 'this vs. this'

The basic meaning and usage of /nee/ is that of the demonstrative 'this', e.g.

\begin{verbatim}
/phyam nee/
person this
'this person'
\end{verbatim}

When combined with the noun meaning 'place', however, it has the locative meaning, eg.

\begin{verbatim}
/nam at taanee/
he lives place-this
'he lives here'
\end{verbatim}

In the former example, /nee/ would, in a discourse, signal old information, i.e. that the person discussed has already been introduced and is now being referred to again. In the latter example, it is not apparent that /nee/ necessarily signals old information each time it is used in this construction. As a locative, /taanee/ 'here' could very well be new information introduced for the first time.

The difference between /nee/ and /ar\textsuperscript{42}/ may be summarised as follows: /nee/, either as 'this' or 'here', is general and non-specific, while /ar\textsuperscript{42}/ is much more specific, especially in focusing, in both meanings.

Normally, according to their respective functions, both words are used interchangeably in a discourse. However, I have one short Mal text describing a healing ceremony in which a piglet was sacrificed. The reason for this was probably the brevity of discourse and the fact that it described sequentially the events that took place during the course of the ceremony. In other words, there was no hierarchical structuring of the description — there were no diversions to fill in background information, flashbacks,

\textsuperscript{6} Other dialects of Tin have cognates to /ee/ but, as far as I have been able to determine, their usage does not include the special functions that /ar\textsuperscript{42}/ has in Mal. This leads me to believe that Set 2 words in Mal may be a later development, perhaps a more explicit development of such a discourse tendency in Proto-Tin.

168
Information processing in Mal discourse structure

loops, etc. Hence nothing (noun or clause) turned out to be the focus or foci of the description. Conversely, everything in the description turned out to be a (or 'the') focus. In any event, there was no need to mark anything specifically; consequently, only /nee/ was used throughout.

As a demonstrative meaning 'this', the specificity of /\dr\ 42/ may be seen in the Mal translation of Mark 8:32. In this verse, Jesus has just finished foretelling his coming Crucifixion. It was not a veiled foretelling, for, as this verse states:

\nee\  \dr\ 42\ yeesus\ mple\ ag\ mac

word this Jesus show to see

'Jesus revealed this plainly...'

so that his disciples would not misunderstand. The demonstrative /nee/ was rejected by the Mal translator; it was not specific enough and thus did not 'cohere' with the adverb 'plainly'. /\dr\ 42/, on the other hand, processed the clarity or 'plainness' of the prediction in a much more straightforward manner and it also marked the foretelling of the Crucifixion as the focus of the discourse, thus turning other pieces of content into background information, a function that /nee/ could not have fulfilled.

However, it is in the locative meaning of 'here' that the difference between /nee/ and /\dr\ 42/ is more apparent. Two examples will help illustrate this.

/at taa-\dr\ 42/ 'stay here!' is the statement of choice with a Mal parent in commanding a child to 'stay put' and not to wander off as children often do. In such a pragmatic context the parent means 'here in the immediate vicinity', such as the child's home and/or compound where the house is located and not in a wider area which may include the whole village. The command /at taanee/ would amount to permission for the child to wander and play in a much wider, less specified area.

On the other hand, in answer to the question 'Where does (someone) live?', one answer may be /nam at taannee/ 'he lives here', i.e. this general vicinity (normally a village) is where he lives. But in answer to the more specific question of 'Where is (someone)?', one answer could very well be /nam at taa-\dr\ 42/ 'he is here', i.e. at the time of the answer it is affirmed that the person is physically located in the immediate vicinity (e.g. inside the house where the questioning took place) and nowhere else.

In these examples the locative /taa-\dr\ 42/ 'here' marks the location of the conversation as a specific focus of the conversation as opposed to /taannee/ 'here', which leaves the location of the conversation unmarked as unspecified or background information. So, the importance or non-importance of location in the conversations above is processed formally in Mal with the use of these two locatives.

In texts, the difference between the locative meanings of /nee/ and /\dr\ 42/ 'here vs. here' takes on an added dimension, especially when the latter is
used in direct quotations. The use of /ərt/ ‘here’ in a direct quote in a text reflects the view of the location from the perspective of the participant(s) in the story. Two examples will help illustrate this dimension.

In a recorded story about a hunting party composed of Mai men, the narrator, who was a member of the party, recounted the instructions given to each regarding his particular place in the party as the men fanned out to ‘sweep’ an area of the forest in order to flush out any wild animal that might be hiding in the underbrush. The instruction to each member was given as a direct quote in the story. For the instruction given to him, the narrator stated that he was told to /at taa-ərt/ ‘stay here’. By using /taa-ərt/, and phrasing it as a direct quote, the narrator was telling the story from his own perspective or position in the hunting party. Positional instructions to other men were given in relation to the position of the narrator.

The second example is taken from Mark 6: 36 during the account of Jesus feeding the 5000. Before the miracle took place, however, the disciples asked Jesus to dismiss the people so they could go to buy food:

\[
\text{təm nyal... at təm ərt}
\]

\[
\text{at village... locate around here}
\]

‘... at the surrounding villages’

The expression ‘surrounding villages’ was not translated directly into Mal; rather, the Mal translator chose the expression ‘villages that are located /ərt around here’. In Mal, the choice of /ərt/ achieves the same meaning but from a different perspective, viz. the perspective of the disciples who were in the same vicinity in which the villages were also located.

In both examples, we should not forget that the location here as expressed by /ərt still specifies a focus of the text. In the former, the position of the narrator in the hunting party was the focus since it served as the point from which other positions were pinpointed and described. In the latter, the area containing the surrounding villages was a focus because it was the area in which the people would have to search for food.

3.2. /een/ and /ee/ ‘that vs. that’

The demonstratives /een/ and /ee/ parallel /nee/ and /ərt/ with the exception of differences in meaning. The former in each case function to process old information while the latter in each case function to mark the focus or foci of discourse.

For example, in the following phrase:

7. It is interesting to note that certain sub-dialects of Prai (the other branch, along with Mal, in Tin) contain the two sets apparently only in the area of locatives, i.e. there are two words meaning ‘here’ and two words meaning ‘there’. The first set is used to denote location only, while the second set is used to establish the locational focus or foci of the discourse. More investigation is needed to pinpoint the difference between the two sets more adequately in the Prai dialect.
Information processing in Mal discourse structure

phyam een^42
person that
‘that person...’

/een^42/ modifies the preceding noun and also signals that the noun is old information which has already been introduced. In addition, when combined with the noun meaning ‘place’, this demonstrative functions as the locative meaning ‘there’.

nam at taa-een^42
he lives place-that
‘he lives there’

However, as in the previous section, it is /ee^42/ that holds more interest for this discussion, as the demonstrative that marks a noun as the focus of a discourse.

This is true in conversation or dialogue as well as in other types of discourse. For example, in a conversation with a Mal man, whom I had never met before, I asked him what village he was from. For some reason his answer was not clear. On my asking again he expanded his answer in order to help me to understand. His expanded answer did help, for when I asked if he lived in a certain village (whose name I also gave), his face lit up as he exclaimed:

ηual ee^42
village that = ‘That’s the village!’

In this dialogue the noun /ηual/ ‘village’ with a certain name was the obvious focus of the interrogation and, when this village was at last named, it was then affirmed formally as the focus.

In longer texts, whose content may include several nouns or items, /ee^42/ ‘that’ has a more complicated role. It not only serves to mark the focus of the text but it also reintroduces the focus (e.g. after background information has been discussed) as well as serving to mark a new focus (i.e. when the focus has changed in a text). To illustrate this, let us look again at the story about the ‘Magic Trail’.

In this story a husband and wife are looking for wild animals, each of which is to be taken back to the village and domesticated. Four such trips are related in the story. After each trip to the village, to take back a wild animal, the husband and wife resume their journey along the trail, which is termed:

yooŋ ee^42
trail that = ‘That trail’
The use of /ee⁴²/ at these junctures of the story signals that the trail under discussion is the 'Magic Trail', the focus of the story, and not some other trail.

However, after some time the focus shifts to other items in the story. For example, a magical snake is introduced and, after some background information is added, the snake is referred to as

\[\text{mar } ee⁴²\]
\[\text{snake that } = \text{ 'that snake...'}\]

to mark the snake as the focus of this part of the story. After this episode, the husband and wife travel on to a magical pond. Again, after background information is added, the pond is referred to as

\[\text{majh}⁴² ee⁴²\]
\[\text{pond that } = \text{ 'that pond...'}\]

to signal that the focus has changed again. The snake instructs the man to splash the pond dry, after which a single fish is found. This fish takes on a human body and becomes a person. Since this fish-turned-human-body now plays a role in the story, subsequent reference to it in the story takes the form

\[\text{naj}⁴² ee⁴²\]
\[\text{body that } = \text{ 'that body...'}\]

to mark it as the new focus of the story.

As pointed out at the beginning of this discussion, focusing in Mal discourse structure also entails a contrastive function, of signalling 'this/that-and-no-other' regarding an item of content in the discourse. This function is clearly seen in the translation of a parable in Mark 4:26-29. In this, Jesus compares the growth of the Kingdom of God upon earth to a farmer who broadcasts his seed upon the ground. Jesus explains that the farmer does not know how the seed grows; nevertheless, he waits patiently for the ground to germinate and grow the seed so that he may harvest the grain.

In the parable, the two main participants are the seed and the ground/earth, and both are introduced in 'the farmer went out to sow seed upon the earth', hence any subsequent mention of these two participants is now old information. In verse 27 'seed' is indeed mentioned again and in the Mal translation it is tagged as:

\[\text{khuq}⁴² ee⁴²\]
\[\text{seed that} = \text{ 'that seed...'}\]

It is also stated that the farmer does not understand how seed sprouts and grows. In verse 28 'earth' is similarly reintroduced; moreover, in contrast, the verse specifies the earth as the cause of the seed's sprouting
Information processing in Mal discourse structure

and yielding its fruits. To convey this contrast, 'earth' is translated into Mal as:

\[
\text{thu: ee}^{42} \\
\text{earth that} \\
\text{'that earth...'}
\]

As the real cause of the seed's sprouting, earth, then, becomes the proper focus of the parable as well as the focus for its correct interpretation as a parable about the Kingdom of God. Seed, furthermore, by being modified by /een\(^{42}\)/, is relegated to a background or supportive role in the structure of the discourse.

The demonstratives /een\(^{42}\)/ and /ee\(^{42}\)/ interact in still another way in Mal discourse. This interaction is associated with the word /kyaw/ 'to be equal to'. In the expression /kyaw een\(^{42}\)/ 'only', the combination is used as a sentence-level modifier without regard to any broader context:

\[
\text{nam uuy piar le? kyaw een}^{42} \\
\text{he has two things equal-to that} \\
\text{'He has only two of them'}
\]

On the other hand, the expression /kyaw ee\(^{42}\)/ is often used as a discourse level modifier in the sense of 'That's the way it is!', or 'That's the summation of the matter!'

In a short recorded text, a Mai shaman explained why he would be unable to convert to Christianity. He explained that his most important role was to be a priest interceding for the villagers before the village deity. In this explanation, he gave a list of religious duties that only he as the village shaman could perform. If he converted, there would be no one to perform these duties for the villagers. Therefore, he could not become a Christian. He then ended his explanation with the final

\[
\text{kyaw ee}^{42} \text{ oc} \\
\text{Equal-to that complete} \\
\text{'and that's the complete summation of the matter!'}
\]

Here /ee\(^{42}\)/ relates not to any single noun as its reference but to the whole discourse or, more properly, to the total content of the discourse. Thus the content is the focus of the discourse and not just a noun or item in that content.

The demonstrative /ee\(^{42}\)/ also has an idiosyncratic but interesting usage at the beginning of discourses which demonstrates its global effect over a discourse. If, for example, a discourse is about an event that took place in a time prior to its narration, it may begin with the time expression:

\[
\text{naam ee}^{42} \\
\text{time that} \\
\text{'at that particular time...'}
\]
in order to set the time frame for the narration that follows. This expression assumes that the listener(s) already know the time to which it refers, and, as such, the time of the event also becomes a focus of the discourse.

3.3. /ta-, 0/ and /ee/ “Clause vs. Clause”
In this section, I shall discuss two common grammatical structures—the relative clause and nominalisation and, more specifically, shall limit the discussion to the words or morphemes that constitute the heading to this section because they are the items that process the difference between focus and non-focus information for relative clauses and nominalisations in Mal.

The items /ta-/ and the null or zero element /0/ may both be used as relative clause markers e.g.

phyam ta-cak een\(^{42}\) kayh aw
person who goes that is father
‘That person who is going is my father’.

phyam 0 cak een\(^{42}\) kayh aw
person goes that is father
‘That person (who is) going is my father’

In addition, /ta-/ may be used before a verb or before a clause, turning it into a nominal.\(^8\) In this structure, it can be understood as ‘that which is’.

nam eem ta-luh
‘he does wrong (that-which-is-wrong)’

On the other hand, the null element /0/ may be used only as a relative clause introducer in Mal,\(^9\) while the word /ee/ may be used both as a relative pronoun and as a nominaliser. The difference is that /ta-, 0/ are used to process non-focused information while /ee/ is used more globally to mark the focus or foci of a discourse. It should be noted at the outset that this difference between /ta-, 0/ and /ee/ as relative pronouns in Mal does not necessarily parallel the distinction between restrictive and non-restrictive relative clauses found, for example, in English, or between /thii/ (falling tone, restrictive rel. pron.) and /sin/ (falling tone, non-restrictive

---

8. Direct quotes are a favourite strategy of story-telling among the Mai. While this is probably more a function of their being members of a pre-literate society, it is nevertheless interesting from the viewpoint of information processing. That is, the content of a Mai story is more often than not processed from the viewpoint of a participant and not from the viewpoint of content as autonomous information unconnected with a narrator. As more Mai become literate it may be expected that content-as-autonomous-information will be used more often as a strategy of story-telling.

9. /ta-/ appears to be the same word as /taa/ ‘place’ which figures in the discussion in previous sections. In this regard it parallels the Thai word /thii/ (falling tone) which also means ‘place’ and can be used as a relative pronoun and as a nominaliser. Indeed, this usage in Thai may have influenced that of /taa/ in Mal, extending it to be used as a relative-clause marker and nominaliser, too. Whatever the facts in this case, I have chosen to keep /ta-/ and /taa/ as separate morphemes for the purpose of this discussion.
Information processing in Mal discourse structure

rel. pron.) in Thai. The information that is processed by way of /ta-, Ø/ and /ee/ is different in nature from restrictive vs. non-restrictive information.

/ee/ as a relative pronoun introducing a relative clause functions to mark the focus of a discourse. In this case, however, it is not the noun (or item modified by the relative clause) that is the focus; rather, it is the relative clause itself or, more properly, the content contained in the relative clause, that is the focus. /ee/ shows that the content of the relative clause, even though encoded in a subordinate syntactic construction, is still a focus in its own right.

In Mark 11: 21 the disciples pass by a fig tree which Jesus had earlier cursed. The disciples draw Jesus’ attention to the tree by exclaiming

look tree fig that you curse that

‘Look at the fig tree that you cursed!’

The focus at this point in the discourse is not the fig tree, as it is modified by the unmarked sentence level demonstrative /ee/ in Mal. The focus is on the cursing, i.e. the embedded relative clause that begins with /ee/ ‘that’, thus showing that the content of the subordinate clause is a focus in its own right.

Earlier, we stated that Set 2 demonstratives in Mal also serve an exclusive function, that of signalling the information of this-and-no-other regarding an item in a discourse. /ee/, as a relative pronoun, performs this function as well. This can be seen in the translation of Mark 3: 13-19 which relates the time when Jesus chose his twelve disciples.

In this list of twelve disciples, two have the same name: Simon. The first Simon Jesus renames Peter. The second Simon is distinguished from the first by a relative clause:

Simon who live side-of group person dangerous

‘Simon who was a terrorist (Zealot).’

The inclusion of /ee/ above clearly processes the information that this second Simon is to be distinguished from the first Simon.

/ee/ may also be used as a nominaliser. In this construction /ee/ occurs before a verb or verbal clause thereby turning it into a nominal, which in turn may function as a subject or object of another verb. The special function of /ee/ as a nominaliser, however, is to signal that the content of the nominal is a focus of a discourse.

In Genesis 2: 17, a tree of the knowledge of good and evil plays a prominent part in the account of the Creation. Of course, the focus of the account is not the ‘tree’, but on the ‘knowledge of good and evil’ that arose after the creation of man. To process this information in a translation of this verse into Mal, /ee/ is used three times.
The tree knowing that-which-is good and that-which-is not good

'The tree (of) knowing (that which is) good and (that which is) not good.'

The first occurrence of /ee/ turns the whole of the following construction into a nominal, and as a nominal directly following a noun ('tree' in this case) it stands in a genitive relationship to the noun (i.e. the 'of' relationship). The following nominals (/ee/ plus verb) are also genitives. In each case, the /ee/ processes the information that the content of the subsequent verb is a focus, more precisely a part of the total focus contained in the construction.

In a usage closely related in meaning to the nominal usage above, /ee/ may stand alone, in which case it functions as a 'pro-nominal' in Mal. That is, /ee/ functions something like a pronoun, except that its reference is not an 'object, place or thing' but rather an event or state of existence; it refers to, and includes, all that happens in an event or to the salient characteristics found in a state of existence. It is as though the /ee+verb/nominal has undergone an ellipsis, which deletes the /verb/, thus leaving only the /ee/ to communicate the full scope of the nominal.

In the story about the ‘Magic Trail’, the husband and wife came to a pond. They splashed all the water out and found a single fish lying in the mud. This fish turned into a human body. At this point the fish-turned-into-human-body becomes a focal point in the story: it (or he) helps the husband and wife in the search for fortune. Consequently, in subsequent reference to this creature the ‘pro-nominal’ /ee/ is used. In this context, /ee/ could be translated as ‘it’, but its complete reference is to the (new) state of existence of the participant, viz. the fish-turned-into-human-body.

However, /ee/ in this usage is not restricted to events. In another story about magic, a brother and sister left home to find their fortune in a similar manner. A snake told them of a tree whose bark had magical properties. The children took this bark and eventually arrived at the king's palace in the city. There they learned that the princess had died. They made their way into the palace and sprinkled some of the bark on the princess. At this, the princess came back to life—and at this point in the story she becomes a focal participant. Not only does the princess return to life but the magical bark causes her to become very pretty. The narrator of the story next states that:

\[
\text{ee \ at ta-ee}^{42}
\]

'she lived like that

'She lived in that (pretty) condition [for the rest of her life].'

/ee/, translated as 'she', refers not only to the princess but also to the event of returning to life and becoming pretty. The full meaning would be something like: 'she-who-came-back-to-life-and-became-pretty'. In this sentence, the existential state of being pretty is signalled by /ta-ee\textsuperscript{42}/, an
Information processing in Mal discourse structure

expression that also means 'there' (see above). However, the full scope of
the expression in this context conveys the information that it was this
state-and-no-other-state or condition that the princess existed in for the
rest of her life.

In each story above /ee/ as a pro-nominal marks the focus. However, it
is not merely the 'person, place or thing' that is the focus. More properly,
the focus is the event or existential state of which the person, place or
thing is a part and, being the focus, it excludes other events or states that
could be associated with the participant in the discourse.

3.4. Special combinations
From the above data, two combinations deserve special comment. They
both involve the nominalisation usage of /ee/ plus a demonstrative. The
two combinations are:

/ee . . [verb] . . een 42/ ‘that [Nominal]’
/ee . . [verb] . . ee 42/ ‘that [Nominal]’

From textual analysis it is difficult to discover any real difference between
these two constructions; they each seem to process the same information
of focus and/or exclusivity regarding the content contained in the
nominal. Both, for example, have been used in translating Mark into Mal.

Mark 10: 35-40 records the time when two of the disciples requested
seats of honour at the time when Jesus assumed complete power in his
kingdom. One asked to sit on Jesus’ right and the other on his left. Jesus’
answer to this request begins with a nominal:

ee paa? soop ee 42
that-which you-two ask that
‘(Regarding) that request of yours . . .’

after which he explains that such positions involve suffering and not
honour. When the two disciples affirmed their willingness to suffer, Jesus
explains that:

ee khuyh phiah sam phiah wi? oǐn een 42
to sit side right side left I that
‘(But) that sitting on my right or left . . .’

was in reality not his to bestow but God’s.

It is difficult to detect any difference in focusing or exclusivity between
these two constructions, unless the former is a higher degree of focusing,
etc., than the latter.

4.0. The /i-/ clitic: ‘This one exclusively
Mal has one clitic, /i-/ , that occurs as a prefix to a few words. As a prefix
/i-/ is stressless but it can be given stress and thus stand as a free
morpheme in a sentence; normally, however, it occurs as a (bound) prefix
in the course of speech. Semantically, though, /i-/ carries no definite meaning in the traditional morpheme sense; rather it functions more as an ‘exclusive enhancer’ for the words on which it may occur as a prefix. Consequently, its productivity in Mai is limited to occurring with only those words that already contain a component of exclusivity as a major part of their semantic composition. On the other hand, given the exclusivity characteristic of Mai discourse, it is not surprising to see these few words plus /i-/ occurring often in both speech and narratives of the Mai people.

/i-/ occurs as an exclusive enhancer of two types of words in Mai, demonstratives and interrogatives. Words in both categories already contain a semantic component of exclusivity, for to modify a noun or nominal with a demonstrative is to exclude other nouns or nominals, and to ask a question often involves a decision to choose one alternative to the exclusion of others. The addition of /i-/ in these cases enhances the exclusive property of such words.

4.1 /i-/ plus Demonstratives

In the case of demonstratives, however, /i-/ occurs only with the sentence-level demonstratives and not with the discourse-level demonstratives, i.e. with Set 1 Demonstratives and not with Set 2 Demonstratives as listed in 3 above. The reason for this is that in processing information, /i-/ as an exclusive enhancer in effect ‘raises the informational level’ from that of old information to that of ‘focused information’ in the discourse. For, with the addition of /i-/, there is a more explicit exclusion of some other referent, whether in text or conversation, thus throwing more attention on the referent modified by the demonstrative in question. There is therefore no semantic or informational reason for combining /i-/ with the Set 2 Demonstratives since this set already performs the same task in discourse.

On the other hand, even though they perform the same task, there is a difference. Consider these examples.

(a) kayh khyaka inee
    is   buffalo   this one
     ‘It’s this particular buffalo.’
(b) kayh khyak or
    is   buffalo   this
     ‘It’s this buffalo.’

Example (a) contains a Set 1 Demonstrative, /nee/ ‘this’, in combination with /i-/ . Example (b) contains the Set 2 Demonstrative /or / ‘this’. The difference between the two is that the former is more negative or constrastive while the latter is more positive and affirmative. Both, of course, make the NP (/khyak/ ‘buffalo’ in this case) a focus. Yet /inee/ turns /khyak/ into a focus by means of excluding other possible buffaloes while /or/ makes it a focus by means of highlighting.
In Mark 10:17-20 a young man came to see Jesus to ask what must be done to inherit eternal life. Jesus told him to keep the following commandments: do not kill, do not commit adultery, do not steal, do not bear false witness, do not defraud, honour (your) father and mother. To these instructions the young man answered:

\[
\begin{align*}
\text{\textit{nee\textsuperscript{42} inee}} \\
\text{word teach these ones}
\end{align*}
\]

'These \textit{particular} teachings... (I have observed from my youth).'

The demonstrative /inee/ in this discourse gives a 'particularising' flavour to the discourse: it was these particular commandments, in contrast to possible others, that he had especially observed in his life. This contrastive characteristic is further brought out in the discourse by what Jesus said next: 'You lack one thing; go, sell what you have and give it to the poor... and come, follow me.' In this context, the demonstrative /inee/ processes the contrastive nature of the conversational exchange between the young man and Jesus in a better way than the more positive demonstrative /\textit{in}\textsuperscript{42}/.

/i-/ does not occur with /\textit{een}\textsuperscript{42}/ 'that' as an exclusive enhancer. The reason for this, however, appears to be phonological and not semantic. As will be seen below, when /i-/ combines with a following vowel of a stressless syllable (which means that the vowel is automatically short in duration), the vowel is assimilated with the /i/. But in this particular case /i-/ would occur with a long vowel of a stressed syllable, /i+een\textsuperscript{42}/, hence no assimilation takes place. That is, */i\textit{een}\textsuperscript{42}/ or */i\textit{in}\textsuperscript{42}/ are both unacceptable.

However, this does not mean that the exclusive enhancer marker /i-/ does not occur with 'that' in Mai. For such occasions, Mai speakers will use the Thai demonstrative /\textit{nan}/ 'that' (high or falling tone) in combination with /i-/ to mean 'that \textit{particular} one'.

In Mark 4:10-20 Jesus explains the parable of the sower who went out to sow seed on various types of ground. Some seed fell among thorns which choked out the plants that grew from the seed. Jesus explained that this part of the parable referred to those people who upon hearing the Gospel accept it but soon fall away because of delights in riches and

\[
\begin{align*}
\text{\textit{so p\textit{oon} inee inan}} \\
\text{want get this that}
\end{align*}
\]

'Wanting this \textit{particular} thing and that \textit{particular} thing.'

In Mai the demonstratives /inee inan/ may occur without accompanying NPs as an idiom in much the same way that 'this and that' is used in English to refer to miscellaneous things or to 'odds and ends' in one's possession. The addition of /i-/ in Mai, therefore, enhances the particularising and constrastive nature of the 'things' referred to.

Before leaving the above example, we should take note of an evident
DAVID FILBECK

relation between phonology and information processing that can be seen in the case of /inan/ in Mal. As in the case of /inee/ ‘this one’, it is also possible, informationally and logically, to have ‘that one’ also. However, a phonological rule blocks the processing, more properly the encoding, of this information using the corresponding Mal demonstrative. Thus the Thai demonstrative /nan/ ‘that’, which has a phonological structure allowing the stressless /i-/ to occur as a prefix, is used to process the enhanced exclusive nature of ‘that one’ in Mal.

/i-/ also occurs as an exclusive enhancer to the word /isoo/ ‘other’. In such occurrences the combination /isoo/ further enhances the exclusive nature of the ‘other’ (thing, person) that is referred to.

In Mark 4:10, after Jesus had told the parable of the sower, many people returned to their homes, but

\[ ah \ isoo \ at \ \text{γρο̣π} \ yeesuu \]

they others locate around Jesus

‘Still another group gathered around Jesus...’

to hear his explanation of the parable. The combination /isoo/ enhances the ‘otherness’ of this latter group in contrast to the former group of people who returned home.

4.2. /i-/ plus Interrogatives

/i-/ also occurs with the interrogative /ehee/ ‘who?’, becoming /ihee/ under the influence of the phonological rule mentioned above. The difference in meaning between /ehee/ and /ihee/ parallels that described above, for example, for /nee/ and /inee/. That is, /ehee/ is an unspecified ‘who?’ while /ihee/ is a more selective, particularising ‘who?’, resembling perhaps the now archaic ‘whosoever’.

In Mark 8:27 Jesus asked his disciples

\[ ah \ soo \ khay \ aĩ \ kayh \ ihee \]

they other say I am who?

‘What particular person do others say I am?’

The enhanced or particularising /ihee/ is more appropriate for this context since there were rumours that Jesus was John the Baptist, or maybe Elijah, or perhaps one of the prophets. The unspecified /ehee/ in the above quote would not process this (particularising) type of information for the context in question.

A second interrogative that /i-/ occurs with is /naa/ ‘which?’. In the same manner as above, /naa/ is an unspecified ‘which?’ while /inaa/ is an enhanced ‘whichever one it is’.

In Mark 12:28, a scribe asked Jesus a question that had caused much argument among scholars of the Torah.
Information processing in Mal discourse structure

\[\text{\text{\(\eta\)\(\epsilon\)\(\eta\)\(\epsilon\)\(n\)\(42\) inaa kayh s\(\epsilon\)k lhia m\(p\)huang\)}}\]

Word teach which is big above group

‘Which particular teaching is more important than all the others.’

Since there were several possible answers, the interrogative /inaa/ processes the particularising nature of the question, something which the unspecified /naa/ is unable to do.

/\(i\)/ occurring with an interrogative in effect throws the focus of the discourse on the question or, more properly, on the cognitive process of choosing among alternatives for the answer. And in choosing an alternative, others are excluded or eliminated. Hence, with the addition of the exclusive enhancer, /\(i\)/, ‘things are kept up front’ in yet another way in Mal discourse structure, i.e. by eliminating or excluding other possibilities.

5. Keeping things up front: ‘That’s the way we talk’

In this paper, I have surveyed several aspects of information processing in Mal discourse, comprising both conversation and recorded narratives. In surveying these aspects we saw where pronouns, demonstratives, relative pronouns, nominalisers, and a clitic, function to process certain types of information regarding content in a discourse or text. Moreover, I stated that this processing gives a particular characteristic to Mal discourse, a characteristic of this-and-no-other, or of ‘keeping things up front’ vs. ‘keeping other things in the background’.

We also included examples from Christian scripture to show how productive this characteristic is even in translating from another language into Mal. After revising a certain text from the Gospel of St. Mark, for example, to include more occurrences of /\(ee\)/ and /\(ee42\)/ according to the discourse principles described in this paper, the text was read aloud to others for reactions. On hearing the text for the first time a Mal woman responded with the comment: ‘That’s sounds just like we talk’. Among others, the revisions met with greater acceptance and satisfaction.

In short, ‘keeping things up front’ is a favourite style of communication among the Mal. Of course, as was also seen above, a complete text or discourse may be unmarked with regard to focus and/or exclusivity, but such a way of communicating appears not to be very exciting. Such an unmarked discourse will transmit content well enough, but for ‘real performance’ in Mal, a discourse should also process information about content.

REFERENCES


DAVID FILBECK


LES DÉRIVÉS DÉSIDÉRATIFS EN KHMER

Saveros Pou

0. Le mot ‘désidératif’ est emprunté à la grammaire sanskrite pour des raisons de commodité et de simplification terminologique. L’on évoque par exemple sk. cikitsā ‘médecine’ < cit ‘observer, penser’, ou mīnāmsā ‘un système de philosophie’ < man ‘penser’. La langue cambodgienne possède des dérivés semblables, mais en vérité ils ne procédaient pas de la même façon qu’en sanskrit tant dans les concepts que dans les formes.

1. Ils sont attestés dans l’usage dès les premiers documents épigraphiques khmers. Relevons quelques exemples avant d’en analyser le type:

| cāy  | ‘disperser’     | > cicāy | ‘détruire’     |
| dāh  | ‘heurter’       | > padāh | ‘encloiner un espace par des murs’ |
| dāṃ  | ‘se percher, se tenir’ | > pdāṃ | ‘dormir’ (lang. princ.) |
| dāp  | ‘barrer’        | > pandāp | ‘enfermer’ |
| dāuy | ‘porter, soutenir’ | > pdāuy | ‘renforcer le soutien, servir un maître’ |
| dūl ~ dval | ‘porter sur la tête, parler à un être sacré’ | > dadval | ‘recevoir’ |
| gāṃ  | ‘joindre, réunir’ | > paṅgāṃ | ‘joindre les mains d’adoration’ |
| gāp  | ‘convenir, plaire’ | > phgāp | ‘désirer obtenir les faveurs de qu’on’ |
| pāt  | ‘perdre, disparaître’ | > pampāt | ‘détruire complètement’ |
| rāc  | ‘usé, ruiné’    | > vṛc | ‘(d’un éléphant) tuer sauvagement’ |
| (sra)sir | ‘écrire’ | > prasir | ‘écrire dans un but élogieux’ |
| tāp  | ‘être bout à bout’ | > stāp | ‘désirer suivre, écouter’ |
| tin  | ‘savoir’        | > tantiṅ | ‘s’informer, s’enquérir’ |
| tval | ‘tomber à la renverse’ | > phṭval | ‘tuer dans un combat’ |
| vat  | ‘faire un geste circulaire’ | > cvat, chvāt ‘borner une terre’ |

1. D’où moderne phdaḥ ‘un espace enclos de cette façon: la maison’. Il convient de rappeler que certains cas de dérivation sont de nature exclusivement sémantique, l’affixation ne jouant aucun rôle dans la production des dérivés, tels que: crat ‘s’appuyer du bras sur’ > crat ‘canne, bâton’ (dès le vieux khmer), ou bien dāp ‘barrer, encloiner’ > (*dāp’) noté maintenant dāb ‘enclos, camp de soldats, armée’.

2. Le dérivé secondaire en est pandāy ‘endroit de refuge; fort, fortification’. Cf. les toponymes Bantéay Mēas, Bantéay Chmar, etc.


183
path 'long' $>$ saven ‘parcourir un long chemin; rechercher’
vyat ‘vrai’ $>$ panvyat ‘s’adonner à la vérité; confirmer’

Au premier abord, il semble qu’on ait affaire à des dérivés de types couramment définis, à savoir des fréquentatifs, des intensifs et similatifs, et des causatifs. Or, en plus de ces traits grammaticaux, nos dérivés ont un caractère fondamental commun très net du fait qu’ils expriment un ‘désir’ de la part du Sujet, où l’on déceler désir proprement dit, souhait, intention et même prétention. Ce trait qui apparaît déjà comme distinctif va s’illustrer de façon copiuse dans le lexique moderne, y compris le lexique moyen (XVe–XVIIIe siècle), et de façon très originale dans ce dernier. Aussi le regroupement de ces dérivés s’impose pour un nouvel examen (à la fois dans le cadre de la langue écrite et orale).

2. En morpho-phonologie, ils ne présentent aucun signe distinctif qui les sépare des dérivés de tout le système.

2.1. Les mots de base sont en général des verbes. Par ex.:
gat ‘mesuré, ferme’ $>$ phgat ‘prendre une attitude réservée, pratiquer une observance’
kat ‘noter’ $>$ sankat ‘presser, opprimer’
paṁṇī ‘barrer’ $>$ paṁpaṁṇī ‘empêcher de voir, rendre invisible’

Mais des substantifs ont aussi servi de base, bien que plus rarement, comme:

bot ‘le maïs’ $>$ paṁbot ‘faire cueillir son maïs par un naïf, lui ravir son maïs, duper qq’un’
cor ‘voleur, brigand’ $>$ pañcor ‘traiter qq’un (surtout une femme) de voleur, l’injurier vilainement’
ruot ‘couche, étage’ $>$ pruot ‘conjuger des efforts dans une entreprise’

2.2. Les exemples du vieux khmer (sup. 1) révèlent trois modes de dérivation, qu’on va retrouver dans toute l’histoire de la langue, à savoir:

– redoublement: cicāy, dadval, tantīn...

Les dérivés désidératifs en Khmer

- préfixation labiale: padəh, phdəy, phtval...
- préfixation de /s~c/: sən, cvat.

Dans le cas précis des dérivés désidératifs, je regroupe ces deux phonèmes, me basant simplement sur leur fréquente alternance rencontrée en khmer comme dans tout le groupe môn-khmer.

3. Jusque là, il semble qu'on puisse classer ces dérivés dans les catégories déjà établies, et évoquées à sup. I, comme des sortes de sous-types; par exemple cicəy comme un fréquentatif de cəy 'épapiller', ou bien sən comme un intensif de sən 'long'. Mais cela ne suffit pas en fin de compte quand on procède à un examen global de ces dérivés dans le contexte dérivationnel entier, et surtout quand on avance dans le temps vers le khmer moyen.

3.1. Par exemple, tantın, dérivé de tin 'savoir', signifie 'désirer savoir, s'informer', de façon univoque. De dau 'aller', pradau, sur les bas-reliefs d'Angkor Vat où il apparaît, n'est pas un simple causatif: 'faire aller' certes, mais dans un sens spécial, intellectuel et moral, d'où 'désirer le progrès, instruire, éduquer'. De vat 'faire un mouvement circulaire', cvat n'est ni un intensif ni un similitatif, car, dans les inscriptions, il signifie 'délimiter un terrain de façon délibérée', dans un but déterminé. Et enfin, cicəy n'est pas simplement comparable à notre moderne kakəy 'gratter à petits coups', dérivé de kəy 'fouiller'.

3.2. Le cas du dérivé à préfixe labial dit 'causatif' demande un examen particulier plus étendu. En effet, si l'on prend le terme 'causatif' au sens propre, i.e. où le sujet fait simplement commettre l'action par un autre:

\[
\text{A fait (Verbe) B,}
\]

on doit noter qu'il y a en gros deux modes d'expression en khmer, soit:

(a) la dérivation, entre autres, par préfixation labiale, comme:

\[
\text{hoer 'voler} \quad \Rightarrow \text{pənhoer 'faire voler'}
\]

(b) une construction syntaxique schématisée par:

\[
\text{oy + Verbe} \quad \text{ou dhvoe oy + Verbe}
\]

Ainsi:

\[
\text{phik 'boire} \quad \Rightarrow \text{A oy B phik, 'A fait boire B'; tandis que le dérivé pəmphik est désidératif, 'faire boire qq'un pour le rendre ivre'.}
\]

De même:

\[
\text{khūc 'abîmé, mourir} \quad \Rightarrow \text{A oy B khūc, 'A abîme B' sans préméditation; tandis que pənkhūc est un désidératif, 'débaucher, calomnier' (cf. inf. 3.3.).}
\]

3.3. C'est en khmer moyen qu'on voit les traits particuliers de ces dérivés ressortir de façon indiscutable et décisive. Prenons d'abord quatre

---

5. Le dérivé secondaire caṇvat est devenu en siamois cànhvāt 'district'.

185
exemples typiques de dérivés à prédéfixé labial /pa-/, /baN-/, répondant aux formes graphiques pa-, pra- ou pa + Nasale-.

Les valeurs généralement reconnues par les linguistes à ces prédéfixes sont soit causative soit réciproque, ce qui n’est pas avéré dans les exemples qui nous préoccupent et que voici:

- cor ‘voleur, brigand’ > pañcor (cf.2.1) ‘traiter qq’un (surtout une femme) de voleur; l’injurier vilenement’
- danďa ‘châtiment’ > ptandă ‘souhaiter le châtiment à qq’un, lui souhaiter beaucoup de mal’
- dăsă ‘esclave’ > piăsă ‘souhaiter la servitude à qq’un; le maudire’
- dos ‘faute, peine’ > pandos ‘imputer une faute à qq’un; le réprimander’

Il est difficile de parler ici d’une expression de causatif car on déformerait la réalité linguistique sur laquelle tous les documents s’accordent sans aucune exception.

3.4. Et que dire de quelques exemples de redoublement qui d’une part prouvent la productivité de certains mots de base, et d’autre part constituent de très beaux spécimens de polysémie? Tels:

- ceh ‘savoir’ > caceh ‘se croire savant, fort; être obstiné’
- -dihn ‘en travers’ > dandihn ‘se tenir en travers du chemin, i.e. attendre l’arrivée de qq’un avec impatience’
- loe ‘au-dessus’ > laloe ‘se croire supérieur; se prendre bêtement au sérieux’
- ţihn ‘être au courant’ > ţațihn ‘s’enquérir, s’informer de la situation d’une jeune fille, la demander en mariage’
- foem ‘le début, principal’ > ţantoem ‘se croire le premier; disputer qqch. à qq’un’

6. La n.4 a laissé pressentir le rôle de la littérature (écrite et orale) dans la formation des mots, d’où le rapport étroit entre littérature et grammaire. Et cela n’est que trop vrai dans le cas des dérivés. Beaucoup d’entre eux furent créés par des écrivains, des poètes et artistes qui—est-il besoin de l’expliciter—maniaient la stylistique et le folklore avec dextérité pour exprimer leurs idées, leurs émotions, parfois au détriment de certaines règles: cas de licences (littéraires et poétiques). Certaines créations plurent rapidement au public, qui les adopte et les propage. Le mouvement part ainsi d’une sphere réservée de locuteurs, mais un consensus était nécessaire pour faire passer les créations dans l’usage national. Dans ces conditions, on saisit que la structure de tels mots échappe parfois aux règles grammaticales qu’on souhaiterait appliquer en linguistique. Pra- n’est pas nécessairement un prédéfixe de réciprocity, ni paN- un prédéfixe de causatif. Ils sont confondus dans notre sujet, et c’est le rôle dominant des arts, au sens large, qui est responsable de cette confusion des formes et d’autres irrégularités. Voir des exemples dans la nomenclature, inf., 4
Les dérivés désidératifs en Khmer

Aucun doute n’est possible sur la nature et la valeur de ces dérivés. Ce qu’ils ont de commun avec vieux-khmer cicây ‘détiruire’, c’est leur expression fondamentale d’un désir orienté vers un but précis, de souhait d’une action à accomplir par soi ou par un autre. Ce trait distinctif clarifié, nous pouvons maintenant examiner des exemples modernes, parfois sémantiquement complexes, qui sont toujours en usage à l’heure actuelle.7

4. Plusieurs modes de classification sont possibles et aussi défendables l’un que l’autre. A côté du critère morphologique proprement dit déjà relatif (sup., 2), le critère sémantique semble s’imposer maintenant, d’autant qu’il met le mieux en relief la productivité de certains mots de base et divers facteurs diachroniques. Suivons donc simplement ces mots de base, arrangés par ordre alphabétique.

bămî /pean/ ‘cacher la vue’ > pambămî /bompean/ ‘calfater’
bân /piam/ ‘grimper sur’ > pambân /bompian/ ‘piétiner volontairement; blesser, abuser de’

bûr /piar/ ‘se cogneter’ > pambûr /bompior/ ‘manquer d’égards, heurter, blesser’

côr /coon/ ‘vouloir’ > phcôr /pcon/ ‘s’appliquer entièrement à une tâche’

câm /cam/ ‘attendre’ > prâcâm /pocam/ ‘être posté à un endroit, s’attacher à un objet’

côp /cap/ ‘saisir’ > pâncôp /bôncap/ ‘mettre un objet en gage’

ceh /ceh/ ‘avoir des connaissances’ > caces /coceh/ ‘se croire fort, être obstiné’

... (suite des exemples avec descriptions de sens)


8. A gloser ainsi: ‘désirer boucher une entrée; empêcher une impregnation ou une infiltration, d’ou calfater les joints, les trous...’ On procédera de la même façon dans l’analyse des exemples qui suivent, à savoir qu’il faut attribuer à chaque dérivé un premier degré, explicite ou implicite, de sémantique représenté par ‘désirer, vouloir, souhaiter’ dans un sens particulier, lequel se développera le cas échéant dans d’autres directions par métaphore ou par métonymie.
SAVEROS POU

dan' /tont/ 'doux, tendre' > dan' /tont/ 'se montrer modeste; s'affaiblir, décliner'
> san' /sont/ 'adopter une manière douce, tendre'

dap ~ dáp /top/ 'aplatis contre' > dandap /tontop/ 'marcher en catimini'
dar /tow/ 'frapper' > pandar /bontow/ 'accompagner en cadence'
daú /tow/ 'aller, avancer' > diúau /tutow > tutow/ 'répandre partout, généraliser, général'

> praúau (3.1)
dák' /teak/ 'poser un piège' > sdák' /steak/ 'poser une embuscade'
dâm /tôom/ '(des coups) nombreux et répétés' > dandâm /tontôom/ 'se montrer gracieusement le corps et les membres'

dáp /tiap/ 'bas' > dandáp /tontiap/ 's'abaisser doucement, se montrer modeste'
dêñ ~ deñ /têñ ~ tôôñ/ 'désirer avidement' > dandeñ ~ dandeñ /tontêñ ~ tontôñ/ 'se montrer hautain, fier, altier'
> pandeñ /bontôñ/ 's'enthousiasmer, s'exalter'

deñ /têñ/ d'une cadence rapide' > dandeñ /tontêñ/ 'réciter rapidement pour mémoriser'
dic ~ dec /tic/ 'petit, inférieur'9 > pradec /patic/ 'avilir qq' un, le couvrir de malédictions'
dim /tüm/ 'atteler, conjuger' > dandim /tontüm/ 'se mettre côte à côte'10
> phdim /ptüm/ 'placer côte à côte (rituellement); comparer'
daes /tôôh/ 'se heurter à un obstacle' > dadas /tataôh/ 'faire obstruction'

> prathuy /pothoy/ 'fencer en avant; tenter la chance'

> dädic /tatuuc/ 'attirer avec insistance, supplier'
> sädic /stuuc/ 'prendre au moyen d'engin: pêcher'

> dädir /tatuu/ 'le sommet'11 > dädîr /tatuu/ 's'envelopper la tête d'une étoffe'

gan' /kon/ 'qui dure, permanent' > phgan' /pkon/ 'subvenir à, pratiquer une observance'

> jañjät' /çôçôt/ 'répéter ce geste au moyen d'un panier pour pêcher, ou pour ramener les esprits vitaux'11

duj ~ du ~ dû /tuy, thoy/ 'être en saillie' > prathuy /pothoy/ 'fencer en avant; tenter la chance'

diac /tuuc/ 'attirer' > dädic /tatuuc/ 'attirer avec insistance, supplier'

9. Cf. le congénère món signifiant 'esclave'.
10. On notera aussi la forme dialectale sandjm.
11. Le corps humain héberge 19 esprits vitaux (brajin) qui s'en échappent partiellement en cas de sommeil ou de maladie, d'où la nécessité de les 'ramener' en vue de la guérison.

*Allomorphe de dul ~ dval (cf. sup., 1)
### Les dérivés désidératifs en Khmer

| jīk /ciīk/  | ‘creuser’ > jajīk /cōciīk/ | ‘interroger de façon insinuante’ |
| jōh /cōh/  | ‘piquer avec une pointe’ > pānjoh /bācōh/ | ‘riposter à des critiques de façon acerbe, blessante’ |
| jōr /cōo/  | ‘s'enfler’ > pānjo' /bācōo/ | ‘flatter, flagorner’ |
| jūn /cuun/ | ‘offrir’ > jajīn /cācuun/ > pānjlīn /bāncuun/ | ‘envoyer vers une destination précise ou dans un but déterminé’ |
| kākā /kōk/ | ‘tremper, être trempé’ > pānkākā /bōnjōk/ | ‘(rite) ondoyer’ |
| kāc /kac/  | ‘casser, couper’ > pānkāc /bōnjōk/ | ‘calomnier, honnir’ |
| kān /kan/  | ‘tenir’ > prakān /pōkan/ | ‘s'attacher aux formes, être à cheval sur les principes’ |
| kār /kaar/ | ‘protéger’ > pānkār /bōnjaa/ | ‘parer à une éventualité’ |
| kāt /kat/  | ‘couper, trancher’ > skāt /skat/ | ‘prendre un chemin court pour retrouver; interrompre un interlocuteur’ |
| kīn /kvn/ | ‘écraser, moudre’ > saṅkin /saŋkvn/ | ‘opprimer, oppresser’ |
| kīn /kvn/ | ‘rigide’ > prakīn /pōkvn/ | ‘exiger sévèrement sans compromis’ |
| la /lō/   | ‘essayer, tenter’ > jāna /cānlō/ > sla /slo/ | ‘éprouver, taquiner’ |
| laēn /lēnj/ | ‘lâcher’ > plēn /pleen/ > banlaēn /pōnleēn/ | ‘noyer en usant d’un lest’ |
| -lāk’ /lōk/ | ‘plonger dans un liquide’ > bhālāk’ /plōk/ | ‘brutaliser par des coups ou des mots’ |
| -lan’ /lōn/ | ‘égaré, désorienté’ > plān’ /plōn/ | ‘attaquer à l’improviste, piller’ |
| lan’ /lōn/ | ‘se noyer’ > banlan’ /pānlōn/ | ‘noyer qq’un délibérément’ |
| lan’ loc /lōn-lōŋc/ | ‘loin-brûlé’ > lanlan’ lanloc /lānlōn lanlōŋc/ | ‘se consumer de nostalgie’ |
| lap /loop/ | ‘agir doucement, secrètement’ > jhlāp /cōop/ | ‘espionner’ |
| lāk’ /lōe/ | ‘en bas–en haut’ > latāk’ lāo /l’lēak laāo/ | ‘se comporter de manière légère, éccrêlée’ |
| lōm /lōm-lōm/ | ‘aux traits indécis’ > panlōm /bānlōm/ | ‘se camoufler, duper’ |

12. On peut gloser par: ‘faire un essai sur plusieurs matières, obtenir une solution à partir de produits dissous ensemble’. C'est là le sens premier de notre dérivé sla, qu'on retrouve dans le fameux dérivé de celui-ci, samla /samlo/, généralement rendu par 'la soupe'.


---

189
lec /lec/ ‘apparaître’ > jhlec /clec/ ‘ressortir, faire sortir la vérité’

len /leːŋ/ ‘jouer, s’amuser’ > panlaen /banlaŋ/ ‘détruire par de la musique’

lic /lic/ ‘être immergé’ > banlic /panic/ ‘enfoncer dans l’eau de force, vilipender, détruire’

loe /ləʊ/ ‘au-dessus’ > laloe (cf. 3.4.) > bhloe /plaː/ ‘se croire important: être niais, bête’

loes /bəʊ/ ‘dépasser, surpasse’ > phloes, banloes /plaːh, panlaŋ/ ‘tenir des propos exagérés, contraires aux faits14

lüt /lu:t/ ‘surgir, croître’ > banlüt /panluːt/ ‘opérer un avortement’

lap’ /lɑːp/ ‘faire un tour, retourner’ > panlap /bənloːp/ ‘confondre qq’un; égarer, détourner’

moel /məʊl/ ‘regarder’ > pramoel /pamaːl/ ‘regarder à distance, observer de loin’

män /mian/ ‘avoir, posséder’ > bramän /pamian/ ‘donner un avertissement’

muc /muc/ ‘s’immerger’ > pramuc, cramuc /pamuc, camuc/ ‘enfoncer qq’un dans l’eau’

ñoey /nɔoŋ/ ‘lever la tête’ > brañoey /pɔnɔoŋ/ ‘se montrer indifférent, insouciant’

paət’ /bəʊŋ/ ‘jeter, lancer’ > pampaṭ’ /bəmbəʊŋ/ ‘abandonner, se débarrasser de’

pāmḥa /bən/ ‘barrer la vue’ > pampāṁh /bəmbən/ ‘empêcher de voir, rendre invisible’15

pit /bət/ ‘fermer, clore’ > pampit /bəmbət/ ‘cacher soigneusement; empêcher de sortir et circuler’

pōh /boh/ ‘jeter’ > pampōh /bəmbəh/ ‘lancer des paroles blessantes’

pos /boʊh/ ‘balayer’ > papos /bəboʊh/ ‘tapoter affectueusement, caresser’

pūl /pʊul/ – puol /pʊuloʊl/ ‘appeler, conjurer’ > papaʊl /bəboʊl/ ‘inciter à une action, proposer’

pih /pəh/ ‘heurter, cogner’ > pampiḥ /bəmpəh/ ‘heurter, se cogner, avec intention blessante’

rat’ /rət/ ‘courir, fuir’ > bairat’ /pərət/ ‘enlever, ravir une jeune fille’

rān /rəŋ/ ‘forme’ > brān, pambrān /priŋ, bəmpriŋ/ ‘donner une autre forme; falsifier16

15. Exemple déjà relaté (sup., 2.1) mais qui gagnerait à être comparé avec pambāṁh (premier ex. de la liste). Les mots de base, quasi synonymes ‘barrer la vue’, produisent de faux synonymes dans les dérivés, en l’occurrence notre pampāṁh est courant en magie pour ‘se rendre invisible’.
16. Les textes en khmer moyen montrent le même usage de ces deux dérivés.

190
Les dérivés désidératifs en Khmer

râp’ /rɔap/ ‘compter, conter’  > prâp’ /prap/ ‘informer’

rây /rɔy/ ‘répandre’  > baňrây /paŋrıy/  > srây /sraay/ ‘faire du désordre’
   > ‘délier, défaire un noeud, résoudre’

rit /ruıt/ ‘serrer’  > prît /prvıt/ ‘se montrer sévère, rigoureux’

rîep /riap/ ‘ranger’  > prîep /pri.ap/ ‘placer ensemble; comparer’

roep /rɔoıp/ ‘remonter, se remettre’  > sroep /sraıp/ ‘s’exciter pour qqch.’

rût /ruuıt/ ‘tirer sur un fil dans une coulisse’  > srût /sroot/ ‘accélérer le pas pour arriver vite’

sâp /saap/ ‘sans saveur’  > pânsâp /bonsaap/ ‘neutraliser un poison, une attaque malicieuse’

soh /soh/ ‘fini, épuisé’  > sâsoh /sasoıp/ ‘démêler; parler pour ne rien dire’

ťâm /dam/ ‘planter dans le sol’  > phtâm /pdam/ ‘faire entrer dans l’esprit; faire des recommandations’

ťeñ /deñ/ ‘poursuivre’  > pântëñ /bdeñ/ ‘chasser, expulser d’un endroit, congédier’

ťøy /daoy/ ‘suivre, accepter une proposition’  > pàntoy /bdaoy/ ‘laisser faire, laisser toute liberté’

ťıc /dooc/ ‘ressembler’  > prâtic /pdooc/ ‘comparer une chose à une autre’

vaen /ween/ ‘long’  > pânaen /bøjaen/ ‘prolonger un trajet; détourner du but, égarer’
   > svaen (cf. 1.)

Cette nomenclature n’est pas exhaustive pour une raison évidente, à savoir l’espace limité d’un article. Et le temps aussi constitue un facteur important, car il reste à approfondir l’exploration de tout l’appareil dérivationnel du khmer si complexe et, je dirais même, sophistique, en interrogeant tous les textes et en écoutant simplement les locuteurs. Mais ces premières notes succinctes, je les offre déjà, ainsi que mes vœux les plus sincères, au grand linguiste-philologue qu’est Monsieur H. L. Shorto.

REFERENCES


A DIACHRONIC SURVEY OF SOME KHMER PARTICLES (7TH TO 17TH CENTURIES)

Judith M. Jacob

Introduction

In modern spoken Khmer, grammatical relationships between the clauses of a sentence may be understood without the use of clause particles. Intonation and pause, together with the context in which the sentence is spoken, make the meaning clear. Thus, when a speaker says /mtay dau phsar, kun dau phan/,\(^1\) intonation and pause mark off two separate phrases and the hearer will know from the context whether he is being told that it is 'because', 'if' or 'when' the mother is going to the market that the child will go too. Particles to express 'because', 'if' and 'when' are available but are not felt to be necessary. In modern written Khmer, the roles of intonation and pause are played by punctuation, in the form of full stops and of spaces between phrases; context still plays an important role, more so than in European languages, and particles occur frequently. When we turn to the inscriptions and the earliest non-inscriptional texts, punctuation is minimal and the contexts are unfamiliar to us in comparison with the communities for whom the texts were composed. On the inscriptions, punctuation occurs rarely, usually in the form of a circle indicating the end of a phrase. In poetry, the only formal pauses indicated are at the ends of verses and stanzas. The function of particles is, therefore, very important to our understanding. This paper is the result of the writer's interest in the form, meaning, grammatical usage and provenance of the particles occurring in Old and early Middle Khmer.

The period under review saw great change and development in the written Khmer language. The early texts have been regarded as struggling attempts to write prose; the complexities of Angkorian sentences seemed to show how difficult it was for the language to express the growing complications of Angkorian life. It is only on the Middle Khmer inscriptions that an easy style was found.

In order to consider the use of particles in these ten centuries, the following texts, (given here with an abbreviation which will be used throughout), have been used:

P.A. Pre-Angkorian inscriptions, seventh–ninth centuries A.D.;
A. Angkorian inscriptions, ninth–fourteenth centuries A.D.;

1. The transliteration used for all Khmer citations is that of Lewitz (1969). For modern pronunciation the writer's transcription is used (Jacob 1968).
Mid. Middle Khmer inscriptions. Special attention was paid to texts dated between A.D. 1574 and 1630 (qqv. Lewitz 1970-72, nos. 1-16c).²

R. *Reamker,* the Khmer literary version of the *Ramāyaṇa.* Special attention was paid to the first five Parts or 2542 stanzas, much or all of which was probably composed during the sixteenth century A.D. (qv. Pou 1977a: 59).

L. *Lpoek Aṅgar Vat.* Poem about the creation of Angkor Vat, dated A.D. 1620.³

C. The old *Chpāp.* Didactic moral poetry, regarded by Pou and Jenner (1979: 134) as dating from before the mid seventeenth century.

Mod. Modern, twentieth century. Examples are either colloquial or from twentieth-century writing.

The P.A., A., and Mid. texts of the seventh to seventeenth centuries are pre-eighteenth-century dated inscriptions on stone. They constitute the whole body of Khmer literature in prose. Although all are associated with religious foundations, the subject matter and style develop from the terse pre-Angkor texts chiefly concerned with lands, duties and provisions, to the Angkor texts in which more information is given about individuals or about historical events, territorial rights, disputes, etc., and to the Middle inscriptions in which the good works and Buddhist fervour of royal and other persons are described in fluent prose.

The *Reamker* represents, along with the old *Chpāp,* the earliest extant Khmer poetry. Though undated, it may be assigned to the early Mid period (at least as far as the early part of it is concerned). The oldest stanzas may well precede the earlier Mid. inscriptions. The first five Parts (i.e. the first 2542 stanzas) have been examined thoroughly and the results used for this paper because it was felt that the text would supply useful additional information, being a continuous narrative with many colloquial or semi-colloquial passages and having a more varied content and style than the inscriptions.

It was not expected that the *Lpoek Aṅgar Vat* would produce as great a variety of syntactic constructions as the *Reamker,* because insofar as it has a story, the narrative style is very simple and the non-narrative part of the poem is, on the whole, descriptive. The reason for its inclusion here as a text for examination is that it is now dated and the date, A.D. 1620, falls within the Middle period chosen for this study.

The old *Chpāp,* though assigned to a suitably early period, have a restricted range of syntactic constructions owing to the gnomic character. They were included for the sake of completeness.

The results of this research are presented in three sections. The first consists of comment, under various headings, on all the particles studied;

---

² The so-called "Middle" Khmer period is from the 16th to the 19th centuries A.D. Only the early Middle inscriptions (for which see Lewitz 1970-72; Pou 1977b; Khin Sok 1978, 1980a, 1980b) are relevant to this paper.
³ See Pou (1975a) for dating, Aymonier (1878) for text.
the second comprises six tables which show the grammatical usages of all the particles at various periods; the third gives citations from a text for each particle with each grammatical usage in each period. The indications of word categories are made in accordance with the writer's analysis of Khmer (Jacob 1968: 330-2).

SECTION I: Comment

Forms of the particles

As may be seen from looking at the examples, many Khmer particles have operated in the language constantly from the seventh to the twentieth centuries. Some, eg. neř, noř, nai and sot are still spelled exactly as they were in the P.A. period. Others have undergone spelling changes in accordance with regular phonetic or orthographic developments.4

However, not quite all the changes of spelling can be explained away. The form njn would not be anticipated from old Khmer nu. It has developed alongside nu ~ nū. There was confusion over the word already in pre-Angkor Khmer, when it was written nau occasionally. Modern Khmer uses the form njn, pronounced according to modern rules, except in formal circumstances, when nūv may be used for 'with' instead of njn. nūv is, however, then pronounced exactly like the verb and particle nau.

The particle pi also develops in a puzzling way. The modern forms of the combinations toempī 'so as to', sūmpī 'even (with following noun)', gappī 'it behoves (one) to', are as expected (all are heralded in Mid. or C.) but, alongside pi in the combination doḥ pi in Mid. (See Table 1: Condition), is the form peh, with the same meaning 'if'. In the Mid. period, open syllables were spelled with final -h. The form we are considering therefore is pe, Mod. poe 'if'. Were the forms pi and pe, future pi and poe, variants of the same word? If so, why did pi have such an unprecedented variant? Lewitz (1972: p. 110, n.3) does not comment on the relationship, if any, between the two words pi and pe, but relates pe to Mid. poe.

Compounded particles

Various combinations of particles occur, seeming to answer a need for reinforcement or greater precision, e.g. the relative pronoun man occurs in the P.A. period in combination with ta, another particle which fulfils the

4. Improved devices, copied from the Thais, for representing different vowels, produced the changes from ley to loey, dep to doep, tempi to toempī, hey to hoey and tel to tael. Vowels in open syllables were probably always pronounced long but were rarely written so; the more modern spellings of nu, ra, ru, pi, and amvi mark the long vowel. The diphthongisation of vowels in Angkorian Khmer, followed by a return to a pure vowel, is shown in the spellings of vōn ~ vvaŋ ~ bunn, svaŋ ~ sīŋ, loh ~ luŋ ~ luŋ. Old Khmer initial i became an imploded, voiced dental consonant, spelled with i. The diacritic ‘ was added in modern times to mark the characters t and k off as complete words. r was frequently written for the initial consonant which was later written b. Finally, the vowel o became an inherent vowel in many modern words and was shortened before h. Hence, phon ~ phan and daḥa ~ doḥ. Old spellings with final h instead of ḥ account for the transcription daḥa. The transcription kāla is due to omission of the killer sign in the Khmer spelling.
relative pronoun function. *man* also occurs in combination with *gi*, both as a relative pronoun and as initiator of discourse, 'i.e., then...'.

*gi* also seems to reinforce *pi* in its occurrences in which it may express either purpose or result. By the Mid. and Modern periods, however, *pi* is combined with *tem* and clearly then expresses purpose. In literary Khmer, *pf* 'as it were' is much used to introduce attributive verbs. Gradually *pf* in this use was combined with *hak* and *tuc*. *hak* *pf* *tuc* became current in modern times as a literary and poetic way to introduce a comparison. In C. *sumpf* 'when; even' and *gappt* 'it behaves' are attested.

The use of *nau* as a final phrase particle, (f.) meaning 'still' is in evidence in modern Khmer only in combination with *foey* and usually in negative statements.

One sequence of particles which must not be construed as a combination of constituents which reinforce each other is *tvrzt* *falt* 'never'. Here, both particles contribute different meanings 'not' and 'one who', respectively.

**Changes of word-order ~ category**

A sphere in which changes seem to have taken place both within one period and from one period to another is that of the word order in which some forms occur. In the case of particles—which are catalysed according to their position in relation to verbs, nouns and phrases—this amounts to a change of word category. Five areas of change (with a possible sixth) have been observed:

1. **Position of the demonstrative particles** *neh* 'this, these' and *noh* 'that, those'

   Tables 3 and 4 and the examples show these particles preceding the noun in the pre-Angkor period but following it from the Angkor period onwards. Although the movement from pre-Angkorian to Angkorian Khmer involves a change of area as well as of time, such a reversal as this does seem unusual (and the P.A. order is uncharacteristic of Mon-Khmer). Jenner (1982) discusses the possibility that *neh/noh* (and often *gi* ~ *gui*) located before the noun are, in fact, the 'phrasal head', while the noun is attributive to it. *neh sre* or *gi neh sre* would then be 'These are the ricefields'. It is tempting to see some such explanation of the word order, though one has also to explain sentences such as *oy gui noh sre sot* (K.79.18) 'gives this ricefield also' where *sre* is required as the object of the verb *oy*. Another possibility would be that *sre*, if given modern punctuation, would be in parentheses or follow a colon: 'gives this (the ricefield) also'. Whatever theory is offered, however, it is very difficult to explain why there is no occurrence at all in P.A. of what we would call the normal order.

5. A change of word order has also taken place in Khmer in the sequence of numeral and classifier. (See Jacob 1965: 161-2).

6. That is to say: a recognisably consistent form with a recognisably consistent meaning has to be regarded as performing more than one grammatical function.
2. Position of ru~ rūva~ rū 'like, as'

Tables 3 and 4 and the examples show that this particle occurs within the pre-Angkor period both preceding and following the noun. Perhaps the English use of the word 'like (such as)' which may also occur before and after a noun, e.g. 'like the moon' and 'flower-like', is comparable?

3. Position of ni~ nī 'with reference to; at'

It seems possible that ni~ nī, which occurs in both Tables 3 (P.A. and A.) and 4 (A. and Mid.) with meanings which could be equated, is another example of changed position. Occurrences of nī following a noun are rare and very restricted. They involve the words muh 'nose' and vnek 'eyes'. It is to Pou (1976: 340-1) that is owed the interpretation of the phrases muh nī and vnek nī as 'before' and 'in the presence' respectively.

4. Position of hey~ hoey 'already; and then, and now'

In both Mid. and Mod. Khmer, as Tables 1 and 2 show, this form carries out two functions: that of marker (m.; occurring clause-initially), and that of final phrase particle.

5. Position of nai 'of'

The particle nai has the same spelt form, the same translation 'of' and the same pre-nominal use throughout Khmer, as Table 3 shows, but it also occurs phrase-finally in R., L., and C. with back reference to the subject of the sentence: 'of him/her/them', as is illustrated in the examples.

In Khmer poetry, sequences of alliterating words which do not convey much additional meaning to the sentence often occur at the end of a verse or stanza, where they are useful as rhymes or to provide the required number of syllables (qv. Jacob 1979: 124-5). Such sequences are ban' bek 'very', kräs' krael 'very', neh nā /noh nā 'this, these'/ 'that, those'. neh nai / noh nai also occur in poetry. There are such occurrences with nai in R., L. and C.

6. Position of ai (?)

R. also provides an example of ai, normally a pre-nominal particle (pre-n.p.), in phrase final position:

R.1.29.1-2. doep an jöe brah staeñ thlai, baňs brah nărây(n) ai, isūr nai lok traitā 'Then will I believe you, honoured sir, to be in the line of Nārāyaṇa, lord of the world, in the traitā'.

Foreign borrowings and Thai comparisons

ukkan < Skt. ukra 'also' occurs phrase-finally in Khmer and has, therefore, been entered on Table 2 as f. It is attested only in the pre-Angkor period, whereas sot performed a similar role throughout all periods.

kāla, kāl < Skt. kāla 'time' is current throughout Khmer as a noun but also may be catalysed as a marker 'when' in the Angkor and modern periods.
ta, ta is a general dependent particle (g.) which may occur before words of any category, and may also act as a marker (m.). It seems just possible that ta was related to Skt. tā, the base of the demonstrative pronoun, as the present writer suggested on an earlier occasion (Jacob 1977: 166-7).

kār, kā. Occurrences on P.A. inscriptions of the form ka show no similarity to the particle kār, kā. Some may be construed as ka ‘to construct’. Some may be a shorthand form of kānum ‘serf’. In A. Khmer, the writer has noted occurrences of ka on only one inscription, K 34, of the tenth century; several very clear examples occur there, however. It seems extremely likely that it was borrowed by the Thais; it is attested on the Sukhothai inscriptions (Ishii et al. 1977). The particle has been treated here as both marker and pre-verbal particle (pre v.p.), both of which functions it performs fully in Modern Khmer. For the Mid. inscriptions under review there is only the one occurrence cited in examples where it is m. In R. it occurs frequently and seems always to be immediately pre-verbal. In L. it occurs pre-verbally, but does not in fact occur very frequently.

ktī. (f.) ‘whether...or’. ktī appears first in Mid. Khmer and was already well established in Sukhothai Thai. There is nothing in its form to prevent the theory that it came from Thai to Khmer. Having co-existed with the Khmer form of the same meaning, lah, during the Mid. period, ktī supplanted it completely.

tī. Table I and the examples show tī ‘earth; direction; focus’ in use as a marker. It has often been described as an indicator of the passive voice but in the opinion of the present writer the passive voice occurs only in the translation and there is no passive voice in Khmer. Clauses following tī have a form no different from other statements translated by the active voice. It seems rather to convey ‘person or thing in respect to which’.

The idea that in this usage tī indicates the focus of attention immediately provokes comparison with Mod. dī in its use, for example, at the beginning of a letter: mak A. jā dī ralīk ‘To A. whom I miss’ (lit. come A. being subject of regret-absence). Lewitz (1971: pp. 115-16, n. 8.) noted the similar usage when editing IMA 4. Thai /thī/ is also used to introduce the focus of attention and was well established at the Sukhothai period. Did the Khmers in the Mid. period borrow Thai /thī/? If so, either it was a literary loan based on spelling or Thai written initial d was not then aspirated when pronounced. (Similar parallels occur with Khmer ghī and Thai khuu ‘that is’, Khmer dān’ and Thai /than/ ‘catch up’, but for these pairs there are Old Khmer occurrences with the expected initial consonants (gi and dan) and one assumes the loan was from Khmer to Thai. Whatever the direction of the borrowing of dī, the Khmers acquired from the Thais the use of it to form an ordinal numeral system. In R., dī seems rather like a new borrowing. It occurs with dis ‘direction’, des ‘country’, uttar ‘north’, and not alone. Although the Khmers may have borrowed dī from Thai /thī/, did the Thais borrow the focus usage from the earlier Khmer tī? Lewitz also suggests the possibility that Thai /thī/ was itself the borrowed form of Angkorian tī.
A diachronic survey of some Khmer particles

tae. This particle seems clearly to be a loan from Thai since the parallel Thai form taa was well established in the Sukhothai period and tae appears for the first time on the Mid. Khmer inscriptions. It is entered as a general particle on Table 6. Its limiting effect on meaning requires many different translations according to context: 'only; just; exactly; simply (and no more)'. It seems to the present writer that tae has very much the same effect on meaning which the final phrase particle guh ~ gus had during the A. and Mid. periods; in this case of replacement, however, a different word category and grammatical role are involved too. tae and guh occurred alongside each other in early Mid. Khmer but gus then became obsolete.

However, there is a need to present a case for the comparison of guh with tae. Lewitz (1970: p. 104, n.10) when translating a very difficult passage, at the beginning of her editions of the Mid. inscriptions, said that, with the negative particle, it means 'not... at all, absolutely not'. For guh in affirmative contexts, in IMA 3, 4 and 6, she gave no translation. There are A. period occurrences of gus ~ guh where the translation 'only, exactly, just' suits very well. Two are cited in the examples. And later, Lewitz herself translated guh as 'seul' and 'seulement', when editing the Mango Grove inscription. She explains (Pou 1978: p. 354, n.3) that she is taking 'l'adverbe guh "exclusivement"' with the preceding verb. (This reference back to the preceding verb is a function characteristic of all particles catalysed as final phrase particles.)

broh 'because'. This word is mentioned here because there are reasons for supposing that it also may have been borrowed from Thai. (In recent times it is also frequently combined with the verb 'to say', as it is in Thai.) Although Thai phrq /phro/ 'due to, because or is attested on the Sukhothai inscriptions (Ishii 1977), it does not occur in Khmer until the modern period. (It does not, for example, occur on any of the Inscriptions modernes d'Angkor nor in the old Cpap'.) Nacaskul (1962: 185-7) held that the direction of borrowing for the comparable words Thai ko/ko/ 'island' and Cambodian koh /koh/ 'island' must have been from Cambodian to Thai because otherwise the Cambodian form would have had the final glottal stop ~k. However, it seems possible to the writer that the loan may have been of a more literary and less colloquial nature. It may have been the spelled rather than the spoken form which was borrowed; both forms have the same written vowel and the final consonant which goes back to the same Indian symbol, the visarga.

Jenner and Pou (1980-81) construe broh as < roh 'form, shape, manner', thus suggesting a Mon-Khmer origin. Did the Thais originally acquire broh from Khmer?

The use of particles in R. (The Reamker Parts 1-5)
The R. occurrences have not been entered on the tables except when they are unusual or constitute the only mid-Khmer entry but they are given in the examples. It was possible to find every particle and usage for which
JUDITH M. JACOB

there was a Mid. inscriptions occurrence except:

*gi pi. pi* occurs in R., as it does on the Mid. inscriptions, expressing purpose. Perhaps it is by chance only that *gi pi* was not also found to occur in the stanzas examined. *gi* does occur with *tà* several times, eg. 1.7.5-6.

*nū ~ nūv. nūv* does not occur as a marker initiating discourse. Owing to the nature of the text, speech is initiated in many other ways: by exclamations of surprise or woe, for example, and frequently by the use of the title or name of the person addressed.

*bī*. No occurrence of *bī* meaning ‘at (past time)’ was found, but its absence may be due to chance.

*be*. This is a non-occurrence which is not surprising because many other ways of expressing superlatives are used in R. and the poetry which followed it, e.g. *kanlân* ‘surpassing’, *krai* ‘very much’, *loes* ‘over and above’. *be* does, however, occur in L. and C.

*man* occurs only with *gi*, and then not as an initiator of discourse. *man* was gradually being less used; one might expect that it would be used in combination. *man gi* in R. may be understood as ‘in that it is’.

*roh* was soon to disappear from the language. Perhaps, too, it was rather formal for poetry? *tǎm* and *tòy* ‘in accordance with’ were soon to replace it.

*lah ~ loh* ‘whether...or’. The newly borrowed *kīl* (see p. 20 above) occurs several times in R. but not *lah*.

*viṇ* occurs as a verb ‘to return’ but not as f. ‘back again’.

There are, on the other hand, occurrences in Parts 1-5 of the *Reamker* of particles or usages which are not be found in the *Inscriptions modernes d'Angkor* 1-16c (Lewitz 1970-72):7

*kūv, nā f.* (Emphasis). The first belongs to poetry, the second, in Mod., to the colloquial.

*nūv, nāi*. These forms fulfil a function, as a link between verb and object, which is new to them and in which they begin to replace *ta*. (See Table 3.) *nūv* continues to behave in this way in Mod. literary Khmer.

*nāi*. f. ‘of’. This is the use with back reference to the sentence subject, discussed under ‘Changes of word order ~ category’ (p. 19 above).

*nau*. f. ‘still, continuously’. R., like L. and C. (see below), provides occurrences in Middle Khmer, a link between the A. period and Mod. *nau toey.*8

7. Further particles which are not on the Mid. inscriptions 1-16c, or in Mod. Khmer, e.g. *lgik* in *lgik poe, leh* (related to Mod. *mleh*. See Pou 1977a: 120-4) occur in R. R. occurrences are not varied or numerous enough to make it possible to establish their word categories.

8. Did this use of the verb ‘to stay, live, remain’ as a final particle meaning ‘still’ go to Thai from Khmer? cf. Thai use of *yūu*.
A diachronic survey of some Khmer particles

... pre n.p. 'together with'. An unusual usage but one which is attested elsewhere.

ra, vvaem fael. It seems surprising that they are in R., and not in the Mid. texts but this may just be by chance.

In addition to the above R. usages, the particles kā and tā have further uses in R. which are not in evidence in the other texts under investigation (i.e. up to A.D. 1630) and which the writer does not remember meeting in later literature. Examples of these uses are given separately here and they are not entered in the tables since it is felt that they are on the border-line between grammar and poetic style or licence.

kā. the normal Mid. use of kā to introduce a slight consequence 'and so' occurs passim. The following two citations illustrate the extra use:

1.14.1-5. yoeō giṭ smān jā randaḥ... neh ḫiu ḫūt ḫūt aṅ thā ṭoeṃ braḥ rūmādhirāj kā loek ṭādhi ḫanū
‘We thought it was claps of thunder... (but) now the envoy tells me about Prince Rām, how he raised the mighty bow’.
the cause: Prince Rām, who

2.2.1. sūrya saeṅ cāṃṅ caeṅ e ambar ḫūc braḥ candr car tracaḥ kū beṅ pūrnāṁī
The light of the Sun shining in the heavens was like that of the bright Moon, when it/which is full'.

kā seems to be a link, like a relative pronoun, between the clauses. No consequence is expressed.

tā is very much in evidence in R. Apart from its use as relative pronoun, as general dependent particle linking an attribute to a noun (the attribute including possessor or a noun in apposition) and as link between verb and object—all of which are illustrated in the examples and entered on the tables—tā also occurs between two nouns or two verbs of similar meaning, e.g.

2.23.1. kal' tā māyā 'tricks, wiles'
1.1.1. prasoe r tā uttam 'admirable, superior'.

Evidence from Lpoek Aṅgar Vat

The following usages which were found in R. and not on the Mid. inscriptions occur in L.: the use of nai as a final phrase particle and as a pre-nominal particle linking verb to object; the use of nā as a final phrase particle adding emphasis and the use of nau as a final phrase particle. (See examples.) L. also corroborates the two further uses of kā and tā just illustrated above from R.:

9. I am grateful to Miss A. Compain, a missionary relief worker, for the information that speakers from Battambang use phañ in this way. She also drew my attention to similar usages in the translation of the Bible, where phañ is used pre-nominally in conjunction with rapas'.
as a kind of relative pronoun. 282.6. chlāk' rūp mahā yaksā vīrād muhimā kā chak’ setā nām rat’.

‘(They) had sculpted a relief of the mighty yaksā Birādh who/as he was seizing hold of Seta to run off with her’.

between two nouns. 288.6 asur ṭā dāhān ‘demon soldiers’.

between two verbs. 289.14 gravās ṭā gravī ‘swinging-up-their-arms and brandishing (them)’.

It is noticeable, however, in reading L. that both ṭā and pī ‘as it were’ are used much less frequently than in R.

Evidence from the old Cpāp’
The composers of the Cpāp’ did not use a great variety of syntactical constructions. They tended to repeat ‘do’s’ and ‘don’ts’ and to prognosticate that ‘if’ certain actions were taken, a good or bad result would ensue. They used a familiar style of discourse which gives a modern colloquial impression. The texts do, in fact, furnish some very interesting occurrences of particles and in one case a non-occurrence of an expected one:

is occasionally used to attach an attribute to a preceding noun. Otherwise it has one, perhaps fossilised, use with lok or paralok, e.g. Hai Mahājan st. 2. oy pān kti gāp’ta lok paralok ‘and so gain advantage in this world or the next.’

toe. m. (Introduces a question). This word, familiar in Mod., was not found elsewhere in the early texts. It occurs in Kūn Cau st. 5. dic toe jīnmān ‘It stings—say, how much?’

min. C. and L. provide what seems to be the earliest evidence for this colloquial negative particle, much used in Mod.

does not occur alone, as it does so frequently in R.; kum pī ‘do not’ is very common. Certain combinations with pī, not found elsewhere in the early texts, occur in C.: sūmpī, translated as ‘lorsque’ by Pou but now meaning ‘even’, and gappī ‘it behoves...to’. Both doh pī and doh poe occur in Hai Mahājan (st. 49 and 80 respectively) which strengthens the theory mentioned in Section I above that they are forms of the same variant word.

m. ‘because’. The only earlier occurrences which have been noted of hetu as a clear marker of cause were those in A., where hetu is combined with man. In C. hetu is also used in combination with tpid, and, as in Mod., with tae.

most interestingly does not occur at all in the old Cpāp’, but it does occur, for example, in the later Cpāp’ Prus, st. 65.

Summary
It will have been observed that, not unexpectedly, some forms occur on more than one table, eg, hey ~ hoey, which functions as both m. and f. (Tables 1 and 2). However, there are also some forms, such as pī and man
A diachronic survey of some Khmer particles

(Table 1) and *ta* (Table 3) which occur in more than one section of a particular table; different meanings, semantic or syntactic, are indicated for the different sections in the left-hand column. It might be argued that two sections should be combined if they contain the same Khmer word, performing the same function, and the English ‘translation’ is the only difference between them. However, justification for the separation is claimed as follows: *pi* may introduce a clause of Cause or Purpose, which should surely be regarded as syntactically separate; *man* occurs in one section with *nu* and in another with *tel*, though *nu* and *tel* do not overlap. *ta*, when following a verb of giving, buying or speaking, links it to the indirect object, but in other contexts provides a link with the direct object.

The apparent confusion in P.A. and A. concerning cause and purpose (both introduced by *pi*) leads one to the conclusion that the meanings which Coedès worked out by reference to the whole inscription and to information from other inscriptions, etc., despite the lack of clarity in the text, were easily understood at the time because people were familiar with the whole context. We therefore return to what was said in the opening paragraph of this article, that Khmer syntax may be inexplicit unless there is a need for precision.

With regard to the borrowing of particles from Thai, Huffman concluded (1973: 491-502) when comparing modern Thai and Cambodian syntax, that Cambodian syntax was strongly influenced by Thai. It is significant too that, at the same early Mid. period considered here, a variety of Thai vocabulary was being borrowed by the writer or writers of the *Reamker* as a whole.

R. and C., and even L., offered likely material, heralding Mod. in one way or another. It was interesting to find that R., though probably predating Mid., had no *roh* or *lak*; that in L., while *tā* and *pī* occur, particularly in the passage describing the Angkor bas-reliefs, linking attribute to noun, their use is much less than in R.; that in L. and C., but not in R., *bek* and *min* are attested; and that *kā* does not occur in C. It is possible that some help towards dating the *Reamker* and early *Cpāp* may come from an examination of the grammatical forms but great circumspection is needed.
SECTION II: Tables

(R., L. and C. occurrences are entered in the tables only if there is no Mid. occurrence from the inscriptions.)

<table>
<thead>
<tr>
<th>TABLE I: Markers (m.)</th>
<th>Indication</th>
<th>P.A.</th>
<th>A.</th>
<th>Mid.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'when (past)'</td>
<td></td>
<td>kāla</td>
<td>R.</td>
<td>kāl</td>
<td></td>
</tr>
<tr>
<td>'when (by the time)'</td>
<td></td>
<td>luḥ</td>
<td>R.</td>
<td>luḥ</td>
<td></td>
</tr>
<tr>
<td>'until'</td>
<td></td>
<td>luḥ</td>
<td>R.</td>
<td>[lāl' (kāl ŋā)]</td>
<td></td>
</tr>
<tr>
<td><strong>Condition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'if'</td>
<td></td>
<td>dāh</td>
<td>R.</td>
<td>dōḥ</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>dōḥ pi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>dōḥ pe</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>(peḥ (=pe))</td>
<td></td>
</tr>
<tr>
<td>'whether'</td>
<td></td>
<td></td>
<td>R.</td>
<td>doḥ (+kti f.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>doḥ</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>doḥ pi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>doḥ jā</td>
<td></td>
</tr>
<tr>
<td><strong>Cause</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'the reason... was because'</td>
<td></td>
<td>pi...gi pi...</td>
<td>(ta) pad</td>
<td>[pān jā...bī broḥ]</td>
<td></td>
</tr>
<tr>
<td>'because'</td>
<td></td>
<td></td>
<td>R.</td>
<td>[tpi]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>[broḥ]</td>
<td>[tadv]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>[toy]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>hetu tae</td>
<td></td>
</tr>
<tr>
<td><strong>Purpose</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'so as to'</td>
<td></td>
<td>pi</td>
<td>R.</td>
<td>pi</td>
<td>kem...oy</td>
</tr>
<tr>
<td>'so that...(not)'</td>
<td></td>
<td></td>
<td>R.</td>
<td>kampi (so that...not)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>pi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>gi pi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>tempi</td>
<td>țoemp (nîn)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>tempi nu</td>
<td></td>
</tr>
<tr>
<td><strong>Result</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'and so, and then...'</td>
<td></td>
<td>gi pi</td>
<td>R.</td>
<td>dep</td>
<td>doep</td>
</tr>
<tr>
<td>'with the result that'</td>
<td></td>
<td></td>
<td>R.</td>
<td>dep</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>kār</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R.</td>
<td>[pān jā]</td>
<td></td>
</tr>
<tr>
<td><strong>Relative pronoun</strong></td>
<td></td>
<td>ta</td>
<td>R.</td>
<td>ta</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ta</td>
<td>R.</td>
<td>ta</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>man</td>
<td>R.</td>
<td>man gi</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>tel</td>
<td>R.</td>
<td>nā</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. tel</td>
<td>R. tel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A diachronic survey of some Khmer particles

Indication

Focus of attention
‘with respect to’
‘which’

Reported speech
‘that...’

Initiation of discourse
‘Now...’

Exhortation and Optation
‘let...(not)...’
‘may... (not)...’
‘do...not...’
‘may...’

Co-ordination
‘and (then)’

TABLE 2: Final phrase particles(f.)

<table>
<thead>
<tr>
<th>Translation</th>
<th>P.A.</th>
<th>A.</th>
<th>Mid.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘already; by now, by then’</td>
<td>lhey</td>
<td>hey</td>
<td>hoey</td>
<td></td>
</tr>
<tr>
<td>‘also’</td>
<td>ukk</td>
<td>sot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘and...too’</td>
<td>viñ</td>
<td>viñ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Emphasis)</td>
<td>ra</td>
<td>R. ra</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hoñ</td>
<td>hoñ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Possession)</td>
<td>gus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘still, continuing’</td>
<td>nau</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘utterly; (not) at all’</td>
<td>ley</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘very much; too much’</td>
<td>lah</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘whether...or...’</td>
<td>lah</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

205
<table>
<thead>
<tr>
<th>Translation</th>
<th>P.A.</th>
<th>A.</th>
<th>Mid.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>'according to'</td>
<td>roḥ</td>
<td>roḥ</td>
<td>roḥ</td>
<td>[tām, āy]</td>
</tr>
<tr>
<td>'all'</td>
<td>iss</td>
<td>is</td>
<td>is</td>
<td>as'</td>
</tr>
<tr>
<td>'As to...'</td>
<td>ṛi</td>
<td>ṛiy</td>
<td>ṛi</td>
<td>ṛi (lit.)</td>
</tr>
<tr>
<td>'at, in (place or time), with...</td>
<td>āi, āy</td>
<td>āi, āy, ta</td>
<td>i, e. ai, ta, ta</td>
<td>e (place)</td>
</tr>
<tr>
<td>regard to'</td>
<td>nā</td>
<td>nā</td>
<td>nā</td>
<td>nā</td>
</tr>
<tr>
<td>'by (the time that)'</td>
<td>loḥ, lvaḥ</td>
<td>lvaḥ ta gi</td>
<td>luḥ</td>
<td>[tal']</td>
</tr>
<tr>
<td>'from'</td>
<td>am̃vi</td>
<td>am̃vi</td>
<td>am̃bi</td>
<td>am̃bi, bī</td>
</tr>
<tr>
<td>'in'</td>
<td>kam̃luḥ</td>
<td>kam̃luḥ</td>
<td>knuŋ</td>
<td>knuŋ</td>
</tr>
<tr>
<td>'like, as, similar to'</td>
<td>ru</td>
<td>ru</td>
<td>ru</td>
<td>ru</td>
</tr>
<tr>
<td>(Link between verb and object)</td>
<td>ta</td>
<td>R. ta</td>
<td>R., L. nai, nūv</td>
<td>nūv (lit.)</td>
</tr>
<tr>
<td>'near'</td>
<td>ŋyaŋ</td>
<td>ŋyaŋ</td>
<td></td>
<td>jīt</td>
</tr>
<tr>
<td>'of'</td>
<td>nai</td>
<td>nai</td>
<td>nai</td>
<td>nai (lit.)</td>
</tr>
<tr>
<td>'this, these/that, those'</td>
<td>neḥ/noḥ</td>
<td></td>
<td></td>
<td>[rapas']</td>
</tr>
<tr>
<td>'to (a place)'</td>
<td>loḥ, lvaḥ</td>
<td>lvaḥ</td>
<td></td>
<td>[dau]</td>
</tr>
<tr>
<td>'to, for, from (re giving, speaking and buying)'</td>
<td>ta, ai ta, āy, ta</td>
<td>ta</td>
<td>[dau]</td>
<td></td>
</tr>
<tr>
<td>'towards (N.S.E.W.)'</td>
<td>ti</td>
<td>ti</td>
<td></td>
<td>[kẖān]</td>
</tr>
<tr>
<td>'with, and'</td>
<td>nu</td>
<td>nu</td>
<td>nu</td>
<td>nū (lit.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>nā</td>
<td>nā</td>
</tr>
<tr>
<td>'with, by means of'</td>
<td>nu</td>
<td>nu</td>
<td>nū</td>
<td>nūv (lit.)</td>
</tr>
<tr>
<td>'with reference to'</td>
<td>nī</td>
<td>nī</td>
<td>nī</td>
<td>nī</td>
</tr>
</tbody>
</table>

206
A diachronic survey of some Khmer particles

**TABLE 4: Post-nominal particles (post n.p.)**

<table>
<thead>
<tr>
<th>Indication</th>
<th>P.A.</th>
<th>A.</th>
<th>Mid.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>'all'</td>
<td>phoň</td>
<td>phoň</td>
<td>phoň</td>
<td>phaň</td>
</tr>
<tr>
<td>Interrog.: 'which?'</td>
<td>nā</td>
<td>nā</td>
<td>nā</td>
<td></td>
</tr>
<tr>
<td>'like, as'</td>
<td>ru</td>
<td>ruv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location, Reference</td>
<td>ni</td>
<td>ni</td>
<td></td>
<td></td>
</tr>
<tr>
<td>'this, these/that,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>those'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 5: Pre-verbal particles (pre-v.p.)**

<table>
<thead>
<tr>
<th>Indication</th>
<th>P.A.</th>
<th>A.</th>
<th>Mid.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Future time,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>possibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negation</td>
<td>voṁ</td>
<td>vvaṁ</td>
<td>bvum</td>
<td>buṁ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L,.C. min</td>
<td>min</td>
</tr>
<tr>
<td>'never'</td>
<td>vvaṁ tel</td>
<td></td>
<td></td>
<td>buṁ țael</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R,.L,. C. buṁ țael</td>
<td></td>
</tr>
<tr>
<td>(Joins plural subject</td>
<td>syaň</td>
<td>syiň, syiň ta</td>
<td>(șiň = almost)</td>
<td></td>
</tr>
<tr>
<td>to verb) 'all'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Result</td>
<td>ka</td>
<td>kăr</td>
<td>kā</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 6: General particles (g.)**

<table>
<thead>
<tr>
<th>Indication</th>
<th>P.A.</th>
<th>A.</th>
<th>Mid.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>'all, complete with,</td>
<td>daň</td>
<td>daňň</td>
<td>dāňň</td>
<td></td>
</tr>
<tr>
<td>as well'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Link between noun</td>
<td>ta</td>
<td>ta</td>
<td>ta</td>
<td>ță (lit.)</td>
</tr>
<tr>
<td>and attribute)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'only'</td>
<td></td>
<td></td>
<td></td>
<td>tae</td>
</tr>
</tbody>
</table>
SECTION III: Citations from texts

The particles entered in the tables are listed here in syllabary order under the heading of the oldest spelling, with citations illustrating the uses.

P.A. and A. citations are given with their Khmer (K.) reference number, as given in Coedès, face and line. Mid. citations are given with their Lewitz (1970-72) number, face and line.

R. citations have a reference to part, page, and line in the Reamker text of the Institut Bouddhique; L. to page and line; C. to the Cpäp' title and the stanza. L. and C. citations are given only if no Mid. inscriptions occurrence has been found.

Where Mod. particles are related to the older forms, an example is given for comparison.

\[ ka, kār, kā \]

[P.A. occurrences of the form \textit{ka} show no similarity to the particle, so may be construed as > \textit{ka} ‘to construct’; some may be a shorthand form of \textit{khiūn} ‘serf’.

‘and so; accordingly’: Khmer m. or pre-v.p.

A. 349 sud 15. \textit{ka gi nōh phdaï krom nōh ta jā daksina hoñ ‘and so it was these very territories presented as gifts’

Mid. 8.30 kār A, B, C, etc. 10 (40) jā saksi ‘And so A, B, C, etc. were witnesses’.

R.1.9.11-12 stec moel as’ bhūdar loek rādhī dhanū kā broes brah bhāktr prim priy ‘As he watched all the kings (trying to) raise the mighty bow, an eager look came into his handsome face’.

Mod. nān nūn khīn knān’ nūn khīum hoey khīum kā khūc khāt mittabhāb rapas’ yōen ‘You will be angry with me and I shall lose our friendship’.

Mod. \textit{poe dau kā dau}. ‘If we’re going, let’s go.’

kāla, kāl ‘when (past time)’ :Khmer m.

A.669.B.3. kāla samrāc homa.. ‘When he had completed the sacrifice,’...

R. 2.67.12. kāl brah riem stec mak phgāh as’ mahā rśi phān sīn ar daduol brah ang ‘When your brother came here and greeted all the great sages and paid them homage, they were delighted to receive him’.

Mod. e nandaka: hoñ, kāl samrec kīṭ prāṭhnā khluon hoey, kā kāp lā‘ ‘As for Nandaka, when he had achieved his own wish, he took his leave respectfully.’

kūv (emphasis): Khmer f.

R.2.65.8. būm tīn toem brah bhīrut jā cneh kūv nā. ‘did not know Bhirut was in this mood’

C. Kūn cau. 62. tae anak muoy kūv ra-ae ra-ūv ‘Now another person, constantly complaining’.

10. In the citations and their translations, A,B,C, etc. are space-savers used to replace the individual names of persons mentioned.
A diachronic survey of some Khmer particles

kum (exhortation): Khmer m.

R. 2.50.2. amespace. ‘And as for me, let me not be negligent’.
C. Hai Mahâjan. 58. kum khin prakaek jhlo. ‘Don’t be angry, arguing and quarrelling’.
Mod. kum dhvoe. ‘Don’t do it’.

kampi, kum pê, kummpi

(i) (Exhortation): Khmer m.
P. A. 724. kampi tve ‘gara le gah. ‘Do not build houses on the grotto’.
Mid. 6. B. 8. kampi oy mok pyen pyen kambujadesa. ‘Let them not come to oppress Cambodia’.
R. 1. 39. 6 kum pê pê git can dos. ‘Do not think of initiating a wrong’.

(ii) (Purpose): Khmer m.
Mid. 3. A. 31. lek yas brah rajasatän... kummpi mân dharmmantarây ley ‘to increase the glory of the royal line... so that there should not be any calamity (for them)’
R. 1. 25. 6 kum pê lanh ‘so as not to be a long time’.

kamlûn ‘in’: Khmer pre-n. p.
P. A. 560. 11. sre kamlûn ml ‘ricefield within (the boundary of) the road’.
A. 291. 22 sre kamlûn vrah mandîra ‘ricefield within the royal palace (grounds)’.

kti ‘even (if)’, ‘whether... or’: Khmer f.
Mid. 5. 2. doh kûn ân kti kmûy aîn kti ‘whether it be my child or a nephew-or-niece or mine’.
R. 2. 75. 7-8 âc dap’ dal’ khyaî’ dâmû 8 dis pok pak’ kti, ‘can even withstand the winds from all 8 directions’.
R. 2. 73. 6-7. doh brah bhurut ras’ lokî(y), luî paralok kti ‘whether Bhirit is alive in this world or has gone to the next world...’
Mod. doh prus kti sî’ kti ‘whether male or female’.

knuñ ‘in’: Khmer pre-n. p.
Mid. 8. 7. knuñ pammros ‘in freedom’ (lit. in (state of) affranchisement).
R. 2. 61. 9 knuñ brai ‘in the forest’.
Mod. knuñ sâstrâ ríoën rám kerti ‘in the manuscript of the Reamker’.

gi pi ‘and that was why’, ‘being an occasion for’, ‘and so’: Khmer m.
A. 222. 16 man khûum noh rat dau gi pi Tân oy Tai B... ‘That servant ran away; that was why Tân A. gave Tai B...’
A. 450. 14-15. pre gruyâga gi pi vrah oy daksîna ‘ordered (someone) to perform-the-sacrifice-to-the-guru so that (i.e. on that occasion,) His Majesty might give offerings’.
Mid. 6. A. 11-12. rantâp kriyâ pîjî sa râp gi pi nîman brah sâgh phôn ‘prepare offerings in readiness for inviting all the monks’.

A. 56. C. 31. çata 1 caññay gus gi ta bhûmi A. ‘at exactly the distance of 100 (measures of distance) is the territory A’.

209
JUDITH M. JACOB

A. 211.6  "let (them) provide only these provisions, letting them be in accordance with the provisions for the god..."
Mid. 3 A 78-9. "Even if all these people err and go and fall into the four-domains-of-suffering..."
R. 3.31.1. "I went to them simply to serve them for the sake of attaining the highest merit'.

"near": Khmer pre-n.p.
P.A. 155.18. "ricefield near Poï’s reservoir'.
A. 239.33. "ricefields near the sanctuary'.

tā. See ta.

(i) ‘who, which (rel. pronoun): Khmer m.
P.A. 561.20. "anyone who spoils (things) here'.
A. 56.B.27. "persons who plant sugar-cane'.
Mid. 2.5. "the lord who is our refuge'.
R. 2.49.1. "the beloved older Prince Rām who is His Majesty’s oldest son'.

(ii) ‘the one who, the one which, the': Khmer g. (linking an attribute to a noun)
P.A.30.C.9. "the males'.
A.19.21. "all the following goods'.
Mid. 6.B.3. "for a long time indeed' (lit. whole-of time particle, long indeed).
R.1.1.3. "decorated pillars and door'.
Mod. (lit.) "this remarkable idea'.

(iii) ‘to, for, from, (re giving, buying, speaking); ‘at, in (re time, place); (a link between verb and object’ Khmer pre-n. p.
P.A. 90.10. "persons offering fruit to the god'.
P.A. 689 A.11 "buy a territory from Yap A'.
A. 259 sud 8. "buy a territory from Yap A'.
A. 249.1 "In the reign of King...’

11. Jenner (1981) demonstrates, by illustrating the Old Khmer uses of ta, that this particle—though found in a variety of contexts which suggest in translation a wide range of grammatical functions in European languages—functions consistently in Khmer as a linking particle. While agreeing that the role of ta is always to be a mere link between other words, I separate the examples here into three groups because for me ta is classified according to whether it links a clause or an attribute or forms an untranslatable connection between a verb and its direct or indirect object.

210
A diachronic survey of some Khmer particles

Mid. 2.4. *dūl ta brah* ‘humbly-inform the Lord’.
Mid. 6.A.36 *sannak ta vrah sāsthā* ‘resorting to the Lord’s teaching’.
Mid. 3.B.2 *cūl ta pad moks* ‘go on to the way of deliverance’.
Mid 8.24. *kāt ta kti neh jā* ‘decide the case as being’.
R. 1.59.5. *pramād ta brah aṅg* attack you, my lord’.

*ta pad*. See *pad*.

*ta man* ‘who, whom, which’ (rel. pronoun): Khmer m.

*ti*.

(i) ‘earth; place; direction’: Khmer pre-n. p.
P.A. 688.4 *ti tīai luc sruk* ‘to the West of the village’.
A. 158.28 *ti purvva... daksiṇa... paccima... uttara* ‘to the East... South... West... North’.

(ii) ‘focus of action; in respect to which’: Khmer m.
P.A. 927.2. *sre ta tel ti kurāk... tve* ‘ricefield which Kurak... worked’.
A. 246-8. 3-4. *dravya... ti aṅ yok* ‘goods...which I take’.

*tempi, tempi nu, ūempoILTER{nīnh* ‘so as to’: Khmer m.

Mid. 6.A.15 (prepare a sacrifice) *tempi panluh do ta brah nāt* ‘to send (merit) to kin’.
Mid. 3 A.28. *tempi nu sān sāsnā brah tathāgat* ‘so as to build up the teaching of the Tathāgat’.
R. 1.5.3-4 *prae basudhī tempi nīnh dhvoe oy ksem ta praǰārāstr* ‘turn the soil in order to bring peace to the people’.
Mod. *khaṃ rien tempi (niñh) pān caṃneh vijjā l-a* ‘study hard in order to achieve a good education’.

*tel, ūael*

(i) ‘who, whom, which’ (rel. pronoun): Khmer m.
P.A. 561.33 *ge tel poñī A. pre* ‘personnel whom Poñī A. orders’.
A. 958. Nord. 29. *ri bhūmi tel sruk ta jmaḥ kaṃnuṅ sruk* ‘as for the territory which (is in) the locality known as Kaṃnuṅ Sruk...’
R. 4.46.5. *trec traṅ’ paṅ’ ūaṅʁ tel tiṅ saṃṭau dau paṅcalīṅg*. ‘He went his way, casting aside the animals, conscious that (lit. who be-conscious-that) he was heading straight towards Paṅcalīṅg’.
Mod. *cor ūael rat’ dau trūv ge cāp*. ‘The thief who ran off was caught’.

(ii) ‘(not) ever’: Khmer pre-v.p.
A. 235.D.27 *kule neh poñī vvaṁ tel cek mūla* ‘the relatives never shared out (the functions of) the original branch’.
R. 4.11.3. *brah bruṅ ūael māṅ hmaṅ* ‘the lord who has never had any imperfection’.
L. 269.10. *bruṅ ūael aṅ* ‘never failing (to do so)’.
C. Hai Mahajan 40. *bruṅ ūael khus* ‘is never wrong’.

211
JUDITH M. JACOB

Mod. gāt' bum tael dau anģar. ‘He has never been to Angkor’.

tae, tae

‘only, but, just, exactly and no more’: Khmer g.

Mid. 9.46. oy paeñ taeḥ anak ta reh pamphros neh ‘strike only those who have taken these freed people’.

R. 1.29.4. stec yak tae braṅ hasth chveṅ bhūdhar loek silp dhanā. ‘With only his left hand, the supporter of the earth raised the magic bow’.

Mod. nau sal’ tae pī de. ‘There are only three left’.

tpit. See pad.

dañ. See don.

daha, doḥ, doḥ pī, doḥ pī ‘if’: Khmer m.

P.A. 1004.7. dāha māṅ kāryya vṛah kamratāṅ añ ‘if there was work to be done for the lord’.

A. 843.C.13-4. dāha ‘yāt kule añ ‘if there-is-no descendent of mine’.

Mid. 4.C.5. doḥ anak anmpāḷ neḥ vaṅveñ dau ai ta catarāpāyabhūṁ guḥ oy anak anmpāḷ neḥ oy rruc lae[n] ‘if all these persons lose their way in the four domains of suffering, even so let them escape’.

R. 3.12.9. doḥ braṅ pād stec taḷ’ oy braṅ dasabal... bidagdḥ āsūr ‘if the prince comes here, let that lord of the ten powers... destroy the demon’.

Mod. doḥ l-a kti ākrak’ kti ‘Whether (it is) good or bad...’

Mid. 6B.4. doḥ pī māṅ rājasatrū... ‘If there should be (any) enemies of the king...’

R. 1.61.6 doḥ pī pā eṅ māṅ dukkh, añ bum bram sranuk ‘If you had troubles, my friend, I could not bring myself to be content’.

Mod. doḥ pī’nak gmāṅ kār “vī, añjoeṅ mak leṅ phdāḥ khūṃ. ‘Even if you do not have any business, do come and visit me at home’.

didai. See dai.

dep (depv, doep) ‘and so; and then’: Khmer m.

P.A. 726.A.8. dep ge cript ti ‘Then (i.e. after the purchase of the ricefield) they reserved the land’.

A. 207.7. man khmi sthāpāṅa āsāṅa kamrāteṅ jagat, dep tve camnam kalpanā ‘He wished to establish the seat of the god and so he made a foundation, furnishing...’.

Mid. 17.35-8. gāpṛī yeṅ khāṃm pamphāṅsu anah A... deby yeṅ khāṃm oy rāntāp phlaeh jheh... ‘It seemed a good idea to us to place Anak A. in a religious community... and so we prepared fruit... (i.e. for the ceremony)’.

R. 4.66.2-4. yal’ ják’ jāti jā braṅ lohit cēṅ mak noḥ doep khūṃ prām proe yak as’ parbat mak... ‘We saw clearly what it was like: it was your blood flowing out. And so I gave the order to bring rocks...’

Mod. trāv pralāṅ jāp, doep āc cūl rień pān. ‘You have to pass the exam, then you may go to the school’.

212
doं, daं, dauं, dāmnì 'all, complete with, as well': Khmer g.\textsuperscript{12}

P. A. 18.2. oy sre daं kīnum 'give ricefields complete with serfs'.
Mid. 8.6. prōs khnumm dāmnì grāv 'free slaves, a complete family (of them)'.
R. 1.17.10. dāmnì bīr ksattrī ksattrā 'both princes'.
Mod. kūn dāmnì pī nāk' 'all three children'.

\textit{doं, doं pī, doं pī. See \textit{dāmnì}.}

\textit{dāmnì See \textit{doं}.}

\underline{nā, nā}

(i) 'at, in (place or time)’: Khmer pre-n. p.
P.A. 79.9. struk nā jīlaं kāol ‘village at Jīlaं Kāol’.
A. 235 passim. nā kamratej jagat ta rāja ‘before the lord of the world who is king’.
Mid. 6.A.6. mok nā sthān brah bisnulok ‘come to the place Brah Bisnulok (= Angkor Vat)’.
R. 2.49.3. sū rat’ dau nau nā brai ‘grimly-determined to flee and live in the forest’.

(ii) ‘with regard to’: Khmer pre-n. p.
A. 231.12. nā vrah jamvan ‘with regard to the royal gifts’.

(iii) ‘who, when’: Khmer m.
A. 878.2. daksīnā mṛataṅ khalō A. nā choh prāśāda ‘honorarium for Mṛatāṅ Khalō A. on-the-occasion-of inaugurating the tower’.
Mid. 6. A.33. mahākṣetr phoṅ nā grōn pranipat brah ‘the gods, mahākṣetr, who protect (the religion of) the Buddha’.
R. 1.25.12. nā brah staeṅ mak tāl’ sthān aṅ neṅ brah staeṅ mān amboe paṃnaṅ prāthnā ‘As you approach my encampment here, you have many demands!’.

(iv) (cf. Mod. nā) ‘which, any, some’: Khmer post-n. p.
A. 348 nord 22. nu sre nā mān ‘with any ricefields (which) they had’\textsuperscript{13}
Mid. 3.B.10. no kāl nā pī ‘at such time as’.
4.30 saṃmer neṅ nā mīy ‘one of the novices’.
R. 2.31.5. it narnaī ‘nak nā rambhīn ‘without any one thinking of...’
Mod. ‘nak nā tīn? ‘Who (which person) knows?’

(v) (emphasis): Khmer f.
R. 4.34.3-4. būm dān’ sabv sāy samīṅ neṅ hoṅ loey nā. ‘before this idea should spread at all’.
L. 269.1. prāśāda neṅ eṅ nā ‘This very palace...’

\textsuperscript{12} Since 1968, I have realised that dāmnì, which may occur immediately preceding numerals (dāmnì bīr ‘both’) and verbs (dāmnì nāy ‘easily’) as well as nouns, should have been classified as g.

\textsuperscript{13} The A. example here might be considered, like the A. example in the previous paragraph, to illustrate nā (m.). However, nā ~ nā seems like natural Khmer and was the way Georges Coedes took it in editing 348.
C. Hai Mahājan 5. lamārik nās’ nā. ‘It will be very hard indeed (for you)’.
Mod. prayātn nāl ‘Look out!’

ni, nī nīy

(i) ‘with reference to’: Khmer pre-n. p., m.
P.A. 44.A.7. ājña vrah kamratān añi nī vrah kamratān añi ċri… ‘Edict of His Majesty with reference to the God ċri…’
A.470.11. phyaṅ ni pre… ‘announced concerning the order to…’

A.245.91. dau vnek nī añ ‘before my very eyes’.
Mid. 2.29. prākaṭ tuc sāmnūm siṭapraṇidhān mūnī viṅ hoṅ ‘exactly in accordance with my earlier prayer’.
R. 8.5. (Pou 1979, st. 3663.) bi muṅ nī ‘before, in times past’.

nin, nīn

(i) ‘with’: Khmer pre-n. p.
Mid. 9.5. brom nin uk ṇā A. ‘were-in-agreement with Uk ṇā A’.
Mod. khīṃ ṇī mitt sāṃlān’ khīṃ ‘my friend and I’.

(ii) ‘shall, will’: Khmer pre-v.p.
Mod. khīṃ nīn dau ‘I shall go’.

nu, nū, nūv, gi nu

(i) (Initiates discourse): Khmer m.
Some P.A. and A. occurrences are spelt nau, eg. 154.A.16; 957.18.
P.A. 493.20-1…canlek yugala yau 3 * nu man poñ A. kiṅ sre… ‘…3 yau of double cloth. Now Poñ A. released the ricefield…’
A. passim (e.g. 925,14) after date, preceding rest of text. (gi nu also occurs similarly.)
Mid.4.B.25. nū tejah phalānisaṅ neṅ camren… ‘Now the power of the fruit of this merit may give prosperity…’

P.A. 424.B.6. man duṅ nu sru vrah ‘which (they) bought with paddy from the foundation’.
A. 239.24 sre neṅ kula A. ti duṅ nu kɾaṉi 2. ‘This ricefield was bought by the family/descendants of A. for 2 buffaloes’.
Mid. 4.C.9. banarāy nī tribidhasaṃmpāt ‘resplendent with the three-fold fortune’.
R. 1.2 hoṅ hoer nūv tej pī banljik ‘go through the air with astonishing power’.
Mod. aṅjoṅ… kamṣaṅ nūv kīḷā praṇāṃ kaṅ’ ‘invite (you) to amuse yourselves with a cycling contest’.

(iii) ‘and’: Khmer pre-n. p.
P.A. 134.20 sre nu pdai karoṁ nu daṁriṅ ‘ricefields and lowlands and orchards’.

214
A diachronic survey of some Khmer particles

A. 215-6.10. yok sruk sre nu khñum phon ‘took the villages, ricefields and all the serfs’.
Mid. 6.A.5. ok hlûn A. nu anak cov bhikkh B. ‘Ok Hluon A. and the bhikkhu B’.
R. 1.3.11. nám râj tâ pabitr nûv brah anuj laksâ ‘brought the august prince and Laksman his young brother’.
Mod. (lit.) stî vâ nûv gû kan dâmñ phunmân ‘addressed them and all their playmates’.

(iv) ‘shall, will, for the future purpose of’: Khmer pre-v.p.
P.A. 341.nord 11. ge gi tâ nu pi niraya ‘they are the-ones-who will be for Heli’.
A. 246-8.3-4. dravya tâ rohî tî aî yok nu ‘ras ‘the goods according to (the above list) were taken by me for a living’.
Mid. 2.6-7. sambau jâ rat rûn útûn tâ jâ nâm chlon ‘a boat, splendid, magnificent, to take (us) across’.
R. 1.43.4-5. tûc nûv cheh chûr chûp vinãs ‘as if it would burn furiously and be destroyed’.

(v) (connector between verb and object): Khmer pre-n. p.
R. 1.9.10. samtaen nûv tej brah dhanû ‘demonstrate the power of his bow’.
L. 289.14. khñh kân’ nûv tampan ‘Some held batons’.
Mod. (lit.) daduol nûv dukkha tâ kluon viî ‘meet with suffering themselves in return’.

neh

(i) ‘this, these’: Khmer pre-n. p.
P.A. 79.12 sre sin panlass neh sre ‘a further ricefield in place of this ricefield’.
(ii) ‘this, these’: Khmer post-n. p.
A. 291.32 pujā neh ‘this cult’.
Mid. 6.A.20. jûn kusal-phal puny neh ‘offer the fruits of this work-of-merit’.
R. 5.20.1-2. mak tal’ sthân brai neh ‘come to this forest-place’.
Mod. kûn neh ‘this child’.

nai, naiy

(i) ‘of’: Khmer pre-n.p.
A. 207.15. thvây dravya... neh jâ nai kammrâtej jagat ‘offer these goods... to be the possession of the lord of the world’.
Mid. 6.A.7. sthân jâmnumm naiy, debatâ ‘meeting place of the gods’.
R. 1.16.9. sthân nai traitrîhsâ ‘the dwelling of the thirty-three’.
Mod. (lit.) knûn nâm nai brah râj-raîthâbhipâl ‘in the name of the royal government’.

(ii) ‘of that (with back reference)’: Khmer f.
R. 2.63.1-2. màn bâl prâm-muoy hmûn nai. ‘He has a force of sixty
thousand (of him). ‘There is a force of sixty thousand of his’.

L. 290.22-291.1. jā sthān brah mahā isī en nai, ‘being the dwelling-place of
a great anchorite (of him)’.

C. Kān cau 19. heu citt cañ’ krep jañjāp klin nai, ‘because it wanted to
taste and drink in the flavour of (it)’.

R. 3.1.4. pi tüc ‘nak tüt nai ṣagī ‘as if someone had lit a fire (in him)’.
L. 283.18. kā phtuol nai yakṣh maraṇā ‘(He) then felled the yakṣha to
death’.

no, nov, nau, a nau (also spelt nūv in R.)

(i) ‘at’: khmer pre-n.p.
A. 249.14-5. nek vrah ‘nau stuk vṛyaṇ ‘before (lit. eyes) the god at Lake
Vṛyaṇ’.

Mid. 4.B.10. do no caturāpāyabhūm ‘go and dwell in the four domains of
suffering’.
R. 1.1.6. pansenō bis nūv selā ‘dilute the poison on the stones’.
Mod. diṁ phdaḥ nau bhām beñ ‘buy a house in Phnom Penh’.

(ii) ‘continuing, still’: Khmer f.
A. 235.C.7. pamre ta vrah pada A. ru noh ‘nau ‘continue serving King A.
in that way’.
R. 1.4.6-7. brah thlaën asūr kākanāsūr ksāy ājīvit kā mar(n) tuol nau. ‘The
lord shot at the demon Kākanāśūr to take her life and so she fell dying
and remained (fallen)’.
R. 1.62.11-63.1. jal netr dhārā sasrāk’ sañjap’ sup nau. ‘His tears flowed
freely and his face was downcast still’.
R. 2.6.10. toem jhāk sāt’ mak jambāk’ dāṃni daṅ doer nau. ‘The lotuses drifted
along and, their stalks entangled, remained resting-on-each-other’.
R. 2.45.1-2. brah pād... nau mān brah jām mān nau loey? ‘Is the King still
alive?’.
L. 283.19-20. pradaḥ sugrāb yam nau ‘came across Sugrib still weeping’.
C. Hai Mahājan 87. kāl blien nau ‘while it is still raining’.
Mod. Only the use with loey is current and this is characteristically in the
negative, e.g. min dān’ dhvoe nau loey ‘has still not done so’.

noh

(i) ‘that,those’: Khmer pre-n.p.
P.A. 44.B.9. pammos ta pos gi noh vnam ‘religious personnel in that
sanctuary.

(ii) ‘that, those’: Khmer post-n. p.
A. 216.10. thve kuti noh pūjā kamraṭeṇ jagat ‘build that cell to carry out
the cult of the lord of the world’.
Mid. 8.11. pammos phoṅ noh ‘that the whole (matter of the) freeing-(of-
slaves)’.
R. 1.14.1. kāl noh ‘that time’.
A diachronic survey of some Khmer particles

Mod. thnai noh ‘that day’.

nau. See no, etc.

pad, ta pad, tpit

(i) ‘for the reason that’: Khmer m.
Mid. 8.3-6. ok hlûn A. mimantr brah ariyasâns phoû jâmnum ta pad okk hlûn B... pros khîunum ‘Ok Hluôn A. invited several noble monks to meet together because Ok Hluôn B... had enfranchised slaves’.
R. 2.62.3-4. neh pad nûn kaikesî tanûom râjy bi brah râm mak oy bhirut. ‘This is because Kaikesî has taken the kingdom from Râm to give it to Bhirût’.
Mod. ilûv nûn dau tok srûp viû e nû kâ pûn tpit kûr vû huos dau hoey. ‘And now, how can one undo what is done (lit. pull out and swallow again), for the matter has gone too far’.

(ii) tpit (Initiates statement): Khmer m.
R. 3.6.9. ang aûlûmneh nai tpit mûn saûnûy ambal’ nûn bhariyû hoey. ‘As for me at present, now there is a problem: I already have a wife to occupy my attention’.

Mod. In undated texts of folktales, printed in the twentieth century, tpit introduces the ‘hero’s’ tale of woe passim, when he meets Judge Hare.

pi, pî ‘for the purpose of, so that; because, the reason why; as it were’: Khmer m.

P.A.21.2 tmo pi vnûk ‘gems for the decoration’.
P.A. 451 sud 6. va cap pi hau ‘Mr. Catch (him) so as to employ (him)’.
A. 207.2. oy srôh kamrâten jągat cři A. pi jvan dravya ta ampall neh ‘arranged for the ablutions of Lord Çři A. so as to (on the same occasion) give all these goods’.
A. 219.8. pi ‘yatr kvan cau ley syûn ta tûc santûna dau phoû ni pi vrâh pûda A. oy prasaûda bhûmi ‘Because there were no children or grandchildren, their line being extinct, that was why His Majesty graciously gave land...’.
A. 207.32 yen vwaû kapata pi thve apavâda tyûn ‘we shall not deceive in such a way as to make a denial of our knowledge’.
Mid. 2.25-6. pân paûcâkhandh âmbi jêt pi andol do mok ai ta sàsàrâbhahbh ‘depart this life so as to move to and fro in the state-of-transmigration’.
R. 5.63.6-7. nau bhjum bal bânûrû pi mak tvây ksat ‘will assemble a force of monkeys to offer to the Prince’.
R. 3.1.11. yaksû pi drahûn ‘A fierce yakṣa’.
Mod. (Only occurs in compounds. See under ‘Compounded particles’, p. 17 above).

peh (=pe), poe ‘if, when (in future)’: Khmer m.
Mid. 8.43. peh mok srati ‘if (such a person) should come and complain’.
R. 1.57.7. poe parâtûm yoeû yâtrû viû mak yal pû. ‘When this is completed we shall come back and see you’.

217
Mod. *poe ǔn kum kān' dos doep khūm hān prāp*. ‘If you won’t hold it against me, then I’ll dare to tell you’.

*poē. See pēh.*

*phañ. See phōn.*

*phōn, phañ*

(i) ‘all’: Khmer post-n. p.

P.A. 127.12 *ampall kule ge phōn* ‘all the relatives of the personnel’.

A. 699.5. *rājakule phōn* ‘all the royal family’.

Mid. 3.A.25. *nu brah ṇāttī phōn* ‘and the whole royal family’.

R. 1.8.9. *dāt deb ksatr phōn* ‘all envoys, gods and kings’.

Mod. *ṇak jīt khān phaṅ* ‘all the neighbours’.

(ii) ‘both... and’: Khmer f.

Mid. 10.4. *kat don satpakar dām 7 hey chloṅ phnūs phōn sot* ‘cut banners for the 7 Books (of the Abhidhamma) and also made (the people) enter the religious life’.

R. *dāl bhlāṅ yam phañ* ‘speaking and weeping at the same time’.

Mod. *pos phāḥ hoey pok ut phañ* ‘sweeps the house and does the washing and ironing’.

(iii) ‘together with’: Khmer pre-n. p.

R. 1.53.4. *ṇak yātṛā phañ brah jēsthā* ‘you journey together with your elder’.

*bī. See amvi.*

*bek, beg ‘too much, very much’: Khmer f.^[14]^

Mid. 2.15. *tryak ar anumodanā nu stac brah rājaputr beg hoṅ* ‘I most delightedly participated in what my son did’.

Mod. *dhñan’ bek ‘too heavy’.*

*bum, bvum. See vom*

*bum tael. See tel.*

*man*

(i) ‘who, whom, which’: Khmer m.

P.A. 561 A.27. *khum man poṅ A. oy' servants whom Poṅ A. is giving’.*

A. 207.29. *nā bhūmi neḥḥ man vraḥ kamrateṅ aṅ A. duṅ ta yeṅ* ‘with regard to this territory which my Lord A. bought from us’.

(ii) ‘that’ (reported speech): Khmer m.

A. 843.C.13 *kule aṅ tyāṅ man bhūmi nū upāya noṅḥ phoṅ vīn āyatta ta santāna aṅ.* ‘My family know that all those territories and their subsistence are returning to the authority of my line’.

Mid. (?) 2.7-8. *sramaddh bol man gi habbh jāt* ‘the ocean: say that it is (taking gi as a verb but see man gi below) existence’.

---

^[14] This is also a change of mind about the classification of a particle as compared with Jacob (1968) when bek was assigned to the adverbial particle class.
(iii) (Initiates discourse): Khmer m.
A. 207.15 °man sruc thvay dravya neh... dep dau ‘Now, having finished the consecration of those goods..., he then went...’

man gi

(i) ‘who, whom, which’: Khmer m.
Mid. 2.10. jai ilineh man gi luh ta 1499 czaka ‘this present life, which is in 1499 czaka’.
R. 5.35.5. neh yal’ man gi kal’ hon ‘this (which) one may indeed see as being a trick’

(ii) (Initiates discourse): Khmer m.
Mid. 3.A.16. man gi anbi kal jra prathamm samtec... pann thlen svey rajasampatt... ‘Now from the beginning of His Majesty’s ascending the throne...’

min

‘not’: Khmer pre-v.p.
L. 273.3 min oy yur luh ‘not being a very long time’.
C. Hair Mahajan 62. min thlas’ dhloy ‘not making careless mistakes’.
Mod. gai’ min dau de ‘He is not going’.

ra, r̄ (Adds emphasis): Khmer f.
P.A. 726.C.8-9 kNUM vNak poN yajamA na dai, klah ra, tel oy ta acrama, klah ra, tel oy ta vrah ‘other servants of the section of PoNi, some of whom he gives to the acrama, some to the god’.
A. 205.8. pek sapp kule ta gi bhaga didai ra ‘divided (it) among all the relatives there, each having his share’.
R. 2.41.1. oH o brah janm skal’ skap’ r̄ pann yal’ hna puny putr cpAn. ‘Alas! Your life was completed, then, when you had the chance to see the merit of your eldest son’.
Mod. mak naeh r̄A. ‘Come here!’.

riy, r̄ ‘as for’: Khmer pre-n. p.
P.A. 518.D.1. r̄ ‘nak ta sok ‘as for anyone who corrupts...’
A. 33.35. riy ‘nak ta parmAt dharmma... riy ta paripala... ‘as for persons who destroy the foundation...; as for those who look after (it)...’
Mid. 4.23-4. r̄ ji dahn pi neh ‘as for these 3 young men’.
R. 5.3.4-6. kAl noh hetu paramitä r̄ brah suriyā buman ac astaṅgat dau. ‘Then, owing to his perfection, the Sun (lit. as for the Sun) was unable to set’.
Mod. r̄ cau adhikār viṁ... ‘Now as for the abbot...’

ru, r̄va, r̄

(i) ‘like’: Khmer pre-n. p.
P.A. 341.sud 5. pre thve pujā vrah kammaratān aṁ ru lingapurvā ‘ordered that the cult of the god should be carried out as (at) Lingapura’.
A. 235.71 rūva nohh "nau 'continuing like that'.
Mid. 2.6. jā buvmnak trey trān pradhān ru sambau 'is a refuge for (our)
protection like a boat'.
R. 5.31.3. sabd srak ru randah 'the sound of roars like thunder'.

(ii) 'like': Khmer post-n. p.
P.A. (Occurs frequently at the end of names.) 689.A.13. ku āl ru 'Ms. In-
a-rush-like'.
A. 231.35 neḥ ruv 'like this'.

roḥ, rohh 'in accordance with': Khmer pre-n. p.
P.A. 904.B.25. gi neḥ ta rohh neḥh phon tel...'It is all these (things), in
accordance with this (list just given), which...
A. 216.10. roḥ man steñ vrah jrai ta 'jī dai kalpanā 'in accordance with
what the other Steñ of the Sacred Fig-tree, ancestors (of this Steñ)
provided'.
Mid. 3.B.20. laen šthit šthir is kāl ta laṅgh roḥ braḥ camnaḥ prāthnā. 'May
it remain firmly established for ever in accordance with the king's
wishes'.

lah, loḥ 'whether... or': Khmer f.
P.A. 51.14. kon prasā lah cau prasā laḥ 'whether it be son-/daughter-in-
law or grandson-/daughter-in-law...'
A. 842.24 caturthaṅça lah şaṭthaṅça laḥ 'whether it be a quarter or a
sixth'.
Mid. 3.A.37. doḥ purus loḥ doḥ čṛī čiṅ 'whether it be male or (further)
female...'

luḥ, loḥ, lvah, luḥ nā

(i) 'to': Khmer pre-n. p.
P.A. passim anvi... loḥ 'from (a place)... to (a place)'.
A. 190.5-6. lvah chok khadira 'as far as Chok Khadira'.
A. 457.8 lvah ta gi camkā 'as far as the garden'.

(ii) 'by the time that, by (a time)': Khmer pre-n. p.
P.A. 259.III 21. loḥ ta gi rājya vraḥ kamratān aṅ 'by the time of/in the
reign of Her Majesty'.
A. 855.5. lvah ta gi 888 čaka 'by 888 čaka'.
Mid. 2.9-10. luḥ jāt iliv 'until this present life'..
R.1.5.2. luḥ bīsākh 'in April-May'.

(iii) 'until, by the time': Khmer m.
Mid. 6.29. luḥ pān cūl ta pad moks 'until (they) manage to enter the way
of deliverance'.
R. 2.5.1. bilāp luḥ lat' ralāy ramsāy sok 'wept until sorrow melted away'.
Mid. 3.A.43. luḥ nā pān pūs 'until he would be able to enter the religious
life'.
R. 2.81.6-7. luḥ sanlap' praḥ praṅ 'until they fainted, writhing'.
A diachronic survey of some Khmer particles

Mod. *luh tralap’ dau thăn devatā viñ... ‘When they arrived back in the dwelling of the gods...’*

*leñ, laeñ* (Exhortation): Khmer m.

A. 219.5. *leñ jā bhūmyupāya nai sruk stuk ransi ‘let the subsistence-of-the-territory belong to (be of) Bamboo Lake’.*

Mid. 6.A.36 *laeñ pān pūs saṃnak ta brah sāsthā ‘May he manage to become a monk, staying in the religion of the Buddha’.*

R. 4.8.8. *laeñ pān khmoc aṅ phoè bhūdar bīr brah dhanū ‘let it be possible for my body to be there for the two supporters of the earth’.*

*leñ, loey* ‘totally, utterly, at all’: Khmer f.

P.A. 557. Est 4. *va ahvān amne ley ‘Mr. Avoid all action’(?)).*

A. 989.B.10-11. *daha ‘yat santāna ta puruṣa ley ‘if there is no male descendant at all’.*

Mid. 4.B.11-12. *kumppi pros anak noh pān ley ‘let (them) not be able to save those persons at all’.*

R. 2.7.2. *pakū khlouon khlau buṃ tīn toem dān jhūk loey ‘the foolish bird had no idea that they were lotus stems’.*

Mod. *cīt anak kamloh... min nāy nīr rasāy pān loey ‘the boy’s feelings... were not at all easily appeased’.*

*loh. See luh
Ivaḥ. See luh.
Ihey. See hey.*

*vii, viñ* ‘again’: Khmer f.

A. 697.B.18. *cat sruk A. viñ ‘restore the village A.’.*

Mid. 2.13. *sān ru purān viñ ‘reconstruct as of old’.*

Mod. *tralap’ dau phdaḥ viñ ‘return home’.*

*vom, vvaṃ, bvoṃ, buṃ* ‘not’: Khmer pre-v.p.


A. 85.2. *nau ‘nak ta vvaṃ thve toy kalpanā ‘now persons who do not act in accordance with the arrangements...’* 

Mid. 2.17. *byun dīen ‘not certain’.*

R. 1.8.10. *buṃ āc nīt loek dhanū pān ‘were unable to raise the bow’.*

Mod. *siēvbhai neh buṃ maen jā pravatī panlām rapas’ jan rūp nā loey ‘This book is not in fact the story in disguise of any (real) person’.*

*vvaṃ tael. See tel.
sot ‘in addition’: Khmer f.*

P.A. 561.38. *oy antyanti sot ‘finally gives in addition’.*

A. 67.C.6. *dār prasāda... dār prasāda sot ‘acquired the royal favour... acquired the further royal favour’.*

Mid. 16a.8. *ri e anak ta jhar lek khmoc noh sot mān A., B., etc. ‘As for those who took charge of the funeral, one may add, there were A., B., etc.’*
R. 2.2.7. campā sot mān phkā ‘and the Michelia too has a flower’. 
Mod. rīoẹn breŋ rī pad camraję sot... ‘legends or songs, too...’

syan, syaŋ ta, syiŋ ta

(P.A. The rare occurrences of siṅ are either in ambiguous contexts or are instances of siṅ ‘to officiate’.)

A. 207.27 “nak neŋ phon syaŋ ta samayuga yok.... ‘all these people together jointly took’.
Mid. 6.A.9. syiŋ jaṁnumm sumukkata (= sam mukha gata) ‘all met face to face’.
Mid. 3.A.64-5. syiŋ ta jaṁnum samuggata (= sam mukha gata) ‘all met face to face’.
R. 1.27.10 ksatr savy prades hoṅ siṅ cuḥ cūl ‘the princes of every country all submitted to him’.

(ii). ‘almost: more or less: generally speaking’: Khmer pre-v.p.

Mod. mān babak... mun ṭampūn sa ruoc praphaēh siṅ khmau ‘there were clouds, white at first, then grey, almost black’.

hetu, hetu man, het tae ‘because’: Khmer m.

A. 348.nord.2. hetu man mān apavāda ‘because there were objections’.

L. 282.10. luṅ lom car cā hetu caṅ’ smoṅ smān braṅ anūg ‘she cajoled (them) with her chatter because she wanted to make friends with them’.

C. Kūn cau 30. hetu min sralāṅ ‘because you do not like it’.

Mod. hetu tae mān saṅgrām ‘because there was war...’

hey, lhey, hoey

(i) ‘and (then)’: Khmer m.

Mid. 6.1. anāc cov bhikkha sāṅ hey cah (= Mod. Khm. cār) cāṛk phoṅ.
‘The monk A. reconstructed (this) and then wrote the inscription as well’.

R. 1.41.9-11. tūc kūn koet bi udar hoṅ prae prās brāt’ phaŋ hoey mak prasab yal’ viṅ ‘as if (they were) their own children (who) had been separated from them and were reunited, seen once again’.

Mod. as’ kamāṁhm hoey jhi grun phaṅ ‘exhausted and feverish too’.

(ii) ‘having already: already; by then: by now; and so’: Khmer f.

A. 413.II 53. adhiṣṭhān pvas lhey dep draṅ braṇ carat cuḥ ambī subarnaprasād ‘Being now ordained (lit. established in the religious life), he descended from the golden tower’.

Mid. 6.A.11. surec hey deby yeṅ khūmm rantāp krayā pūjā. ‘This done (lit. finished already), we prepared the things-needed for the offering’.

15. I feel convinced, after observing for many years that pre-modern occurrences of siṅ - syaṅ are preceded by a plural subject (or, rarely, singular but if so needing the translation ‘wholly’), that siṅ - syaṅ has the function of linking a plural subject to the verb. Frequently, e.g. in R., the verb is separated from the subject by several words and siṅ - syaṅ is also useful in marking the main verb.
A diachronic survey of some Khmer particles

R. 1.57.8-9. kūkhān stāp' brah pandūl phañ srec hoey camlaṅ dāmī pi brah aṅg luḥ tał’ troey. ‘Kūkhān, having listened to the lord’s reply, took the three royal persons across to the other side’.

Mod. diñ ivān’ hoey kā tralap’ dau phdaḥ. ‘Having bought the things, (he) returned home’.

hoñ (Emphasis): Khmer f.

A. 215.35 tāc santānā yen hoñ ‘our line is completely extinct’.

Mid. 8.51. sūm pān dauv kēt nā tuṣidd hoñ ‘pray to be reborn in the Tusita’.

R. 1.57.4. guor hoñ aṅh jhap’ toy pā. ‘It would indeed be good for me to stay with you’.

Mod. (Lit.) kun jriēt jraek gnā tūc kāl grā mun noh hoñ. ‘Do not all push in front of each other like the previous time’.

loey. See ley.
as’. See is.
āy. See ai.
i. See ai.
is, iss, as’ ‘all’: Khmer pre-n. p.

A. 207.27 yok iss dravya noh phon ta jā thlai bhūmi. ‘take all those goods as the price of the territory’.

Mid. 4.A.31. is nāt gār ‘all his relatives’.

R. 1.13.12. as’ dāmī manuss ay tā lok ‘all the men in the world’.

Mod. nissit dāmī as’ ‘all the students’. Also, however, as’=‘to have exhausted (the supply)’, eg. as hoey ‘That’s the lot’ as’ kamlāmn ‘tired’ (out of strength).

ukk ‘also’: Khmer f.

P.A. 493.24. sre A., pradana poñ ukk, man jāhv... ‘the ricefield A., a further gift of Poñ, which was acquired-by-exchange...’

e See ai.
a, āy, i, e, ai ta, āy ta, ai ta

(i) ‘to (re gifts)’: Khmer pre-n. p.

P.A. 44.A.8-9 amnoy tān aṅh kloñ raṅ[k]o doṅ poñ A. ai ta vrah kamratāṅ aṅ ‘gifts of Tān Aṅ, chief of the husked rice, and of Poñ A. to the god’.

P.A. 561.7. ple idaiy āy ta vrah kamratāṅ aṅ ‘other revenues for the god’.


P.A. 38.4. sre ai travāṅ ‘the ricefield at the reservoir’.

P.A. 416.I.5. sre āy stu[k] anrok ‘ricefield at Anrok Lake’.

A. 933.4. ai hari haralaya ‘at Hariharalaya’.

A. 235.D.37. āy viṣaya jeṅ vnaṁ ‘in the district of Jeṅ Vnaṁ’.

Mid. 2.15. sucari sarḍhā ai ta braḥ sāsṅā braḥ tathāgaṭ ‘perfect faith in the teaching of the Tathāgaṭ’.

Mid. 4.B.26. ī leh (= e loe) ‘on high’.

223
JUDITH M. JACOB


(iii) ‘as to’: Khmer pre-n. p.
Mid. 9.32. *e sāgh sot rī mān A., B., C.,* etc. ‘as to the monks, there were A., B., C., etc.’.
R. 1.37.11 *e nān kaikesī nā hetu cītī drijīya dīṭṭhi drust kāc kān’ khjāp*. ‘As for Kaikesī, out of jealousy she wickedly persisted in her wrong attitude’.
Mod. *e khjiṃ viñ, khjiṃ min cūl cītī*. ‘As for me, I do not like (it)’.

amvi, ambī, bī ‘from’: Khmer pre n.p.

P.A. 44.A.11 amvi dik hera loh vraī tamponn ‘from the spring to Cane Wood’.
P.A. 44.A.9. amvi kāla ‘from the time’.
P.A. 149.2. dār amvi ta vrah kamratan añ ‘claim from the god’.
A. 56.A.35. amvi travān trapek dau ti uttara ‘from the Trapek Reservoir to the north’.
Mid. 4.8. ambī chnam ‘from the year...’.
R. 1.58.6-7. pabitr camniāy phīv nā stec nīn yātrā ambī neḥ dau dūragam ‘August One, the length of the road that you will travel from here on! It is far to go!’.
C. Kram 31. bī tūc ‘from childhood’.
Mod. bī khān joen dau khān tpūn ‘from north to south’.

"nau. See nov.

REFERENCES


Ishii, Yoneo, Akahi, Osamu & Endo, Noriko. 1977. *A glossarial index of the Sukhothai inscriptions* Kyoto: Shoukadoh Publ.


A diachronic survey of some Khmer particles

Jenner, P. N. & Pou, Saveros Khin Sok


Lewitz, Saveros


Nacaskul, Karnchana


Pou, Saveros


Pou, Saveros & Jenner, P. N.


THE FORM SYAÑ IN ANGKORIAN KHMER

Philip N. Jenner

In at least one respect the study of ancient languages known only from written monuments is like the study of contemporary languages which have never been reduced to writing: in both cases the investigator encounters as a matter of course lexical items which are not readily amenable to analysis. What is sometimes forgotten is that the linguist working on a living language normally has recourse to informants, whereas the linguist working on a dead language can have no such guidance. Hence, if most of the ancient lexicon is known or knowable, certain forms prove resistant to identification, and alternative means must be brought into play in attempting to account for them.

It is not my intention to claim that the difficulties confronting us in Old Khmer are as numerous or as perplexing as they seem to be, or to have been, in many another ancient language. Since the first inscriptions were published by Aymonier a century ago we have, thanks to the painstaking work of a few dedicated French scholars, seen the gradual elaboration of a tolerably good understanding of Old Khmer. By 1966, when the eighth and last volume of Cœdès’s monumental Inscriptions du Cambodge [= C] made its appearance, most of the texts in the growing corpus of Khmer epigraphy had been analysed and explicated with an acceptable degree of reliability. By that year the bulk of the lexicon was fairly well understood, but a good many specialised terms had not been worked out to our satisfaction. This is still the case today, when a good many lexical and other problems continue to elude us. Among these is a small class of forms which appear to perform grammatical functions. Typical of this class is syan, which I propose to discuss here.

Not found in the pre-Angkorian inscriptions recovered so far, this orthographic form is attested in Angkorian Khmer well over a hundred and fifty times. In addition, its life is extended through the Middle Khmer period, where it appears variously as syan (A.D. 1560-77), syin (1560-1701), and sin (1587-1706), with the apparent nonce-forms sin (1620), siny (1696) and sin (1701). As to the meaning of the Angkorian form, Aymonier (1883: 494, n.1) took it as synonymous with Sanskrit kevalam ‘solely, entirely’ and as the source of modern Khmer ‘sin’ — by which he...
evidently meant *siha* /sŷŋ/ [svŋ] ‘nearly all, almost’. In 1913 Cœdès (Parmentier 1913: 14) understood it in the same sense as Aymonier and rendered it ‘sans exception’. In 1915, however, Finot (1915: 106) re-defined it as a pronoun, glossing *syah* *ta* as ‘qui’. For most of the half-century down to 1966, Cœdès and his contemporaries treated the form in various ways which are impossible to reconcile. Thus in the largest number of cases, representing about 35% of the total, we find it expressed by appropriate forms of *tout*. Yet in roughly 29% of its occurrences it does not seem to be rendered at all, at least overtly, even in those fairly numerous cases where the idea of *tout* would not appear out of place. Several times it is expressed by the related notion of *ensemble*. In other passages, representing a little over 15% of the total, we find it rendered by ‘*comme*’, ‘*tel*’, ‘*réellement*’ and, in one instance, ‘*völlä*’. It must be said, however, that in none of these latter cases is it absolutely certain that *syah* is being expressly rendered at all. Indeed, one has the impression that Cœdès and his contemporaries allowed themselves in some instances to be led into the translation they give under pressure from the supposed contexts. This uncertainty is greatest in passages in which figure appropriate forms of *être*: one simply does not know whether *syah* is being so rendered or is being passed over in silence.

To this diversity of interpretations may be added the treatment of *syah* and its variants in Middle Khmer texts. During the 1970s my learned teacher Saveros Pou redefined *syah* as an anaphoric ‘pronoun or particle’ (*sic*), holding at the same time that the Middle Khmer form or forms were the source of modern Khmer *siha* /sýŋ/ ‘nearly’. As far as I am aware, however, it is nowhere rendered as an anaphoric pronoun in any of her superb translations. Of the fifty-seven Middle Khmer occurrences of *syah* and its variants collected for the present study, twenty-seven (48%) are not overtly rendered at all. Twelve, or 21%, mostly in combination with what we are no doubt justified in now calling the equational copula *jā* /jaa → ciia/, are conceivably assigned a copular function or, if this is not so, are also unrendered. In only three cases is *syah* represented by ‘*comme*’. On the other hand, in eight cases (14.5%), French and my own English translations read as if *syah* marked the apodosis of an ‘if’ or ‘when’ clause, or as an unexpressed ‘then’ introducing the consequence of a prior clause. We also have a small number of instances in which *syah* is, as I might say, plainly none of the foregoing but seems to express a contrastive (‘nevertheless’) or instrumental (‘thereby’) idea. In only two cases, both suspicious, does it appear to be taken as a pronoun.5

3. As will be seen, he may have had in mind modern *siha* /sýŋ/ [svŋ] ‘to rest’.
4. In addition to the occurrences tallied here, in over 9% of its instances the passages in which *syah* occurs are left untranslated for one reason or another. The Angkorian corpus also includes twelve passages with *syah* which are too garbled to be used in the present discussion.
5. I might mention in this connection that I have considered the possibility of a loan relationship, in one direction or the other, between Middle Khmer *syah* – *sih* and Thai /syrJ/ (for the corresponding Lao form see Kerr 1972: 493b), usually treated as a relative pronoun ‘used in reference to a person, an animal, or an inanimate object... in the nominative, objective, or possessive case’ (Sethaputra 1965: I, 349ab, who adds that it is also ‘used in a literary context to introduce a noun in the objective case after a transitive verb, where in ordinary language no preposition (*sic*) is required’; cf. Haas 1964: 157b; McFarland 1944: 309b). Since it seems to yield no useful results, my consideration of this possibility is not included in the present discussion.

228
here that none of the Old Khmer or Middle Khmer occurrences of *syain* allows interpretations suggestive of modern *siha* /σσν/ 'nearly'; as far as I am able to determine, this sense is unattested in the older language.

In view of such radically different interpretations of the meaning and function of *syain*, we have no choice except to re-examine the data. It will be appreciated that the task of doing so has something of the character of a trial at law in which the evidence, abundant though it may be, is entirely circumstantial. This, taken with the diversity of views expressed by a number of eminent scholars, demands that all of the usable evidence be adduced and weighed. With this in mind I have screened all of the Old Khmer inscriptions available to me and collected every occurrence of *syain* with a view to ensuring that no usage escaped notice. For Middle Khmer I have gone through all of the so-called 'modern' inscriptions of Angkor (Lewitz 1970-72); some seventeen cplí'pa /cbap/ or ethical texts; the *Lpêka 'angeravatta* /lbaak qaŋkɔrwost/; the undated metrical *Supîna* /sobáːn/; and nine other inscriptions, namely* K.39 (A.D. 1574), K.177 (1478-1577), K.285 and K.465 (both 1583), K.715 (1586), K.27 (1587), K.261/I (1611), K.261/III (1639), and K.261/II, IV and V (1578-1677). Data in this quantity create a problem of presentation. On the one hand, the reader is entitled to examine the full array of the available evidence; but to provide the latter would occupy from sixty to seventy-five pages of space which could no doubt be put to better use. On the other hand, any significance to be found in the present study lies in whatever conclusions one may draw from the evidence, and such conclusions can be given in a few short paragraphs. One of these extremes being as unsatisfactory as the other, I have attempted to strike middle ground, first, by confining my citations to those from the Angkorian inscriptions and, secondly, by reducing the number of my citations to the absolute minimum consistent with the reader's ability to grasp the problem and its solution. It should be explained at the same time that, if it is true that a strictly objective examination of the evidence would furnish only the Old Khmer passages in question, this would be tantamount to dismissing the work and judgements of my predecessors and to depriving the reader of the

---

6. For a full list of her 'Inscriptions modernes d'Angkor' [IMA], see references below, Lewitz (1970-75).


* K = the inventory number in 'Liste générale des inscriptions du Cambodge', Cœdès (1966: 8, 73-225 and Supplements). (Ed.)
opportunity of weighing their judgements and mine for himself. I therefore give the published translations of each passage cited; my own alternatives to those translations are given later.

The simplest structural contexts we are concerned with are those in which syāṁ follows a demonstrative pronoun (Dem.) and is itself followed by a noun (N) designating a metal—though there seems to be no good reason to take the metallic nature of the noun as in any way obligatory. In each case the Dem. + syāṁ + N sequence is preceded by a more or less lengthy list, which I abbreviate, of objects forming part of an endowment:

(1) vaudi mvāy svok mvāy... 'ārghya pādyā mvāy tampar neḥ syāṁ prāk (K.171: 7-8), 'Un vaudi, un plateau, ... quatre vases pour le lavage des pieds (ārghya pādyā). Tout* cela en argent' (C VI: 166).

(2) ... cancyāṁ 1 ratna ta grī 1 naupura 2 khse chdvāl 1 neḥ syāṁ mās (K.669C: 10), '... 1 bague avec 1 joyau, 2 anneaux de cheville, 2 (sic) chaînes; tout cela en or' (C I: 182).

Only slightly less simple are structures such as the following in which syāṁ, still following a demonstrative pronoun, is itself followed by a noun phrase (NP). The sequences Dem. + syāṁ + N and Dem. + syāṁ + NP may of course be considered equivalent.

(3) ... me 'yak me ṇam me des me sāṁ me dvat neḥ syāṁ sruk 'amarālaya (K.598B: 29), 'Les me Tak, Naṁ, Des, Sān, Dvat, tous (sic)9 du sruk Amarālaya' (Finot 1928: 77).

(4) ... ta duk pṛaṣaṭa neḥ mṛatāṁ śri satayudha nu mṛatāṁ śri ripumatha neḥ syāṁ kvan mṛatāṁ śri prathivinarendra ... (K.956: 58-9), '... Ceux qui conservent cet acte inscrit sont Mṛataṁ Čṛi Satayudha et Mṛataṁ Čṛi Ripumatha(na), enfants de Mṛataṁ Čṛi Pṛthivinarendra' (C VII: 135).

(5) kamarateṁ 'ān yogī ta ṁvās ta neḥh phye phlu pūrvvottara tīrthādyānapūrabāma neḥ syāṁ dharmma kamarateṁ 'ān didai ra... (K.139B: 7-10), 'Les seigneurs Yogin qui sont entrés en religion ici confient le chemin du nord-est, le bain, le parc, le jardin fleuri: ce sont les œuvres pies de chacun des seigneurs' (C III: 179).

In such passages as the following we see that the place occupied by Dem. in the preceding sequences may be filled by an NP:

(6) ... patigraha raupya 2 ḍhās 2 vodi prāk 2 ḍhājana dramvaḥ 1 ḍhājana khpac 1 syāṁ ḍhanra ḍhājana pralvaḥ 3 ḍhājana ta madhyama 6... (K.669C: 15-6), '... 2 crachoirs d’argent, 2 agrafes, 2 vodī d’argent, 1 récipient dramvaḥ, 1 récipient décoré tout en ḍhanra, 3 grands récipients, 6 récipients moyens,...' (C I: 183).

(7) ... deḥ reḥ ta dāi ti syāṁ daksīṇā (K.263D: 44), '... ensuite on en choisit d’autres comme offrande (daksīṇā)' (C IV: 138).

(8) kamsteṁ sāntilākṣmi pāṅkeṭ chloṅ haridāṭta chloṅ somaśarman syaṅ

---

8. Here and hereafter, those forms which I take to be intended to express syāṁ are italicized.
9. One would expect toutes.
bhāgavata pamre (K.989B: 16), ‘Kamsteñ Čaṭilakṣmī donna naissance à Chloñ Haridatta et à Chloñ Somaçaṛman, tous (deux) bhāgavata serviteurs’ (C VII: 183).
(9) ... vaudi 3 kāṭāha 5 svok 10 syañ tap prāṃ jāṇyāḥ padigāḥ 4...
(K.263D: 14), ‘... 3 vaudi; 5 bassines, 10 plateaux, soit 15 jāṇyāḥ; 4 crachoirs;...’ (C IV: 137).

A further step toward structural complexity is seen in cases of the following type in which the NP following syañ is introduced by the subordinating conjunction ta without a grammatical head. As far as the data show, this type is not common.

(10) kamrateñ śivāṣrama nu steñ ’aṁ vnaṃ kansa yok kannvāy 3 striñana syañ ta sahodara ’anvi sruk kuti... (K.235D: 24-5), ‘Le seigneur du Čivāçrama et le steñ aṁ de Vnaṃ Kansa prirent trois femmes, leurs nièces, toutes de même mère, provenant du sruk de Kuti...’ (Cœdès & Dupont 1943-46: 117).

(11) neñ syañ ta gi kalpañañ dau ta vrāh kamrateñ ’aṁ śri cāmpēsvaṭa sap chīna (K.99S: 11-12), ‘...voilà ce qui doit être fourni à V.K.A. Čṛī Cāmpēcvara tous les ans’ (C VI: 112).

(12) nau ’aṁjprall punyaṣrama kuti sruk sre bhūṣyākara kīña (sic) dravya phoñ xxxx dravya syañ ta vrāh rājapunya (K.19: 17-8), ‘toutes les fondations, monastères, cellules, villages, rizières, revenus des terres, esclaves, biens de toute sorte... ces biens sont des fondations royales’ (C VI: 146).

With the foregoing type the NP following syañ ta is realised as N, as Dem.+N, and as N+N. The next structural type, which is the most abundantly represented of all in the Angkorian data, consists of syañ ta followed by a Verb phrase (VP). This is the same as saying that syañ is followed by an NP consisting of ta+VP; viz., a ‘headless ta phrase’ in which the ta subordinates the VP not to the preceding syañ but to some such unmanifested headword as ‘nak ‘person’ (qv. Jenner 1981).

(13) steñ ’aṁ śivasoma nu steñ ’aṁ vāmaśiva syañ ta cat śivāṣrama sthāpanā vrāh noḥ (K.235D: 7-8), ‘Le steñ aṁ Čivasoma et le steñ aṁ Vāmaçiva, ensemble, établirent le Čivāçrama, y fondèrent un sanctuaire’ (Cœdès & Dupont 1943-46: 112).

(14) grāmavṛddha syañ ta sapatha kathā ruva bhūmi ’anin ta nirmūla krau gol (K.598B: 39), ‘Les anciens des villages prétèrent serment et dirent que cette terre d’Anin était sans maître et hors des bornes’.

(15) ’nak neñ phoñ syañ ta samayuga yok iss dravya noḥ phoñ ta jā thlai bhūmi (K.207: 27-8), ‘Tous ces gens ensemble ont pris tous ces biens comme prix de la terre’ (C III: 21).

(16) neñ syañ ta dau dār bhūmi nā the vṛāḥ cāmnāṁ (K.425: 8), ‘Tels ont été ceux qui sont allés demander le terrain pour instituer la prestation’ (C II: 144).
(17) teň bhava x teň kṛṣ teň rudrāṇi teň ke ni teň x te (sic) so teň sa[r]/svatī syāṅ ta dār vṛah karṇāṇāprasā[da]... (K.61B: 7-9), ‘Teň Bhava, Teň Kṛṣ, Teň Rudrāṇi, Teň —, Teň —, Teň Sarasvatī obtinrent toutes de la faveur royale — ...’ (C VII: 22). 

(18) ... nau 'nak ta 'aṅgay ta gi sruk neň nu dharmma sre noň syāṅ ta oy pūjā [ka]mrateń 'aň ti gi dvādaśī phon pratipaksa... (K.100: 3-4), ‘Les gens installés dans ce pays et dans les fondations et sur ces rizières, offrent la pūjā au K.A. le douzième jour de chaque quinzaine...’ (C VI: 215). 

(19) ... vol ekavākya man xxxx [teň hyan]n vasanta nu loň ney vṛah chpār syāṅ 'yat santāna ley... (K.208: 53-4), ‘(Ceux-ci) déclarèrent unanimement que Teň Hyań Vasanta et Loň Ney des saints jardins étaient réellement sans descendance...’ (C VI: 292). 

A related structural type is seen in a few cases in which syāṅ, preceded as usual by an NP, is followed by an NP manifested as the complementiser man+ a VP. In all cases, the man serves as a relative pronoun in the objective case, its antecedent being either animate or inanimate. 

(20) khūṅm šata mṛyā 10 sakarma pūrā, phon neň syāṅ man jvan ta vṛah kamraten 'aň sivalinga thvāy ta vṛah pāda kamraten 'aň śrī sūryavarmmadeva... (K.212A: 11-5), ‘Cent dix esclaves et quatre employés (sakarma), tous ces gens sont offerts au V.K.A. Čivalinga et remis à S.M. le roi Čri Sūryavarmadeva,...’ (C III: 32). 


This leaves us, finally, with a number of passages in which syāṅ, still preceded by an NP, is followed directly by a VP without an intervening ta. Note that in five of the following cases the verb following it is passivised by the marker ti. 

(22) xx khūṅm vṛah neň phon ta daiy ti leh xx vṛah tapasvi bhagavat pāda vṛah kamraten 'aň ta guru jvan kāla sthāpaka syāṅ codita 'aň ti jṇaḥ vṛah tapasvi vidyāspada gi pi mān ta upakal[p]ka jvān ta vṛah sivalinga... (K.523D: 18-24), ‘—tous ces esclaves du dieu et les autres,—Vraḥ Tapasvi Bhagavat Pāda V.K.A. ta Guru me les a offerts au moment où il fit la fondation, en m'incitant, moi qui ai nom Vraḥ Tapasvi Vidyāspada, à faire les préparatifs pour les offrir au saint Čivalinga...’ (C III: 141). 

(23) ľvoḥ ta 1035 šaka pi vṛah pāda kamraten 'aň śrī sūryavarmmadeva ta jā vṛah cau mātpakṣa vṛah pāda karateń 'aň śrī jayavarmmadeva nu vṛah pāda kamraten 'aň śrī dharanaṇḍravarmmadeva svey vṛah dharmmarāja 'aňīeṇ bhagavat pāda kamraten 'aň ta guru śrī divākaraṇḍita jā vṛah guru gi ta thve rājaḥīṣeka man vṛah pāda kamraten 'aň syāṅ thve vṛah dīkṣā ryyan iss siddhānta phon... (K.194: 26-9), ‘En 1035 çaka..., lorsque S.M. Čri Sūryavarmadeva, petit-
neveu en ligne maternelle de S.M. Črí Jayavarmadeva et de S.M. Črí Dharaṇīṇdравarmadeva, accédé à la sainte royauté, il invita le vénérable seigneur Guru Črí Divākarapāṇḍita à remplir les fonctions de Vrahā Guru pour célébrer le sacre royal. Alors Sa Majesté accomplit la sainte initiation (vrahā dikṣā), étudia toutes les sciences (siddhanta),...’ (Cœdès & Dupont 1943-46: 146).

(24) neḥ syañ ti thvāy jā vrahā rājadharma (K.33: 30), ‘Tous ces (dons) sont offerts à titre de fondation royale (rājadharma)’ (C III: 152).

(25) srū ṭa kh’val neḥ phoḥ phle chpār ta nohḥ phoḥn syañ ti jvan ti vrahā kammratan ‘aṅ ekādaśamukha... (K.168: 11-3), ‘Tout ce paddy... et tous les fruits de ces jardins sont offerts à V.K.A. Ekādaśamukha...’ (C VI: 169).

(26) ... nu ‘angāsa nu kaṇje chhāṅ kaṇje kalpita khjėn phoḥn nu duk pāy syañ ti yok dau uk nu sthāli cēḥ dlah (K.353N: 33-4), ‘... nu à distribuer la nourriture, paniers en forme de marmite, paniers kalpita, khjēn, pour mettre la nourriture, tout cela est aussi emporté avec les sthāli, jarres et dlah’ (C V: 142).

(27) ... ti paścima vāyavya ‘aṃyī leṁ sthāpanā sre lvaḥ travāḥ xx dau lvaḥ jen x viṅ uttara prasap vrahā phlu noḥ phoḥn syañ ti jau ta cak svāy ... (K.353S: 31-2), ‘A l’ouest et au nord-ouest, depuis Leṅ Sthāpanā Sre jusqu’au bassin,... revenant jusqu’au bassin au pied de...; au nord, touchant ces chemins sacrés; tout cela a été acheté à Cak Svāy;...’ (C V: 140).

(28) neḥ syañ ti cām cannām ta kammrataṃ jagat srī vrddhesvara (K.33: 23-4), ‘Ces (terres) sont affectées au service des fournitures pour le dieu Črí Vṛddheśvara’ (C III: 151).

These twenty-eight citations illustrate all of the environments in which syañ is found in my Angkorian data; there may be a few more buried in garbled texts. To recapitulate the information just given, we have seven groups of passages distinguished on the basis of the following patterns: Dern. + syañ + N (1-2); Dern. + syañ + NP (3-5); NP + syañ + NP (6-9); NP + syañ + NP (= ta + NP) (10-12); NP + syañ + NP (= ta + VP) (13-19); NP + syañ + NP (= man + VP) (20-21); and NP + syañ + VP (22-28). These seven groups can, of course, be subsumed under the two formulas NP + syañ + NP and NP + syañ + VP. It remains to be seen whether these structural distinctions are useful. Whether they are or not, the number of examples given should be enough to show that not all of the interpretations which have been applied to syañ could be accurate.

Even if it is found that one or more of the above interpretations is correct, any re-evaluation of syañ must begin by challenging all of the interpretations which have been applied to it. In the absence of living informants, our enquiry is essentially a search for common ground underlying every occurrence of syañ. It is reasonable to begin seeking such common ground by testing each of my predecessors’ interpretations against every passage in which syañ is used. I shall spare the reader the
agon of being led through a critical review of the data, and simply assert here that few of the interpretations given above stand up under such a simple test.

For example, the idea of 'solely, entirely' or 'sans exception' would not appear out of place in examples (1), (2), (5), (6), (7) and (11); it would not do in the remaining citations. All other imaginable adverbial ideas have been tried but none seems to fit.

On the other hand, the notion of tout has been particularly beguiling. It is seen in nine or more of the above citations and, as has been said, with these we may group the related idea of ensemble seen in (13) and (15). One thing that seems fairly clear is that in most of these examples the immediate constituents are interpreted as neḥ syaṅ | prāk (1), neḥ syaṅ | mās (2), and so on, in which syaṅ (whether construed as a noun, a pronoun, or a verb) is seen as attributive to neḥ by virtue of its position. However this may be, it is curious that tout would have been no less plausible in nine other of our citations, namely (4), (5), (8), (11), (13), (18), (21), (25), and (28). The fact that it is not used in these cases seems to reflect uncertainty as much as inconsistency. The main point to be considered, however, is that in examples (7), (9), (14), (22) and (23) the idea of tout is pretty clearly excluded. The fact that syaṅ has been taken in senses other than tout, or has been left unexpressed, cannot be ignored, and the more one weighs those cases in which it is understood as tout the more one believes that this idea is contextually derived and that such dubious cases as example (8) were influenced by the apparent preponderance elsewhere of the tout idea.

The same kind of review must be made in weighing the possibility that syaṅ is a relative pronoun, as Finot thought. In this case a cursory examination suffices. While one might be tempted to take it as a pronoun in as many as fifteen of the above citations, to do so does violence to each passage. We now know enough about the role of all the other elements in most of these examples to be on our guard against such interpretations, though our knowledge is still imperfect. I have recently demonstrated the possibility of 'double' demonstratives in Old Khmer (Jenner 1982) and this might suggest that syaṅ man could be something of the kind. But consider the following:

(29) neḥ bhūmi ta roḥh neḥh man vāp 'amarānanta duñ syaṅ man jvan ta vraḥ noh... (K.693B: 20-1).

A close rendering of this would be 'These lands, aforesaid, which vāp Amārānanda bought syaṅ what (he) offers to the sanctuary...' No conclusions can be drawn from this or from our two other syaṅ man citations, (20) and (21), since all three passages might have a zero copula with syaṅ, conceivably, duplicating the office of man. But all the other evidence indicates that this is not so.

The claim that syaṅ fulfils an anaphoric function is as difficult to disprove as it is to prove. Re-examination of our citations yields mixed

10. Cf. CV: 208: ‘... telles sont les terres que Vāp Amārānanda a achetées et qui sont offertes à ce temple...’
The form syaṅ in Angkorian Khmer

results: in six cases the possibility seems very unlikely, while in the remaining cases the possibility is present. In citations (1) to (5) plus (13), (16), (20), (24) and (28) together with citation (21), in which we have respectively neḥ syaṅ and noḥ syaṅ conceivably standing to the left of a binary cut, one wonders why the syaṅ and not the neḥ or noḥ should be anaphoric, and also why the alleged anaphor should be attributive to the demonstrative, as it would be, when syaṅ alone or with a following neḥ or noḥ would be expected. The claim for syaṅ as anaphoric pronoun (or particle) is too weak to be applied to all of our citations, let alone most of the data collected but not given here, and is hence unconvincing.

When we turn to consider the remaining examples given above, we find them nearly equally divided between those in which syaṅ is not expressly rendered at all and those in which it seems to be rendered by appropriate forms of ētre. The former comprise (4), (14), (18), (22), (23), and probably (21) and (28); the latter comprise (5), (12) (19), and probably (21), (25), and possibly (28). If none of my predecessors has advanced the claim that syaṅ is a copula or copula-like verb, it has not been made entirely clear earlier in this discussion that such a claim is almost implicit in a good many of their interpretations both of Old Khmer and of Middle Khmer texts. In a remarkable number of cases it is as if each context forced them into such a position without their being aware of it. Indeed, the moment we take a new look at our twenty-nine citations and consider them apart from their renderings, we are obliged to allow that many of them could equally well, considering the nature of our enquiry, be analysed as neḥ | syaṅ prāk (1), neḥ | syaṅ mās, and so on, in which syaṅ could be functioning as a copula identifying neḥ or another subject with a predicate. To test this hypothesis, I give here and now my own fairly close versions of our twenty-nine examples, my tentative equivalents of syaṅ being shown in italics:

(1) ‘One vaudī, one svok, four... (to hold) water for washing the feet, these consisting of silver.’
(2) ‘... one ring, one jewel (belonging) thereto, 2 anklets, one khse chdval, these being of gold.’
(3) ‘... me ’Yak, me ṇaṃ, me Des, me Sān, me Dvat—these being of the land of Amarālāyā.’
(4) ‘... Those having this edict in their safekeeping (are) the lord Śrī Śatyāyudha and the lord Śrī Ripumatha, these being sons of the lord Śrī Pṛthivīvārendra.’
(5) ‘Our high lords the yogin who have been ordained here give over the northeast road, the garden on the tīrtha [bathing-place] (and) the flower garden, these being their several pious works...’
(6) ‘... 2 silver cuspidors, 2 clasps, 2 silver vaudī, 1 dramvāṅ vessel, 1 figured vessel consisting of hanīra [an unidentified alloy], 3 vessels with spouts, 6 vessels of medium size...’

11. The sequences syaṅ neḥ and syaṅ noḥ are not attested in Old Khmer. This is hardly the place to take up the question of anaphora and its ramifications. Those interested should consult Patricia A. Lee (1981), especially 6.5 on Reference, Anaphora, Deixis.”

235
(7) "... (they) then selected others (to serve) as a fee'.
(8) "The kampana Sāntilakṣmi begat the chloini Haridatta (and) the chloini Somaśarman, (who) were bhāgavata in (divine) service.'
(9) "... 3 vaudi, 5 katāha (with) 10 svok totalling fifteen, an upright panel, 4 cuspidors...''
(10) 'The high lord of the Śivāśrama and the steñ 'añ of Vnaṃ Kansā took three female nephews/nieces (who) were co-uterine (and) from the land of Kuti...''
(11) 'These (items) constitute the endowment gong to the vrañ12 Our High Lord Śri Cāmpeśvara each year.'
(12) 'All (these) gifts of āśramas, cells, villages, ricelands, rents, slaves (and) objects... (all this) valuable property comprises the vrañ the royal gift.'
(13) 'The steñ 'añ Śivasoma and the steñ 'añ Vāmadeva are the ones who founded the Śivāśrama (and) set up the image therein.'
(14) 'The village elders are the ones who declared under oath that the land of 'Anin (was) uninhabited (and) outside (anyone's) bounds'.
(15) 'These individuals are ones who joined together to take all of these possessions as equal to the value of the land.'
(16) 'These are the ones who went and claimed the land (as) a place on which to establish the vrañ the foundation.'
(17) 'The teñ Bhava..., the teñ Krṣ, the teñ Rudrāṇi, the teñ Ke Ni (??), the teñ ...te So (??), (and) the teñ Sarasvatī are the ones who claimed the vrañ the royal favour...''
(18) '... the individuals who are settled on this land and on its gift of ricefields are ones who (shall) offer a sacrifice to Our High Lord on the twelfth days of each fortnight...''
(19) '... (they) declared in one voice that... the teñ hyaṇ Vasanta and the loṇ Ney of Vrañ Čhpaṛ were ones who had no family whatever...''
(20) 'One hundred and 10 slaves (and) four helpers, —these are what (he) offered to the vrañ Our High Lord of the Śivalinga (and) presented to His Majesty Our High Lord Śri Suryavarman.'
(21) 'Those same ricefields are what (they) gave to the vrañ Our High Lord of the Śivalinga and to the vrañ Our High Lord Śivapāda.'
(22) '... these slaves of the sanctuary and others,... the vrañ the tapasvin bhagavat pāda (and) the vrañ Our High Lord the Guru offered (them) on the occasion of (his) consecration of the image (and) enjoined me, named the vrañ the tapasvin Vidyāspada, to make arrangements to offer (them) to the vrañ the Śivalinga...''
(23) 'In the year 1035 of the Śaka era, when His Majesty Our High Lord Śri Suryavarmadeva—who was the vrañ the maternal grandchild of His Majesty Our High Lord Śri Jayavarmadeva and of His Majesty Our High Lord Śri Dharanindravarmaneva—(began to) exercise the just kingship, (he) invited the bhagavat pāda Our High Lord the Guru Śri Divākaranapandita to serve as the vrañ the guru, he (being) the one

12. In these close translations, vrañ is a noun usually but not always marking any divine or royal being or object, and functions as a headword with which the following NP is in apposition. (Note that the purpose of the closeness of the translations is to show the structure of the Khmer rather than the sense.)
to celebrate the king's consecration, (and) His Majesty Our High
Lord *made* his preparations (and) learned all of the *siddhānta*...

(24) ‘These were offered (to serve) as a *vraṭ* a royal good work.’

(25) ‘The paddy in these granaries (and) the yield of these fields are
offered to the *vraṭ* Our High Lord Ekādaśaṃukha...’

(26) ‘... *angāsa* vessels, vessels (in the form of) kettle baskets, *kalpita*
baskets and *khjeṅ*, (and) vessels in which to keep (cooked) rice are
also *taken away* together with *sthāli*, *ceh* (and) *dlah*.’

(27) ‘... westward (and) northwestward from Len Śṭāpanā Sre to the
reservoir.... (and) on to the outskirts of... (and) back northward to
meet the *vraṭ* the roads (which) were *acquired by exchange* from Cak
Śvāy...’

(28) ‘These (lands) are *assigned* to the endowment of Our High Lord of
Creation Śrī Vṛddeśvara.’

(29) ‘These lands, aforesaid, which *vāp* Amarānanda bought are (he)
offers to the sanctuary.

The first twelve of these new versions, together with my interpretation
of the three examples with *man*—(20), (21) and (29)—seem plausible
enough; but my restatement of examples (13) to (19), with *syāṅ ta* + VP,
and of examples (22) to (28), with *syāṅ* + VP, seems less plausible at first
 glance. Having considered these two types for more years than I care to
admit, however, I have come to believe that the difficulty lies not in the
value I assign to *syāṅ* but in the circumstance that *syāṅ* forms part of two
patterns of expression one of which seems to lie just beyond our
comprehension. The seven examples with *syāṅ ta* + VP strike me as not
especially exotic variations played upon the simple VP. Thus I suspect that
example (13) does not mean precisely what it would mean in fluent
English, but is a weak expansion of ‘The *stēṅ* ’aṅ Śivasoma and the *stēṅ*
’aṅ Vamaśiva founded the Śivāśrama (and) set up the image therein.’ The
seven examples with *syāṅ* + VP are sharply divided into two groups. The
last five—*syāṅ ti thvāy* (24) ‘were offered’, *syāṅ ti jvaṅ* (25) ‘are offered’,
*syāṅ ti yok dau* (26) ‘are taken away’, *syāṅ ti jau* (27) ‘were acquired by
exchange’, and *syāṅ ti cām* (28) ‘are assigned to’—seem quite admissible
until it is remembered that my English equivalents would more naturally
be expressed by *ti thvāy, ti jvaṅ*, and so on, without *syāṅ*. These five cases,
then, are as perplexing as the two examples of the remaining group, *syāṅ
codita* (22) ‘enjoined’ and *syāṅ thve* (23) ‘made’. For all that, serial verb
constructions are so widely used during all periods of the language for
which we have documentation that, for the moment at least, I am
unwilling to reject the hypothesis on the grounds that I cannot fathom
these seven cases. It seems more prudent to allow that we are concerned
here with a periphrastic construction the effect of which cannot yet be
determined. This situation, it will be remembered, was alluded to in my
first two paragraphs above.

Rather than leaving the matter here, however, I should, because of its
importance to our understanding of Old Khmer, attempt to account for the form and function of syan. If its function is indeed copular, as I claim, how does it happen that the form itself suddenly appears in A.D. 895 and enjoys a life of nearly seven centuries, only to vanish into thin air? If the function it serves is an innovation, we may suppose that the form itself either has a source in the earlier lexicon or is a loan from some other language. I have not found any cognates of syan in Mon or other Mon-Khmer languages and have, as I believe, considered all the possibilities outside the latter family. In fact, the only light thrown on the question comes from the claims of Aymonier and Pou. As has been explained, Aymonier opined that Angkorian syan is the source of a modern ‘sin’, without specifying whether he meant modern siña ‘nearly’ or modern siña ‘to rest’. On the other hand, Pou has stated categorically that Middle Khmer syañ led to modern siña ‘nearly’. Whether we consider siña or siña, these claims are ones of which I was justifiably sceptical when they first came to my attention, since their authors nowhere troubled to explain the radical line of semantic development involved. In either case, the leap was one that I could not imagine. Yet there is a way of linking modern siña ‘nearly’ with modern siña ‘to rest’, and of linking both with Angkorian syañ—provided that a copular or quasi-copular function is assigned to the latter.

It can be pointed out, first of all, that the semantic range of modern sina is exceedingly narrow: if dictionaries define it as more than ‘nearly, almost’, it is only to ring the changes on this same idea. This is enough to suggest that the modern meaning is a restriction of some broader one. What is more, forms having only adverbial senses in Khmer are rare. This circumstance permits us to suppose that the adverbial sense of sina represents a reduction from a more general verbal idea. In attempting to reconcile sina and sina therefore, we may posit an earlier transitive meaning for sina ‘to be near to, just short of’.

As to modern sina ‘to rest’, this form is attested once in pre-Angkorian (K.44B: 10) with the sense of ‘to dwell in’. In Angkorian it occurs thirty-three times, most often with the meaning ‘to officiate’ but also four times with the meaning ‘to dwell in’ (K.754:13; K.413/IV: 12; K.56C: 37; K.70: 16), and once with the meaning ‘to depend on’ (K.369: 6,7). It persisted through Middle Khmer and into the modern language, where its sense is restricted to ‘(of monks) to sleep’ and ‘(of magicians) to perform (a rite)’. The Middle Khmer form is unquestionably the source of Thai /sín/ ‘to stay, enter and inhabit, possess, (of spirits) haunt’, which tells us much about the range of the Old and Middle Khmer forms. These attested meanings are all we need to show that the idea of ‘nearly’ represented by modern sina /sín/ is an entirely orthographic specialisation of the Old and Middle Khmer verbal idea dimly recognised in modern sina /sâñ/ [svern] as defined above. This recognition, in turn, forces upon us the insight that Angkorian copular syañ is probably a doublet of Angkorian sîn /sîn/, as defined above, and that all of the forms in question here make
The form syañ in Angkorian Khmer

up a fairly compact semantic cluster which can be set up as:

1. (intr.) to stand, hold still, be at rest;
   (a) to rest, lie; to repose, sleep;
   (b) to rest, remain, continue;
   (c) to be inherent or present; to exist, be.

2. (tr.) to remain in or at, inhabit; to dwell in;
   (a) to lie or reside in, be in the presence of, be near to;
   (b) to consist of, comprise, constitute;
   (c) to be present at, preside over, officiate at; to perform or
   celebrate (rite).

The meaning of Angkorian siṅ and of the single occurrence of pre-
Angkorian siṅ having been given above, it is not unreasonable to suppose
that the syañ doublet (if such it is) of Angkorian siṅ was restricted to the
remaining semantic field. All of the data collected show no overlap
between Angkorian siṅ and syañ and suggest, rather, that the latter was
narrowly limited to the intransitive sense of 'to remain, be' and to the
transitive sense of 'to consist of, comprise'.

It is hardly necessary to add that we cannot at this late date know the
genesis of the syañ doublet of Angkorian siṅ. The correspondence
between pre-Angkorian and Angkorian siṅ and what we may now
represent as modern siṅ ~ siṅ is normal, while that between pre-Angkorian
siṅ, Angkorian syañ, and modern siṅ ~ siṅ is seen in only a handful of
cases—most notably in Angkorian tyañ /diŋ/ : modern tiña /dỳŋ/ 'to
know' and Angkorian yat /qìat/ : modern it /qāt/ 'to lack'. If the doublet
relationship is tenable, Angkorian syañ is therefore short /sìañ/ or,
possibly, /suəñ/. That the hypothesis of a doublet relationship explains so
much overrides, as I believe, the apparent impossibility of our ever
knowing how and why these doublets arose.

REFERENCES


1883. Quelques notions sur les inscriptions en vieux khmer. J. Asiat. 1883 (1), 441-505; (2), 199-228.


Haas, Mary
Jenner, P. N.
Kerr, A. D.
Lee, Patricia A.
Lewitz, Saveros
McFarland, G. B.
Parmentier, H.
Pou, Saveros
Sethaputra, So
Supina

The Old Mon (OM) inflectional prefix s-, referred to by Shorto (1971: xxii) as the 'hypothetical', marking in particular 'futurity', has not received the attention it deserves. It has been extensively commented upon by Duroiselle in 1921, while translating the Ananda plaques and a number of Mon nissaya, and has again been mentioned in Jacob's comparative study of 1963, albeit briefly.

The analysis of OM s- is especially challenging for three reasons:

(i) tense, aspect and modality cannot clearly be separated;
(ii) cognates in other Mon-Khmer languages have so far not been identified;
(iii) given the fact that at later stages of Mon the prefix s- has become obsolescent, with no such prefix left in the spoken language today, it can only be analysed on the basis of the attested OM and Early or Epigraphic Middle Mon (EMM) corpus, and must thus be reconstructed internally.

The semantic complexity of OM s- would go unnoticed were it not for the fortunate circumstance that a number of s- inflected forms occur in glosses accompanying frescoes (the Myinkaba Kubyauk-gyi, for instance) or glazed plaques (at the Ananda). While occurrences of s- in the Shwezigon inscription (EB #1)* would strongly suggest the identification of a tense/modal prefix 'future/irrealis', the glosses, on the contrary, would imply the aspect 'ingressive'. This is especially clear in plaque #565 (Duroiselle #28) and plaque #721 (Duroiselle #189), where the action referred to in the verb (#565 scis 'to descend', s-, and #721 slop 'to enter', s-, from the bases cis and lop respectively) has not yet been accomplished, but has already been initiated. It would then appear that art-historical evidence such as frescoes, plaques, or illuminated manuscripts, can no longer be ignored by linguists.

The prefix s- was fully productive in OM. It was prefixed to bases with simple and complex initials. Because of this low selectivity of its host, characteristic for inflectional affixes, it had no allomorph. Whenever syllabic, the vowel in the minor syllable showed allophonic variation a - i in its orthographic form, i being confined to base-initials /s-, k-/.

---

1. Sources for this study: inscriptions have been quoted from transcriptions as published in Epigraphia Birmanica (EB), and, in the case of inscriptions from Thailand, from my own notes. Complete bibliographical references can be found in Shorto (1971), with addenda in 1990 in Bauer (in press) which also includes a glossary.

241
former conditioning, no doubt, the front-colouring of the vowel.

Base types and their inflected forms are given in Fig. 1; all but the last are canonical (native) forms in OM.

OM s- is found attached to Indo-Aryan (IA) loans. In these cases the loans are invariably nouns in the source language, except Pali (P) nimanteti, but are to be analysed in OM grammar as verbs; non-inflected forms are attested as verbs. s- inflected IA loans are given in Fig. 2. In one instance, OM s- inflection has led to re-analysis: sthapana 'to erect, establish', Sanskrit (Skt.) sthāpana ‘founding, storing’, is re-analysed as a simple form OM thapana.2

OM s- may be combined with other verbal affixes such as the causative p- and -a-, and the frequentative -N-. It may, however, not be combined with the attributive infix -m-, yielding non-predicative forms. Typical OM paradigms are set out in Fig. 3.

OM s- inflected forms may occur in declarative sentences, commands, and questions. In serialised constructions only the head-verb is inflected, as in

<table>
<thead>
<tr>
<th>deh scis das</th>
<th>(III.C.9)</th>
<th>'he will come down and be born'</th>
</tr>
</thead>
<tbody>
<tr>
<td>'ey s'or p'ar pun'ar dunhic</td>
<td>(I.G.31)</td>
<td>'I will cause to behave well'</td>
</tr>
<tr>
<td>sāntu binton</td>
<td>(I.B.49)</td>
<td>'[the king] shall instruct [the people]'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>base</th>
<th>s- inflected form</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>nom</td>
<td>'to be'</td>
</tr>
<tr>
<td>sVC</td>
<td>sāp</td>
<td>'to cover the surface of'</td>
</tr>
<tr>
<td>?VC</td>
<td>?ar</td>
<td>'to cover the surface of'</td>
</tr>
<tr>
<td>CCVC</td>
<td>tluq</td>
<td>'to come'</td>
</tr>
<tr>
<td>CaCVC</td>
<td>sāgeh</td>
<td>'to enrich'</td>
</tr>
<tr>
<td></td>
<td>jānak</td>
<td>'to enlarge'</td>
</tr>
<tr>
<td>CaCCVC</td>
<td>tawil</td>
<td>'to attend on'</td>
</tr>
<tr>
<td>CCVCVC</td>
<td>prakāt</td>
<td>'to be clear', Skt. prākatya</td>
</tr>
</tbody>
</table>

Fig. 1. Base-types and their inflected forms in OM

2. Skt. nouns are re-analysed in Khmer as verbs. See Jacob (1977, esp. para. 4.4.2: 163-4).
Old Mon s-

<table>
<thead>
<tr>
<th>base</th>
<th>inflected form</th>
<th>gloss</th>
<th>IA form</th>
</tr>
</thead>
<tbody>
<tr>
<td>prathana</td>
<td>sprathana</td>
<td>'to pray'</td>
<td>Skt. prārthanā 'prayer'</td>
</tr>
<tr>
<td>byapar</td>
<td>sbyapar</td>
<td>'to render service'</td>
<td>Skt. vyāpāra 'service'</td>
</tr>
<tr>
<td>prakot</td>
<td>sprakot</td>
<td>'to be clear'</td>
<td>Skt. prākāya</td>
</tr>
<tr>
<td>nirban</td>
<td>snirban</td>
<td>'to attain Nirvāṇa'</td>
<td>Skt. nirvāṇa</td>
</tr>
<tr>
<td>cuti?</td>
<td>scuti?</td>
<td>'to die'</td>
<td>P. cuti 'death'</td>
</tr>
<tr>
<td>nimon</td>
<td>snimon</td>
<td>'to invite'</td>
<td>P. nimaneti</td>
</tr>
</tbody>
</table>

Fig. 2: s- inflected IA loans in OM

<table>
<thead>
<tr>
<th>base</th>
<th>frequentative</th>
<th>causative</th>
<th>hypothetical</th>
<th>gloss (base)</th>
</tr>
</thead>
<tbody>
<tr>
<td>duk</td>
<td>-N-</td>
<td>*pduk</td>
<td>ponduk</td>
<td>spandonuk</td>
</tr>
<tr>
<td>yuk</td>
<td></td>
<td>ponyuk</td>
<td></td>
<td>sponyuk</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>'to be apparent'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>'to dance'</td>
</tr>
</tbody>
</table>

Fig. 3: OM non-attributive verbal paradigms

Noteworthy is the absence of suppletive forms in negated contexts, a common feature in modern Mon. The OM verbal negative kah may be inflected for the ‘hypothetical’, OM skah (as in #7, Fig. 4), or the verbal negative kum (< Khmer), OM skum (as in #10, Fig. 4). If the verbal negative is inflected for the ‘hypothetical’, the following verb is not.

Problematic are, however, cases where it is not the verbal negative which is inflected but the following verb (as in #9, although there it is sak which is never attested as an inflected form) or the preceding clause or sentence (as in #8).

Purposive clauses (‘in order to, so as to’) are introduced by the clause-head OM tnas or OM dinah; in the latter case, the head is linked to the clause by na, which in some instances is attested without the clause-head (Ananda plaques #715, #780, #798; Kubyauk-gyi gloss #164; see Figs. 5, 7).

Whenever the clause is linked by na the following verb must be inflected for the ‘hypothetical’. In Middle Mon (MM) clause-linkage also occurs with ma, but, except for tnah as clause-head, s- inflection is not obligatory. However, as in OM, na clause-linkage requires s- inflection. Rules of linkage and inflection are listed in Figs. 8, 9. The exact details of clause-subordination in Mon need not concern us here; suffice it to say that s- inflection is obligatory when na linkage occurs in OM and MM.

To conclude: An examination of OM glosses supports the interpreta-
N
-

Translation of OM gloss

565. (Temiya) descending on to the ground

601. They (the two attendants) bring the alms-bowl (and) give it (to the Mahājanaka)

621. Śyāma informs (his parents) that he is going to enter the forest

721. Mahosadha enters the king’s house

756. (Bhuridatta) asks his father for leave to go and observe the silas

816. Vessantara goes to his father

817. Mahasaddha tells (the king) that he is going (to Uttarapancāla)

780. Vidhura tells (Puṇṇa the Yakkha) he is ready to go (with him)

798. (Vessantara) asks for gold to give alms

Comment EB II.1 (text)

... Temiya doubts whether he has still the control of his limbs and gets down from the chariot to ascertain

... he is in the act of descending; his left leg is already outside, his foot resting on a stone

... bids the two servants bring him yellow robes and an alms bowl. In the casket held by the first servant are the robes. The other is holding the bowl.

... he informs his parents that he intends going to the forest

He is asking from his father permission to observe the silas. The king and the queen are seated, he is standing in front of them.

When everything is over, Vessantara, accompanied by his wife and children, goes to visit his father. Vessantara, then Jāli, his son, and near the boy, Maddi carrying Kaṇhajinā on her hip in front of the king, Mahosadha with, being who he is, three attendants

Comment EB II.2 (plates)

... bids the two servants bring him yellow robes and an alms bowl. In the casket held by the first servant are the robes. The other is holding the bowl.

He is asking from his father permission to observe the silas. The king and the queen are seated, he is standing in front of them.

When everything is over, Vessantara, accompanied by his wife and children, goes to visit his father. Vessantara, then Jāli, his son, and near the boy, Maddi carrying Kaṇhajinā on her hip in front of the king, Mahosadha with, being who he is, three attendants

Fig. 6. Duroiselle’s translations, comments, and annotations to the Ananda glosses; text presented in Fig. 5.
Fig. 7-A: *OM s-* inflected forms in the Kubyauk-gyi glosses

<table>
<thead>
<tr>
<th>Kubyaunik-gyi gloss (Myinkaba)</th>
<th>Luce/Ba OM gloss</th>
<th>Shin page ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 'The Lord stays in the Anupiya mango (grove) ... alms-bowl, is about to enter the city of Rājagaha'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>131 'The Lord Buddha is about and become a Buddha'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>164 'Kings of the Asuras, Vepacitti, Succitti, Pahārāda, and Rāhu Lord of Asuras, stay plotting (?) to go and climb up to attack Tāvatīṁsa'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>185 'At the site of Nāgavana garden, all the spirit Yakkhas who dwelt in Lāṅkā were about to fight. The Lord Buddha stations himself in the sky above, produces a hurricane, creates darkness, and frightens them (till) they beg release from terror’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>189 'At Samanakūṭa Deva Samāna who dwelt there, begged the Lord Buddha for something which might make an object of worship. The Lord stroked his head, got one hair, and gave it to him’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>190 'In Nāgadipa two Nāga kings, nephew and uncle, were about to fight. The Lord Buddha came and confronted (?) the two Nāga kings, and warned and reconciled them again to one another’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>195 'When the king Devanampiyatis and the reverend Mahāmāhin were about to plant finally the great Relic, the great Bodha tree ascended and abode in a chamber of clouds for 7 days’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>196 ‘This is when king Dhammāsook wishes to send the great relic of the Tree, with the therī Sanghamitta, to the island of Sinhgal.’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>197 'This is when king Dūthagāmāni plans to build the Ratnaceti’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 7-B: Translation of the Kubyauk-gyi glosses presented in Fig. 7-A.
<table>
<thead>
<tr>
<th>clause-head</th>
<th>clause-linkage</th>
<th>main verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>OM  tonś</td>
<td>mə- nə- s-</td>
<td>AUX* JCT* JCT*</td>
</tr>
<tr>
<td>OM  dyah</td>
<td>nə- s-</td>
<td>V</td>
</tr>
<tr>
<td>OM</td>
<td>nə- s-</td>
<td>V</td>
</tr>
<tr>
<td>MM  tonś</td>
<td>nə- s-</td>
<td>V</td>
</tr>
<tr>
<td>MM  təh</td>
<td>mə- s-</td>
<td>V</td>
</tr>
<tr>
<td>MM  tonś</td>
<td>s- mik</td>
<td>V</td>
</tr>
<tr>
<td>MM  təh</td>
<td>nə s- mik</td>
<td>V</td>
</tr>
<tr>
<td>MM  hot</td>
<td>mə- (s-)</td>
<td>V</td>
</tr>
<tr>
<td>MM  swak</td>
<td>mə- (s-)</td>
<td>V</td>
</tr>
<tr>
<td>MM</td>
<td>mə- (s-)</td>
<td>V</td>
</tr>
<tr>
<td>LM**  tonś</td>
<td>nə- s-V</td>
<td>swak</td>
</tr>
<tr>
<td>LM</td>
<td>s- ku?</td>
<td>V</td>
</tr>
<tr>
<td>SM**</td>
<td></td>
<td>k3?</td>
</tr>
</tbody>
</table>

Fig. 8: Rules for clause-linkage and inflection in Mon

*AUX = auxiliary complex; JCT = junctural complex. (Ed.)
**LM = Modern Literary Mon; SM = Modern Spoken Mon. (Ed.)

<table>
<thead>
<tr>
<th>OM</th>
<th>EMM</th>
<th>LM</th>
<th>SM</th>
<th>nə- s-V</th>
<th>mə (s-)V</th>
<th>AUX</th>
<th>JCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>tonś</td>
<td>təh</td>
<td>təh</td>
<td>EMM</td>
<td>LM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dyah</td>
<td>-</td>
<td>-</td>
<td>OM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>na</td>
<td>( )</td>
<td>-</td>
<td>OM</td>
<td>EMM</td>
<td>LM/SM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hot</td>
<td>-</td>
<td>-</td>
<td>EMM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>swak</td>
<td>swak</td>
<td>-</td>
<td>EMM</td>
<td>MM/LM</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 9: Mon clause-heads (‘in order to, so as to’) and linkage-types
In the process of collecting, sifting and analysing data for an historical grammar of Mon, I cannot yet offer definite grammatical statements in certain areas, OM s- included. It is, however, my hope that this ordered set of data presented here (the lists with s- inflected forms from the Kubyauk-gyi glosses and the Ananda plaques are exhaustive; with the exception of Mg. 2, no s- inflected forms have been found on votive tablets from Thailand and Burma) will be of use to other linguists.

REFERENCES


<table>
<thead>
<tr>
<th>Fig 4-A</th>
<th>Ananda plaques (Fig.5)</th>
<th>Kubyauk-gyi (Myinkaba) glosses to frescoes (Fig.7-A)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>scis</em></td>
<td>565. <em>scis</em></td>
<td>28. <em>slop</em></td>
</tr>
<tr>
<td><em>s'or</em></td>
<td>601. <em>sikil</em></td>
<td>131. <em>spluh</em></td>
</tr>
<tr>
<td><em>sgür</em></td>
<td>621. <em>slop</em></td>
<td>164. <em>s'ar</em> (na +)</td>
</tr>
<tr>
<td><em>sgo'</em></td>
<td>721. <em>slop</em></td>
<td>185. <em>smic</em></td>
</tr>
<tr>
<td><em>spa</em></td>
<td>756. <em>s'ar</em></td>
<td>189. <em>spa</em></td>
</tr>
<tr>
<td><em>spaqot</em></td>
<td>816. <em>s'ar</em></td>
<td>190. <em>smic</em></td>
</tr>
<tr>
<td><em>scuti, stin, scuti, stluh</em></td>
<td>715. <em>s'ar</em> (na +)</td>
<td>195. <em>stil</em></td>
</tr>
<tr>
<td><em>skah sak, skah sak</em> (NEG)</td>
<td>780. <em>s'ar</em> (na +)</td>
<td>196. <em>sclo'</em></td>
</tr>
<tr>
<td><em>sik'im</em> (NEG)</td>
<td>798. <em>skil</em> (na +)</td>
<td>197. <em>skindam</em></td>
</tr>
<tr>
<td><em>sdas</em> (NEG)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>skum</em> (NEG)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 10: List of s- inflected forms presented in their syntactic contexts in Figs. 4-A, 5, and 7-A.