LINGUISTIC COMPARISON
IN SOUTH EAST ASIA AND THE PACIFIC

School of Oriental and African Studies
University of London
1963
Price 42s. net
LINGUISTIC COMPARISON IN SOUTH EAST ASIA AND THE PACIFIC

Editor: H. L. Shorto
LINGUISTIC COMPARISON
IN SOUTH EAST ASIA AND THE PACIFIC

School of Oriental and African Studies
University of London


1963

Price 42s. net
CONTENTS

Introduction, by Eugénie J. A. Henderson 1

Linguistic Comparison, by R. H. Robins 7

The Comparative Method as Applied to Non-Indo-European Languages, by Nils M. Holmer 13

The Morphological Structure of the Austroasiatic Languages, by Nils M. Holmer 17

The Comparative Study of the Austronesian Languages, by E. M. Uhlenbeck 24

Notes on the Comparison of Two Languages (with and without a Genetic Hypothesis), by G. B. Milner 28

The Structural Patterns of Northern Mon-Khmer Languages, by H. L. Shorto 45

Prefixation and Infixation in Old Mon, Old Khmer, and Modern Khmer, by Judith M. Jacob 62

Thai and Vietnamese: some Elements of Nominal Structure Compared, by P. J. Honey and E. H. S. Simmonds 71

Prosodic Analysis, and Phonological Formulae, in Tibeto-Burman Linguistic Comparison, by R. K. Sprigg 79

A Comparison of Arakanese and Burmese Based on Phonological Formulae, by R. K. Sprigg 109

Chinese, Thai, and Miao-Yao, by G. B. Downer 133

The Position of the Munda Languages within the Austroasiatic Language Family, by Heinz-Jürgen Pinnow 140

The Languages of New Caledonia, by A.-G. Haudricourt 153

Contribution to a Consideration of the Present State of Knowledge in the Field of Austronesian Languages, by H. Kähler 156
INTRODUCTION

By Eugénie J. A. Henderson

In July, 1961, a group of scholars from Western Europe, whose work has in one way or another brought them face to face with the special problems of linguistic comparison in South East Asia and the Pacific, met in London at the School of Oriental and African Studies to compare notes on aims and methods and to assess present achievements and future requirements in the area. Those participating were: Mrs. A. J. Allott (London), Mr. E. C. G. Barrett (London), Professor C. C. Berg (Leiden), Mr. J. C. Bottoms (London), Mr. A. H. Christie (London), Mr. G. B. Downer (London), Professor S. Egerod (Copenhagen), Professor J. Gonda (Utrecht), M. A.-G. Haudricourt (Paris), Miss E. J. A. Henderson (London; Chairman), Dr. Hla Pe (London), Professor N. Holmer (Lund), Mr. P. J. Honey (London), Dr. C. Hooykaas (London), Professor H. Kähler (Hamburg), Professor F. B. J. Kuiper (Leiden), M. R. H. Leenhardt (Paris), Miss M. B. Lewis (London), Dr. E. M. Mendelson (London), Mr. G. B. Milner (London), Professor E. G. Pulleyblank (Cambridge), Mr. R. H. Robins (London), Professor N. C. Scott (London), Mr. H. L. Shorto (London; Secretary), Mr. E. H. S. Simmonds (London), Mr. R. K. Sprigg (London), Professor E. M. Uhlenbeck (Leiden). Papers submitted by Mrs. J. M. Jacob (London) and Dr. H. J. Pinnow (Berlin) were read in their absence.

The study group met on three successive days, the morning and afternoon sessions being devoted successively to (i) methodological problems in the application of comparative methods to the area; (ii) the Austroasiatic language family; (iii) other language groups on the South East Asian mainland; (iv) Austronesian languages and their links with the mainland; and (v) the state of descriptive studies and the future prospects for comparative work. Some of the papers read at the meetings are reproduced in this volume; of the lively discussions that followed the papers it is possible to give only the briefest hint in these introductory notes.

The papers here published are arranged so as to carry the reader along a path which progresses naturally from the consideration of general principles of comparative linguistics to the application of such principles to particular comparisons within the South East Asian field. In the opening group of papers, Mr. Robins sketches the historical development of the subject; Professor Holmer, in two papers, examines the limitations imposed on the orthodox interpretation when comparison is carried outside the Indo-European field, and conducts a typological survey of the languages of the area in which he sums up structural similarities and contrasts that cut across genetic boundaries, in order to find features and lines of development valid throughout the region; Professor Uhlenbeck discusses the
application of the comparative method to the particular field of Austronesian. A second group of papers, in which structural systems rather than lexical items are compared, is introduced by Mr. Milner’s contribution, in which he demonstrates the different conclusions to be drawn from the comparison of Fijian and Samoan according as one accepts a genetic hypothesis or not. This is followed by Mr. Shorto’s discussion of the structure of Northern Mon-Khmer languages, with special reference to phonology; Mrs. Jacob’s examination of the morphological systems of Old and Modern Khmer and of Mon; and a syntactic comparison of Thai, Vietnamese, and Modern Chinese by Mr. Simmonds and Mr. Honey. In all these papers the concept of the linguistic area comes to the fore. Moving on to the comparison of smaller linguistic units, Mr. Sprigg’s two papers apply the techniques of prosodic analysis to the phonological comparison of certain Tibeto-Burman languages and of Burmese dialects, showing the unifying effect of this approach on structural statement, while Mr. Downer uses both descriptive and reconstructive techniques to elucidate the parallelism of phonological structure, especially with regard to the tonal systems, between Chinese, Tai, and Miao-Yao. In the two following papers, Dr. Pinnow discusses the criteria for establishing the relationships between the various groups within the Austrasiatic language family, including Munda, and M. Haudricourt outlines a scheme of classification for the languages of New Caledonia. Finally, Professor Kähler surveys the present state of knowledge in the Austronesian field in order to suggest priorities for future research.

In so vast and varied a field of research it was not to be expected that there would be uniformity in either theory or practice, but it was already clear by the end of the first session of the study group that, despite the diversity of the specialized interests of its members, there were certain fundamental problems common to us all. One of the questions in the forefront of the minds of most of us, and one central to all our discussions, was that of the kind of comparison that should be made. After the long predominance of the traditional genetic, historically orientated type of comparison, with the immense prestige of nineteenth-century Indo-European scholarship behind it, the time has come when many linguists are ready to consider seriously the possible advantages of typological or ‘synchronic’ comparison, that is to say, of a comparison ‘not of forms, but of formal principles of organization’ (Robins). This does not mean that there is a move to abandon genetic hypotheses but rather that there is recognition that, once it is conceded that the reconstruction of parent languages need not be the only worthwhile objective of linguistic comparison, other approaches may sometimes prove more fruitful. The comparison of geographically contiguous phonological systems has already given rise to the concept of typologically related ‘linguistic areas’ as contrasted with genetically related ‘language families’, and the suggestion was made that cross-classification by both genetic and synchronic criteria might be an advance on other methods (Shorto). That the results obtained
by applying these two sets of criteria to the same material may differ is demonstrated in the contributions to this volume of Mr. Shorto and Mr. Milner. It was suggested by Professor Berg that there was yet a third type of comparison since, in addition to the traditional genetic comparison, which uses our knowledge of linguistic change to reconstruct a prototype, there is the comparison of linguistic forms in order to study the phenomena of linguistic change as such, taking the genetic relationship for granted. This latter kind of comparison, which has been Professor Berg’s own interest, is exemplified to some degree in the papers submitted by Mr. Shorto and Mrs. Jacob.

Closely linked with the question of the kind of comparison to be made is the question of what the linguist is to compare. The theoretical possibilities are manifold. Are we to compare languages or dialects that are more or less contemporaneous, or the same language at different stages of its history, or different languages at different stages of their history? Are we to compare lists of words or morphological forms or syntactic structures? Should we allow our comparisons to encompass the whole of the linguistic material available or should we focus our attention upon a restricted range within carefully defined limits? Are we justified in comparing written with spoken forms, and if so on what grounds do we postulate a pronunciation for written forms, especially in non-contemporary texts? In practice, of course, the choice of material for comparison is usually considerably narrowed by the lack of reliable linguistic data. Nineteenth-century scholars tended to concentrate to the exclusion of all else upon what Professor Berg terms ‘isocytes’, that is to say, upon forms in different languages which can be shown by phonetic laws to be completely comparable. The inadequacy of this approach as the sole basis of historical reconstruction in the South East Asian field has been recognized for some time (the work of Professor Gonda here springs to mind), and members of the group stressed the importance in this context of the special study of ‘irregular’ correspondences, linguistic variants, innovation, and ‘affective’ language (Egerod, Berg, Holmer, Gonda). There appears to be general agreement that future work should include more comparisons of phonological, morphological, and syntactical systems and that for this purpose many further typological studies are required. In the field of historical, or genetic, comparative linguistics we should concentrate first upon the smaller groups of closely related languages (Berg), and aim at the formulation of reconstructed common languages for such groups, using these as the basis of comparison between the larger groups and, ultimately, between language families (Haudricourt). In some areas, and particularly, perhaps, in the Pacific, where historical data is scarce and our present knowledge very uneven, it may be preferable to postpone further work of the traditional, genetic type until further typological comparative studies have been made (Uhlenbeck). For all comparative work, of whatever kind, the prime requirement is access to full synchronic descriptions of the relevant languages, and these, alas, are still lacking in the great majority of cases. It is recognized
that the descriptive accounts upon which comparison must be based will differ, depending upon the aim and approach of the linguist making the description. It should be possible ultimately to effect a connection between purely synchronic descriptions and historico-genetic statements (Simmonds). A linguist already engaged upon genetic comparison in his chosen field is unlikely not to take genetic considerations into account when making a structural description of a language within this field; others not so engaged, but bearing in mind the possibility that their synchronic accounts may one day provide material for the comparatist, might usefully consider presenting possible alternative statements for him to choose from in suitable cases (Egerod).

Discussion of techniques of comparison raised a number of important points. Professor Egerod stressed the necessity of making a distinction between reconstruction in terms of categories and in terms of specific phonetic values. Failure to make such a distinction has sometimes led to misunderstanding, as, for example, of Dempwolff, whose $g'$ is a reconstruction of the former, not the latter, type (Kähler). Mr. Sprigg pointed out that in linguistic comparison 'the allophone is more important than the phoneme', and suggested that if we are to compare lexical items these should be in such a form that the pronunciation in any given context may be inferred. There was clearly wide general interest in the problems arising out of cultural contacts between languages, in the criteria for setting up a hierarchy of 'layers' or 'strata' in a language, and in the methods by which we handle loans in general. Such matters are touched upon in the papers by Mr. Milner, Mr. Robins, Mr. Downer, and in the joint paper by Mr. Honey and Mr. Simmonds. In a detailed exposition of the various strata he discerns in Tai, Professor Egerod proposed that the term 'contact-word' might usefully be employed for words common to two or more languages when it is wished to avoid prejudging the issue of whether they are loans, or of the direction of borrowing, if they are suspected to be loans in one language or the other. Observing that many of the 'contact-words' common to Chinese and Tai were also found in Vietnamese, Mr. Honey urged further investigation along these lines. Suggestions that the study of the linguistic consequences of cultural contacts should not be limited to vocabulary or to relatively remote periods of a language's history came from many quarters. It was pointed out that important comparative work was waiting to be done on such subjects as the influence of English syntax on twentieth-century Thai (Simmonds), the influence of Dutch upon Bahasa Indonesia, the adaptation of Chinese vocabulary and syntax to contemporary Northern Vietnamese (Honey), the interaction of Malacca Malay and Bahasa Indonesia on the one hand and of Bahasa Indonesia and the regional languages of Indonesia on the other (Kähler). The question of the contact-words in Tai, some of which point to a relationship of some sort between Tai and Chinese while others suggest a link with Indonesian, opened the way for the discussion of two other topics of general interest, namely, that of the concept of 'mixed
languages', and that of the troublesome fact that the linguistic alignment suggested by the comparison of isocytes may conflict with that suggested by a comparison of morphological and structural features. As to the first, M. Haudricourt strongly urged the abandonment of the concept of 'mixed languages' on empirical and practical grounds, since adherence to it had in the past led to serious mistakes in classification. The fact remains that terminology of some sort is needed to describe languages which exhibit signs of an 'overwhelming mixture' of influences (Egerod, Robins). On the matter of the possible link between Tai and Indonesian, M. Haudricourt claimed that the Tai languages of China provide useful evidence that may be taken to support this. Such comparisons as that of Thai nam 'water', beside Dioi lam, Tu (Kham) nam, and of Thai nok 'bird', beside Dioi lok, Tu (Kham) mok can, in his opinion, best be explained by the assumption of a disyllabic prototype.

Professor Berg, in expressing doubts as to the fruitfulness of comparing Tai and Indonesian because of the absence of morphology in the former as contrasted with the complex morphology of the latter, introduced the general question of the relative weighting of phonetic as opposed to morphological and syntactical similarities. If, as some speakers suggested (Berg, Holmer), grammatical elements tend to change more slowly than phonetic elements, it might be argued that historical comparative work should concentrate on morphology rather than on phonetic similarities (Berg), and that more attention than hitherto should be given in this context to syntactic comparisons (Holmer, Honey and Simmonds, Robins). There is a serious difficulty here, however, in the selection of valid material for comparison, and much more purely synchronic typological comparison of morphological and syntactic structures is needed before we can feel on firm ground. Mr. Robins has pointed out the danger of using typological similarities to support genetic classification based primarily on the comparison of isocytes before we have any means of judging how far such typological features tend to be characteristic of language families, or 'linguistic areas', and how far they may be found to be characteristic of languages not suspected of being genetically related and remote from each other geographically. Following upon the question of the relative weight to be attached in historical comparison to phonetic as opposed to grammatical similarities is the special problem of those languages for which the two sets of criteria suggest different patterns of relationship. As a case in point, Professor Berg cited the Mon-Khmer languages dealt with in their papers by Mr. Shorto and Mrs. Jacob, which appear to show a certain phonetic similarity to such languages as Thai and Chinese, but morphological similarity to Indonesian. Such cases surely reinforce Mr. Robins' plea for the working out of isoglosses to encompass linguistic typological areas so that we may see to what extent these can be correlated with generally accepted geographic or language family boundaries. A related problem, raised by Professor Uhlenbeck and Mr. Shorto with reference to their own special fields, is that of languages which show marked
similarity in phonetic elements, phonological resources, and grammatical functions but very little agreement in their distribution. Examples of this state of affairs are to be found in the Mon-Khmer material in the papers of Mr. Shorto and Mrs. Jacob; further examples from Palau and Indonesian and from Malagasy, Malay, and Old Javanese were supplied verbally by Professors Kähler and Berg respectively.

Other points dealt with in the course of discussion concerned the need for care in the selection of vocabulary for comparison (Egerod) and for bearing semantic evolution in mind in such selection (Holmer); the problems inherent in attempts to compare first-hand with second-hand material, and in particular with older material assembled by authors with varying degrees of linguistic training and aptitude; and the need for the re-statement or re-publication with suitable notes of some of the older material (Uhlenbeck). Mr. Christie mentioned the importance in an area where virtually all orthographies are borrowed, and where most historical work rests upon assumptions as to the phonetic values to be attached to orthographic forms in earlier times, of the hitherto neglected study of the implications of borrowed orthographies. The absence of more than passing reference to the considerable volume of lexicostatistical work that is currently being undertaken in the area represented a notable gap in our discussion of comparative techniques, and is to be explained by the fact that no one present had first-hand experience of this work. It is hoped that this omission will be made good at the next meeting of the group, to which it is proposed to invite scholars resident outside Europe.

Looking to the future, the group pressed the need for the compilation of good descriptive accounts, supported by texts of languages inadequately dealt with so far, since, without these, solid advances towards a solution of our problems are unlikely to be made. Certain important gaps in our knowledge were specified, but it was recognized that external circumstances beyond the linguist's control rather than any recommended list of priorities are likely to be the decisive factor in determining where fieldwork is carried out in the foreseeable future. It was pointed out by Professor Berg that certain very important fields of linguistic inquiry, including the study of the social processes involved in linguistic change, and the linguistic (as opposed to aesthetic) study of literature and literary genres, are almost entirely unexplored at present, and that there are unique opportunities in our area for repairing such omissions. It is perhaps by way of work of this kind, which rubs shoulders with that of the social scientist and is directed outward upon the modern world rather than back into the past, that we may hope to enlist the interest of the South East Asian and Oceanic peoples themselves. Future hopes also include a growing community of interest with those working in other disciplines whose special studies lie within our area, and whose researches are to varying degrees contingent upon the proper evaluation of linguistic data.
LINGUISTIC COMPARISON

By R. H. Robins

Some sort of comparison of languages follows as an early consequence of any study or even awareness of languages other than one's own; and it may be made on any criteria, though the results will necessarily vary according to the criteria selected, and in scientific linguistics fruitfulness of the criteria of comparison is the principal metacriterion.

An early example of some sort of comparison between languages was their classification into four groups according to the type of word used for God, made by J. J. Scaliger, which gave approximately the groups known later as Latin and Romance, Greek, Germanic, and Slavonic, though further comparisons were not made at the time by Scaliger in any systematic way.¹

A more refined and productive basis of comparison was that of Wilhelm von Humboldt early in the nineteenth century, with the three well-known language types: isolating, agglutinating, and inflecting,² which, despite its naïveté in some respects, is by no means without use, especially if the three types are regarded as forming a triangular figure within which the typological position of a language may be plotted, and as themselves being three cardinal points towards which languages approach with greater or less purity, rather than as actual classes of languages.

This sort of comparison was dismissed by Antoine Meillet as a mere amusette,³ and after Humboldt and until fairly recently attention tended to turn from the purely comparative study of languages to genetically orientated comparative-historical studies, or, in the English tradition, Comparative Philology, which for a number of reasons dominated the field of nineteenth-century linguistics. The established Indo-European linguistic family, and the system of detailed historical relations between the languages thereof, is tribute enough to comparative-historical linguistics, and its methods have been successfully applied to other genetic groupings of languages.

Apart from the great success that IE comparative linguistics has achieved, the usefulness of this type of comparison is obviously not limited to linguistics alone; cultural and political history, and in less known areas and periods the migrations and diffusions of peoples, are all subjects in which the findings of comparative-historical linguistics may and do throw light, just as, conversely, historical data may be brought in to assist the genetic grouping of languages.

The very success of this type of comparative linguistics, developed within IE languages, and most successfully exploited in this field, has led some to question its necessary utility elsewhere, and even its legitimacy as a purely linguistic activity. In particular one must recognize the specially privileged area covered by IE, with copious texts of all kinds of the three classical languages, Greek, Latin, and Sanskrit, and records of many others, together with detailed historical knowledge of the peoples speaking such languages over a long and continuous period, privileges that are not available to the same extent elsewhere and are wholly lacking in some language groups (e.g. Algonkian and Bantu).

Realization of these limitations and the development of general and descriptive linguistics in the present century have led to a return to the type of comparison implicit in early efforts and in the work of Humboldt. The typological comparison of languages may be made on the basis of phonological structuring, grammatical structuring, or the relationship between these two levels.

Typological classifications based on phonology alone are seen in the classification of phonological systems by Trubetzkoy and by Hockett; and the obvious fact that phonetic and phonological features tend to spread over contiguous areas irrespective of genetic relationships has been treated by Jakobson under the title of 'affinités phonologiques'. In this connection one may instance the area of glottalized consonants in the Caucasus uniting IE Armenian with the non-IE languages of that region, and the area of retroflex consonants in India uniting the Sanskritic languages (IE) of the north with the Dravidian languages (non-IE) of the centre and south; and we may note that these areas are the only ones wherein glottalized and retroflex consonants have come to play an extensive and prominent part in the phonological systems of IE languages.

Classifications can be based on grammatical characteristics alone, as for example on the proportions in which endocentric and exocentric constructions and derivational and inflectional word formations are found, or the degree to which word order is syntactically relevant and with what specific syntactic relations specific word orders correlate.

Humboldt's basis of classification is inter-level and rests principally on the different ways in which languages may represent grammatical relations in sentences by the phonological forms of the words therein. Such inter-level comparison may be the most significant, in that it involves a greater range of formal linguistic

1 cf. F. Boas, 'The classification of American-Indian languages,' Language, 5, 1929, pp. 1–7; W. S. Allen, 'Relationship in comparative linguistics,' TPS, 1953, pp. 52–108, though this latter involves an extreme restriction on what is properly to be termed 'linguistic'.


4 C. E. Bazell, 'Syntactic relations and linguistic typology,' CFS, 8, 1949, pp. 5–20; L. Tesnière, Eléments de syntaxe structurale, 1959, pp. 32–3.
structuring. Among modern linguists Sapir took up and developed the Humboldtian classification, but without Humboldt's suggestions of value judgments to be placed on the different types as means of expression or as indications of the mentality of the speakers, and without the later evolutionary accretions of others.\(^1\) Greenberg has suggested ways in which the Humboldtian comparative classification as adapted by Sapir may be further formalized and quantified, by the introduction of ten indices or proportions of formative elements or processes to words and constructions, resulting in five parameters by which, taken together, languages may be placed on the scales of analytic, synthetic (inflectional), and agglutinative.\(^2\) A similarly based classification of languages has been outlined recently by Bazell, with reference to the morphological forms of words and their relative degrees of determinacy of grammatical segmentation and class indication.\(^3\)

Quite obviously both types of linguistic comparison, which we may call genetic and synchronic respectively, are legitimate procedures if conducted with appropriate rigour and with an adequate body of material. What is not perhaps always sufficiently recognized is that inferences from one type of comparison to the other are not necessarily valid. Genetic relationship means, at the minimum, that two or more languages have during past centuries, through a gradual and unbroken tradition,\(^4\) developed from what was once a relatively homogeneous language area; and the principal criterion for such relationship is a number of words or root morphemes, too many to be dismissed as coincidences, of similar or related meanings, whose phonological forms exhibit systematic correspondences (not necessarily actual similarities)\(^5\) of elements and are not explicable as onomatopoeic creations or as loans. This base statement, of course, is nothing new, and requires considerable elaboration; but in a short paper such as this it is perhaps enough. Now it is clear from numerous examples that structural features of languages at all levels, such as are used as the material of synchronic comparison, may change in the course of time while the basic vocabulary continues to reveal indisputable genetic relationship. Within IE, English syntax and Latin syntax are very unlike, and Pedersen notes some agglutinative developments in Armenian reminiscent of Turkish structures\(^6\); and if B. Karlgren is right, the present isolating structure of Chinese is the result of gradual change from an inflectional type of language.\(^7\) Trubetzkoy's six 'structural marks' of IE status are a case in point here\(^8\); he declared his intention of defining IE without positing

\(^6\) *Linguistic science*, p. 100.
\(^7\) 'Le Proto-chinois, langue flexionelle,' *JA*, 15, 1920, pp. 205–32.
\(^8\) 'Gedanken über das Indogermanenproblem,' *Acta Ling.*, 1, 1939, pp. 81–9.
an _Ursprache_ and of relegating the common vocabulary criterion to a very secondary position, but his six ‘marks’ show every sign of having been selected to fit the pre-existing genetically established IE family as closely as possible, and another six equally fundamental features might have split it up and united parts of it with genetically unrelated languages elsewhere in the world. And it has in fact been shown that at least one of Trubetzkoy’s ‘marks’, the absence of an ergative-type construction, does not apply to all the accepted IE family. It would, of course, have been entirely legitimate for Trubetzkoy to try to redefine IE on a non-genetic, synchronic, basis; but if that had been his intention, there would be no need to choose the particular criteria that should keep the genetically established IE family intact, in preference to a possibly more balanced and equally basic set of criteria. The point is that in view of structural evolution, and structural influences produced by prolonged language contacts, it is illicit to exploit the criteria applicable to synchronic comparison to produce or even to buttress historical genetic groupings.

In dealing with the languages of a geographical area (South East Asia in the present context) it must be emphasized first (though it is obvious enough really!) that there is nothing necessarily linguistic in either of the senses discussed above about a group of languages defined by the location of their native speech communities. But _a priori_ pointers to the fruitfulness of both synchronic and genetic comparison of these languages (apart from immediately obvious typological likenesses and phonetic-semantic lexical correspondences among some of them) are: (1) the tendency for phonological and grammatical (structural) features to diffuse over geographically contiguous areas, and (2) the likelihood that some of the population of at least parts of the area have migrated from unitary centres.

It would seem desirable to keep the following points in mind in a genetic (comparative-historical) classification:

1. The degree to which in certain areas and among certain languages the absence of written records of any great time-depth relatively weakens the effectiveness of genetic grouping, though this may be partly supplemented in due course by the findings of glottochronology. If such groupings are to be supplemented by classifications based on other types of criteria the mixed nature of the resulting picture must be recognized. In this connection the procedure of M. Guthrie in his _Classification of the Bantu languages_ is illuminating. He sets up two sets of what he calls principal criteria, the first being a set of synchronic features mainly relating to the prefix concord systems characterizing the grammar of these languages, and the second being a genetic criterion, namely the possession in a part of the vocabulary of the language of a stock of words referable to a set

---

1 ibid., p. 85.
4 ibid., pp. 11–19.
of hypothetical roots or 'starred forms'. He requires both criteria to be satisfied (as well as some subsidiary, less easily applied criteria which may be ignored here) for him to recognize a language as Bantu; but languages may satisfy only one of the principal criteria, and in this case the genetic criterion receives precedence, as languages having the vocabulary relationship but not exhibiting the prefix system are called 'sub-Bantu' and included in his treatment of the Bantu languages, but those having the synchronic grammatical features of the Bantu languages but not a significant body of common Bantu vocabulary are called 'Bantoid' and not further dealt with.

2. The recognition of different strata within one language with different genetic affiliations, e.g. Sanskritic words in a number of the languages concerned. Traditionally, minority elements of this sort are treated as loans, but where they form a considerable part of the general vocabulary of a language and all come from one or two sources, a separate genetic classification of the different layers or strata composing the language, at least for some speakers, may be needed.

3. More generally, the sources and directions of loanword acquisition, apart from those in whole strata of the type just mentioned.

The synchronic classification of the languages of the area would appear more fluid initially, as the choice of types of criteria is wider. The two major metacriteria of the usefulness of features at every level would seem to be:

1. Features that at the level concerned are more rather than less deeply embedded in the language system will make for a more compelling classification (e.g. features characterizing a whole phonological or grammatical system, as against features relating only to a single consonant or a single word class).

2. Features that together exhibit fairly distinct isoglosses around and within the area are likely to be more significant for areal groupings than features with independent and less distinct isoglosses.

The fact, already mentioned, that structural features do spread over contiguous languages, often across genetic boundaries, favours a priori the search for significant structural traits in a geographically defined region.

Types of articulation, syllable structure, vowel contrast systems, systems of consonants, and pitch exploitation in lexically distinctive tones as against grammatically and semantically significant intonation (or, as in several languages, a combination of each) are all possible lines of comparative investigation at the phonetic and phonological levels. Grammatical level and inter-level classification is possible in terms of word structure, types of affixation or other word-form paradigmatic alternations, and basic syntactic structures. And lexical comparison may be made by studying the way the different languages divide up and structure lexically certain universal and culturally important fields, e.g. kinship terms, colours, social status, natural phenomena, and the like.

It must be emphasized that isoglosses of these kinds will encompass linguistic
typological areas, and not as such geographically significant areas, though there is likely to be a correlation between them. It may well be that the linguistic delimitation of 'South East Asia' by major isogloss coincidences and likewise the division of this area into sub-areas by similar means will define areas that include parts of the world that are not usually classed as South East Asia geographically and politically, and exclude some parts that are, as well as cutting across recognized non-linguistic divisions within the area. In this one must be ready to follow the linguistically significant criteria where they lead. Any selection of criteria primarily on the basis of coinciding with boundaries and divisions not themselves linguistically decided is likely to be, for the reasons suggested earlier, less significant even if they provide some feasible isoglosses.

In an area wherein for many years a good deal of work has been carried on by various investigators and by various methods of grouping and classification, sometimes without it being made explicit just what methods are being employed and principles adhered to, it may be of use at the outset of the present symposium to keep very clearly in mind what data and methods of classification we have available, and on exactly what basis we are working in any particular operation.
THE COMPARATIVE METHOD AS APPLIED TO NON-INDO-EUROPEAN LANGUAGES

By Nils M. Holmer

It has been said from time to time that comparative linguistics as a scientific method does not exist outside the Indo-European languages. It is further not seldom announced that comparative study in the field of non-Indo-European languages requires a different method from the one adopted for Indo-European. Both views are equally wrong. It is, however, easy to see how they have originated.

Regardless of the fact that the comparative study of language on a scientific basis began with the Indo-European languages—in connection with the acquaintance with Sanskrit—we find ourselves in a particularly favourable situation when it comes to the historical-comparative study of these languages. For comparative linguistics is basically historical linguistics. Even apart from the great advantages offered by the discovery of Sanskrit as an Indo-European language we should probably have got just as far, sooner or later, by means of the historical study of, for instance, Latin and Greek. Through it we should have been able to trace the evolution of language and observe the regularity with which it takes place. The idea of the sound law—one of the most fundamental principles of comparative linguistics—would no longer have been far-fetched. Incidentally, the existence of a regular correspondence between the consonant sounds in related languages was observed by a Danish scholar quite independently of the results gained by Sanskrit philology and within quite a different field, viz. that of the Germanic languages.

Comparing the evolution of comparative linguistics in, for instance, Semitic and Finno-Ugric, we find that within the former branch linguistics has tended to take the form of philology, owing to the fact that the integrant languages are too closely related and too little altered in the course of time to stimulate the study of a historical evolution (in the same way the closely related Romance languages have proved to be more interesting from the philological than from the linguistic point of view). On the other hand, the languages of the Finno-Ugric stock are a little too distant from one another to serve as a basis for the establishment of a comparative linguistic science in the same sense as for the Indo-European languages, and this becomes still more noticeable as we proceed further east with a view to including Samoyed, Yukaghir, and perhaps some other languages.

Before the rise of comparative Indo-European linguistics, comparative linguistics consisted in the confrontation of usually isolated words, having a similar sound and meaning, in the most different languages, presumably under the impression that all languages went back to one form of human speech, shattered and split up at the time of the great 'confusion of languages' in Mesopotamia. It is quite surprising that still in our days there are linguists who
apply more or less the same kind of method as soon as the study of languages outside the Indo-European domain is concerned.

The mistake referred to by way of introduction lies in the very conception of a comparative linguistic method. It is clear that if such a method is conceived quite mechanically (as involving for instance ‘sound shifts’, ablaut, umlaut, ‘roots’ and ‘root determinatives’, etc.) this method could hardly have an application anywhere else than in dealing with the Indo-European languages. But this precisely is not the comparative method. The comparative method, as I understand it, is fundamentally a historical one: the study of the various linguistic forms as they occur in different periods and areas, independent of which stock or type of language they may belong to, and the tracing of a system in the evolution, whether similar to what we find in the Indo-European languages or not. Comparative linguistics is ‘comparative’ in a secondary sense only. But a serious problem arises when we are concerned with non-Indo-European languages, for here we often find ourselves deprived of historical documents which might throw light on the earlier phases of the language in question. Or else, where such documents are at hand, they are just as likely as not to give evidence of no evolution whatsoever. In cases like these, we are, however, to some extent aided by an analysis of the dialects of such a language—if they exist or are known to us—or even of related languages or dialects provided the affinities are close enough. The underlying principle in this case is to find among these related languages or dialects certain ones which have the signs of being archaizing and which, therefore, may permissibly serve the purpose of supplying an earlier stage in the linguistic evolution.

Too often we have to depend on the latter method—unfortunately, since there are pitfalls in this procedure. Generally—and this I know from what has been done in comparative linguistics in the field of the Amerindian languages—it is assumed that all forms which are supposed to be comparable, owing to their meaning, have to represent the regular evolution of one ‘primitive’ form. Every case in which the related languages or dialects diverge phonetically has to be projected on to the primitive form, even though the number of categories so obtained or examples representing the supposed phonetic evolution be extremely limited. The danger consists in thinking that the forms compared must needs have an identical origin. Even in Indo-European linguistics the same mistake has not seldom been made: I shall only mention that the English word ‘head’ (Anglo-Saxon héafod) has not strictly speaking the same origin as Latin caput, although many scholars have tried to combine them by various manipulations of the sound laws, and in order to compare and identify the Latin ursus or the Sanskrit rkṣas (‘bear’) with the Greek ἄρκτος (along with a very limited number of analogous words) a special sibilant has been reconstructed of which not a trace is directly perceivable in any of the Indo-European languages (similar objections are expressed in G. B. Milner’s ‘Notes on the Comparison of Two Languages’).
For exactly such reasons, I should seriously question the identity of, for instance, Javanese, etc. *wai* and Malay *ayer* 'water'.

How, then, are we to proceed in order to do comparative linguistic research in the circumstances described just now? I think the first condition is perhaps to investigate the languages—preferably such as are geographically or culturally contiguous¹—without any *a priori* notions as to their relationship; the relationship which might exist between them will then probably announce itself in the course of the research. It is perhaps not out of the way to point to the fact that the demonstration of dissimilarities between any two or more languages also falls within comparative linguistics. As a matter of fact, there is some danger in thinking that comparative linguistics is nothing but tracing words in different languages back to a common origin. 'Relationship' is a very vague term, but it certainly consists in something more than having a number of forms in common. The 'genetic' theory of relationship is concerned with one single phase of linguistic evolution: the mechanical one by which existing forms are gradually worn down, eventually to disappear. Were no other factors at work in linguistic evolution, it is imaginable that language as such would soon cease to exist.

An important aspect is relationship of structure. If we say that previous to the rise of comparative Indo-European linguistics comparative linguistics consisted in the study of isolated words and that the comparative Indo-European linguistic science from its very beginnings became concerned with comparative morphology rather than with comparative lexicology, we might perhaps anticipate comparative linguistics as applied in the field of non-Indo-European languages, at least in the initial stages, as a comparative study of structure. Since, however, structure has to do with 'form' and morphology is the study of linguistic forms, 'comparative structure' is in the first place comparative morphology *in a wider sense*.

Usually the analysis of structure is conceived synchronistically. If possible, however, 'comparative structure' ought to be viewed historically as well. For although it is often found that the main features of morphological structure change less in proportion than do the individual elements, yet there are cases in which we are able to perceive an evolution of the structure, whereby a possibility is furnished to sort out primitive elements from more recent ones. Linguistic patterns no doubt change. Of this we are easily convinced through observation of those Indo-European languages whose history is well known. The change of structure in a language has its greatest interest from the point of view of the *innovations* which are created in the process. The rôle played by innovation is an important factor in linguistic evolution and one which, unfortunately, has been rather neglected at all times. It is important not least because it carries with

¹ cf. G. B. Milner's 'Notes on the comparison of two languages', P. J. Honey and E. H. S. Simmonds' paper on Thai and Vietnamese, and H. L. Shorto's paper on Northern Mon-Khmer, with references to W. S. Allen and to Emeneau's definition of 'linguistic area'.
it the necessary assumption that a rather small part of any language represents elements which are really old. The indubitable fact that much in language depends on innovation is, however, not to be referred to structure only: it affects the individual concrete elements as well\(^1\) and to an extent which may be more considerable than is usually suspected. Nothing, probably, comes from nothing, but as soon as a change takes place we have in a sense something new, and since the actual evolution is not always perceived, I think we might accept the independent evolution of language as a reality in the same way as we have accepted its negative counterpart, or the changes due to the course of use.\(^2\) By trying to sort out the more recent elements—to the extent that this is feasible—we should eventually arrive at an earlier form and possibly get an idea of the earlier structure as well.

\(^1\) That is, elements within the range of 'system', to accept the terminology used by H. L. Shorto in his paper on Northern Mon-Khmer.

\(^2\) I feel much tempted to quote here the final remark in H. L. Shorto’s paper on Northern Mon-Khmer—in itself a quotation—‘The important thing about an influence is not where it comes from but what it turns into’.
THE MORPHOLOGICAL STRUCTURE OF THE AUSTROASIATIC LANGUAGES

By Nils M. Holmer

The term 'Austroasiatic' I shall use here to cover the following groups of languages: (1) the so-called Sino-Tibetan languages (Chinese, Tibetan, Burmese, and Thai, along with minor groups of related languages, of which those are especially interesting which have been described as showing tendencies towards polysyllabicity: Kachin, Bodo, Naga, etc., in Assam and northern Burma, and further those which have been described as 'pronominalizing': Kanawai (kanauri), Limbu, etc., in the Himalayas); (2) the languages of the Munda-Santali group in India and those of the Mon-Khmer group in Further India, which—although they have certain concrete points in common—yet are structurally rather heterogeneous 1; and (3) finally the languages of Austronesian stock (Indonesian, Melanesian-Micronesian, and Polynesian). It will probably strike anyone who is familiar with some or all of these groups to find such a language as Chinese, with its monosyllabic and uninflected words, ranged along with such a language as Fijian, which is a polysyllabic inflected language and one showing even a tendency towards incorporation (or 'polysynthesis' according to the Amerindian pattern). It is evident that such a classification is not one made on synchronic grounds. My purpose is to stress the importance of the historical aspect of the problem of linguistic structure.

As will be seen, I use 'morphological' in a wide sense here. We shall look into the structural relationship of these languages from the following different points of view: (1) with regard to the system of sounds; (2) with regard to the form of the word stem; and (3) with regard to inflexion. However, I shall not enter at all upon the syntactical structure, nor, of course, upon anything connected with semantics. All aspects will be dealt with historically as far as it is possible.

Phonology. What do the phonetic systems and their historical evolution have in common in all these languages? Synchronically, they would not appear upon examination to present any typical traits whatsoever. If, however, one should attempt to trace the evolution of the sounds in the Austroasiatic languages, he might perhaps discern a few significant points. First of all, it is important to try to find the type of language within this heterogeneous complex which might be supposed to come closest to the earlier phases of the evolution. It is highly improbable, for instance, that any one of the monosyllabic languages represents an original state of things, since general trends in linguistic evolution usually entail that longer forms are more primitive, so that monosyllabicity undoubtedly

1 cf. Heinz-Jürgen Pinnow's paper in this volume (5.6 (g)).
Nil M. Holmer

represents a final phase in the mechanical evolution. Along with the fact that the northern continental languages have become syllabically reduced goes the creation of a system of tone differentiation.

Among the languages which we may suppose not to have suffered any considerable reduction of syllables, that is to say some of the languages of the southern groups, we find such as possess an evolved system of sounds and a great phonetic symmetry. I refer to the Austronesian languages. Here we find a simple vowel system (with a tendency for certain vowels to be influenced as to quality by neighbouring vowels; for ablaut alternations, cf. below). If we are to accept Dempwolff’s four fundamental vowels (\(a, i, u, o\), with a tendency for \(i\) and \(u\) to alternate with the more open forms \(e\) and \(o\)) as basic vowels for the whole group of languages with which we are concerned, I think no serious difficulties will arise anywhere.

The polysyllabic southern languages have a richer consonant system: an early opposition voiced-voiceless (if we like, from an earlier opposition low tone-high tone) \(^1\) is perceivable in the stops and affricates. Ancient Chinese is usually reconstructed with at least a triple series of stops; this system in any case mirrors the one found in written Tibetan. \(^2\) For Munda-Santali and Austronesian, however, it seems impossible to trace even a triple series (Melanesian ought not to be an exception); we have everywhere a clear twofold basic opposition of voiced-voiceless stops. It is of course possible—and even probable—that the southern languages have lost one or two of the series typical of the northern groups. These latter reveal a regular alternation of the series of stops and affricates, which has become especially evident in the Tibetan verb. Whether anything like this occurs in Austronesian, I am not prepared to say (cf., for instance, Malay \(gali \sim kali\) ‘dig’).

Apart from the distinctions mentioned, the stops (and affricates) are found to occur in several series differentiated by their point of articulation; the Munda-Santali languages are perhaps representative with the following five series: labials, dentals, cerebrals (or retroflex sounds), palatals, and postpalatals (or velars)—incidentally, these classes are in part characteristic of Dravidian as well. If we consider Old Javanese as typical of Austronesian, we shall find an identical system for this latter group also. The cerebral (or retroflex) series is largely absent in the languages of the northern groups (even the reconstruction of Ancient Chinese does not seem to have revealed any such sounds); the Mon-Khmer languages also seem to lack them. On the other hand, they recur (in various arrangements) in some other Indonesian as well as in some Oceanic languages, notably in

---

\(^1\) H. L. Shorto, however, in his paper on Northern Mon-Khmer, treats tone distinction as having replaced a system of voiced-voiceless distinction in these latter languages.

\(^2\) Søren Egered attempts, partly depending on Karlgrén, a reconstruction of a fourfold system which is reminiscent of the one found in Sanskrit; see ‘Studies in Thai Dialectology’, *Acta Orientalia*, 26, pp. 43 ff.
Malagasy and Melanesian. One significant fact is the poverty of sibilants in the Austronesian languages: Dempwolff reconstructs one only (which he denotes by \( t' \)). This does not hold for the northern languages, even though Munda-Santali and Mon-Khmer, as well as perhaps Thai also, seem to have had—originally at least—but one sibilant. Even in this respect we notice a significant Dravidian parallel: in Tamil, as we know, the only existing sibilant is an s-sound with a tendency towards palatalization (does Dempwolff’s symbol indicate an analogous feature?).

One more detail of the Austronesian phonetic system has often struck me as recalling Dravidian; I refer to the consonant sounds which Dempwolff represents \( l, l', y, \) and \( g' \) respectively. To me it is rather evident that all of these are original liquids of various kinds: Dempwolff’s \( l \) would be the Italian \( l \)-sound, his \( l' \) is evidently in most languages an \( r \)-sound, \( y \) I do not hesitate to place among the \( l \)-sounds (as it actually is in some Melanesian languages), whereas the rarer \( g' \) (the medial consonant in Javanese and Malay \( pira \) ‘how many?’) is possibly a palatalized form. Could it not be that these sounds answer phonetically to the Dravidian \( l \) (as in Tamil \( malai \) ‘mountain’), \( r \) (as in Tamil \( oru \) ‘one’), \( l \) (as in the Tamil plural suffix), and \( L \) (as in Tamil \( tamiL \)), respectively? In any case, the variety of sounds appearing chiefly as \( l \) or \( r \) in the different Austronesian languages rather strongly suggests an original system of liquid sounds similar or related to the one found in Dravidian.

**Structure of the word.** It may be taken for granted that the monosyllabic word of the northern Austroasiatic languages has not been preserved until our day in its original form. In Chinese and Tibetan we can observe, even in historical times, a thoroughgoing reduction of the word forms, either by comparison with reconstructed forms of Ancient Chinese or with Tibetan orthographic forms. In certain languages of Assam (to which I have referred above), which otherwise show Tibetan analogies, vowels appear between certain consonants whereby the word becomes disyllabic; it does not seem to be agreed upon whether these vowels were originally there or have a secondary origin (although the former alternative seems more likely to me). If we now consider the written Tibetan forms as more primitive than the spoken forms and further assume that, by analogy with the above languages of Assam, a vowel had once occurred between the first two consonants of such Tibetan words as e.g. \( kluñ \) ‘river’ (or \( kloñ \) ‘wave’), \( gtin \) ‘bottom’, \( ltun \) ‘fall’, etc., forms would arise which would strikingly suggest the Indonesian type of word: cf. Malay \( kalong \) ‘neck chain’ or \( Kêling \) ‘a Tamil or Telugu’, \( getang \) ‘taut’, \( lutung \) or \( lotong \) ‘a kind of black ape’, etc.

If the Indonesian forms appear most apt to serve as a pattern for primitive Austronesian as far as the elementary sounds go, let us see how these are combined. As a general characteristic we find that consonant groups are absent, with the single exception of certain groups with a nasal, which are permissible
in medial position only (they may, however, occur secondarily in initial position—originally in sandhi—thus giving rise to the so-called prenasalized phonemes in, for instance, Melanesian). The word stem is typically disyllabic, but shorter or longer stems occur occasionally. It might be said that there exists a certain tendency towards three-consonant stems, as in Semitic; for the instability (or non-original character) of one of the consonants, I shall only refer to what is said in the sketch (by Ferrand) in the first edition of Les langues du monde (pp. 419–420; originally put forward by R. Brandstetter, Wurzel und Wort in den indonesischen Sprachen, Lucerne, 1910).

An important characteristic of the Austronesian vocalism is its being subject in a high degree to ablaut variation. The alternation of a, i, u, and the pepēt (where it exists) in Indonesian is too well known to require examples. It appears more regularly in Tibetan (in the verb). In Chinese the apophony is also represented and I presume in the other continental languages as well.

The system of inflexion. Ever since the time of Wilhelm von Humboldt the Austroasiatic languages have been considered largely as uninflected. This is to some extent surprising when we take into account the modifications to which the Indonesian and Oceanic words are subject in order to express different functions. The idea of non-inflexion much depends on what is understood by the term ‘inflexion’. In von Humboldt’s day it was customary to make a distinction between ‘agglutination’ and ‘inflexion’ and in conventional classification ‘Malayo-Polynesian’ is still often described as ‘agglutinative’. Furthermore, these languages (especially those of the northern continental groups) have been considered as uninflected mainly owing to facts which are of an orthographic nature. We may think of Tibetan, which has a declension with seven case forms, whose characteristic elements are, however, usually represented as separate word elements, although they are sometimes very clearly of suffix character.

We have every right, I think, to speak of inflexion in a great number of these languages, although some of them have evidently passed into a stage practically devoid of inflexion (this process has been supposed to be taking place in modern English as well). Another thing is that the Austronesian inflexion is in many respects on quite different lines from what we find in the European languages, long held to be the inflecting languages par excellence.

We are able to recognize two types of inflexion in the languages with which we are concerned: (1) one which we might perhaps call ‘derivational’ (or ‘non-declensional-conjugational’) and (2) one which we might term ‘flexional proper’ (or ‘declensional-conjugational’). Let us begin with the former type.

---

1 cf. Hans Kähler in Afrika und Übersee, 39, pp. 142–3, as well as M. E. J. G. Verstraelen in De bijwoordelijke bepalingen van het werkwoord in enkele indonesische talen (Utrecht, 1955), pp. 63 ff.; incidentally, these authors suggest an evolution similar to the one I should like to propose, in which the Old Javanese ‘article’ ng plays a decisive rôle.
of inflexion. It is well known that the Austroasiatic languages make abundant use of prefixes. We see them in written Tibetan and in reconstructions of Ancient Chinese. In the border languages in Assam and Upper Burma, as well as in Munda-Santali and Mon-Khmer, they are clearly distinguished even in the pronunciation and in Austronesian they have evolved into a most complicated system. Although we may sometimes be poorly informed as to the original function of these prefixes, we notice everywhere a remarkable formal agreement between them.

In Austronesian the 'derivational' prefixes are based chiefly on the consonants \( p-, t-, k-, m-, s-, \) and zero (which latter might be identified with one of Dyen's 'laryngeals'; this, of course, is the element underlying the vocalic prefix \( a- \) in Indonesian), all followed by the vowel \( -a- \) or sometimes \( -i- \) and further liable to enlargement by an additional consonant (usually the one which Dempwolff writes \( *\gamma \)). I do not have to quote any examples of these prefixes from Austronesian, as they are all too well known. The interesting thing is that the same type of consonant reappears in written Tibetan, e.g., the \( b-, d-, g- \) of the verbal inflexion, further the initial ' or \( h- \). In Tibetan words beginning with consonant clusters we often notice \( b-, d-, g-, m-, s- \) as initial consonants. The same is the case in reconstructed Ancient Chinese, where the prefixes take the forms \( *p-, *t-, *k-, *m-, *s-, \) and \( *h- \). Many of these initial consonants are recognizable in the other languages also.\(^1\)

Among the 'derivational' formatives we also have to consider the infixes, chiefly with \( -n-, -m-, -l-, \) and \( -r-, \) which play an important part in the Austronesian languages. They are also found in Mon-Khmer \( (-l-, -r-, -n-, \) and \( -m-) \) and in Munda \( (-l-, -r-, -n-) \). In the other languages they are difficult to trace.\(^2\)

The flexional elements proper (or the 'declensional-conjugational' formatives) are twofold and we may hence distinguish between 'case inflexion' (including the use of prepositions or postpositions) and 'personal inflexion' (including possessive inflexion). Both types are extremely varying in the different languages as they appear to-day (and, as a matter of fact, even in the early stages), and I shall merely attempt to trace the probable evolution. To consider the case inflexion first, let me mention that the Tibetan system is hardly represented anywhere else. In Austronesian, the use of prepositions takes the place of declension forms and prepositions are also used in Thai and Mon-Khmer,\(^3\) although they do not seem to correspond formally to those of Austronesian. Conversely, the other languages have no trace of them, whether or not due to the fact that these might have been lost as easily as the prefixes dealt with above as having a 'derivational' function. Munda and Modern Chinese follow the Tibetan pattern of using

\(^1\) cf. Judith M. Jacob's 'Prefixation and Infixation in Old Mon . . .'.

\(^2\) cf. Judith M. Jacob, op. cit.

\(^3\) On the whole, Thai syntactically approaches Austronesian; cf. the paper on Thai and Vietnamese by P. J. Honey and E. H. S. Simmonds.
postpositions (which do not bear any direct formal relation to the Tibetan 'case suffixes').

The question now arises: since in the course of linguistic evolution one type of language may, by a slow process, evolve into another type, ought we to suppose that the postposition languages in the north have stuck to, and further evolved, an original system of suffix declension (as existing in most parts of Asia), or have the same languages lost an original prefix declension (as found in the Australasian islands), due to the influence of the bulk of 'agglutinative' languages of northern Asia? There may be some truth in each of these theories, which do not really exclude one another altogether. It is quite possible that the Austronesian system of case inflexion (by prefixed elements) is not as ancient as it is general. Although a change from the suffix-postposition construction to the prefix-preposition construction is by no means improbable (this is, as a matter of fact, what has happened in the modern European languages), the lack of agreement on this point is indeed remarkable.

What I have said about the system of case inflexion holds in a great measure for the system of personal inflexion also. The conjugation of the verb in the languages of the world has in certain cases and to a large extent arisen from the possessive inflexion of the noun (as we find it in, for instance, Arabic or Indonesian). When dealing with the Austroasiatic languages, in which formal non-distinction of nominal and verbal stems is a salient and ancient feature, we seem especially justified in treating the personal inflexion of the verb in closest connection with the possessive inflexion of the noun. As far as this latter is concerned, it is, however, difficult to find any kind of uniformity in the various languages. We naturally have to start with a state of affairs in which possession is expressed by suffixes (as in Indonesian). This system is further that of Thai, although the possessive formatives may appear to have less of a suffix character, and of Munda-Santali. In Burmese, Chinese, and Tibetan, however, this construction is unknown. These latter languages use, by analogy with so many other continental Asiatic languages, the genitive (or possessive) form of the personal pronoun (Tibetan nants, Chinese wo ti 'of me, my'), which is placed before the noun.

The evolution of the conjugation (where one may be said to exist) is more doubtful and more complicated. It is rather probable that one has to start with the evolution of elements expressing either the subject of an intransitive verb or the direct object of a transitive verb, from an original system of possessive inflexion (as described above), and the process is perhaps indicated, although sparingly, in Javanese, Malay, and Melanesian. (Parallels are clearly traceable in Arabic; cf. the suffix in yadu-hu 'his hand' and daraba-hu 'struck him', etc.). But it is important to note that Austronesian seems to have early abandoned the objective genitive construction as the normal means of forming a basic conjugation, by contrast with so many 'primitive' languages both in Asia and America. I have, although with some hesitation, been trying to spot traces of the ancient
construction (as I would consider it). The placing of the pronominal elements, especially those expressing the subject, in relation to the verbal stem as well as to a preceding sentence connective, constitutes a problem of its own and one, as I believe, of great interest for 'comparative structure'. Unfortunately, space would not permit even to touch upon this problem.¹

In the main, I think, the evolution is quite on a par with what may be observed in other languages where a possessive inflexion is found. But what are we to do with the so-called 'personalizing' languages in the Himalayas, to which I have referred? In Limbu, for instance, the corresponding formatives seem to be primarily prefixes. If these are not to be explained as the result of a secondary evolution (which I frankly believe cannot be the case), we are forced to assume that these languages in this respect do not represent the common Austroasiatic type, but have conserved features of a more ancient linguistic structure (the one found in the Kashmirian Burushaski or the languages of the Caucasus).

By way of conclusion, I should like to take the case of these archaic Sino-Tibetan languages in order to demonstrate that no type is really pure, and that it is impossible to draw any definite limits between any types or to make a classification which does not suffer from overlapping (cf. the end of the first part of G. B. Milner’s ‘Notes on the Comparison of Two Languages’, with a reference to Professor J. R. Firth). And, furthermore, although the structure of a language, in the sense that I have tried to illustrate in this sketch, is in many respects constant, it may nevertheless happen that it changes on essential points even more rapidly than do certain of the very oldest concrete morphological elements.

¹ cf. the construction ka-m ‘not-you’ in Munda (appearing in an example given by Pinnow in his paper in this volume, 5.4), with, for instance, Tongan ka u ‘and I’, an agreement which I find in a way significant.
1. In the present state of linguistics one cannot begin comparative research without a critical examination of the concepts used in the different types of comparative studies which have been developed in the past.

2. In the Indo-European field the aims of linguistic comparison have become more and more modest. Language comparison in this field started in the nineteenth century with the aim of wholesale reconstruction of the so-called parent language. After Schleicher this aim was more or less tacitly abandoned. Most comparativists in this field will now agree that all that can be done is the setting up of series of formulae by which one can account in a simple way for the existing diversity of a certain number of lexical items belonging to various languages and mostly intuitively collected on the basis of formal and semantic similarities. The merits of the historical-comparative method as applied in this particular field seem to be threefold: (1) a relative chronology between a group of languages can be established; (2) a historically justifiable grouping of these languages can be effected; (3) an insight into the working of the process of divergence can be gained.

The comparative method has been developed in a period of the history of linguistics when general linguistics was still in its infancy, that is, when our knowledge of the phenomenon of language in general was still extremely restricted and such concepts as 'relationship', 'historical development', etc., were handled without much critical sense and were often considered as being self-evident. The terminology of the comparative method was of a crude, heavily metaphorical type, which gave rise to many misunderstandings and unwarranted simplifications. In order to be able to evaluate this method and to determine its usefulness for other language areas, it is necessary to keep in mind the following facts:

(i) The classical comparative method does not deal with languages taken as a whole but with a rather small number of lexical elements.

(ii) The whole set of reconstructed forms bears no resemblance to a real language. There exists an essential difference between the Proto-Indo-European language as it must have been spoken in a distant past and the reconstructed Proto-Indo-European of the comparativists.

(iii) The process of language-divergence is a process which may take place under certain conditions, but which may be absent in other circumstances. This process is only one amongst other possible historical linguistic processes.

(iv) The results of the classical historical-comparative method depend on the exceptionally rich amount of extra-linguistic, historical information available.

3. Beside the genetic type of linguistic comparison there has existed since the early nineteenth century another way of comparing languages. The various
typological classifications proposed from von Schlegel onwards, however, suffered from a lack of method and were vitiated by the very superficial knowledge of languages outside the Indo-European area. Moreover, in setting up classifications one used to pay attention only to certain formal aspects of the morphology of the languages involved.

As Jakobson rightly observed in his report at the Oslo congress, typological studies imply the descriptive technique. In view of the progress descriptive and general linguistics have recently made, linguistic science is now in a much better position to deal with the problems raised by typology. Typological studies seem to be a natural enlargement and expansion of modern monolingual research and one may expect that from such studies insight into historical development may be gained too. As language typology can only be carried out satisfactorily if there is similarity in descriptive techniques, it will be necessary to reach a certain minimum of agreement on, or at least a mutual understanding of, the techniques used. Perhaps it is possible to establish a kind of ‘translatability’ between the various descriptive approaches. It goes without saying that language typology has to be freed from its exclusive preoccupation with morphology in its formal aspect. What is needed is a consistent comparison of linguistic sub-systems (sound system, morphonology, morphology, syntax, intonational system).

4. A survey of the different types of language comparison available at this moment ought to make mention of the glottochronological or lexico-statistical method introduced into linguistics by Swadesh about ten years ago and already applied to Austronesian language data by Elbert and Dyen. It seems difficult at this stage to pass a definitive verdict on the value of this method. Nevertheless it seems certain that if it has a value, this value will be extremely limited. Being by its very nature of a low degree of accuracy, lexico-statistics will only be able to furnish us with an indication of the direction in which it seems profitable to undertake structural comparative research. The danger of glottochronology and lexico-statistics is that it is a method easy to apply. One does not need to become familiar with a great number of intricate linguistic phenomena and in a very short time one may become a specialist in comparing word-lists of practically unknown languages. Therefore it is necessary to stress the fact that lexico-statistical conclusions have always to be followed up by precise comparative research based on a sound knowledge of the structure of the languages involved.

5. In order to be able to decide how comparative studies of the Oceanic languages may best be furthered, it is necessary to survey the Austronesian linguistic scene. Characteristic features seem to be (1) that the number of languages spoken in this area is high, (2) that the size of the speech communities is on the whole small, some communities in Java (Javanese ± 40 million, Sundanese ± 15 million speakers) excepted, (3) that historical data are scarce (notable exceptions being Javanese and Malay), (4) that in many areas there must have been and probably still is a lot of migration, (5) that there is in many areas a considerable
amount of bilingualism, (6) that owing to the development of national languages the language situation is changing rapidly.

As to our knowledge of the Oceanic languages, this is quite uneven. On some languages, as for instance Bare'e, Mori, Roti, Toba Batak, Tagalog, we possess a considerable amount of data, of others we know hardly anything. The situation is certainly not improved by the fact that widely divergent methods are used. Compare, for instance, Bloomfield's description of Tagalog phonology and morphology with Jonker's awkward attempts to furnish us with a picture of the languages of Roti and Bima or with the detailed and much clearer descriptions of Bare'e and Mori by Adriani and Esser. Most descriptions are antiquated, modern descriptive techniques being but rarely applied (a notable exception being Robins' studies of Sundanese). On the whole our descriptive knowledge of Indonesian languages is much more extensive than that of the three other areas (Melanesia, Micronesia, Polynesia).

In addition some modest beginnings have been made with applying the historical-comparative method to Austronesian data (van der Tuuk, Brandes, Kern, and Codrington in the nineteenth century, followed by Schmidt, Thalbitzer, Ray, Brandstetter, Dempwolff, and Dyen), without much insight, however, into the limitations of this method or the fundamental differences between the language situation in this area and the unique Indo-European language conditions. Moreover, these comparative studies were characterized by a certain bias towards Indonesian languages, while sometimes too much value was attached to the well-known geographical distinction into four areas mentioned above. This distinction may be said to have had a retarding influence on the development of a comparative study based on purely linguistic criteria.

6. From existing methods of comparative study on the one hand and the language situation in the Pacific area on the other, one may conclude that—supposing that co-ordination of our linguistic effort is feasible—the best policy seems to be (1) to intensify the descriptive effort on those languages about which information is particularly needed; this can only be done satisfactorily in the field, by well-trained linguists; (2) to try to find out to what extent it is possible to cull from the more or less antiquated but extensive and detailed grammars and dictionaries we possess, descriptions which satisfy modern requirements; this does not need to be done on the spot, but may be undertaken in those places where there are enough library facilities; (3) to reach a certain amount of agreement on the descriptive techniques to be used in future descriptions and in restatements of older descriptions; (4) to start typological comparative studies for those languages or, better, for those sub-systems of which sufficient knowledge is deemed available. From the point of view of linguistics it is in my opinion preferable to postpone further comparative work of the traditional genetic type, in favour of typological comparative studies combined with intensification and further refinement of our descriptive effort.
REFERENCES

ELBERT, S. ‘Internal relationships of Polynesian languages and dialects,’ SWJA, 9, 1953, pp. 147–73.
NOTES ON THE COMPARISON OF TWO LANGUAGES
(WITH AND WITHOUT A GENETIC HYPOTHESIS)

By G. B. Milner

In a suggestive article which appeared some years ago,¹ W. S. Allen, after reviewing the theoretical basis of comparative linguistics, analysing certain of its results in the Indo-European field, and stating the reasons which have led him to question the usefulness and the validity of assuming a genetic relationship between languages, sets out in the form of postulates the principles in accordance with which comparative studies should in future be conducted. Relationship, he claims, is not of languages but of systems. Relationship, moreover, is not a question of either-or but of more-or-less.²

There follows, in the remainder of the article, a number of suggestions proposing other lines of investigation for a new type of comparative study, freed from the limitations of traditional comparativism while remaining systemic and open to examination within the safeguard of rigorous criteria.

A few years earlier Professor Holmer had made the following statement in an article dealing with relationship in the Amerindian field: ‘It is essential to consider that any two languages commonly regarded as related are not comparable to offshoots of a main plant stem. Language is a complex system of far less structural unity than either a plant or an animal, a fact which was entirely disregarded when the idea of the genetic relationship of languages was first propagated.’³

Similar warnings and reservations on the ground that the methods of traditional comparativism might prove to be less rewarding when they were applied to linguistic fields other than the Indo-European had been made at an earlier date. Thus Sweet had said even before the first World War: ‘In applying the results of comparative Aryan grammar to other families of languages, it is evident . . . that we must . . . be ready to widen or modify our methods with the scope of their application.’⁴ Professor Gonda, one of our collaborators, has given the reasons which have led him to the conclusion that the traditional methods of comparative philology are not readily applicable to, or inherently suitable for, the Malayo-Polynesian field, in two articles.⁵

Other authorities, however, in the past and down to the present day, have

¹ 'Relationship in comparative linguistics,' TPS, 1953, pp. 52–108.
² ibid., p. 92.
³ Lunds Universitets Årsskrift N.F. Avd. 1, Bd. 45, 4, 1949, p. 10.
⁴ Collected papers of Henry Sweet, arranged by H. C. Wyld, Oxford, 1913, p. 64.
held up the methods of Indo-European philology as a model for other fields. ¹ Dempwolff recognized that Austronesian linguistics must also develop its own specially adapted techniques if the best use was to be made of the kind of material and of the particular conditions prevailing in that field. ² Within the last ten years several articles and monographs have appeared, including several by Professor Kähler, which are in the tradition of Dempwolff and may be said to continue his studies. ³ They all appear to accept implicitly the validity of the genetic hypothesis of the neo-grammarians.

The purpose of the present paper is not primarily to take up arms and do battle as a protagonist of this or that school of thought. Rather is it inspired by the wish to illustrate the wide discrepancy between the results likely to be obtained by setting out from different premises. For this purpose two languages with which the writer is acquainted at first hand have been selected: Samoan, a representative of the Polynesian group, and Fijian, the exact status of which has long been queried, but which is generally regarded as Melanesian with a strong admixture of Polynesian elements. Both languages have interested comparative philologists since the end of last century. ⁴

PART I: PHONOLOGY

It will be recalled that Dempwolff, in his *Vergleichende Lautlehre*, set up an inductive reconstruction of Proto-Indonesian (*Urindonesisch*) which was based on the results of a detailed examination of Javanese, Toba-Batak, and Tagalog. After testing the validity of his reconstruction by examining three other Indonesian languages (Hova, Ngadju-Dayak, and Malay) he was satisfied that his reconstruction was sound and his original hypothesis justified. He next turned his attention to two Melanesian languages (Fijian and Sa’a) and to three Polynesian dialects (as he called them), Tongan, Samoan, and Futuna. He found both in Melanesian

---

² op. cit., p. 23.
and Polynesian certain regular correspondences and he considered them to be reflexes of the Proto-Indonesian sound system which he had reconstructed. He consequently posited a Proto-Melanesian and a Proto-Polynesian sound system, both of which he regarded as being directly descended from Proto-Indonesian. He thus felt himself entitled to call the latter Proto-Austronesian.  

It is instructive to examine in detail the premises upon which Dempwolff's work is founded as well as the conclusions drawn from its results. Austronesian (or Malayo-Polynesian) languages, he states, do not have a homogeneous grammatical structure like Semitic or Bantu languages, but they do have a common vocabulary (his italics) comprising several hundreds of words. Therefore in his monograph he is only incidentally concerned with grammatical correspondences and its scope is restricted to the comparative analysis of words, and primarily to their sound correspondences.

Having established and confirmed his Proto-Indonesian system, he then investigates approximately fifty other Indonesian languages by the same method, and he is able to establish in each individual language a few hundred words which regularly reproduce equivalents of Proto-Indonesian sounds.  

Also by the same method, having identified the Fijian and Sa’a reflexes of Proto-Indonesian sounds, he can establish between 300 and 400 words in both languages in regular correspondence, out of a total of approximately 1,900 words claimed as being of Proto-Indonesian origin. Since, he claims, the proportion of common Proto-Indonesian words in modern Indonesian languages is often no higher than in Fijian and Sa’a, he is satisfied that it is enough to establish a connection in time and space between Fijian, Sa’a, and Proto-Indonesian. When other Melanesian languages are examined, however, he finds that the number of common words may fall to between 100 and 200 words. There is also the consideration that certain features occur in Melanesian and not elsewhere, notably initial nasal clusters which correspond to medial clusters in Indonesian, and prenasalization as a vestigial element instead of as an active morphological process.  

The low proportion of words ascribable to PI starred forms which he had found in many Melanesian languages, together with their irregular treatment of nasal clusters (irregular, that is, from an Indonesian point of view), would, one might have expected, have made it necessary to re-examine the assumption that present-day Melanesian languages are directly descended from PI and that the latter can be equated with Proto-Austronesian. Dempwolff, however, draws the conclusion that Melanesian is 'aberrant' from the PI norm. That is to say, it

2 IA, p. 13.
3 DA, p. 123.
4 ibid., p. 163.
5 ibid., p. 165.
NOTES ON THE COMPARISON OF TWO LANGUAGES

became separated from Proto-Indonesian before the latter split up into its various modern representatives.¹

The same process of reasoning is applied to Polynesian; i.e., Polynesian 'dialects' may be regarded as a particular branch of Austronesian which enables one to reconstruct a Proto-Polynesian (PPn) sound system.² Moreover, certain similarities in the PPn and Proto-Melanesian (PMn) sound systems make it likely that at one time both groups constituted part of the linguistic stock of a single people.³

Since Dempwolff does not ignore the important divergences between Melanesian and Indonesian, as evidenced by the treatment and distribution of nasal clusters and the relatively low proportion of Indonesian starred forms in many Melanesian languages, it is somewhat strange that he should not have attached more importance to the cases where, as he noted himself, certain Fijian and Samoan sounds cannot be reconciled with his PI hypothesis. It is possible that his purpose might have been better served if, instead of referring each modern Melanesian and Polynesian language directly to PI, he had attempted to reconstruct internally congruent starred systems for both Melanesian and Polynesian. Had he done so, however, it might ultimately have proved very difficult to reconcile PI with a full-scale reconstruction of PMn and PPn. Having set up his PI system, Dempwolff seems at times to be the prisoner of his own creation, particularly when any data, especially in Melanesian, obstinately refuse to conform with the original hypothesis.⁴ Perhaps the best illustration of the results of that tendency is found in the treatment of certain consonants in Fijian and Samoan quā reflexes of PI consonants reconstructed by Dempwolff.

It will be recalled that among the consonants of PI he posited four palatal stops which he writes k' g' t' d'.⁵ Later he establishes the occurrence, under certain conditions, of an optional (homorganic) nasal accretion (fakultativer nasaler Zuwachs) for each of those consonants (written ūjk', ūg', ūt', ūd'), adding a suggestion that the option of a preceding homorganic nasal before medial consonants might originally have been a morphological process (as of course it still is in the case of initial nasal clusters in modern Indonesian),⁶ but no longer functioning as such in modern Melanesian.⁷

In so far as Fijian and Samoan are concerned, the history of those two groups

² DA, p. 191.
³ ibid., p. 193.
⁴ ibid., p. 124, § 119 a, 5 ; p. 125, § 121 a ; pp. 135–6, § 128 b and c. See also Allen, op. cit., p. 53.
⁵ IA, p. 64.
⁷ IA, p. 109.
of consonants is assumed to have been as follows (*Deduktive Anwendung*, pp. 128-41, 166, 170-90, 192):

<table>
<thead>
<tr>
<th></th>
<th>Proto-Melanesian</th>
<th>Proto-Indonesian</th>
<th>Proto-Polynesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fijian :</td>
<td>δ</td>
<td>k' g' t' d'</td>
<td>s</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Proto-Indonesian :</td>
<td>d'</td>
<td>njk' ng' nt' nd'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Proto-Polynesian :</td>
<td>s</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Samoan :</td>
<td>s</td>
<td></td>
<td>φ</td>
</tr>
</tbody>
</table>

Dempwolff regards Fijian δ and Samoan s as the reflexes of the non-nasal set of consonants on the grounds that they both occur in the final syllables (e.g. in the third consonant position in words of type CV.CV, corresponding to the final consonant in Indonesian words of type CV.CV). Fijian δ and Samoan s can therefore hardly be regarded as the reflexes of nasal clusters, which do not occur in final position in Indonesian (cf. *tanit* with *tanida* in Fijian and *tanisra* in Samoan; *Deduktive Anwendung*, pp. 137-8 and 181-2).

Unfortunately (and Dempwolff was well aware of it), this arrangement presents almost as many problems as it solves:

As regards Fijian s (*Deduktive Anwendung*, pp. 137-9):

(a) (i) instead of δ (e.g. *savui*), and (ii) in addition to δ (e.g. *vusa-vuda*), it may be a reflex of *t’* (§ 128 g, 2);
(b) (i) instead of δ (e.g. *sau*, § 128 g, 3), and (ii) in addition to δ (e.g. *musa-uđa*, § 128 g, 2), it may be a reflex of *d’*;
(c) in addition to δ (e.g. *vosota-voda*, § 128 g, 2), it may be a reflex of *g’*;
(d) instead of δ (e.g. *vosa*, § 128 g, 2), it may be a reflex of *k’*.

As regards Samoan o (ibid., pp. 181-2):

(a) instead of s (e.g. *tai*, § 153 d, 2), it may be a reflex of *t’*;
(b) instead of s (e.g. *moe*, ibid.), it may be a reflex of *d’*;
(c) instead of s (e.g. *fia*, ibid.), it may be a reflex of *g’*.

Conversely, Fijian δ may be a reflex of an optional nasal cluster (e.g. *tađi*, § 153 d, 1, and *Austronesisches Wörterverzeichnis*, p. 12), and Samoan s may be a reflex of a nasal cluster (e.g. *tusi*, § 152 a, 6, and § 153 d, 3).

The diagram given above must therefore be considered as being subject to considerable modifications and might be given more realistically, if less conveniently, in the following form:
Nevertheless Dempwolff specifically and more than once rejects the desirability of making a revision of his original hypothesis. If a nasal cluster is attested from Fijian alone, that is not considered a sufficient ground for modifying a non-nasalized form in PI, since Fijian clusters are attested initially as well as medially. Conversely, the occurrence of a non-nasalized form in Fijian as the reflex of a nasalized form in PI is insufficient for a revision of the hypothesis.\(^1\) Likewise in Samoan the words which do not conform do not constitute a sufficiently good reason for revising the starred forms of PI.\(^2\)

We are indebted to an American scholar, Dyen, for a possible solution of the problem which baffled Dempwolff.\(^3\) Dyen proposes that in addition to the series \(k' g' t' d'\) and \(ŋk' ŋg' ŋt' ŋd'\), which he prefers to write \(c j s z\) and \(Nc Nj ñs ñz\),\(^4\) an additional starred phoneme or proto-phoneme \(Z\) should be reconstructed with an optional prenasalized counterpart \(nZ\). The reflexes of these starred forms would then be:

\[
\begin{align*}
\text{for Fijian:} & \quad *Z & \quad *nZ \\
\delta & \quad s \\
\text{for Samoan:} & \quad s & \quad Ø
\end{align*}
\]

This enables him to reconstruct \(*peZem\) leading to Fijian \(moðe\) ‘sleep’ on the one hand and on the other hand \(*penZem\) giving Samoan \(moe\) ‘sleep’. The occurrence of Samoan \(ala\) ‘path, road’ instead of an expected \(*sala\) from a starred form \(*Zalan\), however, immediately raises difficulties, as does the Fijian counterpart \(sala\) ‘path, road’ instead of an expected \(*ðala\). To resolve the contradiction Dyen suggests in a footnote that in that case Fijian and Samoan must reflect a nasal cluster. For an explanation we are referred to another article, where a somewhat abstruse account may be found of the circumstances in which this apparent anomaly has developed.\(^5\) The same suggestion also applies to the fact that contrary to Dyen’s expectation Fijian \(s\) may also be the reflex of \(*Z\).

\(^1\) DA, pp. 135–7, § 128 b, c, and d; p. 182, § 153 d, 3.
\(^2\) ibid., § 153 d, 2 and 3.
\(^4\) See The Proto-Malayo-Polynesian laryngeals, p. 50.
Dempwolff had explained these inconsistencies by assuming that the different reflexes, sometimes found within the same language, sometimes found by comparing one language with another, were evidence of the former existence of proto-phoneme (Nebenformen).\(^1\) Dyen, however, rejects that assumption on theoretical grounds.\(^2\) One may question the usefulness of reconstructing yet another starred phoneme, bearing in mind the already rather unwieldy structure posited for PI by Dempwolff\(^3\) and even allowing for the fact that, as Allen has pointed out, one is dealing with systems set up by linguists for the description of language structures, and not anything inherent in the language as such.\(^4\) Moreover, to posit two starred forms with the same proto-phoneme, one with a nasal cluster and one without, does not tax the imagination unduly in the case of words which might originally have denoted ‘road’ or ‘sleep’ or ‘rain’. It is perhaps less easy to accept the possibility that there were once two forms of the word for ‘nine’ (Fijian ñiwa, Samoan iva) or of the word for ‘who?’ (Fijian ñei, Samoan ai) and that but for the optional nasal element they were identical.\(^5\) Similar difficulties arise in the case of words for plants, which tend to be grammatically even less flexible.\(^6\)

Reading over the materials used by Dempwolff and Dyen, one is struck, first, by the fact that the languages on the evidence of which elaborate reconstructions are made were originally chosen for comparative purposes almost solely as a result of the ‘accident’ that they were better attested than their neighbours, i.e. that more comprehensive dictionaries and grammars were in existence for certain languages than for others which might have served as a basis for the comparison, and on the basis of which different results might have been arrived at. Secondly, it is mainly by the process of culling isolated words from dictionaries and by comparing word against word that results have been achieved. Modern phonological studies, such as that published for Fijian by Scott,\(^7\) have been neglected where they existed, and the comparison of phonological systems as opposed to that of single items has hardly begun. Thirdly, reconstructions often seem to have been made without taking into account other ‘accidents’, such as culture contacts or ‘acculturation’, and other considerations of a synchronic character, the importance of which increases in inverse ratio to the proportion of the vocabulary which in each language may be said to be common Austronesian, that is, relevant to the purpose of the comparative inquiry.

---

\(^3\) \textit{DA}, p. 7.
\(^4\) \textit{TPS}, 1953, p. 90.
\(^6\) cf. \textit{AW}, p. 45, *d’ayanav* ‘turmeric’ (Fijian ñago, Samoan ago); p. 46, *d’alatan* ‘nettle’ (Fijian salato, Samoan salato ‘tree-nettle, \textit{Laportea} sp.’).
Thus, the peoples who inhabit the geographical area of Fiji, Tonga, and Samoa have for many centuries had external contacts (both of a peaceful and warlike character) which have had a profound effect on the mythology, material culture, and social organization of all three island groups. It would be anomalous if those cross-currents had had no effect on the languages and dialects of the area. For instance, it is in a sense 'accidental' that Fijian should have a reflex s for Dempwolff’s *d’ in *d’alan ‘path, road’ instead of an expected ᵁ in certain words such as sala ‘path, road’. To discover that, however, one may need to look at each language from a synchronic point of view.

For what is known to comparative philologists as Fijian is in fact a lingua franca which has developed from a single dialect (Bauan) and which has still not entirely displaced other dialects. It can therefore give only a partial picture of the whole Fijian complex. In some of the dialects of Western Viti Levu, for instance, some of the words which unexpectedly reflect Dempwolff’s non-nasal set with an s are found with a ‘regular’ ᵁ (as Kern predicted might be the case). Thus Western ᵁ lēvũ ‘path, road’, ᵁdũ ‘milk’, and ᵁdǐ ‘snail’ instead of Bauan sala, suõu, and sídi respectively.¹ That is to say, it may well be Bauan which is ‘aberrant’ for purely local or accidental reasons.

Let us suppose, on the other hand, that we abandon the genetic hypothesis, and instead of trying to reconcile the distribution of spirants in Fijian and Samoan with the reconstructed forms of PI, look at those two languages, with the addition of Tongan, as languages which have, at some stage in their history, been in contact. Given that assumption, we can endeavour to see if, to use Holmer’s expression, ‘the conglomeration of phonetical, morphological and syntactic elements constituting a language should perhaps with greater advantage be classed with regard to relative age than to provenance.’²

The following correspondences can be established. In each case the Samoan member is at the apex of the triangle, the Fijian at the left, and the Tongan at the right:

1. Fijian sele ‘knife’, Samoan sele, Tongan hele; Fijian suasua ‘wet’, Samoan susũ, Tongan hũhũ; Fijian salato ‘tree-nettle (Laportea sp.)’, Samoan salato, not recorded from Tongan.
2. Fijian bala ‘be wrong’, Samoan sala, Tongan hala; Fijian ma đa ‘be dry, empty’, Samoan masa, Tongan maha.

² Lunds Universitets Årsskrift N.F. Avd. 1, Bd. 45, 4, p. 11.
(3) Fijian *sala* 'path, road', Samoan *ala*, Tongan *hala*; Fijian *savu* 'waterfall', Samoan *āfu*, Tongan *hafu*.

(4) Fijian *diwa* 'nine', Samoan *iva*, Tongan *hiva*; Fijian *mode* 'sleep', Samoan *moe*, Tongan *mohe*.

The existence of the four tripartite correspondences set out above need perhaps not present any insuperable difficulties and can be accounted for relatively simply, if instead of starting from a genetic hypothesis we regard all three languages as having been in contact and if we concern ourselves only with certain circumscribed features or sub-systems, abstracted from the total mass of each language, and of different age or origin or provenance.

Granted that we accept Fijian as Melanesian, and the other two languages as Polynesian (on any one of a number of possible linguistic grounds such as the fact that Fijian possesses voiced stops whereas Samoan and Tongan do not), the correspondences between Samoan and Tongan can be considered first. After looking at a number of Polynesian consonant systems (on the basis of, admittedly, incomplete and, sometimes, obsolete information) one is inclined to agree (subject to the possible appearance of fresh information) with the view taken in a fairly recent analysis of the problem that 'the most complex and archaic language phonemically is Tongan'¹ and that 'the order of increasing sparseness of forms is then Tongan, Samoan and the East'.²

Let us say, then, that the present-day languages spoken in Samoa and Tonga were at one time in contact (without considering exactly where or when) and that of the two it is Tongan which exhibits the more archaic features. One could then point to the absence of those features in Samoan, together with the close correspondences in the remainder of the two consonant systems, as evidence of that contact between the two languages.

If, however, we put the consonant system of Samoan side by side with that of Fijian, we then no longer find close correspondence, but over a limited sector one can see a correlation between the two which is not shared with Tongan, notably Fijian *s*-Samoan *s*, Fijian *δ*-Samoan *s*. This would seem to constitute evidence of contact between Fijian and Samoan. The question whether Tongan participated in that contact also, although no longer appearing to do so because of its own subsequent history, will be considered presently.

Thirdly, certain correlations can be established for Fijian and Tongan only, namely Fijian *s*-Tongan *h*, and Fijian *δ*-Tongan *h*, giving evidence of contact between Fijian and Tongan.

One can, on the basis of that knowledge, posit contact between Fijian and Samoan at a different level from that already established, i.e. through the 'medium' of Tongan, namely Fijian *δ*-Samoan *ə* and Fijian *s*-Samoan *ə*.

² ibid., p. 164.
The opposite assumption, namely that Tongan was in contact with Fijian through the medium of Samoan, could be argued on the strength of Fijian $s$—Tongan $h$ (via Samoan $h$; e.g. *sele—hele* ‘knife’), but contradicted by the necessity of assuming also Fijian $s$—Tongan $h$ (via Samoan $\emptyset$; e.g. *sala—hala* ‘path, road’) and Fijian $\delta$—Tongan $h$ (via Samoan $\emptyset$; e.g. *mo$\delta$e—mohe* ‘sleep’).

One may therefore posit the following correspondences as being due to ‘immediate contact’:

![Diagram](image)

Consequently one can contrast Fijian $\delta$—Samoan $s$, and Fijian $s$—Samoan $s$, which represent ‘immediate’ or at any rate a more immediate contact, with Fijian $s$—Samoan $\emptyset$, and Fijian $\delta$—Samoan $\emptyset$, which represent ‘mediate’, that is to say a less direct, contact.

The evidence which has been made available so far for Rotuman \(^1\) does not seem to contradict and appears to support the present hypothesis. The Rotuman equivalent of both Fijian $\delta$ and $s$ is in all cases $s$. Thus:

(i) Fijian *su$\delta$u* ‘breast’, Rotuman *susu*, Tongan *hu$\delta$u*, Samoan *susu*;
(ii) Fijian *$\delta$ala* ‘miss, wrong’, Rotuman *sara*, Tongan *hala*, Samoan *sala*;
(iii) Fijian *sala* ‘path, road’, Rotuman *sala*, Tongan *hala*, Samoan *ala*;
(iv) Fijian *mo$\delta$e* ‘sleep’, Rotuman *mose*, Tongan *mohe*, Samoan *moe*.

Certain deductions can therefore be made:

On the basis of (i) and (ii) above one may assume immediate contact between Rotuman and Fijian on the one hand and Rotuman and Samoan on the other.

On the basis of (iii) and (iv) above one may assume immediate contact between Rotuman and Fijian on the one hand and Rotuman and Tongan on the other, Rotuman representing an intermediate stage (through which Tongan may have passed) at which $\delta$ and $s$ had coalesced into $s$ but antedating the shift to $h$ which took place in Tongan.

It is perhaps worth noting that those correspondences can also be illustrated from words which are attested for Fijian and Rotuman only (e.g.

\(^1\) C. M. Churchward, *Rotuman grammar and dictionary*, Sydney, 1940, pp. 159–66 *et passim.*
Fijian yaða ‘name’, Rotuman asa) or for Rotuman and Tongan (with Samoan) only (e.g. Rotuman siŋoa ‘namesake’, Tongan hiŋoa ‘name’, Samoan īŋoa id.; Rotuman soŋe ‘famine’, Tongan hoŋe, Samoan one).

The following diagrams retain the essential features of the tripartite series of immediate and mediate correspondences already set out, while also taking into account the evidence of Rotuman:

<table>
<thead>
<tr>
<th>Fijian</th>
<th>Rotuman</th>
<th>Samoan</th>
<th>Fijian</th>
<th>Rotuman</th>
<th>Samoan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) s</td>
<td>s</td>
<td>s</td>
<td>(iii) s</td>
<td>s</td>
<td>ø</td>
</tr>
<tr>
<td>(ii) δ</td>
<td>h</td>
<td></td>
<td>(iv) δ</td>
<td>h</td>
<td></td>
</tr>
</tbody>
</table>

Such an analysis would enable one to account for the existence of doublets in Samoan without positing proto-doublets as Dempwolff did or additional proto-phonemes as proposed by Dyen. Thus:

- **Fijian**
  - saku ‘sword-fish (of various kinds)’
  - oola ‘live’ (Western dialects)
  - ñuru ‘enter, go through’

- **Samoan**
  - a’u ‘fish (Strongylura sp.) with long snout’
  - uli ‘small corm of a taro sucker’
  - o’ala ‘husk a coconut’
  - ola ‘live’
  -ulu ‘go in, enter’
  - sulu ‘put in, insert’

A number of other doublets exist in Samoan for which no Fijian equivalent can be found. The method of approach suggested above does not, of course, rule out the possibility of mediate and immediate contact between Samoan and other languages. The case of oola, found in a number of Western dialects of Fiji but not in standard Fijian (Bauan), and corresponding to Samoan ola and sola, is a possible indication that both Tongan and Samoan may have in their total vocabulary other evidence of having been in contact with languages regarded as Melanesian. Among other Samoan synonyms and part-homonyms are the following: aga and saga ‘face’, ao and sao ‘collect, assemble’ (cf. ao ‘daylight’ and aso ‘day’, mentioned by both Dempwolff\(^1\) and Dyen\(^2\)), i’u ‘end’; tail of

---

\(^1\) *AW*, p. 11, *a(n)dav and p. 60, *ha(h)g’av.

NOTES ON THE COMPARISON OF TWO LANGUAGES

a fish' and si'u 'tip'; tail of a land animal', ati and sati 'attacked by insects', ui and usi 'dark' and 'greenish blue' respectively, said of plants and animals.

As our synchronic knowledge of each language increases, both qualitatively and quantitatively, it will be possible to pursue other lines of investigation, notably to establish whether there is evidence that the same sound shifts have occurred independently in a number of different areas. In the article referred to earlier on, Holmer takes the almost paradoxical but tenable view that in a sense all languages are mixed languages and all words loanwords. One of the most strongly held beliefs of the late J. R. Firth was that languages were polysystemic and not monosystemic. It cannot be doubted that there is a need for a wider application and a special adaptation to individual problems of new techniques, described some years ago by Miss Henderson, for the identification of loanwords.

PART II: MORPHOLOGY AND SYNTAX

Among the suggestions made in Allen's paper referred to in the preamble of this article is a reference to the desirability of trying to relate systems of pronouns, 'the terms of which would be identified by their function in indicating such contextual features as relative status of speaker and hearer, inclusion and exclusion of the hearer or a third person, and so on.'

In a recent article one of our French collaborators, M. Haudricourt, has already put this suggestion into effect and compared inter alia the pronominal systems of Fijian and Samoan, with particular reference to the interesting problem of a case vide in Fijian which is occupied in Samoan. I am referring to the problem which arises when two grammatical categories appear to conflict with common sense, as for instance when on the one hand there is a distinction between singular and plural, and on the other between inclusion as opposed to exclusion of the person addressed.

One would have expected on grounds of 'common sense' the latter distinction to be 'neutralized' in the singular. In point of fact Fijian has one form only but Samoan has two. It is interesting to note that Allen, being at that stage of the argument no longer concerned with the need to establish genetic identity, does not even require a necessary phonemic identification of the pronominal forms, and declares himself satisfied with symbols such as x and y. An attempt has accordingly been made to represent two of the pronominal subsystems (one in Fijian and the other in Samoan) by means of symbols, with a value established

4 'La première personne inclusive du singulier en Polynésie,' BSLP, 54, 1959, pp. 130-5.
5 loc. cit.
by contextual situation, and abstracting two grammatical categories only, person and number. The justification for equating two systems which might be said to have been arbitrarily abstracted from two different grammatical complexes will be discussed presently.

Let \( x \), then, represent the person(s) addressed, \( y \) the person(s) referred to, and \( z \) the speaker(s). Let 1 represent the singular and 2 the dual. Let 3 represent the plural, 3a being the trial, that is, a plural appropriate for three or a small number of people, and 3b being an unlimited plural. The following could then be used to represent the two sub-systems:

\[
\begin{align*}
\text{Fijian} & \\
x1 & y1 & zy1 & \emptyset & \text{Samoan} & \quad & x1 & y1 & zy1 & zx1 \\
x2 & y2 & zy2 & zx2 & x2 & y2 & zy2 & zx2 \\
x3a & y3a & zy3a & zx3a & x3 & y3 & zy3 & zx3 \\
x3b & y3b & zy3b & zxl & x3b & y3b & zy3b & zxl
\end{align*}
\]

If one has postulated a genetic relationship, one will presumably be concerned with the history of the divergences of two systems such as these from an original to be reconstructed. Haudricourt, for instance, takes the view that Samoan has lost category 3b with the exception of the exponent for zx3b, which survives in a vestigial form as zx1.\(^1\) As he points out, discrepancies such as may be found between the Fijian and Samoan systems illustrated above can be brought into greater clarity by considerations of a synchronic as well as of a diachronic nature. One might add that as an illustration of that point in Fijian, it is not only 'contextual' number which determines the choice between categories 2, 3a, and 3b, but the status of the speaker(s) relative to that of the person(s) addressed or referred to.

In Samoan, where the linguistic expression of the relative status of speaker and hearer (or person referred to) is to be found mainly at the lexical level (i.e. in the existence of an elaborate system of terms of respect), that consideration applies only to forms zy1 and zx1, and to a lesser extent to forms x1 and x2. That is to say, one can elevate a person whom one addresses (a) by using an x2 form, 'you two,' or (b) achieve the same effect by 'lowering' oneself, i.e. by using a form zx1 which has a connotation of self-disparagement together with an appeal for sympathy or pity (i.e. 'I in common with you', 'I who have a claim on your interest, affection, or sympathy' as opposed to 'I alone' (zy1) and 'you and I together' (zx2)).

One can see therefore that, bearing in mind these considerations, the two systems are in fact in closer agreement than would appear if one was concerned merely with finding evidence of phonemic identity, or, failing that, with a one-to-one agreement for person and number. In other words, where the two systems.

\(^1\) BSLP, 54, pp. 134–5.
NOTES ON THE COMPARISON OF TWO LANGUAGES

differ one can, instead of positing a 'starred system', ascribe those differences to considerations of a synchronic order which have an effect on the equilibrium of each system.

Perhaps an even more suitable problem for comparison (without the prerequisite of a genetic hypothesis) is that provided by the existence of four nominal classes in the Fijian pronominal system and two in the Samoan. There is no obvious and immediate similarity between the two, either from the point of view of a search for phonemic identity or from their distribution and what one might call the notional and sociological correlation of those classes.

Thus, while in Fijian classes 1–4 are associated respectively with people and things in general; food and appurtenances; drink; parts of a whole, such as parts of a plant, parts of the body, terms of kinship, in Samoan classes 1 and 2 are associated with people, objects, and activities over which the speaker (or person addressed or person referred to) respectively does or does not exercise control. Thus in Fijian:

<table>
<thead>
<tr>
<th>Class 1</th>
<th>na nomu wanga</th>
<th>your boat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 2</td>
<td>na kemu uvi</td>
<td>your yam (to eat)</td>
</tr>
<tr>
<td>Class 3</td>
<td>na memu tī</td>
<td>your tea (to drink)</td>
</tr>
<tr>
<td>Class 4</td>
<td>na liŋamu</td>
<td>your arm</td>
</tr>
</tbody>
</table>

and in Samoan:

<table>
<thead>
<tr>
<th>Class 1</th>
<th>'o lau lāuŋa</th>
<th>your speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 2</td>
<td>'o lou nu'u</td>
<td>your village.¹</td>
</tr>
</tbody>
</table>

To compare a four-term system with a two-term system would, prima facie, appear to be a somewhat questionable undertaking. Moreover, the notional and semantic categories one can establish, with greater or lesser exactitude, may seem to be of very limited usefulness, and furthermore applicable only to the language for which they have been devised. Thus terms of kinship (class 4 in Fijian) are in Samoan divided between classes 1 and 2. Having 'justified' their membership of class 4 in Fijian on the grounds that they can be considered to be 'parts of a whole', one would then have to account for the fact that in Samoan they are subject to the dichotomy between control and lack of control. For instance, one might even have to argue that the word āvā 'wife' (class 1) implies control exercised by a husband over his wife, while the word to'alua 'spouse' (class 2), when used by or apropos of a husband, implies lack of control over his wife. Similarly many words of class 1 in Fijian have semantic or lexical equivalents which are distributed between classes 1 and 2 in Samoan.

Yet in a special grammatical sense classes 1 and 2 in Fijian and classes 1 and 2 in Samoan correspond exactly. In addition to the nominal particles which occur

in some of the pronominal systems of Fijian and Samoan, there exist also ‘genitive’ particles which have the same distribution. Thus in Fijian:

- na nona wanga his boat or na wanga nei Tui Tui’s boat
- na kena madrai his bread or na madrai kei Tui Tui’s bread
- na mena tī his tea or na tī mei Tui Tui’s tea
- na lii’sa his hand or na lii’sa Tui Tui’s hand

The same feature is found in Samoan:

- 'o lana lāuŋa his speech or 'o le lāuŋa a Pai Pai’s speech
- 'o lona nu’u his village or 'o le nu’u o Pai Pai’s village

It is possible for a given word to belong to more than one category according to the relationship of the two terms of each phrase:

<table>
<thead>
<tr>
<th>Fijian</th>
<th>Samoan</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>na nona itaba</td>
<td>'o lana ata</td>
<td>his photograph (i.e. taken by him or belonging to him)</td>
</tr>
<tr>
<td>na kena itaba</td>
<td>'o lona ata</td>
<td>his photograph (i.e. the photograph of him, his likeness)</td>
</tr>
<tr>
<td>na nomu itukutuku</td>
<td>'o lau tala</td>
<td>your story (i.e. the story you have told)</td>
</tr>
<tr>
<td>na kemu itukutuku</td>
<td>'o lou tala</td>
<td>your story (i.e. the story about you) ¹</td>
</tr>
</tbody>
</table>

This correlation between classes 1 and 2 in Fijian and Samoan is of importance in the analysis of certain features of the verbal systems of the two languages.

As will be recalled, both in Fijian and Samoan certain words of type CVCV and CVCVVCV (or other patterns which may be subsumed under the same formulae) may in certain circumstances be followed by monosyllabic or disyllabic suffixes, generally of the pattern CV or CVCV (also CVV in the case of Samoan). The first consonants of these suffixes have been of particular interest to comparative philologists, who wanted to establish whether or not those consonants could be regarded as the reflexes of Proto-Indonesian final consonants.²

In Fijian the final vowel may be either -a or -i. -a denotes active transitivity between a (usually) verbal form and its object when the object is ‘common’. Thus:

- au nanuma na koro I remember the village
- e nanumi ko koya he is remembered

- i denotes, inter alia, (i) active transitivity between a (usually) verbal form and its object when the object is ‘proper’. Thus:

- au nanumi koya I remember him

(ii) passive transitivity between a verbal form and its subject:

- e nanumi na koro the village is remembered
- e nanumi ko koya he is remembered

An important reservation in the use of the word 'passive' is that no 'agent' can be expressed. It is a characteristic of Fijian that one can nominalize a great number of verbal forms by substituting for the verbal pronoun a nominal pronoun of the corresponding person and number. It is here that one can observe a kind of concordance between, on the one hand, class 1 nominal pronouns and active forms (in -a), and on the other hand class 2 nominal pronouns and passive forms (in -i). Thus:

\[
\begin{align*}
\text{na noqu nanuma na koro} & \quad \text{my remembering the village} \\
\text{na kena nanumi na koro} & \quad \text{the 'being remembered' (i.e. the remembrance, the memory) of the village.}
\end{align*}
\]

In Samoan, however, where there is also a system of verbal suffixes and where it is also possible to nominalize verbal forms, efforts to establish a concordance between suffixes and nominal pronouns as in Fijian have not been successful, much of the evidence being apparently conflicting. The central difficulty would seem to be whether the CVV suffixes of Samoan can be described as active or passive.

The question whether those two terms are applicable to, and meaningful for, Malayo-Polynesian languages is of course a very complex one. After giving rise to a prolonged controversy among Dutch scholars at the end of last century, it was re-examined in considerable detail by Professor Gonda.\(^1\) Many persistent illusions and misconceptions were then removed. A similar controversy apropos of Samoan and Polynesian in general developed in 1928 in Australasia.\(^2\) Some of the chief protagonists were evidently unaware that what was essentially the same problem had already been discussed at length apropos of Indonesian languages.

A re-examination of this problem in so far as it affects Samoan would fall outside the limits of the present study. The present writer has made some progress towards a solution, and has discussed it elsewhere.\(^3\) It would seem that a comparative approach to syntax, freed from the preoccupations that stem from a desire to account for differences by reference to a common origin, would throw a good deal of light on the grammatical problems of each language. Even if in the last resort each syntactical problem proves to be soluble mainly by reference to a synchronic set of relations, the comparison of integrated systems of morphology and syntax (as opposed to the comparison of separate morphemes linked

---


by a sometimes tenuous phonemic identity symbolized by starred morphemes) may also be rewarding.

It would, for instance, be possible to devise a grammatical framework of reference within which both the Fijian and the Samoan verbal systems might be studied synchronically. For that purpose some of the transformation techniques which have been suggested recently by Chomsky ¹ could be adapted to the problems outlined above. One could, for instance, study the effect of substituting a pronoun for one of the two members in a Samoan transitive or intransitive construction. One could also study the different syntactical relations which result from the possibility, where it arises, of using a pronoun before a verbal form (with or without a suffix), after the verbal form, or both before and after, as well as the limitations to which those transformations are subject. It is probable that an independent analysis of the verbal system in both Fijian and Samoan is necessary before the two suffixation processes can be studied as a whole and before the understanding of one may be said to throw the other into better perspective.

THE STRUCTURAL PATTERNS OF NORTHERN MON-KHMER LANGUAGES

By H. L. Shorto

In this paper I attempt, following a suggestion of Miss E. J. A. Henderson, to explore the relevance to comparative and historical studies generally of systematic comparison of the total structural patterns of a number of cognate languages. These patterns are established for each language at a series of levels of analysis—syllable, word, phrase, sentence—as a manifold of syntagmatic structures and the paradigmatic systems of commuting terms which constitute them. The method differs from that of classical comparative philology, which uses the formal confrontation of selected lexical items to reconstruct an omnivalent set of hypothetical anterior patterns. Nevertheless the two procedures are not alternative but complementary, for the classical method presupposes the establishment of the patterns of each language compared, while the specification of cognate languages for the study now envisaged requires that a comparison of the conventional kind should first have been made.

Such an enlargement of the basis of comparison gains added significance in the light of recent studies of the linguistic area or Sprachbund. A linguistic area has been defined as ‘an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families’,¹ and both India and continental South East Asia have been put forward as examples of linguistic areas; relevant traits including, in the former case, the occurrence of retroflex consonants, and in the latter that of back unrounded vowels. Areally conditioned phenomena may embrace the favouring of one inherited feature in a language at the expense of another where the first is analogous to features in contiguous non-cognate languages. A comparative linguistics which, as in a recent approach by H.-J. Pinnow,² assigns the elements of a language to two inventories, areal and genetic, might overcome the gravest limitation of the traditional method, that it operates only on a partial and arbitrary selection of items.

The phonological patterns of Palaung, Riang-Lang, and Praok, all languages of the Northern Mon-Khmer group, which are examined in the first section of this paper, display a parallelism which cannot be attributed to the archetypal pattern that they share, since the lexical correspondences oblige us to postulate a more complex one in which some of the relevant features do not appear. This suggests an amendment to the definition of the linguistic area quoted above; for it seems reasonable to regard the parallelism as an areal phenomenon, even though no language of another family is cited.

The analysis of Palaung in this section is based on my article 'Word and syllable patterns in Palaung',¹ the material for which was collected in the field in 1957; that of Riang-Lang on my phonologization of material recorded in a broad phonetic transcription, and generously made available to me, by Professor G. H. Luce of Rangoon; and that of Praok, a Wa dialect of Kengtung State, on material which I collected in 1957.

This study is restricted to the primary patterns to which the great majority of words in the three languages are referable; an exhaustive treatment of loanwords would entail the statement of fragmentary secondary patterns in addition.² As a measure of the legitimacy of this procedure, it may be noted that in a sample Palaung text of 2,800 words the primary pattern accounted for 98·2 per cent of occurrences.

Word patterns. For general purposes it is sufficient to define the word in each of the languages under discussion as a structure of syllables. Two classes of syllable can then be established in each case, to which the terms major and minor are applied. The ultimate definition of major and minor syllables is in terms of the patterns specified below, and so differs in detail from language to language, though there is a broad general equivalence. For the present it may be said that a minor syllable is one which contains no vowel other than an anaptyctic one.

The word consists, in Palaung, of a major syllable which may be preceded by one or two minor ones; in Riang-Lang, of a major syllable which is accompanied by either term of a two-tone system, and may be preceded by one or two minor syllables; and in Praok, of a major syllable which may be preceded by a minor one. These structures may be symbolized as follows, putting Ma for a major syllable, Mi for a minor one, T for tone, and using parentheses to indicate that a constituent may be present or absent:

- Palaung: (Mi) (Mi) Ma;
- Riang-Lang: (Mi) (Mi) MaT;
- Praok: (Mi) Ma

The possibilities may be illustrated by the following examples:

- Palaung: ta’ ‘to weave’; kərta’ ‘tongue’; rəkərta’ ‘loom’;
- Riang-Lang: tək ‘to adhere to’; tək ‘to predict’; təktak ‘to attach’;
- tətətəkkhran ‘to be troubled’;
- Praok: ka ‘fish’; sika ‘to speak’.

In Palaung and Riang-Lang there is a high degree of correlation between the categories of minor and major syllable on the one hand—the latter taken together with tone in Riang-Lang—and those of prefix and root, as established by morphological analysis, on the other. This does not, however, hold good for Praok.

**Syllable patterns: major syllables.** The word having been defined as a structure of syllables, the two syllable classes can in turn be defined as structures of sounds, which commute in systems at the appropriate places in structure. The patterns for each language are tabulated below.

In the symbolizations of structures, C indicates a consonant system, V a vowel system, and h a one-term aspiration system, parentheses being used as in the symbolizations of word structures above. Thus the formula C(h)(C)V((V)C) implies a minimal structure CV and a maximal structure ChCVVC, the sequence VV- occurring only with a consonant following.

The structures of the major syllable are:

- Palaung:  
  \[ \text{C(h)(C)V((V)C)} \]
- Riang-Lang:  
  \[ \text{C(h)(C)V(V)(C)} \]
- Praok:  
  \[ \text{C(h)(C)V(V)(C)} \]

They thus differ only, as structures, in respect of the possibility of VV occurring without a following consonant in Praok.

The systems are set out in the tables which follow. Columns 1–5 correspond to places in structure, V(V) being treated as one unit for this purpose.

Places 1 and 4 are always filled (i.e. the minimal structure is CV in every case). Other places may be filled or not, subject to the restrictions indicated; no restrictions operate as between places 1–3 (the initial elements) and places 4–5 (the nuclear and final elements). In Palaung and Riang-Lang, when place 4 is filled by a member of the group marked (ii), place 5 is also filled; except that in Riang-Lang there are some suspect instances of a in open syllables.

### Palaung:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>n</td>
<td>m</td>
<td>i</td>
<td>k</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>y</td>
<td>e</td>
<td>n</td>
</tr>
<tr>
<td>m</td>
<td>r</td>
<td>l</td>
<td>o</td>
<td>p</td>
</tr>
<tr>
<td>y</td>
<td>i</td>
<td>e</td>
<td>a</td>
<td>m</td>
</tr>
<tr>
<td>l</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>r</td>
</tr>
<tr>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
</tr>
</tbody>
</table>
Comparison. I do not propose to consider all the similarities and dissimilarities of these patterns in detail, but single out for discussion three points of particular interest.

(1) Plosive distinctions. The terms of the three place 1 systems have for convenience been listed in different orders, the more readily to show their correlations (which vary from language to language) with those of the place 2 and 3
systems. However, it will be seen on examination that apart from this varying capacity for combination with following consonants, the place 1 systems of Palaung and Praok are identical, while that of Riang-Lang differs from them only in lacking the voiced plosives g and j and in possessing a distinction between v and w.

That this similarity is not simply attributable to the persistence of the archetypal system will appear from the following representative sets of cognates:

- Pal par, RL par, Pr pa 'you two';
- Pal gar, RL kar, Pr ke 'they two';
- Pal ta, RL ta, Pr ta 'grandfather';
- Pal ya, RL ya, Pr ye 'grandmother'.

These point to an old NMK voiced/voiceless distinction in initial plosives which has survived as such only in Palaung, being replaced by a distinction of low and high tone in Riang-Lang and lost in Praok except in the context of certain vowels, where it has been replaced by one of vowel quality.\(^1\) It seems likely that the partial filling of the resultant gap by voiced b and d in Riang-Lang follows from the introduction of glottalized plosives into that language at some stage, as an areal diffusion phenomenon; while the voiced plosives of Praok can be shown to derive from old prenasalized plosives: cf. Pr dak, Umpai and Mapä Lawa ndaak 'tongue'.\(^2\)

(2) Palaung and Riang-Lang share a threefold distinction between open syllables, syllables with final k, and those with final glottal stop which is comparatively rare in Mon-Khmer.\(^3\) Like the plosive distinction just examined, however, it is of complex origin. This may be illustrated by the examples:

- Pal ka, RL ka, Pr ka 'fish';
- Pal ti, RL ti, Pr tay 'hand';
- Pal toh, RL tu, Pr to 'to pound';
- Pal leh, RL le, Pr li 'to go out';
- Pal muk, RL mak 'ox';
- Pal a, RL ak, Pr ak 'bow';
- Pal ka, RL kak, Pr kak 'bough'.

Open syllables in Palaung correspond to final glottal stop in Riang-Lang and sometimes to open syllables in Praok, whereas the open syllables of Riang-Lang have similar reflexes in Praok and correspond to final h in Palaung. A final glottal stop in Palaung regularly corresponds to k in the other two languages; Palaung

\(^1\) In Khmer and Mon a similar distinction has been replaced by a complex one embracing both 'register'—a feature analogous to tone—and vowel quality.

\(^2\) Lawa examples cited in this paper are from Sanidh Rangsit, 'Beitrag zur Kenntnis der Lawasprachen von Nord-Siam', *Anthropos*, 37-40, 1942-5, pp. 689 ff.

\(^3\) It is found in Mon and Sré. In Khmer, however, final [k] and [?] are in complementary distribution.
final k occurs mainly in loanwords and apparently results from interlinguistic and interdialectal borrowing.

The evidence of Theng, the easternmost of the Northern Mon-Khmer languages, and of other Mon-Khmer languages outside the group, justifies referring Palaung -h and the Riang-Lang open finals to a proto-NMK *-h: cf. e.g. Pal kah, Theng kah ‘to untie’; Pal, Th koh ‘to cut’. Whether Riang-Lang -' should be regarded as an isolated survival of the proto-NMK pattern, or as a neologism, is less certain. It corresponds to Mon -' in some cases, but not in others (where the Old Mon reflexes have variously -y or -h). Whichever solution is adopted, we can postulate for the proto-NMK final system a twofold distinction—of k and either open final or -’—but not the threefold one found in Riang-Lang and Palaung.

(3) a-vowels. The vowel nuclei of Palaung and Riang-Lang fall into two distributionally distinct classes, the smaller set—marked (ii) in the tables above—occurring only in closed syllables. The vowels which have this limited distribution are, in Palaung, a, ia, ea, ua, and in Riang-Lang a, ia, ua. (There are no diphthongs in the larger class.) In Palaung they are in addition phonetically characterized by the shortness of the nuclear a-element; I have elsewhere given them the name ‘a-vowels’.

No similar contextual distinction is found in Praok. But the vowel system of that language does include eight vowels—three front, three back, one open central, and one back unrounded—which in quality and in the nature of their oppositions correspond nearly enough to the eight open-syllable vowels of Palaung and Riang-Lang. The five which remain, a, ia, i:e, ua, and u:e, may then tentatively be compared with the Palaung a-vowels and with the corresponding class in Riang-Lang, their occurrence in open syllables being linked with the absence of r and l from the Praok final consonant system.

It is possible that the anomalous length distinction in these vowels residually reproduces the more general one attributed to Common Austroasiatic and Mon-Khmer; vowel length is present as a distinctive feature in the related languages to the eastward, Lawa and Theng, as these are recorded by Sanidh Rangsit and Maspero. And since the distribution of Palaung ia and ea suggests that they were in origin non-distinctive variants, the difference between the systems of the three languages is not such as to debar us from postulating a set of a-vowels for proto-NMK: say *a, *ya, *wa. Once again it is the lexical material which engenders doubt; the a-vowels show the irregular correspondences characteristic of features propagated by diffusion.

3 Whose transcriptions are phonetic rather than phonological.
4 Shorto, op. cit., p. 551, n. 2.
The same vowel is sometimes found in all three languages, as in Pal rior, RL _riis, Pr ria 'root'. More often the correspondence is of a to a non-a- vowel, or, less frequently, of a (Riang-Lang a) to a a-diphthong, as may be seen from the following examples:

(Praok non-a) Pal kuan, RL _kuan, Pr kon 'child';
(Riang-Lang non-a) Pal lear ~ RL _les, Pr ple 'spear';
(Palaung a) Pal siem, RL _sem, Pr sem 'Shan';
(Riang-Lang a) Pal goq, RL _kuan ~ Pr goq (<*ngog) 'stalk';
(Praok a) Pal jum, RL _em, Pr eu 'leg';
(a-a-diphthong) Pal tuar, RL _tal 'to find';
RL _han ~ Pal fhan, Pr phuan 'five'.

**Minor syllable patterns.** These show greater divergence than the patterns of major syllables, the structures being:

<table>
<thead>
<tr>
<th>Palaung</th>
<th>Riang-Lang</th>
<th>Praok</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(C)</td>
<td></td>
<td>C</td>
</tr>
</tbody>
</table>

The systems are set out in the tables below. Place 1, in Palaung and Riang-Lang, is always filled. No consonant may follow itself in a two-place minor syllable, *kok-, *rar-, and *lal- being excluded.

In Palaung, when place 1 is filled by b, place 2 is always filled; no syllable *ba- occurs. In Riang-Lang, m in place 1 occurs only before a major syllable with initial r, as in _maran 'horse'; N_1, N_2 in place 2 represent a twofold nasal distinction the exponents of which vary according to context.

<table>
<thead>
<tr>
<th>Palaung</th>
<th>Riang-Lang</th>
<th>Praok</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>k p b r</td>
<td>s t</td>
<td></td>
</tr>
<tr>
<td>k p b r</td>
<td>s t</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>k e t p r</td>
<td>k N_1 N_2</td>
<td>r l</td>
</tr>
</tbody>
</table>

These systems are notably less symmetrical than those of the major syllable, and in view of the correlation mentioned above between minor syllable and prefix, and of the restricted structural possibilities, it may be thought that their limitations should be treated as lexical rather than strictly phonological; that is, that they form part of wider ideal systems some possibilities of which are not realized, just
as major syllables can be constructed within the terms of the phonological pattern which do not actually occur.

*Loanwords.* Contiguous non-cognate languages have been deliberately excluded from this study. Nevertheless, pattern analysis throws a good deal of light on the loan process. The comparison already undertaken suggests that within the realm of phonology structure is less subject to change than are systems. Thus differences of system are probably assimilated more readily in borrowing than differences of structure, so that where the structure of the borrowing language includes that of the language from which words are borrowed—as is largely the case with Northern Mon-Khmer and Tai—words are readily naturalized. Moreover, when prefixes are added to borrowed words, extensive borrowing may take place without the frequency distribution of patterns being upset. Such prefixes may function partly as grammatical labels; cf. Pal kārcu', RL tārcu' 'to assemble', both of which contain the reciprocal verbal prefix together with a root derived from Shan sā.

**Morphology**

Phonology is only one field in which the comparison of structural patterns can be illuminating; Mrs. Jacob’s paper in this volume applies a similar technique to the study of morphological patterns in Mon and Khmer, with equally instructive results. In this and the succeeding section I discuss some of the considerations which are relevant to the comparative study of morphological and grammatico-lexical patterns in Northern Mon-Khmer, without attempting an exhaustive analysis.

The material cited is drawn from the same sources as in the preceding section, except that the Palaung data are taken directly from my 1957 field notes. Lawa data are from Sanidh Rangsit, op. cit.

In the phonology different classes of syllable were recognized having specific functions in word structure, and structures and systems were then stated for each class of syllable. Similarly in morphology structures and systems are to be stated for each of a number of word classes having specific functions in the phrase or sentence, the constituents of structure being systems of prefixes and roots. The structures for the word classes noun and verb, which are identical, may be symbolized as follows, putting P for a prefix system and R for a root system:

\[
\begin{align*}
\text{Palaung} : & \quad (P)(P)R \\
\text{Riang-Lang} : & \quad (P)(P)R \\
\text{Praok} : & \quad (P)R
\end{align*}
\]

There are resemblances between this table and that of syllabic structure on p. 46. It should be observed, however, that morphological and syllabic structure do not always coincide even in Palaung and Riang-Lang, nor are morphological
units necessarily statable in terms of those used in phonological analysis. Examples of morphologically complex monosyllables are Pal *pyum* 'to set up' ~ *yu$m* 'to rise', RL *ple* 'to drive out' ~ *le* 'to go out'; in the latter case prefixation affects the tone, which on p. 46 was assigned to word and not to syllabic structure. Again, in Praok it is convenient to describe as a prefix a morpheme whose exponent is voicing of the initial plosive of the root, as in *glom* 'load' ~ *klom* 'to carry on one's shoulder'.

The next set of tables shows the prefix systems for nouns and verbs in the three languages. It is based on the analysis of a vocabulary of about 1,700 words in the case of Palaung and Riang-Lang and about 1,850 words in that of Praok, the number of instances noted being given in brackets after each prefix. Only those words are counted the roots of which, occurring themselves as words in their simple form, can be assigned to classes.

The symbol $V$ in the first column denotes the Praok initial voicing; / symbolizes complementary distribution. The second column shows the classes of root with which a given set of prefixes colligates, $N$ here standing for nominal, $V$ for verbal, and NS for noun suffix roots. For the sake of simplicity no account is taken of the lexical meaning which may be assigned to a given prefix (as with Palaung *kar-* and Riang-Lang *tar-*; which in most of their occurrences may be classed as 'reciprocal' prefixes).

<table>
<thead>
<tr>
<th>Prefix systems: nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Palaung:</strong></td>
</tr>
<tr>
<td>e- (2)</td>
</tr>
<tr>
<td>kon- (1)</td>
</tr>
<tr>
<td>kar- (1)</td>
</tr>
<tr>
<td>pa- (1)</td>
</tr>
<tr>
<td>ra- (12)</td>
</tr>
<tr>
<td>i- (2)</td>
</tr>
<tr>
<td>kon- (1)</td>
</tr>
<tr>
<td>kar- (1)</td>
</tr>
<tr>
<td>pan- (18)</td>
</tr>
<tr>
<td>ra- (39)</td>
</tr>
<tr>
<td>ren- (1)</td>
</tr>
<tr>
<td>i- (4)</td>
</tr>
<tr>
<td>se- (1)</td>
</tr>
<tr>
<td><strong>Riang-Lang:</strong></td>
</tr>
<tr>
<td>ke- (3)</td>
</tr>
<tr>
<td>tak- (1)</td>
</tr>
<tr>
<td>tan- (1)</td>
</tr>
<tr>
<td>tar- (6)</td>
</tr>
<tr>
<td>pan- (2)</td>
</tr>
<tr>
<td>ra- (3)</td>
</tr>
<tr>
<td>sa- (1)</td>
</tr>
<tr>
<td>san- (1)</td>
</tr>
<tr>
<td><strong>Praok:</strong></td>
</tr>
<tr>
<td>g-/V (4)</td>
</tr>
<tr>
<td>si- (6)</td>
</tr>
<tr>
<td>g- (5)</td>
</tr>
<tr>
<td>b- (1)</td>
</tr>
<tr>
<td>V (5)</td>
</tr>
<tr>
<td>si- (2)</td>
</tr>
</tbody>
</table>

$V$ in combination with verbal roots is in complementary distribution with $g$- and $b$-, which occur before roots with initial $r$, $l$, but does not maintain the same twofold distinction.
The most striking observation to emerge from these tables is the great divergence of the morphological as compared with the phonological systems, although, as one might expect, this was partly reflected in the systems stated for the minor syllable. Riang-Lang has the most complex overall pattern, with eighteen prefixes occurring in thirty-six functions, against eleven prefixes occurring in twenty functions in Palaung; Praok, which is altogether more idiosyncratic, has four prefixes occurring in nine functions. Two-consonant prefixes with final k are peculiar to Riang-Lang, as are those depending on nasal oppositions (raJ-, ram-; san-, sam-), while one-consonant syllabic prefixes, which account for almost three-fifths of all instances in Palaung, play a very small part in the other language. Indeed, the divergence is decidedly more marked if relative frequency is taken into consideration. The following points may be noted:

1. Syllabic prefixes account for 98 per cent of instances in Palaung and 97 per cent in Riang-Lang, against only 41 per cent in Praok.

2. Verbs form a small proportion of polymorphemic words in Palaung (22 per cent) as compared with Riang-Lang (61 per cent) and Praok (43 per cent).

3. The most productive prefixes, accounting for 5 per cent or more of instances, in Palaung are re- in nouns, colligating with nominal roots; pan- and re- in nouns, colligating with verbal roots; and kar- in verbs, colligating with verbal roots. These four together account for three-quarters of the instances.

1 a-, i- are regarded as consisting of /ʔ/, /y/ plus a feature of syllabicity.
(To them should perhaps be added i- in nouns, colligating with noun suffix roots, since the four words in which I recorded it—i’m ‘this’, itay ‘that’, imum ‘far upstream’, and ima ‘which’—have collectively a high frequency of occurrence.) The most productive in Riang-Lang are ra:n- in nouns, colligating with verbal roots, and tak-, tar-, and sak- in verbs, colligating with verbal roots; they account for rather less than three-fifths of the instances.

Comparison thus indicates that, in the morphological pattern of Northern Mon-Khmer, structure is more stable than systems; while these are peculiarly subject to change. It remains to consider how far this change is due to influences from outside the languages in question, and how far to spontaneous innovation. It seems prima facie unlikely that areal factors, which in phonology were seen to result in a uniformity greater than that attributable to genetic inheritance, should in the morphological field produce a diversity. Contact may indeed operate selectively on the genetic inventory of a language, as well as through the borrowing of items, but in this case it seems perverse thus to explain the innovatory tendencies which appear most clearly when questions of frequency are taken into account. Thus, of the productive prefixes singled out above, only Palaung kar- and Riang-Lang tar- can be confidently equated from a genetic point of view, for the correspondence between Palaung ra:, with verbal roots, and Riang-Lang ra:n- is probably partial only. Innovations peculiar to Riang-Lang are the extension of the reciprocal tar- to form collective nouns from nominal roots, as in tarti ‘(in) various places’ ~ ti ‘place’, and the prefixes with final k, which appear to have no analogue in any other Mon-Khmer language. In Praok, si- probably results from the generalization in almost all prefixial contexts of a prefix which originally corresponded to those with initial s in the other two languages, although we have noted only one instance of such a prefix in Palaung.

The view of the morphological facts given by a comparison of patterns is quite different from that derived from a genetic comparison such as is set out in summary form in the next table. This is based on correspondences between words containing a given prefix and not on the functioning of the prefix in a system; it is a synoptic one, and does not distinguish between word classes. Data for Lawa, a language of intermediate type, have been added to those for the three languages hitherto considered.

This table does not include all the prefixes listed on pp. 53–54, and others appear in it which are not listed there. But it does not necessarily follow that prefixes omitted from it are innovations, or that those included reflect ancient ones; the data which make up the two universes of discourse are not the same. The purpose of such a genetic comparison is reconstructive; but the procedure adopted has inherent dangers.

It should be noted first that, as with other kinds of reconstructive comparison, the number of instances on which the equivalences shown in the table are founded is small in proportion to that of the words, in any given language, exhibiting the
prefixes in question. Once new formations, and the workings of analogy on morphologically complex words, effect a shift in the pattern within which morphemes function, the historical antecedents even of the words which survive unchanged lose their relevance. It may help to make this clear if for a moment we leave the immediate field of this paper and consider an example from a flexional language.

The -s ending of the plural of French nouns is sometimes said to be derived from the Latin -s endings. In Old French, however, the function of -s varied according to the declension: it was either the mark of the plural (corones), of the nominative singular and oblique plural (murs), or of the nominative singular and both cases of the plural (flors). All these uses can readily be traced to Latin antecedents. The generalization of -s as a solely plural ending took place in the fifteenth century. It seems legitimate to say that with that further divergence of the two systems the apparent historical connection between certain Latin and French forms (having very different extensions) lost most of its significance.

The danger of relying on a table such as that above to establish a starred

<table>
<thead>
<tr>
<th>Riang-Lang</th>
<th>Palaung</th>
<th>Lawa</th>
<th>Praok</th>
</tr>
</thead>
<tbody>
<tr>
<td>kən-</td>
<td>kən-</td>
<td>a-</td>
<td>V?</td>
</tr>
<tr>
<td>kem-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>can-</td>
<td></td>
<td>a-</td>
<td></td>
</tr>
<tr>
<td>tan-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pən-</td>
<td>pən-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kər-</td>
<td>kər-</td>
<td>ra-</td>
<td>si-</td>
</tr>
<tr>
<td>tər-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pər-</td>
<td>pər-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sək-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sən-</td>
<td>sə-</td>
<td>sa-</td>
<td></td>
</tr>
<tr>
<td>səm-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tok-</td>
<td>kə-, kən-, kər-</td>
<td>?</td>
<td>zero?</td>
</tr>
<tr>
<td>pək-</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>ŋən-</td>
<td>ŋə-</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>ŋəm-</td>
<td>ŋə-</td>
<td></td>
<td>?</td>
</tr>
</tbody>
</table>
prefix system for the ancestral language should be apparent. Nor can it necessarily be taken as demonstrating the conservative character of Riang-Lang, with Palaung, Lawa, and Praok representing progressive reductions of an archaic system approximating in complexity to that of Riang; a wider basis of comparison would suggest that the prefixes with final k, at least, are Riang-Lang innovations. Indeed, it may be thought that it does little more than illustrate the varying morphological complexity of the four languages, without the precision which pattern comparison permits.

Most hazardous of all, perhaps, is the reconstruction of prefixes supported by lexical comparisons alone and not by any opposition existing within a single language. It is tempting to cite such 'prefixes' in explanation of partial correspondences such as Pal katam, RL 'tam 'egg', alongside the more usual type of Pal kate, RL 'kate' 'earth'. But such differences should not be projected into the parent language as oppositions unless some positive contributory evidence is available. (It is fair to add that certain phenomena in Khasi might possibly be used to support the assumption of a prefix in this case.)

LEXICON

Grammar and lexicon, in the conventional sense, may together be regarded as constituting a set of patterns analogous to those discernible in phonology and morphology; grammar providing the structures, and lexicon the systems. Indeed, grammar—specifically syntax—and lexicon may be seen as the prototypes of these two concepts; one may as legitimately speak of a 'lexicon' of phonemes as of one of words, and the case and tense endings of flexional languages are in reality a specialized type of lexical sub-system. Such sub-systems, found in a specific grammatical context, bear the same relationship to the total lexicon as do the phonological systems appropriate to particular places in structure to the 'lexicon' of phonemes.

These considerations are relevant when we seek kinds of lexical comparison which will have the force, in our very different field, of morphological coincidence in the Indo-European field. The emphasis accorded to this in Indo-European comparative linguistics is to be justified by the low probability of random coincidence in such a congruent set of items as the endings of a tense provide, when compared with an equivalent number of lexical items arbitrarily selected. It is the total or preponderant correspondence of such flexional elements as a set which enables them to be given priority in determining linguistic relationships, and makes it possible, for instance, to classify English as Germanic in spite of the mixed origins of its vocabulary.

1 Lexicons of Burmese by convention list morphemes, not words.
It might be thought that, since in the languages with which we are concerned there are classes of particles which can be said to discharge a similar function to that of declensional and conjugational affixes in Indo-European, correspondence between those particles might provide the required criterion. However, both in the comparison of indisputably cognate languages and in the few cases where historical evidence is available, investigation shows that, on the contrary, the Mon-Khmer particles are peculiarly unstable. This may in part be due to the frequent development of weak forms of words which occur in relatively unstressed positions, leading to homophony, and to extreme reduction followed by replacement by an originally reinforcing element. It is undoubtedly in part due to the greater combinatorial and positional variability of these particles, and the words with which they are collocated, compared with the affixes of other languages.

It is therefore highly desirable to establish whether in Mon-Khmer and languages of similar structure there are any lexical sub-systems so coherent that their coincidence may have some of the diagnostic value of Indo-European affix systems. Two possibilities which come to mind are the sub-systems formed by numerals and pronouns. The first of these has the disadvantage that not only is borrowing of numerals widely attested, but, as is well known, they are subject to alliterative deformation as a result of their use in counting. There is evidence of such irregular development among the numerals of Northern Mon-Khmer.

I propose here briefly to examine the Northern Mon-Khmer pronominal systems, as a pointer to the type of investigation that may prove fruitful in this field. It is first of all necessary to consider the nature of the oppositions between terms of the systems, in order to demonstrate the comparability of those systems taken as a whole.

Palaung, Riang-Lang, and Praok all have pronouns which may be assigned to first, second, and third persons singular, dual, and plural. This categorization accounts for the whole inventory of Riang-Lang. Palaung has in addition in both dual and plural two forms for the first person, which may be termed endo- and exodeictic respectively: ‘I and my fellow(s) here’, ‘I and my fellow(s) elsewhere’. Praok has the same number of forms in these categories, but, so far as I have been able to determine, the two forms of the first person dual and plural are in free variation; certainly their use is not endo- and exodeictic as in Palaung. Praok has a further peculiarity in the use of the ‘dual’, which may refer to a single female; or, to put it differently, one male is referred to by the singular, one female

---

1 The Mon negative particle (derived from Old Mon sak) in some contexts has as its only exponents aspiration and labialization of the initial consonant of the verb with which it is collocated. It is now frequently accompanied by the reinforcing particle phuh, borrowed from Burmese.

2 cf. four, five, and quatuor, quinque.

3 No such arrangement is discernible in the Lawa pronouns as recorded by Sanidh Rangsit, but it is unlikely that his investigation was exhaustive enough to be decisive.
may be referred to by either the singular or the dual, and two of either sex are referred to by the dual. None of the languages has oppositions referable to status or to grammatical function (case, pre- and post-verbal forms, etc.).

This state of affairs may be set out schematically as follows:

<table>
<thead>
<tr>
<th></th>
<th>Palaung</th>
<th>Riang-Lang</th>
<th>Praok</th>
</tr>
</thead>
<tbody>
<tr>
<td>S:</td>
<td>1 2 3</td>
<td>1 2 3</td>
<td>1 2 3</td>
</tr>
<tr>
<td>D:</td>
<td>1a 1b 2 3</td>
<td>1 2 3</td>
<td>1/1 2 3</td>
</tr>
<tr>
<td>P:</td>
<td>1a 1b 2 3</td>
<td>1 2 3</td>
<td>1/1 2 3</td>
</tr>
</tbody>
</table>

The differences between the systems are minor, although interesting, and we may proceed to consider how far the parallelism of categories extends to the forms. These are:

<table>
<thead>
<tr>
<th></th>
<th>Palaung</th>
<th>Riang-Lang</th>
<th>Praok</th>
</tr>
</thead>
<tbody>
<tr>
<td>S 1 :</td>
<td>o</td>
<td>&quot;o&quot;</td>
<td>aq</td>
</tr>
<tr>
<td>2    :</td>
<td>mi</td>
<td>&quot;mi&quot;</td>
<td>may</td>
</tr>
<tr>
<td>3    :</td>
<td>en</td>
<td>&quot;hn&quot;</td>
<td>no</td>
</tr>
<tr>
<td>D 1 :</td>
<td>endo. ay</td>
<td>&quot;ay&quot;</td>
<td>{ a, ye }</td>
</tr>
<tr>
<td></td>
<td>exo. yar</td>
<td>&quot;yar&quot;</td>
<td></td>
</tr>
<tr>
<td>2    :</td>
<td>par</td>
<td>&quot;par&quot;</td>
<td>pa</td>
</tr>
<tr>
<td>3    :</td>
<td>gar</td>
<td>&quot;kar&quot;</td>
<td>ke</td>
</tr>
<tr>
<td>P 1 :</td>
<td>endo. e</td>
<td>&quot;e&quot;</td>
<td>{ e, yi }</td>
</tr>
<tr>
<td></td>
<td>exo. ye</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2    :</td>
<td>pe</td>
<td>&quot;pe&quot;</td>
<td>pe</td>
</tr>
<tr>
<td>3    :</td>
<td>ge</td>
<td>&quot;ke&quot;</td>
<td>ki</td>
</tr>
</tbody>
</table>

The great majority of these exhibit regular correspondences, the first person dual and plural pronouns of Riang-Lang agreeing with the endodeictic forms of Palaung; the divergent vocalism of Praok pa; ye, ke and e, pe; yi is attributable to the old distinction between voiceless and voiced initials, preserved in Palaung, which is referred to on p. 49 above. Five forms are irregular, those for the third person singular; the Praok first person dual alternant a; and the third person plural of Riang-Lang.

Dealing with these in the reverse order, Riang-Lang "ke" is irregular only in its vocalism, where its cognates lead one to expect *ke’. No ready explanation suggests itself, unless the pronoun be taken as a weak form, as the third person singular appears to be (see below). Praok a is perhaps an analogical formation based on the paradigm a : pa : : e : pe. The expected form is *ay; it may be relevant that with this a homophony would arise between *kon ay ‘ our child ’, kon ay ‘ son ’. The third-person forms, no two of which agree, are the obscurest. In the first place, Riang-Lang "hn’ is a unique instance of the anomalous phonological structure CCC. It is apparently a weak form; but although Palaung en
has the weak contextual variants [n], [n], it would be impossible to connect the
two words without making unprovable assumptions. ən has at least one cognate
outside Northern Mon-Khmer, though with a different meaning (Stieng vn
‘who? ’), and there is some temptation to take it as representing the original
form; but its cognates within the group, RL _an, attributive particle—as in
_an_day tru ‘(the one) who will come’—and Pr an ‘that, the’, show that the use
of its proto-NMK reflex is most likely to have been deictic.

With these facts we seem to be on safer ground in making inferences to the
parent language. (It must be stressed that this would not be equally true of assem­
blages of words which failed to constitute lexical sub-systems.) While the
parallelism in the systems of oppositions is not in itself conclusive, and the
possibility of areal influences would have to be investigated before it could be so
taken, the agreement of twenty-six out of thirty-one forms reinforces that of the
systems. The occurrence of two forms of the first person dual and plural in both
Palaung and Praok suggests that this was a feature of proto-NMK also, one
form falling into disuse in Riang-Lang; but the original distinction need not have
been, as in Palaung, endodeictic/exodeictic.

The possession of dual pronouns is a special characteristic of Northern Mon­
Khmer. As an innovation of the group, their formation (except for that of the
‘endodeictic’ first person) is transparent when they are compared with the word
for ‘two’: Pal ar, RL _ar ~ Pr ra < *r’ar, Lawa la’aa. Palaung yar, par, gar
and the corresponding forms in the other languages are to be derived from old
compounds *yee-ar ‘we two’, etc. The prefix underlying Pr ra, Lawa la’aa
postdates this formation and belongs to a period after the separation of the Wa
languages.1

Whether other lexical sub-systems of comparable stability can be found
remains a valid subject for investigation. The essential requirement seems to be
a more or less symmetrical set of internal oppositions.

The foregoing may serve to indicate the possibilities inherent in a comparison
of structural patterns in the several fields of phonology, morphology, and lexicon,
this last being taken with syntax as an organic whole. Phonological patterns cast
most illumination on areal phenomena, grammatico-lexical patterns on genetic
ones. Both these classes of facts seem to me relevant to a truly comprehensive
comparative linguistics. Languages, considered as related entities, may be viewed
either in the orthodox way as the matrices of a largely arbitrary residue of elements

1 cf. also Pr rəm ‘water’ < *r’əm ~ -əm in pay’əm ‘cloud’, Pal om, RL əm.
which testify to the former existence of a single parent language,\(^1\) or, as Professor Holmer has suggested elsewhere,\(^2\) as records of the whole history of the contacts of their successive speakers. The virtue of a structural approach is that it allows borrowed elements the weight appropriate to their frequency. And here I may perhaps quote the words of no linguist, but a literary critic, J. G. Weightman: 'The important thing about an influence is not where it comes from but what it turns into.'

\(^1\) cf. e.g. Dyen, *Language*, 32, 1956, p. 612. This is an extreme view, of the justice of which I am not wholly convinced. The concluding paragraph of G. B. Downer's paper in this volume suggests an alternative interpretation.

\(^2\) *Lunds Universitets Arsskrift*, N.F., Avd. 1, 45/4, 1949, p. 11.
PREFIXATION AND INFIXATION IN OLD MON, OLD KHMER, AND MODERN KHMER

By Judith M. Jacob

First I must acknowledge my indebtedness to my colleague, Mr. H. L. Shorto, for most kindly supplying me with information on OM prefixes and infixes. All the data on OM here used have been taken from his notes. For OKhm the body of material is smaller and the language much less diverse than OM (many of the OKhm texts being lists of slaves' names and their duties). It has therefore provided me with many problems and fewer examples than I should have liked.

In attempting this comparison between the three languages I looked at the collected data with a view to answering three questions:

I. What graphic/phonetic and phonological elements may be (a) prefixed, (b) infixed in each language?

II. In what graphic/phonetic and phonological contexts do these elements occur?

III. What grammatical functions do they perform?

The answers which I found are presented in the three following sections.

Section I. The graphic/phonetic and phonological elements which may be prefixed or infixed.

A table of these elements is given on pages 65–67.

In so far as OKhm and MKhm are concerned, prefixes and infixes may be treated phonologically as being entirely consonantal, consisting of one or two consonants. Such vowels as are written or pronounced before or after or between two consonants of a prefix or infix (short 'inherent' vowel or short neutral vowel, a, in MKhm; short 'inherent' or occasionally i in OKhm) do not show sufficient variation to be regarded as essential phonological elements and may be satisfactorily treated as prosodic features of the junction between consonants. Cf.:

Khm chl֪wah¹ 'to quarrel' prəchl֪wah 'to squabble together' (Prefix pr and prosodic ə)

rəs 'to rake' rənəs 'a rake' (Prosodic ə and infix n)

OKhm jəhv 'to barter' pamjəhv 'bartered goods' (Prefix pN and prosodic a between consonants)

pre 'to use' pamre 'person used, servant' (Prosodic a and infix m)

¹ The phonetic transcription here used for MKhm was evolved by Miss E. J. A. Henderson. See 'The main features of Cambodian pronunciation', BSOAS, 14, 1, 1952, pp. 149–74. For OM and OKhm a straightforward transliteration is used, with the symbol ' representing the vowel-base.
(Contrast som 'to beg', smom 'beggar', where no junctional short vowel is felt to be necessary (Infix m).)

It may also be conveniently mentioned here that aspiration is another Khm feature of junction which is recorded in later OKhm spelling and in MKhm: e.g. OKhm sam 'together', phsam 'to add together'; MKhm kóap 'pleasing', phkóap 'to please'. (In both cases prefix p and prosodic aspiration.)

In OM more frequent use seems to have been made of a short vowel between prefixes and the initials of roots to which they were attached and a slightly wider variation of vowel is noted (short a, i and u all occur). Thus the prefix p (causative) has at least three forms:

- lop 'to enter'
- duk 'to be poor'
- piñ 'to be full'
- plop 'to bring in'
- paduk 'to oppress'
- pupiñ 'to fill'

and the hypothetical prefix s has two forms, s and si:

- kir 'to dig'
- dmoñ 'to stay'
- skir 'shall dig'
- sdmön 'shall stay'

and, before another s,

- sül 'to write'
- sisül 'shall write'
- sisgih 'shall be rich'
- sisip 'shall embellish'

and

- k'ím 'to smile'
- sik'ím 'shall smile' (an isolated instance of si before a consonant other than s)

I have entered in the graphic column of the tables the exact spelling of OM words (as for OKhm also), but in the phonological column I have taken the liberty of singling out the M consonants, for the purpose of comparison with Khm.

Some OM infixes are preceded or followed by a short vowel and here Mr. Shorto takes the consonant as the phonological element, e.g.:

- gruñ 'to laugh'
- ginruñ 'laughter'
- p'är 'to put into practice'
- pun'är 'conduct' (Infix N in both cases)

These are entered in the phonological column as N.

The M infix u/i, for which Mr. Shorto uses the symbol u, occurs as follows. Either

(a) it is infixed between the two consonants of an initial sequence

(i) with no change taking place in either consonant, e.g. kcit 'to die', kucit 'death'; or

(ii) with voicing of the first consonant, e.g. phic 'to fear', buhic 'to frighten'. Or
(b) when the root has a single initial consonant

(i) this is reduplicated and the infix occurs between the two consonants,
   e.g. moy 'one', mimoy 'each'; or
(ii) reduplication occurs but the first consonant is deglottalized, e.g.
   ḫat 'to compare with', biḥat 'to measure, test'.

Such occurrences as (b) (i) and (ii) are phonetically so like the Khm words
formed by the occurrence of the reduplicative prefix (e.g. kəkary 'to scratch
continually', from kary 'to scratch'; ḫəbaos 'to sweep diligently', from baos
'to sweep'), that I have entered them, in square brackets, in the graphic column
of the prefix table to demonstrate that a phonetic similarity may conceal a
morphological difference.

The M infix N may similarly occur either between the two consonants of an
initial sequence or between the two consonants formed by the reduplication of
a single initial in the root, e.g. pi 'three', pumpi 'triad'; jwui 'leg', jiṉjiṉi 'post
of building'. Again the resulting phonetic, as opposed to morphological, similarity
between this phenomenon and some Khm prefixed forms (e.g. tɔm 'to be in time,
to catch up', twɔntɔm (with prefix + N) 'to be just behind') seems sufficiently
striking to be entered in square brackets in the prefix table, I A 2, as well as in the
infix table, I B 2, where they belong morphologically. pun, pum, in the M prefix
column have also no place in the phonological column, since they represent a
prefix (p) and an infix (N), and not two prefixed consonants.

Summary: Section I

No particular likenesses are observed in connection with the single consonants
which may be prefixed in the three languages, but when two consonants are
prefixed these may in all cases be either C + r or C + N.

Where one consonant is infixed it may be n, m, or N in all the languages. In
addition M has r and ḫ infixes while Khm has p, b. Where two infixed consonants
are concerned, all three languages have only nasals and liquids. One two-
consonant infix, mn, occurs in all the languages.

M is unique in having a vowel infix, u.

Section II

1. The graphic/phonetic contexts in which prefixes and infixes occur.

   It was found

   (a) that the prefixes of all the languages (single consonant and two consonants)
may all be prefixed to roots having a single initial consonant, thus forming words
in which two or three consonants occur in close succession. Some prefixes occur
before words beginning with a two-consonant sequence, e.g. s in OM (dmon 'to
stay', sdmôn 'shall stay'); kn in OKhm (stên (title), kamstên (title)); pr in
MKhm (chliwâh 'to quarrel', prâchliwâh 'to squabble together').
## I. A. Prefix Table

### 1. Single Consonant prefixed

<table>
<thead>
<tr>
<th>Phonetic nature of prefix or infix</th>
<th>OM</th>
<th>Graphic</th>
<th>Cons. only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velar</td>
<td>(gu)</td>
<td>(g)</td>
<td>—</td>
</tr>
<tr>
<td>Palatal</td>
<td>[ji + j]</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Dental</td>
<td>[ti, tu + t]</td>
<td>(ta) [+ t]</td>
<td>Rd. t</td>
</tr>
<tr>
<td>Bilabial</td>
<td>p, pa</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>Liquid</td>
<td>(r) (ru/lu)</td>
<td>(r)</td>
<td>—</td>
</tr>
<tr>
<td>Sibilant</td>
<td>s, si</td>
<td>s</td>
<td>—</td>
</tr>
<tr>
<td>Glottal</td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OKhm</th>
<th>Phonol.</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>—</td>
</tr>
<tr>
<td>k</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MKhm</th>
<th>Phonol.</th>
</tr>
</thead>
<tbody>
<tr>
<td>k</td>
<td>—</td>
</tr>
</tbody>
</table>

### 2. Two Consonants prefixed

<table>
<thead>
<tr>
<th>Phonetic nature of prefix or infix</th>
<th>OM</th>
<th>Graphic</th>
<th>Cons. only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velar</td>
<td>[kin + k]</td>
<td>—</td>
<td>kan, kan, kam</td>
</tr>
<tr>
<td>Palatal</td>
<td>[jīn + j]</td>
<td>—</td>
<td>(cam)</td>
</tr>
<tr>
<td>Dental</td>
<td>—</td>
<td>(tam) [d]</td>
<td>(dN)</td>
</tr>
<tr>
<td>Bilabial</td>
<td>pan, pun (pri)</td>
<td>pN</td>
<td>pañ, pan, pam</td>
</tr>
<tr>
<td>Liquid</td>
<td>rin (r) (lum)</td>
<td>rN</td>
<td>(ram)</td>
</tr>
<tr>
<td>Sibilant</td>
<td>[sir + s]</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Glottal</td>
<td>—</td>
<td>—</td>
<td>'an, 'am</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OKhm</th>
<th>Phonol.</th>
</tr>
</thead>
<tbody>
<tr>
<td>kN</td>
<td>kN</td>
</tr>
<tr>
<td>kN</td>
<td>kN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MKhm</th>
<th>Phonol.</th>
</tr>
</thead>
<tbody>
<tr>
<td>kr</td>
<td>—</td>
</tr>
<tr>
<td>cN</td>
<td>cr</td>
</tr>
<tr>
<td>tN</td>
<td>tr</td>
</tr>
<tr>
<td>pN</td>
<td>pr</td>
</tr>
<tr>
<td>rN</td>
<td>rN</td>
</tr>
<tr>
<td>sN</td>
<td>sN</td>
</tr>
<tr>
<td>'N</td>
<td>'N</td>
</tr>
</tbody>
</table>
## I. B. INFIX TABLE

### 1. Single non-nasal Consonant infixed

<table>
<thead>
<tr>
<th>Phonetic nature of prefix or infix</th>
<th>OM</th>
<th>OKhm</th>
<th>MKhm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphic</td>
<td>Graphic</td>
<td>Phonol. (HLS)</td>
<td>Graphic</td>
</tr>
<tr>
<td>Bilabial Liquid</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ur, ra, ri</td>
<td>(l)</td>
<td>w</td>
<td>(l)</td>
</tr>
<tr>
<td>ir, ar, ur, ra, ri</td>
<td>r</td>
<td>p, ap</td>
<td>---</td>
</tr>
</tbody>
</table>

### 2. Single nasal Consonant infixed

<table>
<thead>
<tr>
<th>Nasal</th>
<th>m</th>
<th>m</th>
<th>m, hm, am</th>
<th>m</th>
<th>m, hm</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>(All occur when root has a single initial consonant)</td>
<td>(Occurs between two cons. of init. sequence also)</td>
<td>(All occur when root has a single initial consonant)</td>
<td>(Occurs between two cons. of init. sequence also)</td>
<td>(Occur between two consonants of initial sequence of root or between two consonants formed by reduplicated initial of root)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n, hn, an</td>
<td>n</td>
<td>n, hn</td>
<td>n</td>
</tr>
<tr>
<td>in, in, un, aň, am</td>
<td>N</td>
<td>N</td>
<td>am, an, an</td>
<td>N</td>
<td>short 'inherent' + any nasal</td>
<td>N</td>
</tr>
</tbody>
</table>

### 3. Two Consonants infixed

<table>
<thead>
<tr>
<th>OM</th>
<th>OKhm</th>
<th>MKhm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphic</td>
<td>Phonol. (HLS)</td>
<td>Graphic</td>
</tr>
<tr>
<td>inm, aňm</td>
<td>Nm, Nn</td>
<td>---</td>
</tr>
<tr>
<td>amn</td>
<td>Nn</td>
<td>---</td>
</tr>
<tr>
<td>irn</td>
<td>rn</td>
<td>---</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>umw</td>
<td>Nw</td>
<td>---</td>
</tr>
</tbody>
</table>

(All occur only when root has single initial consonant)
4. Vowel, or Vowels, infixed

<table>
<thead>
<tr>
<th>OM</th>
<th>OKhm</th>
<th>MKhm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphic</td>
<td>Phonol.</td>
<td>Graphic</td>
</tr>
<tr>
<td>(u/i)</td>
<td>(u)</td>
<td></td>
</tr>
</tbody>
</table>

(Occurs between two consonants of initial sequence of root or between two consonants formed by reduplicated initial sequence of root. May be accompanied by voicing or deglottalization of initial.)

Meaning of abbreviations, etc. in tables:

( ) (OM) only one or two occurrences known. (OKhm) meanings not completely established.

[ ] Placed for interest in graphic column but have no place in phonological column of the table.

Rd. Reduplicated.

N Nasal which may vary in point of articulation according to the context, or even at the discretion of the scribe.

(b) That in all the languages infixes may occur with roots having either a single initial consonant or a two-consonant initial sequence.

2. The phonological contexts in which prefixes and infixes occur.

Here there are some striking features of resemblance. In all three languages both prefixes and infixes are chiefly consonantal and as such either

(a) they occur without adding a syllable to the structure of the root, e.g.
OM cow ‘to return’, scow ‘shall return’; goû ‘to be brave’, gmoû ‘brave’; OKhm gûl ‘to present oneself to a superior’, phûl ‘to have oneself presented to a superior’; raû ‘to dance’, rmâm ‘dancer’; MKhm rûːs ‘to pick out, choose’, crûːs ‘to take one’s choice’; doː ‘to trade’, thnoː ‘price’. Or

(b) they are linked to the root by means of a short vowel. In this case, the resulting word often has an iambic construction, length of syllable and, for MKhm at least, stress occurring in the utterance of the vowel and final of the root, e.g. OM puwiû ‘play’, from wiû ‘to play’; jirhû ‘love’, from chûn ‘to love’; OKhm ūntûm ‘plantation’, from tûm ‘to plant’; camnûr ‘engraving’, from cûr ‘to engrave’; MKhm boûrûn ‘to teach’, from rûn ‘to learn’; trûnûm ‘perch’, from tûm ‘to perch’.

It is in all the languages the initial or initial sequence of the root to which an infix (as well as a prefix) is attached, the vowel and final of the root being unaffected. It is partly because of this that phonetically similar material has seemed grammatically different to Mr. Shorto and myself. Mr. Shorto, for whom \(u/i\) is an infix in its own right (kûn ‘to be king’, kûmûn ‘to enthrone’; etc.), arrives at the conclusion that the same infix occurs in the initials of the roots jûp ‘each, all’, moû ‘one’, necessitating the reduplication of those initials: jîjûp ‘all’,
mimoy ‘each’. Similarly the high-frequency occurrence $C_1NC_1VC$ (e.g. $jīnjuñ$) is to him an instance of infix $N$ in a root with a single initial consonant ($jun$). To me the high-frequency occurrence in MKhm: $C_1SC_1VC$, represents a reduplicating prefix, as in $koka:y$ ‘to scratch constantly’, from $ka:y$ ‘to scratch’. (I have no reason to suspect the presence of a vowel-infix.) The comparatively rare occurrence $C_1NC_1VC$ represents a prefix $CN + \text{root}$ with accidentally similar initial.

Section III. The grammatical functions performed by the prefixes and infixes.

The collected data are presented in a table on p. 69.

It will be observed that the nominalizing and causative functions are common to all three languages and are carried out by both prefixes and infixes. It is only for OKhm that no instance of a causative infix has been noted and this may be due to lack of evidence only. The nominal function may be further analysed into types of noun produced, and here one is struck by the occurrence in all three languages of the double nasal infix ($Nn/Nm/mn$) for general nouns, the utensil-infix $n$, and the agent-infix $m$.

The frequentative function is common in both OM and MKhm and the absence of instances in OKhm may be due again to lack of evidence and also to lack of the right kind of text; frequentative words tend to occur in a rather vivid style of language in Khm. Only prefixes are used in Khm to express this function, whereas M uses both prefixes and infixes. A glance at the Khm prefixes, however, shows that the same kinds of consonant (i.e. nasals and $r$) which are infixed into the beginning of the word in M are present in several of the Khm prefixes. The latter have been divided in the table according to whether they tend to give a straightforward repetitive meaning to the root or have an intensifying or restricting effect on the meaning of the root. Reduplicated $c$, $m$ and $s$ fulfil both functions, e.g. $thèak$ ‘to kick’, $tathèak$ ‘to kick repeatedly’; $kaon$ ‘to bend’, $comkaon$ ‘to bend with effort’ (both intransitive).

The entry for OKhm in the quantifier section was made on the strength of only two occurrences noted by me, $tanloñ$ (capacity measure; quantifier), from $tloñ$ (capacity measure), and $sanre$ ‘rice-field’ (quantifier), from $sre$ ‘rice-land’. These words I have elsewhere $^1$ regarded as nouns which have taken the $n$-infix (utensil, means), but it seems worth while to take them out here for comparison with M. In MKhm the two words seem to have disappeared, and among the quantifiers which I have listed I find only a few words in which the presence of an infix of this kind might be suspected: $comriðok$ ‘strip’ (quantifier for pieces of fish or meat), from $criðok$ ‘to tear in strips’; $sonrap$ ‘suit, set’, from $sráp$ ‘finished’; $kondap$ ‘sheaf’, from $kdap$ ‘to grasp in the hands’. In all these

words, the infix is \( N \), which may represent the general nominalizer or the agent infix just as well as the \( n \)-infix.

OM and OKhm are again comparable in using prefixes or infixes to express a verbalizing function, although evidence for OKhm is slender ('ampăn ' to arrest ', from păn ' to hold, keep ').

OM is alone in using prefixes and infixes for two sets of functions:

(i) The hypothetical and preparative function (prefixes \( s \) and \( sa/su \)). These
I have put together as they are similar in meaning as well as possibly identical phonologically: e.g. *gap* 'to please', *sgap* 'shall please'; *rap* 'to hold', *surap* 'to put ready at hand'. MKhm has a separate word for this grammatical function (*mʊŋ* 'shall, will, having the idea of'). In OKhm, the word *pi* 'in order to, with the intention of', seems to have a very wide variety of usages and might be regarded as at least partially fulfilling this function. No evidence of a prefix has been found.

(ii) Two functions which relate to numerals, i.e. the formation by means of infixes of distributive and group numerals. These seem to be entirely non-existent in Khm.

OKhm alone has a prefix to indicate a proper name. MKhm is alone in having a series of prefixes which turn a transitive verb into an attributive verb, with 'passive' sense, e.g. *cat-cay* 'to spend', *khcat-khca:y* 'dispersed'; *do:k* 'to uproot', *radxk* 'uprooted'. The OM infix N₂, which also forms an attributive verb, may, however, be regarded as comparable in function.

In M more than one prefix/infix may be used at the same time with a root, e.g.:

- *dow* 'to run away'    *spdow* 'shall chase away' (Prefixes s and p)
- *ca:äh* 'to be pure'    *scu:äh* 'shall purify' (Prefix s and infix u)

This is very rare in Khm. Even where a 'family' of words are current, e.g. *kɔr* 'pleasing', *phkɔr* 'to please', *bɔŋkɔr* 'to give an order', one does not regard *-ŋ* (N) as infixed into the prefixed form *phkɔr*, since *bɔN* is itself so commonly used as a prefix with many roots for which there is no intermediate form comparable with *phkɔr*, e.g. *pɛ* 'full', *bɔmpɛ* 'to fill'; *tum* 'ripe', *bɔntum* 'to ripen' (transitive).

It is possible, however, that a diachronic study of Khm would reveal that some apparent two-consonant prefixes developed from two separate prefixed or infixed units.

I would summarize the result of this inquiry as follows. There is very little in OKhm which has not continued and developed in MKhm. Where important contrasts are found, they are chiefly between M on the one hand and Khm, O and Mod, on the other. One is faced by close resemblances, in which are combined all aspects, phonetic, phonological, and functional, and by complete contrasts. Nevertheless it would seem to me that a comparison of MK on these same points (elements prefixed or infixed, contexts in which they occur, functions performed) with other non-MK languages would demonstrate the resemblance rather than the difference between M and Khm.

---

1. Mr. Shorto has suggested that the OM preparative function is in effect a 'causative of the hypothetical'. 'To put ready at hand' might then be paraphrased 'to cause (someone) to be about to hold'.
THAI AND VIETNAMESE: SOME ELEMENTS OF NOMINAL STRUCTURE COMPARED

By P. J. Honey and E. H. S. Simmonds

Professor Allen has remarked on the 'definition of languages as brothers if and only if they were once their father'.¹ This is a comment of wit upon Professor A. S. C. Ross's famous dictum which expressed, in somewhat extreme form, a fundamental assumption of traditional comparative philology.² Those who work in the field of South East Asian languages, dealing with problems of comparison of special complexity, will not be inclined to accept a simple hypothesis of genetic relationship as a basis for work. There is, rather, an attraction to the belief that useful work can be done by studying the relationships of linguistic systems and sub-systems without making an immediate assumption of relationship in genetic terms between language and language.

W. S. Allen has laid down certain fundamentals for such studies, one of the chief of which is the aim 'to reduce the stock of categories by the erection of more general systems than those established for the single languages'.³ In applying such methods it must be remembered that we are studying linguistic systems set up by investigators and not languages. Nevertheless the writers believe that it is not incorrect to reconnect with languages at an appropriate stage of the investigation. Work on system relationships ought eventually to be related to relationships between languages. By accepting the convenience of avoiding the intrusion of the genetic hypothesis at every turn, that concept need not be dismissed for ever.

Professor Nils M. Holmer has referred to languages where 'foreign elements often represent a heavy portion of the entire word material'.⁴ This statement is relevant to the languages under discussion here, where a very high proportion of lexical items in Vietnamese is attributable to Chinese loans (c. 50 per cent), and where Thai contains much lexical material of Khmer and Indic origin. We should perhaps be wise, therefore, to enter a warning against the immediate assumption that the existence of sets of cognates in two languages in our area affords evidence of genetic relationship between those languages. Even if the material concerned appears to be of considerable age we cannot dismiss the possibility that it results from a loan relationship.

In South East Asia we are in touch with receptive cultures in more or less

³ Allen, op. cit., p. 90.
⁴ Nils M. Holmer, 'Lexical and morphological contacts between Siouan and Algonquian', Lunds Universitets Årsskrift N.F. Avd. 1, Bd. 45, 4, 1949, pp. 1–36.
close contact for prolonged periods at varying times and we have to deal with a remarkable and highly complex process of synthesis. This applies, in fact, not only to language as such but also in the fields of literature, religion, art, government and so on. By comparing in systematic typological terms we are not at once committed to a decision in terms of genetic relationship. We can in this way pay full regard to, for instance, the Sprachbund hypothesis,\(^1\) which has interesting possibilities in our region. The idea of the 'mixed' language as expressed by H. Maspero with reference to Vietnamese may have its application at any rate at an intermediate stage of the investigation.

In an attempt to class phonetic, morphological, and syntactic elements with regard to relative age Professor Holmer has tentatively assigned syntactic elements to an early level. In the belief that the comparative study of syntactic structures may, in this sense, eventually prove of value the writers offer, in this paper, some remarks on nominal phrase structures in Thai and Vietnamese. The study is synchronic, based on the modern spoken language of the educated in Bangkok and Saigon. In making statements of this sort a common method of analysis and a common terminology are essential. No full and satisfactory statements of the Thai and Vietnamese structures under discussion have yet been published. The writers have found it useful to refer to the method and terminological framework used by H. F. Simon in his study of 'standard Chinese'.\(^2\) Moreover, 'standard Chinese,' or rather the system constructed by H. F. Simon for the description of the nominal complex, is taken as a system of reference for the other two languages.

Simon makes a statement of classes within the noun complex. The writers find that this method of analysis is meaningful in the cases of Thai and Vietnamese, though they themselves prefer to adopt a method which takes account of dependent–non-dependent relations between the classes.

The Simon analysis applied to the three languages would give:

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determinative</td>
<td>(Dave)</td>
<td>Determinative</td>
<td>Determinative</td>
</tr>
<tr>
<td>Determinator</td>
<td>(Dor)</td>
<td>Determinator</td>
<td>Determinator</td>
</tr>
<tr>
<td>Noun</td>
<td>(N)</td>
<td>Noun</td>
<td>Noun</td>
</tr>
<tr>
<td>Determinate</td>
<td>(Date)</td>
<td>Determinate</td>
<td>Determinate</td>
</tr>
<tr>
<td>Substantival suffix</td>
<td>(-S)</td>
<td>Preposition</td>
<td>Preposition</td>
</tr>
<tr>
<td>Substantival desegment</td>
<td>(de-seg)</td>
<td>Attributive linking particle (ALP)</td>
<td>Attributive linking particle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-noun particle (PNP)</td>
<td>Post-noun particle</td>
</tr>
</tbody>
</table>

\(^1\) Trubetskoy; see M. B. Emeneau, 'India as a linguistic area', *Language*, 32, 1956, pp. 3–20.
Class correspondences between the systems may be tabulated as follows:

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Thai</th>
<th>Chinese</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totals:</td>
<td>6</td>
<td>7</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Totals:</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

Thus by creating generic systems from pairs of individual language systems it appears:

(a) that a significant reduction of the number of classes is effected;
(b) that the reduction as between Thai and Vietnamese is greater than that between either of these systems and that of Chinese.

Furthermore, Thai–Vietnamese correspondences are again most numerous when sub-classes are examined. The Simon analysis of the noun complex in Chinese requires three sub-classes of Determinative and Determinator, and two sub-classes of Determinate.

The three systems include unrestricted determinates, but restricted determinate is not a useful sub-class in the case of Thai or Vietnamese.

The sub-class restricted determinator is not applicable to Thai or Vietnamese, though it is useful to state Generic and Measure determinators as sub-classes for all three systems.

A close examination of class members, the paradigmatic rather than the structural aspect of the analysis, might reveal closer correspondences between Thai and Vietnamese. For instance, it is interesting to note that within the paradigm of demonstrative determinatives Thai and Vietnamese exhibit three-term deictic systems as opposed to a two-term system in Chinese.

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>jeh</td>
<td>nǐi</td>
<td>này</td>
<td>this here</td>
</tr>
<tr>
<td>nah</td>
<td>nān</td>
<td>dō</td>
<td>that there</td>
</tr>
<tr>
<td>nōon</td>
<td>kia</td>
<td></td>
<td>that yonder</td>
</tr>
</tbody>
</table>

**ORDER IN THE NOUN COMPLEX**

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>jeh</td>
<td>D-Dave</td>
<td>G-Dor</td>
<td>N G-Dor D-Dave</td>
</tr>
<tr>
<td>nān</td>
<td>lēm nǐi</td>
<td></td>
<td>quyên sách này</td>
</tr>
<tr>
<td>this item book</td>
<td>book item this</td>
<td>item book this</td>
<td>'this book'</td>
</tr>
</tbody>
</table>

In Thai and Vietnamese D-Dave occurs in final position.
The Simon analysis does not include a class of noun attribute because 'forms of an adjectival kind' are regarded as an integral part of the determinator or noun. The writers would prefer to set up such a class for Thai and Vietnamese, regarding the 'stative verbs' when used as 'adjectives' as being down-graded forms within the nominal complex. In the following examples a class attribute is included for each language. The nature of the relationship of the generic systems set out above is not affected by this since one term would be added to the class count of all three systems.

*With attribute*

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Dave</td>
<td>N-Dave</td>
<td>G-Dor</td>
<td>N-Dave G-Dor N N Att D-Dave</td>
</tr>
<tr>
<td>D-Dave</td>
<td>N-Dave</td>
<td>G-Dor</td>
<td>N Att D-Dave</td>
</tr>
</tbody>
</table>

In mooted contexts Thai and Vietnamese exhibit an exact structural correspondence:

**Thai**

G-Dor Att D-Dave

lèm yẩy nǐi

item large this

**Vietnamese**

G-Dor Att D-Dave

quyên lớn này

item large this

*With numerator*

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Dave</td>
<td>N-Dave</td>
<td>G-Dor</td>
<td>N-Dave G-Dor N N-Dave G-Dor D-Dave</td>
</tr>
<tr>
<td>D-Dave</td>
<td>N-Dave</td>
<td>G-Dor</td>
<td>N-Dave G-Dor D-Dave</td>
</tr>
</tbody>
</table>

In mooted contexts Thai and Vietnamese exhibit exact structural correspondence:

**Thai**

G-Dor Att D-Dave

jeh leang been shu

this two item book

**Vietnamese**

G-Dor Att D-Dave

hai quyên sách này

two item book this

*With attribute and numerator*

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Dave</td>
<td>N-Dave</td>
<td>G-Dor</td>
<td>Att N A / B / C</td>
</tr>
<tr>
<td>N-Dave G-Dor</td>
<td>N Att D-Dave</td>
<td>G-Dor / N Att B / C / A</td>
<td></td>
</tr>
</tbody>
</table>

\[1\] ibid., pp. 338, 348.

\[2\] Continuation expressions where the topic of discussion has already been raised either verbally or by non-verbal action, e.g. gesture.
Abtracting the common element N-Dave G-Dor (B) we have the formulae:

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai-Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AC</td>
<td>CA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>jeh leang been dah shu</td>
<td>nánsúu yãy sôọ lêm nii</td>
<td>hai quyên sách lớn này</td>
</tr>
<tr>
<td></td>
<td>this two item large book</td>
<td>book large two item this 'these two large books'</td>
<td>two item book large this</td>
</tr>
</tbody>
</table>

Also to be noted is the exact structural correspondence: N Att in Thai and Vietnamese as opposed to Att N in Chinese.

A point of difference in nominal structure between Chinese and Thai/Vietnamese is the existence of the substantival suffix in the former system. A roughly equivalent set of structures in the latter is effected in prepositional terms.

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V N-S</td>
<td>V Prep N</td>
<td>V Prep N</td>
</tr>
</tbody>
</table>

Examples

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tzay utz-llí</td>
<td>yůu nay họọ</td>
<td>ṭ trong bương</td>
</tr>
<tr>
<td></td>
<td>to be [in place] room-inside</td>
<td>to be [in place] in room</td>
<td>to be [in place] in room</td>
</tr>
<tr>
<td></td>
<td>'is in the room'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>baa shu getzay juotz-shanq</td>
<td>waan nánsúu bon tọ̀ năn</td>
<td>dề sách trên bàn kia</td>
</tr>
<tr>
<td></td>
<td>take book(s) put on table-on top</td>
<td>put book(s) on table that</td>
<td>put book(s) on table that</td>
</tr>
</tbody>
</table>

Structurally the latter example being:

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V-N-V N-S</td>
<td>V-N Prep N D-Dave</td>
<td>V-N Prep N D-Dave</td>
</tr>
</tbody>
</table>

The Chinese structures collected under the head of De-seg in the Simon analysis are several. In Thai/Vietnamese a series of different but mutually corresponding structures has to be considered. When an attributive expression occurs the concept of the down-graded sentence can apply in all three languages. The linking details differ.

<table>
<thead>
<tr>
<th>Language</th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>De-seg D-Dave G-Dor</td>
<td>G-Dor ALP seg</td>
<td>G-Dor ALP seg</td>
</tr>
</tbody>
</table>
Example (mooted context)

Chinese  
nii tzwo-tian mae-de ney been  
you yesterday buy-de that item

Thai  
lèm thii khun sùu mìøwaan nii  
item ALP you buy yesterday

Vietnamese  
guỳên mà anh mua hôm-qua  
the one you bought yesterday

However, the application of De-seg extends, for example, to the ‘possessive’.

Chinese  
De-seg N  N PNP Pr  
wo-de shu  nàntsùu khọơn phóm  
I-de book  book object I

Thai  
(G-Dor) N PNP Pr  
[muò] sâch cùa tọi  
[item] book object I

Vietnamese  
[guỳên] sách của tôi  
‘my books’

In this case the Thai/Vietnamese structures show closer correspondence.

The place of determinate structures in all three systems is comparable but structural differences within the expressions again show Chinese opposed to Thai/Vietnamese.

Chinese  
D-Dave Date  
ney tian  that day

Thai  
Date D-Dave  
wan nàn  day that

Vietnamese  
Date D-Dave  
hôm áy  day that

Chinese exhibits special characteristics in certain expressions, e.g.:

Chinese  
tian-tian  day-day

Thai  
thük wan  every day

Vietnamese  
mọi ngày  every day

and

Chinese  
woomen yijjing tzooule [ellshyr lii luh] dahjia dou-ley-de-yawsyy  
we already walked [twenty mile road] all all-de tired very

Thai  
rau drxνn [yiiς p kiiloomèef] maa lëethük khon kʊø nụoi màak  
we walk [twenty kilometres] come already every person also tired very

Vietnamese  
ta di [hai muroi çay só] ròi và ai cùng mét lăm  
we go [twenty kilometres] already and every one also tired very

‘we have already walked twenty miles (kilometres) and everyone is tired out’

In Thai and Vietnamese the structure N-Dave Date is not marked by a particle. Note also with reference to the verbal system, not treated in this paper, the Chinese // Thai-Vietnamese structures in relation to the place of the aspect particles.
A Chinese // Thai-Vietnamese opposition is also shown in certain extended structures. An example of an extended D-Dave structure in mooted context is given below:

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Thai</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Dave</td>
<td>D-Dave</td>
<td>D-Dave</td>
<td>D-Dave</td>
</tr>
<tr>
<td>N-Dave</td>
<td>jeh-leang</td>
<td>been</td>
<td>yiuway</td>
</tr>
<tr>
<td>ext</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>this two item except</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai</td>
<td>nɔcɔŋkɔak</td>
<td>sɔŋ lɛm nii</td>
<td></td>
</tr>
<tr>
<td></td>
<td>except two item this</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnamese</td>
<td>trứ hai cǎi nay</td>
<td></td>
<td>except two item this</td>
</tr>
<tr>
<td></td>
<td>'apart from these two'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The comparison of descriptive systems for the noun complex in Chinese, Thai, and Vietnamese shows that considerable general similarity exists in terms of class, order, and, to some extent, in class content. However, as a more detailed analysis proceeds, Thai and Vietnamese are revealed as exhibiting closer correspondences than Chinese can show with either of them. It would be premature to draw any firm conclusions from such facts, but it would be interesting to extend the work by (a) comparing descriptive systems of the verbal complexes of the three languages, and (b) by attempting to apply the method of description to the nominal systems of, say, Khmer and Malay. The existence of special processes of word formation in the latter languages, making necessary, probably, a morphological level of analysis not appropriate for Chinese, Thai and Vietnamese, complicates the comparison. Syntactically, however, it is obvious at first glance that some interesting correspondences can be adduced in terms of order in the main nominal classes. As far as Malay is concerned Gustav Schlegel long ago drew attention to syntactical correspondences between that language and Siamese (Thai).¹ H. Maspero has referred briefly to syntactic correspondences between Vietnamese, Thai, and Mon-Khmer languages.² Further synchronic work in classifying structural features would produce results which might, at length, be compared with the results of diachronic investigations within and between language systems even though the lack of historical material creates grave difficulties in the case of a number of languages.

In discussion on this paper Professor Søren Egerod drew attention to the greater extent of correspondence between the 'nominal classes' of Thai/Vietnamese and earlier states of Chinese. He suggested that such diachronic comparisons involving several languages would be valuable if reconstruction were the end in view.

The fact of syntactic change in Chinese is obviously of great interest in itself, and synchronic analyses of the present type can at once point to the existence of

syntactic change in languages of comparable type, the historical nature of which can best be examined, in the view of the writers, by diachronic method within language and dialect groups in the first instance. The linguistic history of Thai and Vietnamese is relatively short in terms of reliably dated documents and reveals, in fact, a slow rate of change.¹

Mr. H. L. Shorto said that ‘word-class’ and ‘function’ should be distinguished as guiding notions. If it is recognized that some languages are more suitably dealt with by ‘function systems’ and others by ‘word-class’ systems an important difference in language types may be indicated thereby.

In the absence of morphological systems of the inflective type in Thai and Vietnamese, classes are certainly established by syntactic function, in which order is an important criterion. The fact that large numbers of homonymous forms can be seen to exist between certain classes, e.g. nominal attribute and stative verb, is irrelevant at the grammatical level because the differing grades of such forms can be determined by grammatical means. The concept of the down-graded sentence, phrase, and class must be recognized as valuable, perhaps essential, in the analysis of languages of the type under discussion.

¹ Except in the language of modern newspapers where, in the present century, the influence of Western languages, notably English and French, has induced syntactic change.
PROSODIC ANALYSIS, AND PHONOLOGICAL FORMULAE, IN TIBETO-BURMAN LINGUISTIC COMPARISON

By R. K. SPRIGG

GENERAL

In his article 'Relationship in Comparative Linguistics' W. S. Allen draws attention to 'a current and growing tendency to move away from the traditionally phonemic type of analysis in the direction of analyses having two outstanding characteristics: (1) They are "prosodic", in the sense that they are orientated with a view to syntagmatic implications rather than segmental oppositions...; (2) They are "phonological", in the sense that their systems are relevant to the structural positions for which they are established, and are congruent with other levels of analysis, notably the grammatical... This does not raise any special new difficulties for comparison—more likely the reverse; for the categories established by these techniques often show a wider range of application than the traditional phonemic classes'. That Allen has not exaggerated the advantages to be gained from prosodic analysis, the non-phonemic analysis, promulgated by J. R. Firth, to which he refers, I hope to demonstrate in this paper by applying it to phonic data from a selection of Tibeto-Burman languages, each of which presents rather different problems to the comparatist, Burmese and two dialects of Tibetan, the Lhasa and the Sherpa (pp. 99–108).

In particular the purpose of this paper is to propose as the most suitable forms for linguistic comparison in Tibeto-Burman languages, and possibly in other languages of South East Asia as well, a phonological formula for each lexical item such that it shall summarize all recorded variant phonetic forms of that lexical item. Such a phonological formula is not to be identified with any particular one of the variant phonetic forms considered as a norm from which the remaining forms deviate (the citation form, or form used in speaking the item as a one-word sentence, or the form used in absolute position, is commonly given priority over other forms), but is equally representative of, and equally remote from, every one of the variant phonetic forms. Each phonological formula is, then, invariable, and exemplified in utterances only through one or other of its associated phonetic forms; and the test of adequacy of each is what J. R. Firth

1 TPS, 1953, pp. 84–5.
4 As, for example, by William Cornyn, in his 'Outline of Burmese Grammar', Language, 20, 4, 1944, Suppl., p. 9 (25, 28–9), and by Raven I. McDavid, Jr., in his 'Burmese Phonemics', Studies in Linguistics, 3, 1, 1945, pp. 6 (1), 11 (8.1), 15 (9.2).
has called *renewal of connection*, via an exhaustive statement of phonetic exponents; indeed, each such phonological formula serves as an index of the total exponency of each of its components, whether prosodic or phonematic.\(^1\)

The terms prosodic and phonematic refer to the two major phonological categories distinguished by Firth in prosodic analysis. The prosodic category is primarily concerned with the analysis of the phonic data into sequences of syntagmatically associated features extending over two or more segments; these sequences of related features are stated as exponents of terms in prosodic systems applicable to (prosodic) units, or Pieces.\(^2\) In extent prosodic Pieces range from the monosyllabic Piece (or Syllable Piece) at one extreme to the polysyllabic Sentence Piece, for which intonation systems are statable, at the other; between these two extremes such prosodic units as disyllabic Piece, Word Piece, Clause Piece can be distinguished.

Although the primary concern of the prosodic category is with sequences of features extending over two or more segments, this category can also be applied to features characterizing single segments, provided that there are good grammatical grounds for associating the forms characterized by these single-segment features. Examples of this use of prosodic statement to group together at the phonological level two different grammatical forms of one Verb, each of which forms is distinguished from the other by phonetic features limited to a single segment, appear in the Tibetan section (pp. 102, 105–7), e.g. short vowel duration and backness (\(\text{o}, \text{\text{o}/\text{a}}, \text{\text{\text{a}/\text{a}}})\) and long vowel duration and frontness (\(\text{y}, \text{\text{\text{e}/\text{e}}}\)\) are stated prosodically, through the Quantity system, rather than phonematically, in order to emphasize the identity in grammatical function of the phonetic distinctions obtaining between the two forms of each Verb regardless of vocalic differences, e.g. \(\text{so}\) and \(\text{sy},\) \(\text{srub}\) and \(\text{srubs},\) \(\text{churn},\) \(\text{so/so}\) and \(\text{so/\text{e}},\) \(\text{bzo}\) and \(\text{bzos},\) \(\text{make},\) \(\text{ta/\text{ta}}\) and \(\text{ts};\) \(\text{b\text{\text{a}}t\text{a}}\) and \(\text{b\text{\text{\text{a}}t\text{\text{a}}}}\), look at.

As a matter of procedure prosodic analysis gives priority to the prosodic category: phonematic analysis is not undertaken until analysis of the phonic data in terms of prosodic systems has been completed. At this stage, when no further syntagmatic associations of features are observable, and there are no grammatical grounds for further prosodic statement, any phonetic differentiation that remains unaccounted for is subjected to analysis into phonematic units, on the basis of lexically significant minimal phonetic distinctions within prosodically comparable Syllables. The phonematic aspect of prosodic analysis

---

\(^1\) cf. also A. E. Sharp: 'The word as such has no audible features: it exhibits, for instance, no attributes of stress of a kind that may be “lost” or “modified” in sentences. Rather it may be given what I call a “conphonational formulà”, ... to summarize the totality of its exemplification in utterance' (‘Stress and Juncture in English’, TPS, 1960, p. 108).

For *renewal of connection* see F. R. Palmer, ‘“Openness” in Tigre’, BSOAS, 18, 3, 1956, p. 577.

\(^2\) When used technically, as formally established phonological and grammatical terms, Piece, Syllable, Sentence, etc., are distinguished by capital letters.
is thus paradigmatic, as opposed to syntagmatic, and in this respect resembles phonemic analysis. The components of phonological formulae reflect this phonological distinction: prosodic components refer to terms of prosodic systems, and phonemic components to terms of phonemic systems; and the two types of component are differently symbolized below.

A phonological formula such that it summarized all phonetic variant forms of a given lexical item, a lexical-item phonological formula (p. 79), is not the only possible type of phonological formula: the degree of generality of phonological formulae varies with the purpose that they are intended to serve. A formula such as Burmese Vz, for example, is generalized to a high degree, at which it applies not to an individual lexical item but to all monosyllabic lexical items that are prosodically classifiable as z (p. 89) provided that they are also of the type for which V systems can be stated (p. 98), e.g. pe pez give, we way buy, thwa swās go; pju(āē) phrū(sāñ) white, pō pui’ send, sozo coco early. At the other extreme there is the phonological formula specific enough to indicate a particular variant phonetic form from among several variant phonetic forms of a given lexical item. The variant phonetic forms of the Burmese Verb sok, drink, for example, are numerous (p. 90), and include thao’ sok drink it, (tu)thao(je) lūsokre drinking-water, and (ma)thao(phu) masokbhūz does not drink; the variant phonetic form in maθao(phu), for example, would be formulated as lbpysOkphf, in which lowercase letters refer to terms of prosodic systems, and capitals (only O here) to terms of phonematic systems; and components that are concerned with Syllable-final features of the lexical item are placed last, after the phonematic component, if any.

---

1 The phonetic transcription, in the International Phonetic Alphabet but with the addition of three symbols, is not more detailed than is necessary for the immediate purpose. Pitch features have, therefore, been generally left unsymbolized; duration features are symbolized only on pp. 87–8, 99–107; and θ and δ have been used to symbolize dentality accompanied not only by friction but also by the Word-initial occlusion that has been observed to alternate with friction for one and the same Burmese lexical item. θa symbolizes dentality when followed by ō, but otherwise alveolarity; t and d symbolize alveolarity, together with dorsality when followed respectively by g or z, but otherwise apicality. The three additional symbols are: 'e, which symbolizes glottal-trill (or ‘creaky’) voice quality, and V and C, which have been used to symbolize vowel and consonant respectively in preference to the symbols suggested on p. 19 of The Principles of the International Phonetic Association (London, 1949).

2 The present statement, of phonological formulae in which phonematic components V are recognized for some Syllables, including sok, while others have no phonematic components, and are wholly prosodic, supersedes my statement in 'Junction in Spoken Burmese' (pp. 109 ff.), in which structures CV and CVC are recognized. This latter statement is defective in that it does not fully take into account the syntagmatic implications of certain sequences of inter-related phonetic features (those dealt with through the Quality, Juncture, Aspiration, and Labialization systems; pp. 85–90, 91–5).

It is unfortunate that the resources of the roman alphabet are so limited as to make it necessary to use the same letter for more than one term, e.g. f (from fronting) as a term of the Labialization system, and f (from fast) as a term of the Tempo system; but there is in practice little chance of confusion: some of the terms, e.g. f, of the Tempo system, are not needed in the
While all phonematic components of phonological formulae, together with those prosodic components which refer to prosodic systems that are stated for a monosyllabic Piece (or Syllable), could be regarded as elements of Syllable structure, the remaining components, all those which refer to prosodic systems statable only for polysyllabic Pieces (in Burmese, Tone, Juncture, Quality, etc., pp. 85–96) are not elements of Syllable structure but classifications of the lexical item in question in accordance with its power of inclusion in the relevant types of polysyllabic Piece. Thus, the component ‘1’ does not mean that the lexical item in whose phonological formula it occurs is a Tone-1-Syllable lexical item but that that lexical item is exemplified exclusively in Tone-1 Words (of from one to six Syllables, in the case of Words in which monosyllabic Verbs are exemplified), e.g. the Tone-1 monosyllabic-Verb lexical item *swās* can be exemplified in the hexasyllabic Tone-1 Verb-and-Particle Word *(mə)θwa(bajažine)* *maswāsparacenhaṅ* do not let’s go, equally with the monosyllabic Tone-1 Verb Word *θwa* *swās* go away.1

The phonetic-variant phonological formula given above (p. 81), and many other phonological formulae of whatever degree of generality, can be simplified; for certain components are implied by other more specific components, with the result that the more general components can be omitted from the formulae. In the phonetic-variant formula 1b̥pysOkphf, for example, (i), the component 1 (Tone 1) is implied by k; for k Verb Syllables are restricted to Tone-1 Verb-and-Particle, and Verb, Words (p. 89); (ii), b (from back, of a five-term Labialization system, b, s, ə, c, f; p. 96) is implied by the phonematic unit O; for O is one of the two terms (O, U) of the V system statable for kb Verbs (p. 98); (iii), p (non-p, of a two-term Palatalization system, p, ŋ; p. 97) is implied by the component s (of the thirteen-term Juncture system; pp. 91–6); and so also is y (non-y) of the two-term Yodization system (p. 97). This formula could therefore be abbreviated to sOkphf, which would be sufficient for renewal of connection with the particular phonetic-variant form of *sok* that occurs in *maθaophu* *masokkhūs* does not drink, via a complete statement of the exponents of each term referred to in the formula, implied or explicit.

For linguistic comparison, though, a formula so detailed as to indicate a particular variant phonetic form would be too specific, while a formula as generalized as Vz (p. 81) might well not be specific enough. Formulae as generalized as this last are appropriate to broad comparisons, of Tone systems and other prosodic systems; but what is required for comparing particular lexical items is a formula sufficiently specific to distinguish the individual lexical formulae devised for the unambiguous symbolization of lexical items (as opposed to phonetic-variant forms; p. 81); and it is as a last resort always possible to specify the system to which a particular term belongs.

items from all but homophones belonging to the same grammatical category, but not so particularized as to indicate one out of several variant phonetic forms of that lexical item. The following Burmese phonological formulae are at a suitable degree of abstraction for the comparison of individual lexical items (components rendered omissible by the implications of other more specific components are enclosed in brackets, and could be dispensed with provided that those implications have been adequately stated):

(1bp)gphyOz po pu' send; (1bp)mhyOk mjao- mhrok(sañ) raises; (bgpy)2khUm kho khun jump; (bgp)2phyUz pju- phru(sañ) is white; (1bp)gchΩz sozo coco early.

The prosodic and phonematic systems to which each of the components of these five lexical-item phonological formulae refers are described, and the exponents of at least one term of each system stated, in the Burmese section below (pp. 85–99).

The lexical-item phonological formulae established here for modern spoken Burmese do not supersede Old-Burmese or modern-Burmese orthographic forms in Tibeto-Burman linguistic comparison; they have been devised for comparison with phonological formulae similarly arrived at for lexically comparable items in modern Tibeto-Burman spoken languages and dialects for the purpose of producing common, or comparative, formulae (asterisked forms) valid, in the first instance, for all the dialects and languages of the Southern Unit of the Burma Branch of the Burmish Section of the Burmic Division, e.g. Burmese, Arakanese, Tavoyan, resting on the prosodic analysis of each member language.¹

For this first step, the constructing of asterisked forms for the Southern Unit, the lexical-item phonological formulae described here, each of which summarizes all observed variant phonetic forms of each lexical item, will, I believe, prove to be more suitable than, for example, the diversity of phonemic forms given for certain Burmese lexical items by Cornyn and by McDavid. The Verb sok (drink), for example, appears in at least five different phonemic forms (/θaup/, /θaum/, /θaut/, /θauθ/, /θauc/) in Cornyn’s analysis, and the Verb koñë (good) in four (/kâun/, /gâun/, /kâun/, /kâm/) in McDavid’s, one of which would have to be arbitrarily chosen for this purpose.²

As compared with Cornyn’s analysis McDavid’s achieves a reduction in diversity of phonemic forms by phonemicizing syllable-final phonetic features for the most part independently of syllable-initial features, thereby giving his

¹ This classification is Shafer’s (‘Classification of the Sino-Tibetan Languages’, Word, 11, 1955, p. 103).
² Cornyn, ‘Outline’, especially pp. 9–10. On the dominant role of the phoneme in linguistic comparison see Allen, ‘Relationship’, p. 60: ‘The formal correspondences are stated monosystemically in what are to all intents and purposes phonemic terms’. cf. also Henry M. Hoenigswald, Language Change and Linguistic Reconstruction, Chicago, 1960, in which the phoneme theory is assumed throughout.
phonemic analysis a prosodic flavour.¹ By this means McDavid gains the same advantage as my prosodic analysis, the summarizing of the diverse phonetic features of two types of syllable each in a single unit: (i), his tone phoneme /'/, my prosodic term k (pp. 85–90); (ii), his phoneme /n/, my prosodic term m (ibid.); but his analysis is less successful in dealing with syllable-initial features: like Cornyn he recognizes an alternation of syllable-initial voiced phonemes, through 'sandhi change' in 'close juncture', with the phonetically corresponding voiceless phonemes of 'open juncture', an alternation that will give him two phonemic forms, differing in their initial phoneme, for not a few lexical items, e.g. /hn/ and /n/, as in /hnêy/ or /nêy/, 'slow'; /p/ and /b/; /ph/ and /b/; etc.²; like Cornyn he recognizes a change of tone for certain lexical items, from a basic Tone I or II to Tone III ³; like Cornyn he assumes a particular phonemic form, the form whose initial and final phonemes are those appropriate to 'open juncture', to be basic.

A further argument against phonemic analysis as a basis for linguistic comparison is that, through its preoccupation with grouping allophones into one phoneme in order to secure economies in number of units (and letters), it necessarily obscures the importance of environmental difference as a source of change. Hoenigswald emphasizes the point that 'the positional allophones grouped into one phoneme change in different ways, governed by the very similarities with neighbouring phones . . . which are likely to determine their phonetic differentiation in the first place. Conversely, if the historian finds that a phoneme has been split by conditioned change, he will conclude that allophonic variation has preceded it' ⁴; but prosodic analysis, which gives maximum prominence to environmental difference by treating spans of syntagmatically associated features each as an aspect (exponent) of a unit, or Piece, would seem to provide a better basis for comparison than a theory that suppresses environmental difference in the interests of phoneme unity.

**LEXICAL-ITEM PHONOLOGICAL FORMULAE IN BURMSE AND TIBETAN**

The examples of phonological formulae at a suitable degree of generalization for linguistic comparison are given in the order Burmese (pp. 85–99), Tibetan (pp. 99–108). They are drawn mainly from lexical items that can be classified grammatically as Verbs, with a further restriction to monosyllabic Verbs in order to avoid having to deal with those Burmese Verbs which are disyllabic (cañscâz consider, ruise respect, kacâs play, etc.).⁵

¹ 'Burmese Phonemics', pp. 7–8, 12; the phonemic overlapping that arises from his partially independent phonemicization of initial and final sounds is condemned by Bloch in 'Phonemic Overlapping', *American Speech*, 16, pp. 278–84, reprinted in *Readings in Linguistics*, Baltimore, 1957, pp. 93–7.
³ ibid., p. 17, n. 4.
⁴ Hoenigswald, *Language Change*, p. 73.
⁵ The term Verb of course differs in signification from one language to another in accordance
I. General

Burmese has been chosen as the source of the first examples of lexical-item phonological formulae because it is a language in which there are no differences in Verb root (in this respect it differs from Tibetan; p. 99) and also because it is possible to contrast the Burmese phonological formulae directly with phonemic forms, those of Cornyn and of McDavid.¹

II. Quality System

In the passage quoted on p. 79 Allen refers to the characteristics of prosodic analysis as being "'prosodic' in the sense that they are orientated with a view to syntagmatic implications rather than segmental oppositions;" this characteristic is illustrated by the following prosodic treatment of certain syntagmatic associations of features within Verb Syllables together with junction features that serve to unite the Verb Syllable with the following (Particle) Syllable. Since it has to do with inter-related vowel qualities and consonant qualities, this disyllabic Verb-and-Particle unit, or Piece (p. 80), is termed the Quality Piece.² The focus of interest in the following examples of the Quality Piece is differences in the phonetic form of the final part of the Verb Syllable considered in conjunction with matching features in the following Syllable (these features are enclosed in round brackets):

(i) \[
\begin{align*}
\text{mædwaðu} & \quad (-ab-) \\
\text{mædwaðsu} & \quad (-aβ-) \quad \text{mædwanæ} \quad (-an-) \quad \text{ðwaʃi} \quad (-ai-)
\end{align*}
\]

(ii) \[
\begin{align*}
\text{mijaɔmbu} & \quad (-aomb-) \\
\text{mijaɔmmu} & \quad (-aomm-) \quad \text{mijaɔnnæ} \quad (-aonn-) \quad \text{jaɔʃi} \quad (-aɔʃi-)
\end{align*}
\]

(iii) \[
\begin{align*}
\text{mæθaɔp'phu} & \quad (-aop'ph-) \\
\text{mæθaɔphu} & \quad (-aoph-) \quad \text{mæθaɔnnæ} \quad (-aonn-) \quad \text{θaɔʃi} \quad (-aɔʃi-)
\end{align*}
\]

with the number of major grammatical categories distinguished for each; and the Verb category in Burmese, which is non-Noun and non-Particle, cannot be identified with the Verb in Tibetan, which is non-Noun, non-Adjective, non-Postposition, and non-Particle; but the impossibility of equating grammatical categories is no bar to comparing components of the phonological formulae of their members.

At the suggestion of Dr. Hla Pe, Reader in Burmese at the School of Oriental and African Studies, ₅ has been transliterated as ʰi.

¹ Cornyn, ‘Outline’; McDavid, ‘Phonemics’.

² Some disyllabic Quality Pieces are co-terminous with a (disyllabic) Word, e.g. swaipā please go (for the phonetic criteria used in delimiting Words see Sprigg, ‘Junction’, pp. 108–24); others are contained in a Word, of three or more Syllables, e.g. swaíũ(maĩnõ) I must go now, (ma)swaíũp(a ICMPacraen̩) let me not go.

³ For the Verb swaíũ, both voice (.sendStatus) and voicelessness (_voiceless) have been observed as initial features for my informant, U Tin Maung, of Sagaing, the former being apparently more common, but for all other dental-initial monosyllabic Verbs voicelessness (_voiceless) only.

⁴ Limited to Fast-Tempo utterances.
(i) maswā:ḥū:z, he does not go; maswānhaːn', do not go; swāːlhyaːn, if you are going; (ii) maroːn:ḥū:z, he does not sell; maroːnhaːn', do not sell; ronːlhyaːn, if you sell; (iii) masokbhuːz, he does not drink; masoknhaːn', do not drink; soklhyan, if you drink.

The Verb Syllable of the examples at (i) (swāː) is characterized by a different vowel quality (a) from those at (ii) and (iii), and by non-nasality, while the Particle Syllable bhuːz is characterized throughout by voice, and initially by either plosion (b) or friction (β); the Verb Syllable at (ii) (ronː) is always characterized by nasality (m n ~) together with voice, while the Particle Syllable bhuːz is again characterized by voice throughout, and initially, as a Fast-Tempo alternative to plosion (b), by nasality (m) (in which case nasality is a feature of both Syllables: mm), but not by friction (cf. (i)); the Verb Syllable (sok) at (iii) may or may not be characterized by oral occlusion; if it is so characterized (p', n), the occlusion can be accompanied either by voice (n) or by voicelessness (p'), and either by nasality (n) or by non-nasality (p'), while the Particle Syllable bhuːz is characterized, initially, not only by voicelessness of consonant (p) but also by partial voicelessness of vowel (h, i.e. ʊ); in which case both Syllables may be characterized by voicelessness (p'ph).

The examples on p. 85 show sequences of interdependent phonetic features of three types (i), (ii), (iii); these figures could therefore be used each as an index corresponding to, and summarizing, a particular type of sequence, type (i), type (ii), and type (iii), or the disyllabic Piece characterized by each of these sequences, type-(i) Piece, type-(ii), or type-(iii). In other words a three-term prosodic system, named the Quality system, can be set up for disyllabic Pieces, the exponents of each of the terms of which are drawn not only from the Verb Syllable, the first of the two Syllables, but also from the remaining, or second, Syllable as well. The figures (i), (ii), and (iii) that were used above in classifying the examples by sequence of phonetic features could be retained each as a name of one of the terms of this three-term prosodic system; but it seems preferable to replace them with names that shall have some mnemonic value. (i), (ii), and (iii) are therefore replaced by z, m, and k respectively, which serve as a reminder each of some prominent orthographic feature of the first (Verb) Syllable in each type of Piece: the first Syllable of a z Piece is generally represented by zero-consonant-final letter, e.g. kriːz, big (and, less commonly, by -y, e.g. way, buy, and by -n, e.g. kraːn', look at), that of an m Piece by -m, e.g. kramːz, rough (but also by -n, -m, -n, and -n, e.g. pin, thin; sumːz, use; mrən, see; caːns, mince), and that of a k Piece by -k, e.g. sok, drink (but also by -p, e.g. ip, sleep, -t, e.g. sat, kill, -c, e.g. phraːc, become).

At this point it seems desirable to illustrate the terms exponency and exponent by a complete list of the phonetic exponents of z, m, and k.
**First Syllable (Verb)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>z</td>
<td>long duration + clear/glottal-trill ('creaky') voice quality + (i) close/half-close/half-open front/back vowel ; or (ii) open back/frontish vowel</td>
</tr>
<tr>
<td>m</td>
<td>long duration + clear/glottal-trill voice quality + (i) diphthong, or (ii) open backish/frontish vowel, or (iii) close-to-half-close front/back centralized vowel</td>
</tr>
</tbody>
</table>

**Second Syllable (Particle)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>voice + plosion/affrication/friction/nasality/flap/laterality/semi-vowel/vowel</td>
</tr>
<tr>
<td>(ii)</td>
<td>voicelessness + palatalisation + friction</td>
</tr>
</tbody>
</table>

| (i) | b d g ; dz ; z dz ; ž ž ž ž f ; r ; m n ; l j ; v |
| (ii) | e |

1 The degree of frontness (a ā ā) varies with the presence or absence of initial palatalisation and labiovelarisation.
2 Ž, ž, ž, and ž are peculiar to Fast-Tempo utterances.
3 Varies in degree of frontness in accordance with (i), presence or absence of initial palatalisation, (ii), consonantal versus vocalic nasality, e.g. mja:mmja: mranmran quickly, mja:mmaja: mhanmnhan correctly.
4 A single example of labial friction (ž) has also been noted in an m exponent, in the Fast-Tempo utterance kac:mba:bi: kōtepāpri all right (cf. kac:mba:bi:).
5 In careful speech diphthongs ta and to appear to be free variants of t ; and o ; respectively.
6 Limited to Slow-Tempo utterances ; cf. (a) above.
**First Syllable (Verb)**

k

- short duration + clear voice quality +
  - (i) diphthong, or
  - (ii) half-open/open front vowel, or
  - (iii) close-to-half-close front/back centralized vowel;

(a) ¹

or (b) voicelessness + stop

- (c) + friction
- (d) voice + nasality
- (e) + laterality
- (f) voicelessness + nasality
- (g) + laterality
- (h) glottality + plosion

\[
\begin{align*}
(a) & \quad \{ (i) \text{ voicelessness } + \text{ plosion/affrication/friction/nasality/laterality}^2 \\
& \quad \{ (ii) \text{ voice } + \text{ nasality/laterality/semi-vowel/vowel} \\
& \quad \{ (i) \text{ voice} + \text{ homorganic plosion/affrication} + \text{laxness}^3 \\
& \quad \{ (ii) \text{ voicelessness} + \text{ homorganic friction} \\
& \quad \{ (i) \text{ voicelessness } + \text{ homorganic nasality} \\
& \quad \{ (ii) \text{ voicelessness } + \text{ homorganic laterality} \\
& \quad \{ (i) \text{ voice } + \text{ vowel} \\
& \quad \{ (ii) \text{ voice } + \text{ vowel} \\
& \quad \{ (i) \text{ p ph th k kh tch th s m}^2 \\
& \quad \{ (ii) \text{ m n l j V} \\
& \quad \{ (i) \text{ b/d/g/dz} \\
& \quad \{ (ii) \text{ b/d/g/dz} \}
\end{align*}
\]

A comprehensive series of examples of k, m, and z exponents is given on pp. 91–2, 93–4, and 94–5 respectively.⁴

¹ Restricted to Fast-Tempo utterances.
² Occasional Fast-Tempo utterances have been observed in which the appropriate diphthongs and pure vowels have been followed by voice + plosion/friction, e.g. **wods**: **watsañ** wears (cf. **wotde**), **paødela**: **poksalat** does it go as far as (cf. **paødela**).
³ Confined to Slow-Tempo utterances.
⁴ In Burmese orthography the z term of the Quality system is regularly represented, in the first of the two Syllables of the disyllabic Piece, by **i, e, ai, ā, ā, -y, -ñ, e.g. sir, ride; wet, far; lā, come; way, buy; and by i, a, u, o, and ui, provided that they are syllable-final, e.g. **ka**, dance; **rhi**, be; **pru**, do; **m** is represented by -n, -ñ, -ñ, -m, and -m, e.g. **mrañ**, see; **cañ**, mince; **sumt**, use; and k by -k, -c, -t, -p, e.g. **khak**, difficult; **phrac**, become; **wat**, wear.
The k Verb-and-Particle Piece differs from the z and the m in its relations with one of the Burmese Tone systems.\textsuperscript{1} A detailed account of the Burmese Tone systems would be out of place here; it is sufficient to mention that for Words that contain a monosyllabic Verb a two-term Tone system can be stated, whence Tone-1 and Tone-2 Words are distinguished, and that the k Verb-and-Particle Piece can be exemplified only in Tone-1 Words, e.g.

\begin{align*}
\text{z Piece} & \quad \text{m Piece} & \quad \text{k Piece} \\
\text{Tone 1 (\text{\textasciitilde})} & \quad \{\text{thud}\} & \{\text{thaond}\} & \{\text{\textasciitilde}aotd\} \\
\text{Tone 2 (\text{\textasciitilde})} & \quad \{\text{thud}\} & \{\text{thaond}\} & \\
\end{align*}

(thūsañ differs, thon̄sañ pounds, soksañ drinks, thusañ beats, thon̄sañ bends; thūsañ is thick, thon̄sañ traps).\textsuperscript{2}

The Particles bhūz, nhañ', and lhyañ, which appeared as second Syllable of the disyllabic Verb-and-Particle-Piece examples on page 85, and all other Particles that, like them, can immediately follow a Verb (and therefore be contained in a disyllabic Verb-and-Particle Piece), are clearly not exclusive to any one of the three types of Quality Piece considered there; for they appear in all three. For Verbs it is otherwise: granted that the disyllabic Piece is of the grammatical type Verb-and-Particle, swāz can be exemplified, as first Syllable, only in a z Piece, roñz only in an m Piece, and sok only in a k Piece. swāz, and all other monosyllabic Verbs that resemble it in this respect, can therefore be classified prosodically, in terms of the Quality system, as z-Piece, or z, Verbs, roñz, and similar Verbs, as m-Piece, or m, sok, and similar Verbs, as k-Piece, or k. To classify a monosyllabic Verb as z, as m, or as k not only summarizes its range of Syllable-final phonetic variation (sixfold for the m, thirteentofor the k, nil for the z; p. 90), but also implies a corresponding limitation on the range of Syllable-initial phonetic variation of any following Syllable associated with it as second Syllable of a disyllabic Quality Piece.

In other grammatical types of disyllabic Piece too, e.g. the disyllabic Verb, it is only in the m Piece, for example, that an m Syllable can be the first Syllable of the Piece, though an m Syllable is unrestricted in its occurrence as second Syllable, e.g. the m Verb Syllable koñz as second Syllable of the z Piece swākoñz (əwəgaō) swāžlim'mañ he may well go, of the m Piece koñzkoñz (kaonggaō) well, of the k Piece sokkoñz (təoggaō) soklim'mañ he may well drink, and of the Weak-Stress-Syllable Piece makoñz(bhūz) (makaom-) it is not good. Corresponding relations hold between the z or the k Verb and the z, m, or k disyllabic Piece.

These symbols, though orthographically speaking part of the Verb Syllable, provide a key to the pronunciation of both Syllables of the disyllabic Quality Piece, and can therefore be considered to have a prosodic function.

\textsuperscript{1} Sprigg, 'Junction,' p. 128.

\textsuperscript{2} The pitch patterns shown are appropriate to one-Word Sentences; they are not the only possible patterns for these Tone-1 and Tone-2 Words.
Though the Particles *bhūs*, *nhañ’, and *lhyan* cannot be classified, as z, m, or k, from their membership of the Verb-and-Particle disyllabic Piece (p. 89), *bhūs* and *nhañ’* can be classified as z-Piece, and *lhyan* as m-Piece, on much the same grounds as Verbs, from their membership of disyllabic Particle-and-Particle Quality Pieces in which each of them is followed by another Particle, e.g.:

- **z**: *nhañ’* (*maswā)*nhañ’to’ -ēd- ;  
  *bhūs* (*makoň)*bhūslās -ul-

- **m**: *lhyan* (*swās)*lhyanlā/ns -ū- ;  
  *ūs* (*swās)*ūsto’ -ōond-

- **k**: *tat* (*mapro)*tatpā(bhūs) -apb- ;  
  *phrac* (*swās)*phracsañ -ūtd-

### III. Variation in the Phonetic Form of Verbs: Junction, Juncture, Tempo, and Aspiration Systems

The examples on p. 85 have been used to illustrate sequences of interconnected features (p. 86); they also show that some lexical items, both Verb (*roñs*, *sok*) and Particle (*bhūs*), have variant phonetic forms. The Particle *bhūs*, for example, has four, one (bu) appropriate to both the z and the m Piece, one (~u) to the z Piece alone, but only in Fast-Tempo utterances, one (mu) to the m Piece only, also only in Fast-Tempo utterances, and the fourth (phu) only to the k Piece. The Verb *roñs* has three phonetic forms: *jaom*, *jaon*, *jaō* (to which could be added *jaon*, *jaon*, and *jaon*), each appropriate to an m Piece, but to a prosodically different type of m Piece; and the Verb *sok* has three forms: *θaop’, *θao, *θaon* (to which could be added *θaoθ, *θaok’, *θaot’, *θaot’, *θaom, *θaom, *θaom, *θaom, *θaol, *θaol, and *θaol’), each appropriate to a prosodically different type of k Piece.

The difference in phonetic form of the Particle *bhūs* can be completely accounted for by reference to the three-term Quality system (z, m, k) and to the two-term Tempo system (f, from fast, and s, from slow); but in order to account for the thirteenfold difference in the phonetic form of *sok*, and the sixfold difference in the phonetic form of *roñs* (in association with phonetic differences in the initial of the following Syllable, *bhūs*, *sañ*, *lhyan*), etc.) four further prosodic systems are needed: Junction, Juncture, Tempo, and Aspiration.

The Junction system has been dealt with in detail elsewhere; and a brief mention is all that is needed here.\(^1\) It comprises two terms, Interverbal Junction and Intraverbal, so named because the exponents of the Interverbal term correlate with grammatical Word boundaries, and those of the Intraverbal with absence of grammatical Word boundaries. The phonetic features that provide the exponents of either term must be drawn from two Syllables for the Intraverbal term (disyllabic Intraverbal Piece); for the Interverbal term too they are drawn from two Syllables (disyllabic Intraverbal Piece), except for utterance-initial and utterance-final Interverbal Junction, in which of necessity only one Syllable can be characterized by Interverbal-Junction features.

The phonetic forms of *roṁs* and *sok* appropriate to Interverbal Junction are *jaō*, with Syllable-final nasality of vowel, and *θao’,* with Syllable-final glottal occlusion, respectively; all six phonetic forms of *roṁs*, and all the phonetic forms of *sok* except *θao’,* are appropriate to Intraverbal Junction (the phonetic form *jaō* is, thus, common to either; but the features immediately following this form are not for the most part the same for *jaō* in Interverbal Junction as in Intraverbal Junction); e.g.:

```
<table>
<thead>
<tr>
<th>Inter.</th>
<th><em>jaō</em></th>
<th><em>roṁs</em></th>
<th>*jaō pījī</th>
<th><em>roṁs pījprī</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra.</td>
<td><em>jaōlāttdābī</em></td>
<td><em>roṁsluiikkraprī</em></td>
<td><em>jaombi</em></td>
<td><em>roṁsprī</em></td>
</tr>
</tbody>
</table>
```

(Int. : sell it ; he has sold it ; Intra. : they have sold it ; he is selling it);

```
<table>
<thead>
<tr>
<th>Inter.</th>
<th><em>θao’</em></th>
<th><em>sok</em></th>
<th>*θao’ pījī</th>
<th><em>sok pījprī</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra.</td>
<td><em>θao(m)me</em></td>
<td><em>sokmaṅ</em></td>
<td><em>θaopbi</em></td>
<td><em>sokprī</em></td>
</tr>
</tbody>
</table>
```

(Int. : drink it ; he has drunk it ; Intra. : he will drink it ; he is drinking it).

The Juncture system comprises the thirteen terms *p, m, t, n, c, l, j, n, y, k, η, s, v* (the names of the majority of the terms are, for convenience, taken each from the letter, or the letters, used to indicate that term in Burmese orthography, except that typographically unsuitable letters like ń and ń, which have diacritics, and ky, which is a digraph, have been passed over in favour of η, η and j respectively; *v* is named from vowel); the Aspiration system comprises the two terms *h* and *Ł* (non-*h*). A complete statement of exponency is not attempted here; but the sequences of phonetic features that would be drawn on for such a statement are illustrated in the sets of examples below and on pages 93-5 (the left-hand column contains the thirteen terms of the Juncture system, the next the two terms of the Aspiration system; *f* and *s* are here the two terms of the Tempo system). The examples in the following table also illustrate the exponency of the k term of the Quality system (p. 88):

```
p | { h -p’ph- ; (f) -Vph- -θao(p’)phu | (ma)sokbhūs does not drink
h | -p’bh- ; (, ,) -p’b- -θaop’h/ba | sokpā please drink
m | h -mm- ; (, ,) -Vm- -θao(m)ma | sokmaṅ drinking
h | -mm- ; (, ,) -Vm- -θao(m)me | sokmaṅ will drink
t | h -t’q- ; (f) -t’d- -θaot’q/đe | soksaṅ drinks
n | h none for the Verb-and-Particle Piece
h | -nm- ; (f) -Vn- -θao(n)me | (ma)sokhan’i do not drink
s | h -ssh- ; (, ,) -Vsh- -θao(s)shē | sokchāi while drinking
h | -ss- ; (, ,) -Vs- -θao(s)st- | sokca(rā) drinkable
l | h -l’l- ; (, ,) -Vl- -θao(l)lì | sokhu(prī) about to drink
h | -l’l- ; (, ,) -Vl- -θao(l)lò | soklu⁹” on drinking
k | h -k’kh- ; (, ,) -Vkh- -θao(k’)khē | (ma)sokkhaṇ before drinking
h | -k’g- ; (, ,) -k’g- -θaok’g/gə- | sokka(taṅska) since drinking
```

---

1 ' Juncion', pp. 119–21, 123–4.
The above examples show that (i), where labiality and occlusion (p') are a final feature of the first Syllable, labiality and plosion are initial in the second (p Piece: -p'ph-, -p'b-, -p'b-), and vice versa except in the h Piece in Fast Tempo; and the same sort of statement holds good for occlusion and plosion combined with alveolarlty (-t'd-, -t'd-), velarity (-k'g-, -k'g-, -k'kh-), and palatality (-t'tllh-, -t'tllh-, -t'd-); (ii), where friction and either alveolarlty or dentality, nasality and either labiality or alveolarlty, or laterality and alveolarlty, are final features of the first Syllable, the same combination of features is initial in the second (-ss-, -ssh-, -ss-; -mm-, -mm-, -mm-; -ll-, -ll-), and vice versa except in Fast-Tempo utterances; (iii), where glottality ('), or a syllabic or non-syllabic vowel (V, i), is initial in the second Syllable, the final feature of the preceding Syllable is vocalic; (iv), where voicelessness combined with nasality or laterality is a feature of the first Syllable, it is also a feature of the second (-ss-, -ss-, -ss-), and vice versa except in Fast-Tempo utterances; (v), where, in Fast-Tempo utterances, partial voicelessness is a feature of the vowel of the second Syllable (V, alias h), the final feature of the first is vocalic (-Vph-, -VTH-, -Vsh-, -Vkh-, -Vtllh-). It is this last sequence of associated features, extending over two Syllables, that provides grounds for treating the Aspiration system as statable for a disyllabic Piece rather than for a monosyllabic.

The table above not only gives examples of each term of the Juncture, Tempo, and Aspiration systems; it also shows that the s term of the Juncture system correlates with neither of the terms of the Aspiration system: the behaviour of dental-initial Syllables is such that it does not lend itself to classification in terms of the latter system (h, ħ); and the solution adopted in ‘Junction in Spoken Burmese’ (pp. 111–16) still seems more satisfactory than any other.

The Junction, Tempo, and Aspiration systems provide the means of accounting for the twelvefold Syllable-final phonetic diversity of sok and other k-Piece Verbs (p. 90) in the Verb-and-Particle type of Piece. The k Verb is characterized by the following Syllable-final features in accordance with differences in the type of Junction Piece, Aspiration Piece, or Tempo Piece:

---

Prosodic Analysis in Tibeto-Burmese Comparison

<table>
<thead>
<tr>
<th>Features</th>
<th>Type of Juncture Piece</th>
</tr>
</thead>
<tbody>
<tr>
<td>occlusion +</td>
<td></td>
</tr>
<tr>
<td>labiality p'</td>
<td>p</td>
</tr>
<tr>
<td>alveolar t'</td>
<td>t</td>
</tr>
<tr>
<td>palatal t'</td>
<td>j</td>
</tr>
<tr>
<td>velar k'</td>
<td>k</td>
</tr>
<tr>
<td>labial m m</td>
<td>m</td>
</tr>
<tr>
<td>alveolar n</td>
<td>n</td>
</tr>
<tr>
<td>laterality +</td>
<td></td>
</tr>
<tr>
<td>alveolar l</td>
<td>l</td>
</tr>
<tr>
<td>friction +</td>
<td></td>
</tr>
<tr>
<td>alveolar s</td>
<td>c</td>
</tr>
<tr>
<td>dentality θ</td>
<td>s</td>
</tr>
<tr>
<td>nasal +</td>
<td>y (also h)</td>
</tr>
</tbody>
</table>
| labiality p , m , t, j, n, c, l, k, y, s, and v of the Juncture system (there being no examples of η or η in the Verb-and-Particle type of Piece), and of the terms of the Aspiration and the Tempo systems, nearly all differ from those given above (p. 91-2); and certain terms have identical exponents:

Where the disyllabic Piece is not also k but m (of the Quality system; p. 87) the exponents of the terms p, m, t, n, c, l, k, j, y, s, and v of the Juncture system (there being no examples of η or η in the Verb-and-Particle type of Piece), and of the terms of the Aspiration and the Tempo systems, nearly all differ from those given above (p. 91-2); and certain terms have identical exponents:

p \{ h -mh- ; (f) -mm- -jaomb/mu \ (ma)roñzhūs does not sell
m \{ h -mm- jæomb/mu roñzpā(sa) sells
h \{ h -nd- -jaοmb/ma roñsmhā selling
n \{ h -nn- -jaοmme roñsmañ will sell
t \{ h -nd- -nn- -kaοnd/ne koñsañ is good
c \{ h -nz- -jað(n)ze roñzchai while selling
l \{ h -zl- -lű (muighz) lañslhu about to dawn
k \{ h -̃z- -jaŋg/ŋī -jaŋg/ŋa roñskhān before selling
j \{ h -̃g- -jaŋdzum roñskhyān(sa) since selling
h \{ h -ndz- jaŋdzum roñskhra(sa) they sell
y \{h no collocationally suitable example (but see p. 95)
\{h \-\Vj- ja\djt ron\=ilhyan if he sells
s \-\n\d- ja\p\d\d (ma)ron\=i\d\ez(bh\u010du) has not yet sold
v \h \-\V\V- \(s\) \-\V\v- sa\d\(\u0101\)oom- con\=\uz\(\ma\) I will go on waiting

From the above table it will be seen that in Verb-and-Particle Pieces that are also m (of the Quality system), (i) the exponent of h nowhere differs from the exponent of h (but compare pp. 91–2); (ii), in Fast-Tempo utterances p and t can have exponents that do not differ from the exponents of m and n respectively (-mm-, -nn-) \(^1\); (iii), where labiality is a feature of the first Syllable, so also is it of the second, and vice versa (p Piece : \-\mb-, -\mm-; m Piece : -\mm-); a corresponding statement can be made for velarity and dentality (-\n\g-, -\n\n-; -\n\d-), but not necessarily for alveolarity and palatality (t Piece : -\nd-, -\nn-; n Piece : -\nn-; j Piece : -\n\d\n-; but c Piece : -\V\V-; l Piece : -\V\d-; y Piece : -\V\i-), and not at all for glottality (-\V\v-), though in Slow-Tempo utterances the Syllable-initial alveolarity of the c Piece is matched by Syllable-final alveolarity (-\n\n-).

The sixfold Syllable-final diversity of ro\n\d and other m Verbs (p. 90) can be accounted for by reference to the Juncture and to the Tempo system (but not the Aspiration system; cf. the k-Piece examples, p. 93). m Verbs are characterized by the following Syllable-final phonetic features according to the type of Juncture and Tempo Pieces in which they are exemplified:

<table>
<thead>
<tr>
<th>features</th>
<th>type of Juncture Piece</th>
</tr>
</thead>
<tbody>
<tr>
<td>labiality</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td>p, m</td>
</tr>
<tr>
<td>alveolarity</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>t, n</td>
</tr>
<tr>
<td>velarity</td>
<td>\n</td>
</tr>
<tr>
<td>palatality</td>
<td>\n</td>
</tr>
<tr>
<td>dentality</td>
<td>s</td>
</tr>
<tr>
<td>nasality of vowel</td>
<td>\V</td>
</tr>
<tr>
<td></td>
<td>c if not also s (Tempo)</td>
</tr>
<tr>
<td></td>
<td>l, y, v</td>
</tr>
</tbody>
</table>

In the z Piece the exponents of the terms of the Juncture, Tempo, and Aspiration systems are:

p \{h \-\Vb-; (f) \-\V\b- \-\twab/\Bu (ma)sw\=\d\bz does not go
\{h \,, \,\, \-\V\b- \-\twab/\Bu sw\=\zp\=\a please go
m \{h \-\Vm- \-\twama sw\=\zm\h\= going
\{h \,, \,\, \-\V\m- \-\tw\ams sw\=\zm\a\=\n will go
s \h \-\Vd-; (,,) \-\V\d- \-\twad/\e sw\=\zs\a\=\n goes
\h \-\Vn- \-\twan\ez (ma)sw\=\zn\a\=\n' do not go

\(^1\) Provided that the exponents of two terms of the same system differ in at least one context, there is, in prosodic analysis, no bar to their having identical exponents in some or all of the remaining contexts. Phonetic exponents are not to be confused with phonetic criteria (cf. 'Vowel Harmony in Lhasa Tibetan', BSOAS, 24, 1, 1961, pp. 117–18, 131–32).
PROSODIC ANALYSIS IN TIBETO-BURMAN COMPARISON

Where the Verb-and-Particle Piece is also z, (i), the exponents of h and of .TYPE are identical (but cf. the k, pp. 91–2); (ii), the features labial friction, velar friction (β, y), and flap (r) are part of the exponency of p, k, and t respectively in Fast-Tempo utterances; (iii), in Verb-and-Particle Pieces that are also z, unlike k and m (pp. 91–2, 93–4), the localization features labiality, alveolarity, velarity, etc. are confined to the second Syllable of the Piece.

Not only does the Verb-and-Particle type of Piece provide a means of accounting for the Syllable-final phonetic diversity of k Verbs and m Verbs, through the Juncture, Tempo, and Aspiration systems (pp. 92, 93, 94); it also provides a means of classifying prosodically those Particles which can immediately follow the Verb within the Word. Certain of those Particles have already been classified prosodically in terms of the Particle-and-Particle Quality Piece, which associates each with the following (Particle) Syllable (p. 90); they can also be classified in terms of the Juncture and Aspiration Pieces, which associate them with the preceding Syllable, whether Verb or Particle. As second Syllable bhūz, for example, is restricted to the p Juncture Piece (and is therefore characterized by initial labiality, together with plosion, friction, or nasality: b/β/m/ϕ), and to the h Aspiration Piece (whence initial aspiration or non-aspiration) (pp. 91, 93, 94); nhan' is restricted to the n Juncture Piece (whence initial alveolarity together with nasality: n) and to the h Aspiration Piece (whence initial voice); lhyan is restricted to the y Juncture Piece (initial palatality together with either friction or a non-syllabic vowel: j) and to the n Aspiration Piece (initial voice). The three Particles bhūz, nhan', and lhyan can therefore be formulated, in part, as ph, nh, and yh respectively.

Particle Syllables, except for ma (Negative) and a (Nominalizing), are necessarily in Intraverbal Junction with a preceding Syllable (Verb or Particle), e.g. makoŋ:bhūz:laː:z, is it not good, in which the Particle Syllable bhūz is preceded by the Verb Syllable koiːz, and the Particle Syllable lāːz by the Particle Syllable bhūz, and there is therefore no difficulty in providing examples in which they are
contained in disyllabic Juncture and Aspiration Pieces, in the light of which they can be given a Juncture classification, though not always an Aspiration classification (it is only in disyllabic Aspiration Pieces that are also k Quality Pieces that there are criteria for distinguishing h from h, pp. 91–2). Verb Syllables, on the other hand, are not necessarily in Intraverbal Junction with a preceding Syllable; on the contrary they are commonly Word-initial, and therefore in Interverbal Junction with the preceding Syllable, if any, e.g., swāślhyañ, if you go, in which the Verb Syllable swāś is not preceded by another Syllable within the Word. When Word-initial, Verb Syllables clearly cannot be contained in a (disyllabic) Juncture Piece or Aspiration Piece; for there is no preceding Syllable within the Word to combine with them; but Verb Syllables are not, however, exempt from classification in terms of these two systems; for there are types of Word in which another Syllable does precede them: (i), disyllabic Verb, e.g., punban-, pahpan(sañ), he gets tired out, from which pan(s) can be classified as p (Juncture), but is indeterminate as regards the Aspiration system (p. 93); (ii), doubled Verb, e.g., phrep’hje’, pyakpyak, ruinous, from which pyak can be classified as p (Juncture) and h (Aspiration); (iii), disyllabic or trisyllabic Noun, e.g., luđao-, lūsok(re), drinking-water, from which sok can be classified as s (Juncture); (iv), Negative-Particle Piece or Nominalizing-Particle Piece, e.g., mjaon-, maröns(nhañ), do not sell; ojañ, aröns, selling, from which röns can be classified as y (Juncture) and h (Aspiration).

IV. Prosodic Systems statable for Monosyllabic Pieces:

Labialization, Palatalization, Voice Quality, Yodization

The remaining four prosodic systems, termed Labialization, Palatalization, Voice Quality, and Yodization, apply to monosyllabic Pieces (or Syllables). The first of these, the Labialization system, is designed to associate particular Syllable-initial features with the appropriate vowel features; it comprises the five terms b (so named from backness), s (from spreading), f (from fronting), c (from centralized), and s (shwa). This system is dealt with elsewhere in this volume; and the present description is limited to an account of one term only, the b. The b type of Syllable has been chosen partly because it is one of the few types of Labialization Syllable for which V systems can be stated, and partly because the majority of the Verb examples introduced so far (sok, röns, pui’, etc.) are members of the b category.

The b term of the Labialization system associates lip-rounding as a vowel feature (u, o, a, oo, ao, though only partial in the case of ao) with certain other Syllable-initial features, and particularly with (i), lip-rounding (except for Syllables in ao and palatal-initial Syllables), e.g., ku kūs cross, kha kho (but òao’ sok drink, ju pru do), and (ii), potential initial glottality (‘), e.g., ə ð shout, əə’ up

1 'A Comparison of Arakanese and Burmese Based on Phonological Formulae'.

R. K. SPRIGG
cover, 'aō oû overcome. This latter feature is enough to exclude from the b
category monosyllabic Verbs in o, e.g. wō' wat wear, piō prwan is accomplished,
θo' swat put into, in spite of their having lip-rounding as a vowel feature; for
the vowel o cannot be preceded by ''; and there is therefore no Syllable-initial
sequence **'o-**.

The b term has the additional function of associating with the vowels u,
o, o, oo, and ao such Syllable-initial features as velarity (k η) and both palatalized
and non-palatalized labiality (pî mj ; p m), features that are excluded from certain
types of Labialization Syllable (for examples of the b Syllable see p. 98).

For every bm and bz Verb a further distinction, in voice quality, can be
made, and all such Syllables classified as either g (from glottal-trill) or ġ,
in accordance with a two-term Voice-Quality system; for bk Syllables, on the
other hand, no such distinction can be made. 2 The exponent of g is glottal-trill
(or 'creaky ') voice quality, e.g. pju pru do, sāō con' wait, which are invariably
characterized by glottal-trill voice quality; that of ġ is clear voice quality, except
in the first Word of certain two-Word Noun phrases, in which it is glottal-trill,
e.g. the Verb Syllable kon; in mākāō dāgaō makoṇ; takoṇ; fairly good, and mhī in
mamī tāmī mamhī(') tamhī not quite reaching.

Some b Verb Syllables have a single initial consonant, e.g. p, te, j; others have
a Syllable-initial sequence, in which a non-syllabic front vowel (or semi-vowel ; j)
or a voiceless palatal fricative (i) is associated exclusively with labiality (pî pî
mj mi). The syntagmatic association of labiality with this front, or palatal, feature,
in contrast with Syllables not characterized by any such sequence, is stated through
the two-term Yodization system (y, ĕ), the exponent of y being labiality (n m m)
and a following non-syllabic front vowel or a palatal fricative (i j), while the
exponent of ĕ is labiality/dentality/alveolarity/palatality/velarity/glottality and
an immediately following syllabic vowel (pV θV nV, etc.), e.g.

\[
\begin{align*}
y : & \quad \text{pju- phrū(sañ) is white} ; \\
\bar{y} : & \quad \text{pù- pū(sañ) is hot} ;
\end{align*}
\]

\(\bar{y}\) correlates with the Juncture components t, n, c, l, k, η, j, n, y, s, and v, and can
therefore be omitted from a formula that includes any of them (cf. also the Initial
system, p. 100).

Every b Syllable can also be further classified, by reference to the two-term
Palatalization system, as p or as ā. The exponent of the p term associates the

1 b is regularly represented in Burmese orthography by a(ı), u, ui', uī, o(ı'), and ŏ, and by
ui when syllable-final (and not, therefore, followed by k or n).
2 There is some evidence that presence or absence of syllable-initial glottal plosion in
Intraverbal Junction is to be associated with glottal-trill or with clear voice quality respectively
in the preceding Syllable, e.g. (glottal-trill), pō'oon- puitian send it, then, sāō 'oom- con'āsmañ
I shall go on waiting; cf. (clear), neoon- neüttō wait for me, then, naiaō nuinōn so as to win.
If regular, this syntagmatic association of features drawn from two Syllables would make it
necessary to state the Voice-Quality system for a disyllabic rather than a monosyllabic Piece.
following three features: (i), Syllable-initial palatality \((pj mj tc n c j)\); (ii), advanced vowel \((u o o o ao)\) as compared each with the phonetically corresponding type of vowel in a \(p\) Syllable; (iii), less rounding of vowel \((u o o o oo)\), again as compared with the phonetically corresponding types of vowel in the \(p\) Syllable. The exponent of the \(p\) term associates Syllable-initial non-palatality \((p m t n s l k n \theta h 't)\) with vowel features the reverse of those stated for \(p\) at (ii)–(iii) above \((u o o o ao)\), e.g.

\[p : pj u\] pru do, jaö roön sell; \[p : pu pü\] hot, ògo' sok drink.

\(p\) correlates with the Juncture components \(j, n, y, \) and with the Yodization component \(y, \) and can be omitted from any formula containing any of these more specific components; \(p\) can, on corresponding grounds, be dispensed with in any formula containing one of the Juncture components \(t, n, c, l, k, \eta, s, v, \) or the Yodization component \(y; \) in fact the Palatalization components \(p\) and \(\tilde{p}\) are always omissible under those conditions.

V. Phonematic Systems

The Syllable-initial and Syllable-final phonetic features of most types of Labialization Syllable are to such an extent interdependent that their phonetic features are completely accounted for by prosodic statement; and there is no need to have recourse to phonematic analysis at all (p. 80). Thus, all the components of the formulae of \(s\) Syllables, \(f\) Syllables, \(c\) Syllables, \(\epsilon m\) Syllables, and \(\epsilon k\) Syllables (p. 96) are prosodic; and the only types of Syllable that require post-prosodic (i.e. phonematic) analysis are the \(b\) and the \(az\). The examples given here are of the \(b\)-Piece, or \(b\)-Syllable, phonematic systems.

The \(V\) systems of \(b\) Syllables differ according as the \(b\) Syllable is also classifiable as \(z, m, \) or \(k\) (of the Quality system, p. 89). For \(zb\) Syllables there is a three-term system: \(U, O, Ω; \) for \(mb\) and \(kb\) Syllables the \(V\) system is two-term: \(U, O.\)\(^1\) The exponents of the terms in these \(V\) systems are:

\[
\begin{align*}
U & : \text{closeness } \quad pu \quad pūsān is hot; \quad pjū \quad prusañ does \\
O & : \text{half-closeness } \quad po \quad pu'\text{sān sends; } \quad lo \quad luisān wants \\
Ω & : \text{half-openness } \quad pjo \quad prosañ says; \quad so \quad cosañ is early \\
mU & : \text{half-closeness } \quad ŋoon \quad ŋuṅ'sañ stoops; \quad toon \quad tunsān trembles \\
O & : \text{openness } \quad jaon \quad roṅ'sañ sells; \quad saon \quad con'sañ waits \\
kJ & : \text{half-closeness } \quad loot' \quad lupsañ works; \quad joot' \quad yutsañ is inferior \\
O & : \text{openness } \quad θaot' \quad soksañ drinks \quad jaot' \quad roksañ arrives
\end{align*}
\]

\(^1\) Despite identity of symbol it is not, of course, legitimate to attempt to identify either of the two terms \(U\) and \(O\) of the two-term system (mb, kb) with the homographic terms of the three-term system (zb): they differ in commutability, and, therefore, in systemic value. Cf. Allen, ' Relationship ', p. 84: ' They are "phonological" in the sense that their systems are relevant to the structural positions for which they are established '.

PROSODIC ANALYSIS IN TIBETO-BURMAN COMPARISON

(since the V term $\Omega$ is restricted to the $z$ Syllable, the component $z$ can be omitted without ambiguity from any formula that contains $\Omega$; and $b$ can, similarly, be omitted from any formula that contains $U$, $O$, or $\Omega$, p. 82).

*Tibetan*

In the modern spoken dialects of Tibetan the comparatist has to face a problem that does not arise in Burmese, the problem of grammatically distinct forms for certain Verbs. For some Tibetan Verbs two or even three such forms have to be distinguished; and Tibetan Verbs of this type thus offer the comparatist a choice of grammatical forms, Present, Perfect, and in some cases Imperative, as well as a choice of phonetic variants of a single grammatical form.

Shafer solves this problem with regard to the verbal forms of classical Tibetan (his Old Bodish), in which as many as four forms are distinguished for some verbs, e.g. *skem*, *bskams*, *bskam*, *skom*, dry, by taking as basic either the perfect root itself or the perfect root modified by removal of the $s$ suffix ¹; but for the dialects of spoken Tibetan the aim should, in my opinion, still be to set up for each Verb through prosodic analysis a lexical-item phonological formula that shall subsume all the variant phonetic forms, regardless of whether these variants do or do not reflect differences of grammatical category.

I. Lhasa Dialect

This is not the place for a detailed account of the various prosodic systems that make it possible to construct such a formula; but I will briefly illustrate some of these systems, first from the Lhasa dialect, and principally from the following three Verbs: (a), *srub/srubs*, churn; (b), *bzo/bzos*, make; (c), *bltaj/bltas/ltos*, look at.²

The complete lexical-item phonological formulae for these three Verbs (with omissible components enclosed in brackets; p. 82) are:

(a) (h1)1wc$pz$Sz so/sy:  
(b) (h1)2wo$pb$Sz so/so/sb:  
(c) (o$p$1r)1$h_b$Tz ta/ta/te: (p. 82).

The components 1 and 2 of these formulae refer to the terms of a two-term Tone system applicable to the Verb, or to the Verb-and-Particle, Word: Tone-1 Verbs are exemplified only in Tone-1 Words, and Tone-2 Verbs only in Tone-2 Words.³

¹ e.g. ‘Newari and Sino-Tibetan’, *Studia Linguistica*, 6, 1952, p. 95.  
³ For the orthographic problem presented by the modern spoken Tibetan dialects see ‘Verbal Phrases in Lhasa Tibetan—I’, *BSOAS*, 16, 1, 1954, p. 135, n. 1.
a and w are terms of a three-term Labialization system (a, w, y) that unites rounding-spreading, and such associated features as velarity, with the appropriate vowels (rounded, non-rounded) of the Verb Syllable, and, in the case of certain Particle Syllables, extends to the following Syllable, e.g. w, soɣry- srub-kyi-red he will churn; a, taga-bla-gi-red he will look at; y, suɣi- zer-gyi-red he will say. In these three examples the Particle Syllable kyi/gyi/gi is characterized by either lip-rounding (ɣy), lip-spreading + frontness (git), or lip-spreading + non-frontness ([..]) according as it is in a w, a y, or a a disyllabic Piece (cf. also Burmese, pp. 96–7).

p (non-ŋ) refers to the Palatalization system, the two terms of which (p, ŋ) relate Syllable-initial palatality or non-palatality to relative frontness or backness of vowel respectively (cf. Burmese, pp. 97–8). Since the phonematic unit ŧ is restricted to p Syllables, ŋ can be omitted from formulae that contain T.

h and ŋ (non-h) refer to the two terms of the Aspiration system, which associates partial voicelessness, or complete voicing, of vowel (ŋ, v) with the appropriate Syllable-initial consonant features (cf. Burmese, pp. 91–95). Ń implies ŋ.

b, and the component z preceding one of the phonematic components (example (a) only), are terms of the Initial system, a three-term prosodic system (b, n, z) established to account for Syllable-initial single consonants and Syllable-initial sequences in Intraverbal Junction (for Intraverbal Junction see p. 90), and in this instance the Syllable-initial sequence ps that characterizes bzo/bzos in e.g. nɔpsɔ; dngo-bzo-ba, shoemaker, in Intraverbal Junction (cf. the Syllable-initial single consonant s in Interveral Junction, e.g. sɔdži; bzo-rtsis, intending to make, and the other examples at (b) on p. 101 below), and the Syllable-initial sequence pt that characterizes blta/bltas/ltos in Intraverbal Junction in e.g. soptə, bzo-blta, appearance (cf. Syllable-initial t in Intraverbal Junction in the examples at (c) on p. 101 below). There is no evidence of anything other than a Syllable-initial single consonant s for srub/srubs in Intraverbal Junction, whence its classification as z (from zero).

The component z that follows the phonematic components (S, S, T), on the other hand, is a term of an eight-term Quality system (z, g, ŋ, d, n, b, m, r) that associates the vocalic features of Verb Syllables with their final consonantal features, if any, and, further, with both consonantal and vocalic features of certain following Particle Syllables (cf. Burmese, pp. 85–88). The following examples illustrate the sort of sequences of syntagmatically associated features that provide the justification for this system; these examples are all disyllabic Verb-and-Particle Pieces in which the Verb Syllable is w (and therefore characterized by lip-rounding; above) and in which the Particle category is represented by pa/ba (Nominalizing):

---

1 See also 'Verbal Phrases', BSOAS, 16, 2, pp. 320–2, 338–9; BSOAS, 16, 3, pp. 566–8.
These examples show the interdependence of (i), a closer central vowel (ə) with preceding labiality (b), in the d, n, b, or m Piece, but a more open central vowel (u) with non-labiality (g η r), in the g, η, or r; (ii), long vowel duration (ə:), and velar plosion (g), in the g Piece, as compared with short vowel duration (u η) and nasality (η m), or friction (r), in the η, n, m, or r, and either long or short vowel duration (ə: η) with labial plosion (b), in the d or the b; (iii), backness (a: η) with velar plosion (g), in the g or in the η Piece, and also with absence of intervocalic consonant (a:η), in the z, as compared with either backness (v) or frontness (a) with labiality or alveolarity (b m r), in the d, n, b, m, or r; (iv), half-openness (a:) with velar plosion (g) and with absence of intervocalic consonant, in the g or the z Piece, but openness (v) with velar nasality (η), with labiality (b m), and with alveolarity (r).

All three of the Verbs chosen to illustrate lexical-item phonological formulae in Lhasa Tibetan ((a) srub/srubs; (b) bzo/bzos; (c) blta/bltas/ltos; p. 99) are exemplified in Verb-and-Particle Pieces only of the z type. z-Piece, or z, Verbs have been chosen because of their relative complexity: the majority of them have at least as wide a range of variation in phonetic form as the g-Piece, η-Piece, or other types of Verb, and also differ from them in generally requiring more than one grammatical form to be established (the g, the η, and the remaining non-z types of Verb, except for some of those which are also classified as ə—p. 102—are single-form).

The variant phonetic forms of the three Verbs are:

(a) so sy: ; (b) so/so so: ; (c) ta/ta te:

the first of each pair of phonetic variants is appropriate, for example, to Present and Future Clauses (columns (i)–(ii) below), except for Imperative-Affirmative, while the second is appropriate to Past Clauses (col. (iii)), and to Imperative-Affirmative Clauses (col. (iv)) except for ə Verbs ((c) (iv); p. 102 below) (crucial features are given in brackets immediately after the example from which each has been abstracted):

(i) (ii) (iii) (iv)
(a) sogyre: (so-) sobare (so-) sy:bare (sy:-) sy:
(b) sodzi: (,) sobare (so-) sobare (so:-) so:
(c) tagdu: (ta-) tabare (ta-) te:bare (te:-)

1 The n and the m examples are not distinguished by syntagmatically associated features in the Verb-and-Nominalizing-Particle type of Piece; but they are in Pieces in which the Particle category is represented by song, e.g. (n) khö:s5, gon-song, he wore; (m), zoms5, 'dzoms-song, they assembled: the n Piece is distinguished from the m by long vowel duration, frontness, and absence of oral occlusion (-5:s- ; cf. -dms-).
R. K. SPRIGG

((a) srub-kyi-red they will churn, srub-pa-red they churn, srubs-pa-red they churned, srubs churn; (b) bzo-rtsis intending to make, bzo-ba-red they make, bzos-pa-red they made, bzos make; (c) blta-gi-'dug they are looking, blta-ba-red they look, bltas-pa-red they looked).

As between columns (i) and (ii) example (a) does not vary (so-); but (b) and (c) show a variation in degree of aperture between a closer degree, in (i) ((b) o ; (c) a), and a more open degree, in (ii) ((b) o ; (c) a), according as each is contained in a c (from close) or in an o (from open) disyllabic Closure Piece, in harmony with vowel features of the following Syllable (c-Piece : y i: t ; o-Piece : ø).¹ This phonetic variation in examples (b) and (c), and Verbs like them, can be formulated as c/o, i.e. exemplifiable in either the c or the o disyllabic Piece, as distinct from c Verbs, like type (a), which can be exemplified in the c disyllabic Closure Piece but not the o.

As between columns (i)/(ii) and (iii)/(iv) the difference in duration ((a) o v. y ; (b) o/o v. ø ; (c) a/a v. e;), in association with a backness-frontness difference (the front vowel a of example (c) (ii) is backer than the front vowel e of (c) (iii) and could be represented in a more detailed transcription as a), can be accounted for by setting up a Quantity system comprising s (from short : (a) o ; (b) o/o ; (c) a/a) and l (from long : (a) y ; (b) ø ; (c) e;), whence s/l Verbs.

The phonological formulae given above for srub/srubs and bzo/bzos (p. 99, (a) and (b)), and their variant phonetic forms, have now been accounted for; but there remains one of the forms of example (c) blta/bltas/ltos ; te:(-), confined to Imperative-Affirmative Clauses. Certain o-Piece Verbs, of all eight Quality-Piece types (z, g, η, d, n, b, m, r), except that there is only one b-Piece example, have Imperative forms that are characterized by lip-rounding, to account for which at the phonological level of analysis a Rounding system is established comprising the two terms r (from rounding) and ð (non-r), e.g. : 

\[
\begin{array}{ll}
\text{r} & \text{ð} \\
\text{æz : tæ} & \text{tæ/ta} & \text{te} : \text{blta/bltas/ltos} & \text{(oy)l sr/ð} \text{lqipTz look} \\
\text{œd : do} & \text{de/de} & \text{bsdad/sdod} & \text{(oy)r/ð\text{æh}zTd stay} \\
\text{or : tœo} & \text{tœa/tœa (tœar)} & \text{bcar/gcor} & \text{(oy)r/ð\text{æh}zJr visit}
\end{array}
\]

The only phonematic systems in Lhasa Tibetan are C systems ; there are no V systems (cf. Burmese, in which there are V systems but no C systems ; pp. 98–9).

One of the results of classifying Verbs in terms of the Quantity, Closure, and Rounding systems (above) has been the introduction of trigraphic components into the phonological formulae : certain Verbs have been classified as c/o, as opposed to c (Closure system), as s/l as opposed to s (Quantity system), and as r/ð as opposed to ð (Rounding system), according as each could or could

¹ For a detailed study see ‘Vowel Harmony’. 
not be exemplified in the disyllabic o Piece as well as the c, as an 1 Syllable as well as an s, and as an r Syllable as well as a ī. The trigraphic components c/o, s/₁, and r/ī are clearly more cumbersome than a monographic component would be; and, in particular, the role of the oblique stroke, which associates the immediately preceding component with the immediately following component as alternatives, might mistakenly be understood as signifying that all the components preceding the oblique stroke were alternative to all those following it. Since, however, only two classes of Verb are distinguished by the c/o and the c, the s/₁ and the s, and the r/ī and the ī components, i.e. the c-Piece-only, or c, class for example, as opposed to the non-c-Piece-only class (hitherto c/o), or the s-Syllable-only, or s, class as opposed to the non-s-Syllable-only class (hitherto s/₁), and similarly with the ī class of Verb as opposed to the r/ī, each of the cumbersome trigraphic components could without ambiguity be replaced by a monographic component formulating the potentialities of the alternative class. Thus, either c (non-c) or o, for example, would serve the purpose of distinguishing what has hitherto been termed the c/o Verb class from its alternative, the c; but o has in fact been adopted here as being typographically more convenient than c.

The component o in the phonological formula of a Verb therefore indicates that that Verb can be exemplified not only in the o Piece but in both the o Piece and the c Piece equally, e.g. the o Verb bzo/bzos (formerly c/o) is exemplified in the c Piece sodzi; bzo-rtsis (p. 101, (b) (i)) as well as in the o Pieces sōba- and sō:ba- (bzo-ba and bzos-pa; p. 101, (b) (ii)-(iii)). By the same principle s/₁ is from this point onwards replaced by ₁, and r/ī by r, with ₁ signifying the possibility of being exemplified as either an ₁ or an s Syllable, and r as either an r or a r Syllable. In this way the formulae on p. 102 are simplified to (oŷ)₁rlghpTz, (oŷ)r₂ghzTd, and (oŷ)r₁hzhJr (the r component preceding each phonematic component refers to the r term of the two-term Rounding system, and the r following it, in the third formula, to the r term of the eight-term Quality system).

In certain formulae o (formerly c/o) could be dispensed with: it is implied by the component a, except only in the case of a Verbs that are also classified as b, s Verbs, e.g. bslabs teach, lab tell, rgyab place. When conjoined with z, g, η, d, n, m, or r, a therefore implies o; all az, ag, aη, ad, an, am, and ar Verbs are also o (formerly c/o), e.g. the azo Verb blta/bltas/ltos, p. 101 (c), col. (i): c Piece; cols. (ii)-(iii): o Piece. It is for this reason that the component o has been included in brackets in the formulae of the three a Verbs (az, ad, ar) on p. 102.

It is only for z Verbs (p. 100) that the Quantity system (p. 102) needs to be established; but, though a clear majority of z Verbs are classified as ₁ (formerly s/₁), there is not here the correlation that there is between a and o ¹; nor, even though a majority of a Verbs have r Imperative forms, does the component r

¹ In the available material twenty-six z Verbs out of thirty-five are classified as ₁ (formerly s/₁); the remaining nine Verbs, s Verbs, include, for example, at shi die, tche/tche phyə open, na/na na ill, and kho/kho go hear (there are no forms *siː, *tcheː, *neː, *khoː).
(formerly r/ī) correlate with s to the exclusion of r.¹ 1 (formerly s/l) and r (formerly r/ī) cannot therefore be omitted from phonological formulae on the same grounds as o is omitted from formulae containing s (except əb) (p. 103) : that they are without exception implied by another component.

Nevertheless, it is possible that the components s and 1, and r and ɨ, referring to the terms of the Quantity system and the Rounding system respectively, might be omitted from the sort of phonological formula that is suitable for language comparison on other grounds. There is no question of disregarding the sub-classification of ə Verbs in terms of the Quantity system, as s or 1, or the sub-classification of s Verbs in terms of the Rounding system, as r or ɨ (formerly r/ī), at the phonological and grammatical levels : no grammatical statement of the Verbal Phrase could be made without reference to these phonological sub-categories, which indeed provide phonological exponents of certain grammatical categories (p. 99) ; but these two prosodic sub-categories seem less important in the context of language comparison, and not as relevant to Tibeto-Burman comparison as the major prosodic categories, the Quality system (z, g, ə, d, n, b, m, r ; pp. 100–1) and the Labialization system (y, w, s; p. 100), within which the Quantity and Rounding systems function. For a ə Verb (e.g. shi die, blta/ bltas/ltos look) the component z is indispensable in comparison ; but the sub-classification of that Verb in terms of the Quantity system appears not to be indispensable. The fact that e.g. srub/srubs and bzo/bzos have the two grammatically distinct phonetic forms so or sy; and so/sə or so: respectively (p. 101), while e.g. phye and na have only the forms təh/teh and na/na respectively, without grammatical distinction (p. 103, n. 1), seems in the light of experience to date to be of little consequence for comparison ; what is important is that all four Verbs are z, not, for example, g, or ə ; and their further sub-classification into s or 1 can, it would seem, be treated as a subsidiary implication of z : z implies sub-classification into s or 1 ; the non-z terms g, ə, etc., do not.

Similarly, the fact that such s Verbs as blta/bltas/ltos and bsdad/sdod have each the grammatically distinct phonetic forms əo: (Imperative) or tə/ta and tə: (non-Imperative), and də: (Imperative) or də:/də: (non-Imperative) respectively (p. 102), while e.g. the ə Verbs 'jag (settle) and gnang (grant) have only the forms də:/də:(g) and nən/nən/nə: respectively, with no distinction of Imperative from non-Imperative form, also appears not to be significant for language comparison ; the significant thing for all four Verbs is that they are ə Verbs, not y or w ; and further classification in terms of the Rounding system, with a consequential

¹ Fifty-four out of seventy-eight ə Verbs in the available material have an r form (Imperative), though my informant was prepared to accept a ɨ alternative form for twenty-three of these fifty-four Verbs, e.g. rnga/brngas/rngos reap : (r) nə:, (f) na ; bzhag/zhog put : (r) əo:, (f) əa ; bkaŋs/khongs fill : (f) kə:, (r) kə; ; bsdad/bsdod kill : (r) əo:, (f) əe ; btab/thob sow : (r) təp, (r) təp ; bskams/skoms dry : (f) kəm, (r) kom. He preferred the ɨ Imperative form to the r in eleven of these twenty-three examples (the alternative Imperative forms given above are in order of preference).
distinction at the grammatical level of analysis, is an exclusive characteristic of the œ Verb class.

It is because of the assumption that the Quantity components s and l, and the Rounding components r and Ʌ, are potentialities of z and of œ respectively that they have been included in brackets in the complete lexical-item phonological formulae (p. 99) as omissible. To include them would be to make that type of phonological formula more detailed than seems necessary for language comparison.

II. Sherpa Dialect

The method just illustrated for dealing with the problem presented to the comparatist by grammatically distinct phonological forms of the same lexical item in Lhasa Tibetan is open to the objection that whatever success it has achieved can be attributed not to its own merit but to choice of dialect; for it is well known that fewer grammatical forms ('roots') need to be distinguished for Lhasa-Tibetan Verbs than for those of certain other spoken-Tibetan dialects, and for classical Tibetan; but in fact the method proposed here has been found equally applicable to a grammatically more complex spoken-Tibetan dialect, Sherpa, which more closely resembles classical Tibetan in diversity of grammatical forms than does the Lhasa dialect.1

In certain respects the Sherpa accords with the Lhasa dialect: (i), a two-term Quantity system (s, l) needs to be established in order to deal with alternative duration features of Sherpa œ Verbs; e.g. (s) la v. (l) la:, la:; or (s) zo v. (l) zo:, as in

<table>
<thead>
<tr>
<th>s</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>la (la-)</td>
</tr>
<tr>
<td>(ii)</td>
<td>zo (zo-)</td>
</tr>
</tbody>
</table>

(the s form is appropriate to Present and to Future Clauses, the l form, or forms, to Past and to Imperative-Affirmative Clauses); cf. Lhasa dialect:

<table>
<thead>
<tr>
<th>s</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s)</td>
<td>ta/ta v. (l) te:, te:</td>
</tr>
<tr>
<td>(,)</td>
<td>so/so v. (,) so:</td>
</tr>
</tbody>
</table>

1 'The student is warned against using the different roots for the present, future, perfect, and imperative given in general dictionaries. These hold good only for the literary language and could often lead him astray in the [Lhasa] colloquial' (C. A. Bell, Grammar of Colloquial Tibetan, Alipore, 1939).

2 The Sherpa Quality system, which discharges much the same prosodic functions as the Lhasa (pp. 100–1), comprises not eight but ten terms (z, g, n, d, n, b, m, r, l, s), of which two (d, n) are not represented in the disyllabic œ Piece. The terms of the Lhasa and Sherpa Quality systems regularly correspond as follows:

<table>
<thead>
<tr>
<th>Sherpa:</th>
<th>z g n d n b m r l s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lhasa:</td>
<td>z g n d n b m r r d</td>
</tr>
</tbody>
</table>

The Sherpa examples are of the Khumbu dialect; my informant was Tenzing Lodroo (bstan-'dzin blo-gros), a young monk of Thyangboche monastery, born in Namche Bazaar.
and, (ii), a two-term Rounding system (r, \( \check{r} \)) needs to be established in order to deal with alternative labial features of the vast majority of Sherpa \( \sigma \) Verbs; e.g. (\( \check{r} \)) \( \text{la}, \text{la}: \) v. (r) \( \text{lo}, \) or (\( \check{r} \)) \text{teor} v. (r) \text{teor}, as in

\[
\begin{align*}
\text{(i) } & \text{la: (la:-) I shall look} \quad \text{lo: (la:-) I looked} \quad \text{la: look} \\
\text{(ii) } & \text{teor (teor:-) \,, \,, visit} \quad \text{teor visit}
\end{align*}
\]

(the r form is appropriate to Imperative-Affirmative Clauses, the \( \check{r} \) forms to all other types of Clause); cf. Lhasa dialect:

\[
\begin{align*}
\text{(r) } & \text{ta/ta, te: v. (r) te: } \text{blta, bltas v. ltos look} \\
\text{,} & \text{teor/tear v. ,} \text{teo: bcar v. gcor visit (p. 102).}^1
\end{align*}
\]

The Sherpa differs from the Lhasa dialect in that for certain \( \sigma \) Verbs, further classifiable as g, \( \eta \), b, m, r, and l in terms of the Sherpa Quality system (p. 105, n. 2), two grammatically distinct \( \check{r} \) forms (non-Imperative) need to be recognized, one of them appropriate to Present and to Future Clauses (excluding Imperative-Affirmative) and the other to Past Clauses. Not all Sherpa \( \sigma \) Verbs have two \( \check{r} \) forms; but for the sake of those which do, two further prosodic systems are established, the Neutrality, and the Fronting.

The two-term Neutrality system (n, \( \check{n} \)) is needed to deal with an alternation of labial features, lip-rounding (\( \check{n} \)) as opposed to a non-rounded (or neutral) lip position (n); e.g. (\( \check{n} \)) \text{zog} v. (n) \text{za:}, or (\( \check{n} \)) \text{tong} v. (n) \text{tang}, as in

\[
\begin{align*}
\text{(i) } & \text{zogt (zog-) I shall put} \quad \text{za:qi (za:-) I put} \\
\text{(ii) } & \text{tongt (tong-) \,, \,, send} \quad \text{tangi (tang-) \,, sent}
\end{align*}
\]

(the \( \check{n} \) form is appropriate to Present and to Future Clauses, the n to Past Clauses); the number of Verbs to which this system applies is twenty-eight.

The two-term Fronting system (f, \( \check{f} \)) applies to nine Verbs of types \( \sigma \eta \), \( \sigma \beta \), \( \sigma \mu \), and \( \sigma \lambda \) (p. 105, n. 2), and is concerned with an alternation of frontness (f) with backness (\( \check{f} \)); e.g. (f) \text{deb} v. (\( \check{f} \)) \text{t\( \check{a}\)s\( \check{b} \)}, or (f) \text{gjem} v. (\( \check{f} \)) \text{kam}, as in

\[
\begin{align*}
\text{(i) } & \text{debqunk (deb-) he sows} \quad \text{t\( \check{a}\)bs\( \check{b} \) (t\( \check{a}\)s-) he sowed} \\
\text{(ii) } & \text{gjembi: (gjem-) I shall dry} \quad \text{kamb: (kam-) I dried}
\end{align*}
\]

1 Of sixty Sherpa \( \sigma \) Verbs fifty-nine have both r and \( \check{r} \) forms. The sole exception is the \( \sigma \eta \) Verb \text{nak}, grant (cf. the Lhasa dialect, pp. 103-4).

2 The r forms (Imperative) of these two Verbs are: (i), \text{zdk}; (ii), t\( \check{a}\); the corresponding classical-Tibetan forms are:

\[
\begin{align*}
\text{(i) } & \text{Present 'jog, Past bzhag, Imperative zhog} \\
\text{(ii) } & \text{, , gtong , , btang , , tong}
\end{align*}
\]
Thus, to classify a Sherpa Verb as \( \sigma \) carries with it the implication of further classification in terms of the Quantity system for \( \sigma z \) Verbs (p. 105) and the Rounding system for \( \sigma z, \sigma g, \sigma \eta, \sigma d, \sigma n, \sigma b, \sigma m, \sigma r, \sigma l \) and \( \sigma z \) Verbs alike, with further sub-classification of the \( \dot{f} \) forms (non-Imperative) in terms of the Neutrality and Fronting systems as follows:

1. **Neutrality** (\( n, \tilde{n} \)):
   - \( \sigma g, \sigma \eta, \sigma b, \sigma m, \sigma r, \sigma l \);

2. **Fronting** (\( f, \tilde{f} \)):
   - , , , , ,

3. **Neither**:
   - , , , , , \( \sigma s \)

but, again, these sub-classifications, though highly significant for grammatical statement, seem to be of no account in language comparison. It is enough that all the Sherpa Verbs in the preceding two paragraphs are classifiable as \( \sigma \); there is no need to formulate their sub-classes in lexical-item phonological formulae.

The examples of \( f \) and \( \tilde{f} \) forms (p. 106) do more than show a difference of frontness versus backness; for these \( f \) and \( \tilde{f} \) forms also differ from each other in other respects: the \( f \)-form examples are also examples of Tone-2 Words, with initial low pitch, and of the associated complex of Word-initial features voice + occlusion (\( \dot{d}, g \)), while the \( \tilde{f} \)-form examples are also examples of Tone-1 Words, with initial high pitch, and of the associated complex of Word-initial features voicelessness + occlusion + non-aspiration (\( t, k \)). In fact the two forms of each Verb differ not only in their Fronting classification, as \( f \) versus \( \tilde{f} \), but also in their Tone classification, one form being a Tone-1, and the other a Tone-2. The two Verbs in these examples, and other Verbs that resemble them in this tonal alternation, including the associated alternation in voice-voicelessness, can therefore be classified as \( 1/2 \); and it is not only \( \sigma \) Verbs that can be classified as \( 1/2 \) but \( \sigma, y, \) and \( w \) Verbs equally; e.g.

\[
\begin{align*}
\text{Tone 2 (voice)} & \quad \text{Tone 1 (voicelessness, non-asp.)} \\
(i) \sigma : \_\text{\textit{ph}u} & \quad \overset{\text{\textit{k}}}{\text{\textit{ph}u}} \quad \text{I shall load} \quad \text{I loaded} \\
(ii) y : \_\text{\textit{b}u\textit{di}} & \quad \_\text{\textit{p}u\textit{di}} \quad \text{take off} \quad \text{took off} \\
(iii) w : \_\text{\textit{b}o\textit{g}u} & \quad \_\text{\textit{pu\textit{w}u}} \quad \text{pierce} \quad \text{pierced}
\end{align*}
\]

(the Tone-2-Word form is appropriate to Present and to Future Clauses, the Tone-1 to Past Clauses).²

¹ The \( \text{\texttt{r}} \) forms (Imperative) of these two Verbs are: (i), \texttt{top}, (ii), \texttt{kwom}. The corresponding classical-Tibetan forms are:

\[
\begin{align*}
(i) \text{Present} & \ '\text{debs}, \ \text{Past} & \ '\text{btab}, \ \text{Imperative} & \ '\text{thob} \\
& \ '\text{skem}, & \ '\text{bskams}, & \ '\text{skom}.
\end{align*}
\]

² cf. classical Tibetan:

\[
\begin{align*}
(i) \text{Present root} & \ '\text{gel}, \ \text{Perfect} & \ '\text{bkal} \\
& \ '\text{bud}, & \ '\text{phud} \\
(iii) & \ '\text{bug} & \ '\text{phugs}
\end{align*}
\]
By reference to the two-term Tone system (1, 2) three classes of Sherpa Verb can be distinguished: (i), Tone-1 only; (ii), Tone-2 only; (iii), alternating tonally as above (1/2). In this case, there being three categories, there is no means of simplifying the trigraphic component 1/2 (cf. pp. 102–3); and the three categories would have to be formulated as 1, 2, and 1/2, if, that is, the relationship of component to prosodic term is to be preserved.

The greater grammatical and phonological complexity of Sherpa as compared with Lhasa Tibetan is reflected in two additional prosodic systems (Neutrality, Fronting; pp. 106–7) and in the three Tone components 1, 2, 1/2 (cf. the two Lhasa components 1 and 2; p. 99); but this difference between the two dialects does not affect the efficacy of the method of language comparison proposed here, comparison on the basis of regular correspondences between components of the lexical-item phonological formulae established for each language or dialect by prosodic analysis.
A COMPARISON OF ARAKANESE AND BURMESE BASED ON PHONOLOGICAL FORMULAE

By R. K. SPRIGG

INTRODUCTORY

In ‘Prosodic Analysis, and Phonological Formulae, in Tibeto-Burman Linguistic Comparison’ (pp. 79–108) I stated my belief that lexical-item phonological formulae, arrived at by applying to each language or dialect the techniques of prosodic analysis, would provide the most satisfactory forms for the linguistic comparison of Tibeto-Burman languages (p. 79); in this paper I attempt to prove my contention by using Arakanese and Burmese lexical-item phonological formulae as a basis for constructing formulae for Arakanese-Burmese (A.-B.), a first step towards formulae for what Shafer has termed the Southern Unit of the Burmese Branch, comprising Burmese, Arakanese, Tavoyan, Taungyo, Intha, Danu, and Yaw, of which only Burmese is now a written language.¹

Since Verbs are less susceptible to borrowing than Nouns, I have restricted myself to Arakanese and Burmese Verbs, with a further restriction to monosyllabic Verbs.² My material comprises 1,002 Burmese, and 667 Arakanese, phonologically distinct Verbs; the former are a complete inventory of such Verbs for one speaker.³ The corpus of Arakanese examples is much smaller than the Burmese because, having elicited an example of a Verb with an aspirate initial, I did not always inquire for an otherwise phonetically identical Verb with an unaspirated initial, and vice versa.

For a very few of the Arakanese Verbs no regular correspondence can be shown with a Spoken-Burmese Verb, e.g. ʔett throw a shadow, ʔeit shiver, ʔet string [a necklace].⁴

MAJOR PROSODIC SYSTEMS

There are four major prosodic systems in terms of which Arakanese and Burmese monosyllabic Verbs can be classified: Tone, Quality, Voice Quality, Labialization.

In both Arakanese (A.) and Burmese (B.) Verb Words (comprising a Verb Syllable only) and Verb-and-Particle Words (comprising a Verb Syllable and one or more Particle Syllables) can be classified in accordance with differences

---

² When used as formally established phonological or grammatical terms, Verb, Word, Syllable, etc., are distinguished by a capital letter.
³ My informant for the Burmese material was U Tin Maung, an Upper Burman, from Sagaing, and for the Arakanese Saw Hla Pru, an aircraft-engineering student, from Akyab.
⁴ Verbs are cited in the phonetic form appropriate to junction with the Particle Syllable re/te (A.) or te/de (B., san).
in pitch behaviour as Tone-1 or Tone-2 Words, e.g. (r symbolizes a fricative, i.e. the r of the I.P.A.):

Tone 1 (‘—, etc.) : A. pōre thaibre lootte ; B. póde thands lootdē ;
Tone 2 (—, etc.) : A. pore thaibre ; B. podē thande

(pui’sañ sends ; thiün’sañ is damp ; lupsañ works ; puisañ exceeds ; thiün’sañ sits). ¹

Those monosyllabic Verbs which can be exemplified only in Tone-1 Words are therefore classified as Tone-Word-1, or Tone-1, Verbs, while those which are exemplified only in Tone-2 Words are, correspondingly, classified as Tone-2. On the basis of identity, or close similarity, of translation-meaning of particular Arakanese and Burmese lexical items, supported by regular correspondences between the components of their phonological formulae, the Tone components 1 and 2 of Burmese can be shown to correspond to 1 and 2 respectively of Arakanese, whence *1 and *2 can be established for A.-B., with 1 as the reflex of *1 in either language, and 2 of *2.

The Quality systems of both Arakanese and Burmese are three-term : z, m, k. ² The exponents of these three terms are drawn from features of both the Verb Syllable and the following (Particle) Syllable, e.g. (f = Fast Tempo ; s = Slow):

(i) (ii) (i) (ii)

z : A. pro:re pru:re: B. pio:de pju:de:
   {pio:re pju:re: (f)

m : A. sā:re kho:re: B. sā:nds kho:nds:
   {sā:nds kho:nds: (f)

k : A. kṛaotte tshootte B. tsha:nds tshootdā
   {tsha:nds tshootdā

(prosañ speaks ; phrūsañ is white ; con’sañ waits ; khunsañ jumps ; khroksañ fears ; khyupsañ sews).

The exponents of z, m, and k are stated in full in my other paper (pp. 87–8); the above examples are enough to show that they comprise syntagmatically related Syllable-final features of the Verb Syllable (vowel quality, nasality versus non-nasality, presence versus absence of final stop) and Syllable-initial features of the following Syllable (plosion versus flap or friction, nasality v. non-nasality, voice v. voicelessness).

Verbs are classified as z-Piece (or z), m-Piece (or m), or k-Piece (or k) according as they are exemplified in the z, the m, or the k, disyllabic Piece (‘Prosodic Analysis’, p. 89).

¹ For Words in which the Verb is disyllabic, see ‘Junction in Spoken Burmese’, Studies in Linguistic Analysis, Oxford, 1957.
² 'Prosodic Analysis,' pp. 85–8.
There is a regular correspondence of each of the terms z, m, and k of the Burmese system to z, m, and k respectively of the Arakanese, whence *z, *m, *k can be established for A.-B., with reflexes z, m, and k respectively in either language.

A z or an m Verb can be either Tone-1 or Tone-2 in either language, but a k Verb must also be Tone-1. In phonological formulae the component k therefore implies the component 1; and an economy can be achieved in the Arakanese and Burmese, and in the A.-B., formulae, through a convention whereby k does duty for the component 1 (in A.-B., *1).

The third major prosodic system is two-term: g (from glottal-trill), ĝ. A g Syllable is invariably characterized by glottal-trill (or 'creaky') voice quality, a ĝ almost invariably by clear voice quality, e.g.:

$\begin{align*}
g & : A. z' \text{pö} ; m' \text{prao} ; \\
ĝ & : A. z' \text{po} ; m' \text{praö} ; k \text{paot} ; B. z' \text{po} ; m' \text{pjaon} ; k \text{paot}
\end{align*}$

(po' is light; phröñ' is straight. po is plentiful; pröns moves; pok is pierced).

There is a regular correspondence of g and ĝ in Burmese with g and ĝ respectively in Arakanese, whence *g and *ĝ can be established for A.-B., with reflexes g and ĝ respectively in either language.

All Tone-2-Word Verbs are also ĝ, and so also are all Tone-1-Word k Verbs (lk); since, therefore, the components 2 and k imply ĝ, the component ĝ can be omitted from any formula containing either of them. Indeed, it is possible to achieve further economies in symbolization: since every Tone-1 z or m Verb (lz, lm) must also be further classified as either g or ĝ (lzg, lzĝ; lmg, lmg), a similar convention to that whereby the component k does duty for the Tone component 1 in both Arakanese and Burmese, and in A.-B., formulae, can also be adopted for g and ĝ, and the component 1 omitted from any formula containing either of them.

Thus far, there has been a term-for-term correspondence between the terms of Arakanese and Burmese systems (Tone, Quality, Voice Quality); in the case of the fourth major prosodic system to be considered, the Labialization, this identity in number of terms as between the two languages breaks down. The Burmese Labialization system comprises the five terms s (from spread), c (from centralized), a, f (from fronting), and b (from back); the Arakanese comprises the four terms s, a, f, and b. Since all Burmese b monosyllabic Verbs regularly correspond to Arakanese b monosyllabic Verbs, and vice versa, whence A.-B. *b, b-Syllable correspondences are considered first.

1 'Prosodic Analysis,' p. 97, above.
2 The only non-correspondence to be noted was of Arakanese ĝ in 'oö, is overcast, with Burmese g, 'oon um'.
3 'Prosodic Analysis,' p. 96.
Burmese b Syllables can be distinguished from Burmese s, f, c, and ə Syllables, and Arakanese b Syllables from Arakanese s, f, and ə Syllables, by the following, among other features: (i), complete or partial rounding of vowel (u, o, ə, a), and (ii), rounding of initial consonant if non-palatal (except Syllables in ə), e.g. A., khoõ jumps, jə is reduced; B., khoat folds khok, əon bows nun”, except for one sub-category of ə Syllable, in either language, that is characterized, like the b Syllable, by rounding of vowel, e.g. B., wat wears wat, əon harrow thwan; A., wō overflow, kwnt is spotted; from this sub-category of ə Syllable the b Syllable is distinguished in either language by potential initial glottal plosion (‘), e.g. A., ’u lays an egg, ’aə passes (an examination); B., ’oon swarms round un, ’aot is musty ok (*’o- is not a possible Burmese, or *’on- a possible Arakanese, sequence).

Burmese

Burmese b Syllables can be further classified in terms of three prosodic systems: Juncture, Aspiration, Palatalization.

The fourteen-term Juncture system (p, m, t, n, c, l, w, k, ọ, j, n, y, s, v) and the two-term Aspiration system (h, ʰ) are set up in order to be able to associate initial localization and posture features of the Verb Syllable (labiality and plosion, labiality and nasality, dentality and friction, etc.) with the corresponding final features of Syllables preceding them in Intraverbal Junction, i.e. within the same Word, e.g. -k'g- and -ŋg- in the two k-Juncture Piece examples sokkoны (θəoʊk'gaʊ) soklim'maŋ he may well drink, koŋkoŋ (kæʊŋgaʊ) well, and -t' th- in the t-Juncture Piece lakthok (le(t')θaʊ') chará assistant.¹

The two-term Aspiration system (h, ʰ) is also set up in order to associate voice or voicelessness as a feature of the vowel of the Verb Syllable (v; əv, or h), or, in the case of non-occlusive-initial Syllables, as a feature of the initial consonant (m versus m, l versus l, etc.), with the corresponding features of a preceding Syllable in Intraverbal Junction, e.g. h, -v(t')th-; ʰ, -k'g-; as in le(t')θaʊ’ and θəoʊk'gaʊ respectively in the immediately preceding paragraph.²

Each b Verb is then classified as e.g. p, m, or s, or as h or ʰ, according to which type of Juncture Piece or Aspiration Piece it can be exemplified in (as second Syllable of the Piece) except that no b Syllables can be classified as w (p. 113).³

The two-term Palatalization system (y, ɨ) needs to be stated in order to account for syntagmatically related features of the initial consonant or consonants of b Syllables and the vowel: in y Syllables initial palatality (j, p, ɨ, etc.) is associated with Syllable-initial lip-spreading, with relative frontness of vowel (u, o, əo, etc.),

¹ For Intraverbal and Interverbal Junction see ‘Prosodic Analysis’, pp. 90–1, and for the Juncture system, ibid., pp. 91–5.
² ibid., pp. 91 ff.
³ ibid., p. 96.
and with relative lack of rounding (u o oo, etc.); in ŋ Syllables initial non-palatality (\'ph \(\theta\), etc.) is associated with initial lip-rounding (except in ao Syllables) and with relative backness of vowel (u o ao etc):

\[
\begin{align*}
\text{y:} & \quad pj pj mj mj t(h) j e n \hat{n} & \quad \text{u o o o ao, etc.} \\
\hat{y}: & \quad p(h) m m t(h) n s(h) l \hat{l} k(h) \eta \theta \quad h & \quad u o o oo ao, e.g. \\
y: & \quad pj \hat{u} pru does; & \quad jaot rok arrives; & \quad mj\hat{y} myo floats; \\
\hat{y}: & \quad pu \hat{p}u is hot; & \quad 'gon on matures; & \quad mo mo is tired.  \\
\end{align*}
\]

Since, however, the j, \(\eta\), and y Juncture components correlate with the Palatalization component \(\hat{y}\), while the Juncture components t, n, c, l, k, \(\eta\), s, and v correlate with the Palatalization component \(\hat{y}\), either \(\hat{y}\) or y is invariably implied by one or other of these more specific components, and can without ambiguity be omitted from a phonological formula containing any of them (but not from formulae containing p or m).

The Syllable-initial features of b Verbs as classified in terms of the Juncture, Palatalization, and Aspiration systems, are as follows, when Word-initial:

<table>
<thead>
<tr>
<th>Juncture</th>
<th>p</th>
<th>m</th>
<th>t</th>
<th>n</th>
<th>c</th>
<th>l</th>
<th>k</th>
<th>(\eta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palatalization</td>
<td>(\hat{y})</td>
<td>y</td>
<td>(\hat{y})</td>
<td>y</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
</tr>
<tr>
<td>Aspiration</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Juncture</th>
<th>j</th>
<th>(\eta)</th>
<th>y</th>
<th>s</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palatalization</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
</tr>
<tr>
<td>Aspiration</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
</tbody>
</table>

(from this figure it will be seen that the Aspiration system does not apply to the (disyllabic) s Piece).

There being no further prosodic systems in terms of which the b type of Syllable can be classified, it remains only to state the b-Syllable phonematic systems, of which there are two. Both are V systems (there are no C systems in Burmese). A three-term V system (U, O, \(\Omega\)) is stated for those b Syllables which are also classifiable as z by reference to the Quality system (p. 110), zb Syllables, and a two-term (U, O) for those which are also classified as m or as k (mb and mk Syllables). The phonetic exponents of these terms are:

\[
\text{z} \quad \text{m, k} \\
U: \quad \text{closeness u}; \quad U: \quad \text{half-closeness oo} ; \\
O: \quad \text{half-closeness o}; \quad O: \quad \text{openness ao} ; \\
\Omega: \quad \text{half-openness o}^2
\]

\(^1\) ibid., pp. 97-8. \(^2\) ibid., p. 98.
Examples of Burmese b-Syllable lexical-item phonological formulae are given on p. 116 below. The component b can be omitted from formulae containing U, O, or Ω: they are restricted to, and therefore imply, the b Syllable. On like grounds z (of the Quality system) can be omitted from formulae containing the z-Verb phonematic component Ω.

**Arakanese**

Arakanese b Syllables can also be classified in terms of four further systems, Juncture, Rhotacization, Aspiration, and Palatalization, two of which, Aspiration (h, ħ) and Palatalization (y, ĭ), are identical in number of terms with the Burmese. The Arakanese Juncture system, on the other hand, comprises not fourteen but fifteen terms, to which the same names have been given as for the Burmese (p, m, t, etc.) but with the addition of r ¹; and the Arakanese b Syllables can be classified, as p, m, n, etc., by reference to this system, but not as w (there are no such sequences as *wu, *waō, *woot.

The Arakanese Rhotacization system comprises two terms (r, łą). This system serves to associate alveolar friction with labiality and velarity in initial-consonant sequences (pr mr nr, etc.), but not with alveolar friction, for example (e.g. t n s).²

The exponents of these terms are, then:

r: labiality/velarity and alveolar friction pr pf mr nr kr kř nr

łą: labiality, velarity, etc. and no alveolar friction, e.g. p(i) m(j) kh θ

e.g.
r: pru is white; křaot terrifies; nro resents;
łą: pu is hot; mjo floats; raot arrives.

As in Burmese, h and ħ both correlate with every Juncture component except s; the Juncture components j, n, and y, correlate with the Palatalization component y, while t, n, c, l, r, k, η, s, and v correlate with ĭ, and so does the Rhotacization component r (as in Burmese, y and ĭ therefore need to be explicitly formulated only in bp and bm Syllables); only the Juncture components p, m, k, and η correlate with the Rhotacization component r; but all fourteen Juncture components correlate with łą (and the ten components t, n, c, l, r, j, n, y, s, and v exclusively with łą).

¹ The time available for the Arakanese research was short, and the number of Arakanese examples that require the setting up of the Juncture system, e.g. the p Piece pambē(re), is tired, with labiality and plosion as its phonetic exponents (-mb-), and the t Piece tatte, climbs, with alveolar friction and occlusion followed by plosion (-tt-), is consequently small; but no examples have been noted that are inconsistent with the stating of this system.

² r has been used to symbolize an Arakanese alveolar fricative (I.P.A. x) as being more convenient typographically.
The Syllable-initial phonetic features of each type of Arakanese b Verb classified prosodically in terms of the Juncture, Palatalization, Rhotacization, and Aspiration systems are, in Interverbal Junction (i.e. when Word-initial), as follows:

<table>
<thead>
<tr>
<th>Junc.</th>
<th>p</th>
<th>m</th>
<th>t</th>
<th>n</th>
<th>c</th>
<th>y</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal./Rhot.</td>
<td>˘y˘</td>
<td>y˘(˘r˘)</td>
<td>r(˘y˘)</td>
<td>˘y˘</td>
<td>y(˘r˘)</td>
<td>(˘y˘)</td>
<td>(˘y˘)</td>
</tr>
<tr>
<td>Asp.</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
</tbody>
</table>

The phonematic systems statable for Arakanese b Verbs, and the corresponding components, are the same in number of terms and in exponency as for Burmese (p. 113): z, U u, O o, Ω o; m/k, U oo, O ao.

A.-B.: *b Correspondences

Regular correspondences can be established between the Burmese and the Arakanese Aspiration components h and ū, whence *h and *ū, with h and ū as their respective reflexes in either language.

Regular correspondences can also be established between the Burmese and the Arakanese Juncture components, in combination with the Palatalization and Rhotacization components, as follows:

A. p˘y˘ py(˘r˘) pr(˘y˘) m˘y˘ my(˘r˘) mr(˘y˘) t(˘y˘) n(˘y˘) c(˘y˘)
B. p˘y˘ py py my my t(˘y˘) n(˘y˘) c(˘y˘)
A.-B. *p˘y˘ *py *pr *m˘y˘ *my *mr *t(˘y˘) *n(˘y˘) *c(˘y˘)

A. 1(˘y˘) r(˘y˘) k(˘y˘) k(˘y˘) r η(˘y˘) η(˘y˘) j(˘y˘) n(˘y˘)
B. 1(˘y˘) y(˘y˘) k(˘y˘) j(˘y˘) η(˘y˘) η(˘y˘) j(˘y˘) n(˘y˘)
A.-B. *1(˘y˘) *r *k˘y˘ *kr *η˘y˘ *η˘ r *j(˘y˘) *n(˘y˘)

A. y(˘y˘) s(˘y˘) v(˘y˘)
B. y(˘y˘) s(˘y˘) v(˘y˘)
A.-B. *y(˘y˘) *s(˘y˘) *v(˘y˘)

From the above correspondences it will be seen that (i), the Arakanese Palatalization component y(˘r˘) regularly corresponds to the Burmese Palatalization component y, whence *y; (ii), Burmese ū regularly corresponds to Arakanese ū˘, whence *ū; and (iii), Arakanese r(˘y˘) corresponds to Burmese y, whence *r.
The Arakanese phonematic components U, O, and Ω correspond each to its Burmese homograph, whence *U, *O, and *Ω.

The following are examples of b-Syllable correspondences (implied, and therefore omissible, components are enclosed in parentheses):

A.-B.  

<table>
<thead>
<tr>
<th>A.</th>
<th>B.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b1)gpỳhOz</td>
<td>(b1)gpỳhOz</td>
</tr>
<tr>
<td>(b1)gmyhΩ(z)</td>
<td>(b1)gmyhΩ(z)</td>
</tr>
<tr>
<td>(b1)y̆ghUz</td>
<td>(b1)y̆ghUz</td>
</tr>
<tr>
<td>(b1)y̆ghnOz</td>
<td>(b1)y̆ghnOz</td>
</tr>
<tr>
<td>(b1)chOm</td>
<td>(b1)chOm</td>
</tr>
<tr>
<td>(b1)y̆2hUz</td>
<td>(b1)y̆2hUz</td>
</tr>
<tr>
<td>(b1)y̆1hOk</td>
<td>(b1)y̆1hOk</td>
</tr>
<tr>
<td>(b1)y̆nOm</td>
<td>(b1)y̆nOm</td>
</tr>
<tr>
<td>(b1)y̆vOm</td>
<td>(b1)y̆vOm</td>
</tr>
</tbody>
</table>

(sends, floats, is different, wakes, carries, offers, shrinks, terrifies, resents, sews, is exhausted, loosens, uses, passes [examination]).

**s Syllable**

Burmese and Arakanese s Syllables (B. -i, -en, -ett; A. -i, -et, -ett; so named from spreading) are distinguished from Burmese and Arakanese e and f prosodic types of Syllable by the fact, among other criteria, that in neither language can velarity (k kh ƞ) immediately precede the syllabic vowel, and there are therefore no sequences *kV or *ƞV. In fact, in Burmese, initial velarity is not a feature of s Syllables at all ¹; but in Arakanese, velarity is a feature of the initial sequences kr and Ωr, and, though only in Nouns, kW, e.g. kri, nrei, kwi (B. tci krē, nēn nrim, kwe khwes), but does not immediately precede the vowel.

Arakanese s Syllables are further distinguished phonetically from b, e, and f Syllables by the initial combination of features labiality, plosion, and voice, e.g. bri be finished, run.

**Burmese**

Burmese s are distinguished from c Syllables (-tt, so named from centralized) by the fact that while, in s Syllables, initial labiality is either (i), non-palatalized

¹ U Tin Maung volunteered two exceptions: keun kyin/kint sleep [of monks, buddhas] and khi [obscene], only the former of which appears in Judson's Burmese-English Dictionary (Rangoon, 1953); keun is a learned alternative to the more usual teen.
(p m), in sm or sk Syllables (s Syllables that are also classifiable as m or as k by reference to the Quality system), e.g. pet, pit shut, mën mhin's shut [eyes], or (ii), palatalized or non-palatalized equally (pj mj ; p m), in sz Syllables, e.g. pj phre untie, mji mrañs, pi prë finish, in c Syllables, on the other hand, initial labiality must be palatalized, e.g. pjit pac throw, mjît mrac obstruct.

s Syllables can be classified, like the b (p. 113), in terms of the Aspiration, Juncture, and Palatalization systems, but with the difference that no s Syllables can also be classified as k or ñ (Juncture), a difference that reflects the absence of velarity (k ñ) from the permitted initial phonetic features of the s Syllable, and that the Palatalization system associates initial palatality (te n j e, and, in sz Syllables, pj mj) with relative closeness of vowel (i ë), through the y term, and non-palatality (e.g. p n ð s) with relative openness, through the y term. In s Syllables that are also classifiable, by reference to the Quality system, as m or k, sm and sk Syllables, the Juncture components p and m correlate with y (there are no sequences *pjët and *mjët).

The Syllable-initial features of Burmese s Syllables as classified in terms of the Juncture, Palatalization, and Aspiration systems, are, in Interverbal Junction:

(i) sm and sk Syllable (-en, -et)

<table>
<thead>
<tr>
<th>Junc.</th>
<th>p</th>
<th>m</th>
<th>t</th>
<th>n</th>
<th>c</th>
<th>l</th>
<th>j</th>
<th>n</th>
<th>y</th>
<th>s</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal.</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
</tr>
<tr>
<td>Asp.</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>ph</td>
<td>m</td>
<td>m</td>
<td>t</td>
<td>th</td>
<td>n</td>
<td>ñ</td>
<td>s</td>
<td>sh</td>
<td>1</td>
</tr>
</tbody>
</table>

(ii) sz Syllable (-i)

<table>
<thead>
<tr>
<th>Junc.</th>
<th>p</th>
<th>m</th>
<th>t</th>
<th>n</th>
<th>c</th>
<th>l</th>
<th>j</th>
<th>n</th>
<th>y</th>
<th>s</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal.</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
<td>(y)</td>
</tr>
<tr>
<td>Asp.</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>ph</td>
<td>pj</td>
<td>pj</td>
<td>m</td>
<td>m</td>
<td>mj -</td>
<td>t</td>
<td>th</td>
<td>n</td>
<td>ñ</td>
</tr>
</tbody>
</table>

Only one possible vowel is appropriate to the s Syllable, whether sz, sm, or sk: sz, i ; sm/sk, ë. It is not, therefore, possible to set up a V commutation system; and the appropriate vowels are, accordingly, adequately formulated by the prosodic component s, qualified by the Quality components z, m, and k.

Examples of s-Syllable formulae are given on pp. 120–1 below.

**Arakanese**

Arakanese s Syllables can be further classified by reference to the Aspiration, Rhotacization, Palatalization, and Juncture systems on the same grounds, in general, as the b Syllable, but with certain differences as regards the Juncture and Palatalization systems.
sz Syllables can be classified in terms of the Juncture system not only as being also p, m, t, n, c, l, r, k,  readFile.png
The Syllable-initial features of s Verbs classified in terms of the Juncture, Palatalization, Rhotacization, Labiovelarization, and Aspiration systems, are, in Interverbal Junction:

(i) sm/sk Syllable (-et, -ett)

\[
\begin{align*}
\text{Junc.:} & \quad \frac{p(\bar{y})}{\bar{r}} \quad \frac{m(\bar{y})}{\bar{r}} \quad \frac{t(\bar{y})}{\bar{r}} \quad \frac{n(\bar{y})}{\bar{r}} \quad \frac{c(\bar{y})}{\bar{r}} \quad \frac{l(\bar{y})}{\bar{r}} \quad \frac{r(\bar{y})}{\bar{r}} \\
\text{Rhot.:} & \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \\
\text{Asp.:} & \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \\
& \quad p \quad ph \quad pr \quad m \quad m \quad mr \quad t \quad th \quad n \quad s \quad sh \quad l \quad r \quad -
\end{align*}
\]

(ii) sz Syllable (-i/-i)

\[
\begin{align*}
\text{Junc.:} & \quad \frac{p(\bar{y})}{\bar{r}} \quad \frac{m(\bar{y})}{\bar{r}} \quad \frac{t(\bar{y})}{\bar{r}} \\
\text{Rhot.:} & \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \quad \frac{\bar{r}}{\bar{r}} \\
\text{Lab.:} & \quad \frac{\bar{w}}{\bar{w}} \quad \frac{\bar{w}}{\bar{w}} \quad \frac{\bar{w}}{\bar{w}} \\
\text{Asp.:} & \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \quad h \\
& \quad n \quad n \quad nw \quad nw \quad s \quad sh \quad - \quad sw \quad l \quad r \quad - \quad rw \quad kw \quad kr \quad -
\end{align*}
\]

Since, like Burmese (p. 117), there is only one possible vowel in an Arakanese s Syllable (sz : i/-i; sm/k : ei), there are no V systems, and no phonematic components.

Examples of Arakanese s-Syllable formulae are given at pp. 120–1 below.

1 There are, however, Noun examples: prwh pwh ringworm, mrwh mwh earth, mrwh mwh snake, (w)khr kw dog, jwh tswi sweat.

2 gi, be, the only Verb example in his dialect, Saw Hla Pru considered to be a Burmese loanword (rhi), for which the correct Arakanese form was ri in Akyab but hi in the country districts. Cf. also the Noun example ri [personal name].
In the prwñ Syllable, and in this type of Syllable alone, voice is a Syllable-
initial feature, even in Interverbal Junction, e.g. bri, finish, run (B. prñs and pres
respectively).

A.-B. : *s CORRESPONDENCES

Burmese s Syllables regularly correspond to Arakanese s Syllables (but the
converse does not hold: some Arakanese s Syllables, including all sw Syllables,
correspond to Burmese s Syllables; p. 124); an A.-B. category *s can therefore
be established, with reflexes s in either language.

The same correspondences of Palatalization, Rhotacization, and Aspiration
components can be established as for the *b:

A. h ɦ ɣ(ɣ) r ŭɬ
B. h ɦ y y ŭ
A.-B. *h *ɦ Ɪ *ɣ *r *ɬ

The Juncture components correspond as follows:

(i) sm and sk Syllable

A. pᵒ pr mᵐ mr t n c l r k(r) ṯ(r) j ṭ j ɲ
B. p Ɪ Ɪ m Ɪ t n c l y j ɲ j ɲ
A.-B. *pɬ Ɡ Ɡ mɬ Ɡ t Ɡ c Ɡ l Ɡ k Ɡ ṭ Ɡ j Ɡ ɲ

(ii) sz Syllable

A. prw mᵐw tw nᵐ cᵐ 1 k(r)w ṯ(r)w jw ɲ y Ɡ sw v
B. py Ɪ Ɪ t n c l j ɲ j ṭ j y s v
A.-B. *pr Ɪ Ɪ t Ɪ t Ɪ c Ɡ l Ɡ k Ɡ Ɡ Ɡ j Ɡ ṭ Ɡ y Ɡ s Ɡ v

1 One exception has been noted: A. h, phet shut; B. ŭ, pett pit.
2 The sole Arakanese example of pr, ‘pref, shiver, has no corresponding Burmese Verb
form; but compare phín’phín’ (phēmbeṭ) tunsañ, it trembles.
e.g.:

<table>
<thead>
<tr>
<th>A.-B.</th>
<th>A.</th>
<th>B.</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(g)2sprhz</td>
<td>(gŋ)2sprwzh</td>
<td>pʃi</td>
</tr>
<tr>
<td>*(1)gsmýh́z</td>
<td>(1n)gsmf́wzh</td>
<td>ṃi</td>
</tr>
<tr>
<td>*(g)2snhz</td>
<td>(gŋ)snwzh</td>
<td>ṇi</td>
</tr>
<tr>
<td>*(1)gškį́h́z</td>
<td>(1n)gškrẃzh</td>
<td>ḳri</td>
</tr>
<tr>
<td>*(1)gšnj́h́z</td>
<td>(1n)gšnр́wzh</td>
<td>ṇ̄i</td>
</tr>
<tr>
<td>*(g)2sį́h́z</td>
<td>(gŋ)2sį́wzh</td>
<td>ṭ̄į́</td>
</tr>
<tr>
<td>*(1)gšýh́z</td>
<td>(1n)gšýwzh</td>
<td>ṭ̄́į́</td>
</tr>
</tbody>
</table>

| *(g)2sprhz | (gŋ)2sprwzh | pʃi |
| *(1)gsmýh́z | (1n)gsmf́wzh | ṃi |
| *(g)2snhz | (gŋ)snwzh | ṇi |
| *(1)gškį́h́z | (1n)gškrẃzh | ḳri |
| *(1)gšnj́h́z | (1n)gšnр́wzh | ṇ̄i |
| *(g)2sį́h́z | (gŋ)2sį́wzh | ṭ̄į́ |
| *(1)gšýh́z | (1n)gšýwzh | ṭ̄́į́ |

There are six Burmese syllables that correspond not to Arakanese syllables but to θ syllables in the V term H (Gk. ɛta):

\[
\begin{align*}
ṭi & \quad krañ̃ (A. kr̂è) ; \\
ji & \quad ṛay (A. re) ; \\
ṃ̄ji & \quad mrañ̃ (A. mre) \\
ṃ̄ji & \quad mrañ̃ (A. mre) ; \\
ṣi & \quad cañ̃ (A. se)
\end{align*}
\]

On the other hand there are a number of Arakanese syllables, including all Arakanese sw syllables, that correspond to Burmese θ syllables, in the V term E. The next stage must therefore be to give an account of the Burmese and the Arakanese θ syllable.

θ Syllable

In order first to dispose of the correspondence of Arakanese syllables to Burmese θ syllables (-e, -ɛ, -a ; -tn, -an, -on ; -et, -at, -ot) mentioned in the preceding paragraph it is convenient to consider first those θ syllables which are also classifiable in terms of the Quality system as z, θz syllables (-e, -ɛ, -a) ; for these outstanding correspondences are all correspondences involving θ syllables (A. sz, B. θz) ; the θm and the θk syllable are dealt with subsequently (pp. 126-30).

θz Syllable

To distinguish the Burmese θz syllable from the two remaining prosodic types of syllable, the f (from fronging) and the c (from centralized), it is only necessary to point out that f syllables can also be classified, in terms of the Quality system, as m or as k, and c syllables as k, but neither as z ; cf. θz, pe pes, we way, θwa swãs ; fm, than thuĩ̃, f̣k, latt luik ; c(k), pțt pac. The Arakanese θz syllable

1 Page 119, note 2. The Akyab form ri and the rural Arakanese hi would give correspondences:

\[
\begin{align*}
*(1)gšrh̃ & \quad (1n)gšrw̃h \quad ṛi & \quad (1)gsỹh̃ qi ṛi \quad \text{is} \\
*(1)gsṽh̃ & \quad (1n)gsṽh̃ \quad ħ̣i & \quad " \quad " \quad "
\end{align*}
\]

2 Arakanese 'bri, is finished, an example of the Rhotacization component r, corresponds not to a Burmese y syllable *(p̣ji) but, irregularly, to a y syllable, ů pij prit ; but the Burmese reading-style pronunciation pij would give a regular correspondence *(1)gspiracy.
is similarly distinguished from the sole remaining Arakanese prosodic type of Syllable, the f, by the fact that the latter can be also classified as m or as k but not as z; cf. əz, la goes, sỳe pulls; fm, that sits; fk, la tốc follows, paı́t throws.

**Burmese**

Burmese əz Syllables can, like the b and the s, be classified in terms of the Aspiration (h, ḥ) and the Palatalization (y, ỹ) systems; but, unlike the b and the s Syllable, they can also be classified by reference to a two-term Labiovelarization system (w, ñ).

The Burmese Labiovelarization system is set up on the same grounds as the Arakanese (p. 118): to associate a non-syllabic rounded vowel (w u) with initial lip-rounding (except in y Syllables, which have initial palatality: j e n te) and with a relatively back vowel quality (e a) in w Syllables, but to associate consonantal and vocalic features the reverse of these with absence of labiovelarity or labio­palatality in ñ Syllables; e.g.:

\[
\begin{align*}
\text{w: } & \text{we} \quad \text{we} ; & \text{teñ̂} \quad \text{khyway}' ; & \text{θwa} \quad \text{swās} ; \\
\text{ñ: } & \text{je} \quad \text{re} ; & \text{ñe} \quad \text{nay} ; & \text{ja} \quad \text{ra} \\
\end{align*}
\]

All the Juncture components except w correlate with ñ; all the Juncture components except n and v correlate with the Labiovelarization component w.

This new system, Labiovelarization, affects the statements of exponency of the y and ỹ terms of the Burmese Palatalization system in much the same way as it affected the Arakanese (p. 118): in ñ Syllables the y term associates initial palatality (pi mj l te j e n) with lip-spreading and relative frontness of vowel (e a), while in w Syllables its role is to associate initial palatality (te j e) with labio-palatality (u); in the ñ Syllable the ỹ term associates non-palatality (e g a) with relative backness of vowel (e a), and in w Syllables it associates non-palatality (e.g. p m s, but not ỹ) with labiovelarity (w), e.g.:

\[
\begin{align*}
\text{y: } & (ñ) \quad \text{pià} \quad \text{pra} ; & \text{ja} \quad \text{ra} ; & (w) \quad \text{teqà} \quad \text{krwa} ; & \text{juə} \quad \text{rwā} ; \\
\text{ỹ: } & \text{pà} \quad \text{pā} ; & \text{thā} \quad \text{tha} ; & \text{pwa} \quad \text{phwās} ; & \text{θwa} \quad \text{swās} ; \\
\end{align*}
\]

but because of their correlation with the Juncture components, y and ỹ need to be formulated independently of them only in formulae containing the Juncture components p and m (cf. the b Syllable, p. 113).

Classified prosodically in terms of the Juncture, Palatalization, Labiovelarization, and Aspiration systems the Syllable-initial features of əz Syllables, as they appear in Interverbal Junction, are as follows:

<table>
<thead>
<tr>
<th>Junc.</th>
<th>p</th>
<th>m</th>
<th>t(ỹ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal./Lab.</td>
<td>ñw</td>
<td>ñ(ñ)</td>
<td>ñ(ñ)</td>
</tr>
<tr>
<td>Asp.</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
</tbody>
</table>
A phonematic system, a three-term V system (E, H, A), is statable for əz Syllables, the exponents of each term being:

E : half-closeness (e e e) ; H : half-openness (e e e) ; A : openness (a a a).

The prosodic components ə (Labialization) and z (Quality) can be omitted from a formula containing any of the (more specific) phonematic components.
Examples of Burmese əz formulae are given at pp. 124–6.

Arakanese

Arakanese əz Syllables (-e, -a) are also further classifiable, like sz Syllables, in terms of the five prosodic systems Aspiration (h, ʰ), Palatalization (y, ʲ), Rhotacization (r, ɾ), Labiovelarization (w, ʷ), and Juncation, but differ from them phonetically in having the following Syllable initials : pj mj mj lw lw k kh ju and sq, and in not having voice and plosion (br) as an initial feature. The phonological differences reflected in these phonetic differences are too great for the prosodically classified list of sz Syllable initials (p. 119) to stand for the əz Syllable, even with modifications ; and the following əz list is therefore given, in which the Juncation components p and m correlate with y as well as with ʲ (pj mj mj), l and y with w as well as with ʷ (lw lw ju sq), and k with w as well as with w (k kh) :
A two-term phonematic V system (H, A) is established for Arakanese øz Syllables, the exponents of which are:

\[ H : \text{half-closeness (e e ø e)} ; \quad A : \text{openness (a ø ø ø)} \]

Since the corresponding phonematic components H and A imply both ø and z, these two prosodic components can be omitted from formulae containing H or A.

A.-B. : ARAKANESE-s-SYLLABLE AND BURMESE-ø-SYLLABLE CORRESPONDENCES (*E)

Arakanese sw(z) Syllables (pp. 117–8) regularly correspond to Burmese øw Syllables in E, whence *(ø)wE(z), with s as the Arakanese reflex of *E, and (ø)E as the Burmese; some Arakanese sw(z) Syllables, other than those considered on pp. 120–1, correspond to Burmese (ø)wE(z) Syllables, whence *wE.

The correspondences of the Aspiration (h, ñ), Palatalization (y, ſ), and Rhotacization (r, ſ) components are as stated for the *s correspondences (p. 120); the Arakanese Labiovelarization components w and ſ correspond to the Burmese w and ſ respectively, giving *w and *w.² The Juncture components too do not differ significantly in their correspondences from those stated for the *s correspondences: there are no examples of *ŋ (A. ſr, B. ſ), *ŋ (A. ſ, B. ſ), or *y (A. y, B. y); but there are in addition *ř (A. r, B. y) and *w (A. w, B. w); e.g.:

<table>
<thead>
<tr>
<th>A.-B.</th>
<th>A.</th>
<th>B.</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(1)gpposição</td>
<td>(1)įspposição(z)</td>
<td>*pi</td>
</tr>
<tr>
<td>*(1)gpposição</td>
<td>(1)įspposição</td>
<td>*bri</td>
</tr>
<tr>
<td>*(g)2mywH</td>
<td>(g)y2smwH(z)</td>
<td>*mwi</td>
</tr>
<tr>
<td>*(1)gcwH</td>
<td>(1įy)gscwH(z)</td>
<td>*swi</td>
</tr>
<tr>
<td>*(1)grwH</td>
<td>(1įy)gswH(z)</td>
<td>*ri</td>
</tr>
</tbody>
</table>

¹ There are Noun examples: *(1y)ŋnwH(z) *pe whistle; *(g)ęy)2ŋnwH(z) *we silver.

² In Arakanese mrów, snake, the Rhotacization component r corresponds not to the Burmese Palatalization component y (in *mży) but, irregularly, to ſ: mwe mrwe.
ARAKANESE AND BURMESE PHONOLOGICAL FORMULAE 125

A.-B. A. B.
*(1) grwhE (lry)gsrwh(z) ’rwi (lỳ)ygywhE(z) ’gwe rhwe’ moves
*(g)2kwhE (grỳ)2skwh(z) kri (gòy)2jwhE(z) _tee kre is crushed
*(g)2jwhE (grỳ)2sjwh(z) _tei ” ” kye gets even
*(g)2whe (gỳw)2swh(z) _we (gòyw)2whE(z) _we we we distributes

The regularity of the *E type of correspondence rests on twenty-eight examples in the corpus of material, to which more could certainly be added (p. 109); the following six examples, in which a Burmese (Ə)E Syllable (-e) corresponds not to an Arakanese s Syllable (-i) but to an Arakanese ə Syllable in H (-e), are therefore taken to be irregular:

- *(g)2wiiE (gi’w)2swii(z) wi (ggyw)2wiiE(z) we we distributes
- *(g)2jwiiE (gry)2sjwii(z) kre (ggy)2jwiiE(z) tee kri is crushed
- *(g)2kwiiE (gry)2skwii(z) kri (gòy)2jwE(z) _tee kre is crushed

One would have expected Arakanese (Ə)H Syllables like these six to correspond to Burmese (Ə)H Syllables (pp. 125-6; -ę).

There are also in the Arakanese material six (Ə)H Syllables, like those in the preceding paragraph (-e), that correspond, irregularly, not to Burmese (Ə)H Syllables (-ę) but to s Syllables (-i):

- *(g)2wiiE (gi’w)2swii(z) wi (ggyw)2wiiE(z) we we we distributes

The fact that nine of the twelve exceptional examples given in the last two paragraphs are, in Burmese, y Syllables may well be significant; it is also noteworthy that eight of them have final ų in the orthography: Burmese forms in -ņ are notoriously inconsistent, as these eight examples show, and variable (praŋ’ and praŋ, for example, are pronounced with e by some speakers and with i by others). The Arakanese pronunciation of the corresponding forms seems to be remarkably uniform (-ę).

A.-B. : *(Ə)H(z) AND *(Ə)A(z) CORRESPONDENCES

Apart from the exceptions considered in the last three paragraphs Arakanese əz Syllables in H and A (-e, -a) correspond respectively to Burmese əz Syllables in H and A (-ę, -ą) whence *H and *A.

The correspondences stated for the terms of the Aspiration, Palatalization, Rhotacization, Labiovelarization, and Juncature components above (*E, p. 124) hold good for comparisons of Arakanese and Burmese (Ə)H and (Ə)A Syllables too; these last also afford examples of the correspondences lacking in the *E type: *ŋ (A. ŋ, B. ɲ), *n (A. ŋ, B. ɲ), *y (A. y, B. y); e.g.
**A.-B.** | **A.** | **B.**
---|---|---
*(1)gpypwhH (1η)gpypwhH(z) | 'phē (1η)pypwhH(z) | 'phē phai'  
*(1)gpypwhA (1ηγy)gpwhA(z) | 'pwā (1ηγy)gpwhA(z) | 'pwā pwa  
*(1)gpfrH (1ηγw)gpfrH(z) | 'prā (1ηγw)gpfrH(z) | 'prā pra  
*(1)gmypwhH (1ηγy)gmwhH(z) | 'mwe (1ηγy)gmwhH(z) | 'mwe mwai  
*(1)gmypfrA (1ηγw)gmfrA(z) | 'mjā (1ηγw)gmfrA(z) | 'mjā myāś  
*(g)21whA (gəγy)21whA(z) | _lwa (gəγy)21whA(z) | _lwa łhwā  
*(g)2rwhH (gəγy)2rwhH(z) | _rwe (gəγy)2rwhH(z) | _rwe rway  
*(g)2kźwhA (gəγy)2kźwhA(z) | _kwa (gəγy)2kźwhA(z) | _kwa khwā  
*(1)gkrajwhA (1ηγy)gkrajwhA(z) | _krwā (1ηγy)gkrajwhA(z) | _teqā krwa  
*(1)gŋrwhA (1ηy)gŋrwhA(z) | _ŋa (1ηy)gŋrwhA(z) | _ŋa nhāś  
*(1)gŋrwhH (1ηy)gŋrwhH(z) | _ŋe (1ηy)gŋrwhH(z) | _ŋe khyāś  
*(g)2pH (gəγy)2pH(z) | _ŋa (gəγy)2pH(z) | _ŋa nhāś  
*(1)gjwhH (1ηγy)gjwhH(z) | _jwe (1ηγy)gjwhH(z) | _jwe ūyai'  

(breaks off, is swollen, shows, is bankrupt, are a lot of, slices, intends, separates, proceeds, borrows, marries, widens, is considerate, is distorted).

_om and ok Syllable_

The criterion used to distinguish Arakanese and Burmese œz Syllables from Burmese and Arakanese f (from fronting) Syllables and from Burmese c (from centralized) Syllables (pp. 121–2) is not valid for Burmese and Arakanese œm Syllables (A. -ē, -ā, -ō; B. -un, -an, -on) and Burmese and Arakanese ok Syllables (A. -et, -at, -ot; B. -et, -at, -ot), for all f Syllables are also either m (A. -āt; B. -an) or k (A./B. -atk), and c Syllables can only be also k (B. -tt). Burmese œm and ok Syllables can, however, be immediately distinguished from c(k): initial velarity (k ŋ) is a potential feature of the œ Syllable but not of the c (there are no such sequences as *ktt or *nt)\(^2\); but Burmese and Arakanese œm and ok Syllables are less easily distinguished from fm and fk Syllables. In fact no one single criterion suffices to distinguish these œ Syllables from the f; but three separate criteria have to be stated, each for a sub-class of œ Syllable. The stating of these criteria is therefore postponed until after the relevant prosodic sub-categories have been distinguished and illustrated (p. 130).

For both Burmese and Arakanese œm and ok Syllables a further prosodic system needs to be stated, the two-term Front-Back system (ŋ, ŋ), the terms of which are so named from the letter ŋ of Burmese orthography (but in the

---

1 There is an Arakanese Noun example of ly: `lja, length (B. əlja alyā); I have no note of any corresponding Arakanese form to the Burmese lyi Verb lja lya (is flimsy), which is in any case somewhat literary.

2 The sole exception, Burmese ktt, pinch, is assumed to be a recent loan, perhaps from Mon; it does not appear in Judson's.
typographically more convenient form ‘ŋ’); for ɲ symbolizes some of the inter-related vocalic and consonantal features concerned.

The function of this additional prosodic system is to associate close and half-open front vowels (A. -ë, -et; B. -un/-tån, -et) with initial-consonant sequences comprising consonant and non-syllabic rounded vowel (Çw/Çq, i.e. pw rw mju çu) on the one hand, and, on the other, to associate open vowels and back vowels (A. -ń, -at; -ń, -ot; B. -un/-ân, -at; -on, -ot) with the impossibility of any such Çw/Çq sequence.

There are advantages in formulation in stating as the exponents of Burmese ɲ the same type of phonetic feature as is stated for Arakanese ɲ, and vice versa. Arakanese ɲ and Arakanese ɲ phonetic features are, thus, the reverse of Burmese ɲ and ɲ phonetic features respectively:

\[
\begin{align*}
\text{ɲ} & : \text{A. } \text{mrû, B. } \text{mju } \text{mran sees. A. } \text{prut, B. } \text{pwet prwak swells.} \\
\text{ɲ} & : \text{A. } \text{mrë, B. } \text{mian } \text{mran speeds. A. } \text{prwet, B. } \text{pjot prwat clusters.}
\end{align*}
\]

Apart from the Front-Back system the other prosodic systems in terms of which Arakanese and Burmese əm and ək Syllables can be classified are those familiar from the previous section (əz, pp. 122–124); but there is one important feature that distinguishes the Burmese əm and ək from the əz, and indeed from all other prosodic types of Syllable: in their formulae alone is the w Labiovelarization component combined with the y Palatalization component and one of the Juncture components p, m, or ŋ, e.g. ŋ(y)w, ŋot ŋhwat bend; pyw, pjoat prwam: mix together; myw, mjœt mrwak utter. In Arakanese, similarly, it is only in the əm and the ək Syllable that the w Labiovelarity component correlates with the Juncture component ɲ, e.g. mjœt bend; ŋû ŋe bend.

In the Arakanese and the Burmese əm and ək Syllable the y term of the Palatalization system again associates initial palatality (ɲj mj te ŋ j e) with relative frontness of vowel (A. ę şt ă şt; B. ŋn şt ŋn at ŋn ot) except in Arakanese wị and Burmese wị Syllables (A. ę şt; B. ŋn şt), in which it associates initial palatality (tę ŋ e, and j too for Burmese) with labiopalatality (ŋ); further, in Burmese əm Syllables, y associates initial palatality with pureness of vowel (un on). ŋy, on the other hand, associates non-palatality (e.g. km th) with the reverse of those features: (i) relative backness of the phonetically corresponding type of vowel (A. ę şt ă at ŋ bt; B. ŋe şt ŋn at ŋe ŋn ot), except in Arakanese wị and Burmese wị Syllables, in which the associated feature is labiovelarity (w), and (ii), in Burmese əm Syllables, potential diphthongization (uə ŋe).
In the əm and the ək Syllable the w term of the Labiovelarization system associates rounding of consonant (except palatals), e.g. ə k l, with either (i), in front-vowel Syllables (A. ɪ : ɛt; B. ɪ : ʊən et), a non-syllabic rounded vowel (w ʊ) and a degree of retraction (A. ɛGTK; B. ʊən et), or, (ii), in open- or back-vowel Syllables (A. ɪ : ʊat, ʊut; B. ʊ : ʊən at, ʊən ot), rounding of syllabic vowel (A. ʊ; B. ʊ); the w term associates non-rounding of initial consonant, e.g. ʊk l, with absence of rounding as a vowel feature (syllabic or non-syllabic), and, in Arakanese ɪ and Burmese ɪ Syllables, with absence of retraction (as compared with the w Syllable) (A. ɛGTK ɛt; B. ʊən ɛt / ʊən at).

Arakanese əɪm and əɪk Syllables and Burmese əɪm and əɪk Syllables (A. ɛTK; B. ʊən et) do not differ from the respective əz Syllables as regards classification in terms of the Juncture system; indeed they differ from the əz Syllable in Syllable-initial features only to the extent that the Burmese əɪm and əɪk Syllable are without examples of ɪ, li, sw, and nw, but have the advantage over the əz Syllable in one respect, in the additional initial ʊɪ (ɪəq / ɪəq / mwt sw / ɪəq), while for Arakanese əɪm and əɪk Syllables no initials ʊɪ, ʊ, mɪ, mi, ɪ, p, kw, k, lw, or ɪ have been noted, in comparison with Arakanese əz Syllables, though the additional initials ʊ, rw, li, lw, and nw and ɪw have; but the list of Arakanese Verbs is admittedly incomplete (p. 109). Since the only important respect in which these Arakanese and Burmese Syllables differ from the əz is the Arakanese ɪ and Burmese ɪ Syllable, there would seem to be little point in giving a separate list. The əz lists (B., pp. 122–3; A., pp. 123–4), as modified by the comparison in this paragraph, can therefore stand for both.

Like the Burmese and the Arakanese əwz Syllable (B. -e, -e, -a; A. -e, -a) no Burmese (ə)wəm (B. -an / an) or (ə)wək (ət) Syllable, and no Arakanese (ə)wəm (-ə) or (ə)wək (-ət) Syllable, can be classified in terms of the Juncture system as w (there are no sequences (B.) *wan, *wat; (A.) *wə, *wat; and in their Syllable-initial features too these sub-classes of the Burmese and the Arakanese əm and ək Syllable do not differ significantly from Burmese, or from Arakanese, əwz Syllables: the Burmese (ə)wəm Syllable, for example, has no ʊ-, mɪ-, or li-initial, and has in addition li-; but the differences are not sufficiently great for it to be worth while to list the Syllable-initial features of the Burmese (ə)wəm or (ə)wək Syllable, or the Arakanese (ə)wəm or (ə)wək Syllable, independently of the əz (B., pp. 122–3; A., pp. 123–4).

The remaining sub-class of Burmese əm and ək Syllables, the (ə)wəm (-ə(ə)n) and (ə)wək (-ət), unlike the ɪ Syllables of the preceding paragraph, do contain Verb Syllables that can be classified in terms of the Juncture system as w, e.g. wət wat, wən wəm', but none, on the other hand, that can be classified as v (there are no sequences *əo, *ho), or, it so happens, as ɪ (* ho, *ho); but, apart from the

1 The sole exception, Burmese ɪwən lhma, should perhaps be treated as a reading-style pronunciation, appropriate to this literary word.
prosodically significant absence of initial ‘- and (and the fortuitous absence of
n-, n-, j-, t-, t-, m-, and n-), and apart from the prosodically significant
potential presence of initial w-, the Burmese (a)wjm and (a)wjk Syllables do not
differ from the initials listed for aż (pp. 122-3) as w; and this difference is not
great enough to warrant a separate list of (a)wjm/k initials.

With Arakanese (a)wjm (a)wjk Syllables, on the other hand, the position is quite different: they are distinguished from all other Arakanese
Syllables by the fact that in terms of the Palatalization system they are all y; and there are therefore no examples of such sequences as *pjn, *mijn, *tc(h)n,
*mm, *jv, *mm. Consequently, it is clear that no Arakanese Syllables of this type
can be classified, in terms of the Juncture system, as j, n, or y. On account of
these individual features Arakanese (a)wjm Syllable initials are listed separately,
as follows:

(a)wiJm (-u)
(a)wiJk (-nt)

June.: p m
t(r) c(i') l(i') r(i')
k w(i') s(i')

Rhot.: I I I I I I
Asp.: hh hh hh hh hh hh hh hh hh hh hh

There is never more than one vowel in each sub-category of Arakanese or
Burmese am or ak Syllable:

A. źń : a; źń : ė; wń : o;
B. źm : u(ə); źk : ė; ńm : a/a; ńk : a; ńi : o;

no V commutation system can therefore be stated, the requisite vowel being
adequately indicated by the appropriate term of the Labiovelarization (w, ź)
and the Back-Front (ń, ź) systems, and, for Burmese, the Quality system
(m, k) too.

A.-B. : am and ak CORRESPONDENCES

All Arakanese am and ak Syllables regularly correspond to Burmese am and
ak Syllables respectively (the converse is not true: some Burmese am Syllables
correspond to Arakanese f Syllables; p. 131); an A.-B. *ə category, with
reflexes a(m/k), can therefore be established.

On the basis of a regular correspondence of Burmese ź and ź to Arakanese ź
and ź respectively, *ź and *ź can also be established.

The correspondences stated for the other prosodic components, Aspiration,
Palatalization, Rhotacization, Labiovelarization, and Juncture, in previous
sections (pp. 120, 124, 125) can stand equally for the present section (ə); and
to them can be added:

*wh (A. wh; B. pwh)
A single exception to the *r correspondence (A. r; B. y) has been noted: Arakanese prwt (prw), swell, corresponds not to a Burmese form *pjyt, but, irregularly, to pwst (pŷw) prwak/pwak (but cf. Burmese mijst below).

α correspondences include:

A.-B. A. B.
*(l~)prwh̩jk (l̩g̩o)yprwh̩jk 'prwet 1 (l̩g̩o)pywh̩jk 'pjyt prwat
*(l~)w̩nk (l̩g̩o)y̩w̩nk 'wn̩t (l̩g̩o)pywh̩jk 'pwet whak-
*(l~)mrwh̩jk (l̩g̩o)ymrwh̩jk 'm̩mt (l̩g̩o)mywh̩jk 'mjst mrwak
*(l~)g̩1w̩mj (l̩g̩o)y̩g̩1w̩mj 'lw̩t (l̩g̩o)y̩g̩1w̩mj 'lo(ə)n lwam-
*(g̩)2rwh̩mj (g̩o)y̩rwh̩mj r̩̩ (g̩o)y̩2ywh̩mj ət̩̩ rhwaŋ
*(l~)r̩wh̩jk (l̩g̩o)y̩r̩wh̩jk 'r̩̩t (l̩g̩o)y̩r̩wh̩jk 'jat rap
*(l~)g̩krrwh̩mj (l̩g̩o)y̩g̩krwh̩mj k̩̩ (l̩g̩o)y̩j wh̩mj ət̩̩ khrwaŋ-
*(l~)j̩̩jk (l̩g̩o)y̩j̩̩jk 't̩eht̩ (l̩g̩o)y̩j̩̩jk 't̩eht̩ khyak
*(l~)jn̩w̩h̩jk (l̩g̩o)y̩jn̩w̩h̩jk 'n̩̩j̩ (l̩g̩o)y̩n̩w̩h̩jk 'nj̩ phwat
*(g̩)2y̩w̩mj (g̩o)y̩2y̩w̩mj ə̩ (g̩o)y̩2y̩w̩mj kan rhaŋ

(clusters, hides, utters, regrets, rejoices, stands, leaves, cooks, bends, passes over).

Those Burmese αm Verbs which do not correspond to Arakanese αm Verbs correspond to f Verbs; the next stage must therefore be to define the f Syllable.

f Syllable

The criteria by which, in Arakanese and Burmese alike, the f Syllable (A. -at̩, -at; B. -an̩, -at) is distinguished from the αm and the αk Syllable vary according to the following three sub-categories: (i), Burmese αŋ (-un, -et), Arakanese αŋ (-ə, -et); (ii), Burmese αwŋ (-an̩/un̩, -at), Arakanese αwŋ (-ū, -ut); (iii), Burmese αwŋ (-on, -ot), Arakanese αwŋ (-ū, -ut). The first of these three sub-categories is distinguished from the f Syllable by a potential initial sequence comprising consonant and rounded non-syllabic vowel (Cw-, Cq-) (*Cwat is an impossible sequence), the second by the impossibility of having an initial labiovelar semi-vowel (w) (*wan/wun and *wat are not possible sequences; but cf. wan, wat), and the third by the impossibility of having a glottal initial (') (*'on and *'nt are impossible sequences; but cf. 'an, 'att).

From the Burmese c Syllable (-at), the f Syllable is distinguished, among other criteria, by potential initial velarity (k ŋ) (*ktt and *nt are impossible sequences; but cf. khat, nät).²

The Burmese and the Arakanese f Syllable closely resembles the b Syllable: Palatalization systems (y, ý) need to be stated for both Arakanese and Burmese f Syllable, and a Rhotacization system (r, ř) for the Arakanese; but it differs from the b Syllable in having initial labiovelarity (w) as a potentiality, a phonetic difference that is reflected in the need to set up a Labiovelarization prosodic

¹ Also řsyt.
² But see p. 126, n. 2.
system (w, ū) that is not needed for the b Syllable. As an additional difference, there are f Syllables that can be classified as w in terms of the Juncture system, e.g. B. waw, A. waw.

The exponents of the terms of the Labiovelarization system are:

- w: initial labiovelarity, relative retraction of vowel
  - A. waw, B. waw
  - A./B. waw

- ū: initial non-labiovelarity, a relatively advanced vowel
  - A. waw

E.g.:

- w: A. waw, B. waw
- ū: A./B. waw

Since, apart from the prosodically significant labiovelarity feature (w), the Burmese f-Verb Syllable is distinguished from the b only by the fortuitous absence of initial Ŝ, ph, ŕ, and ť, and the Arakanese, similarly, by the absence from the available material of Ŝ, ŕ, M, and ḫ, there is no need to list the f-Syllable initial features: the b Syllable list (p. 113) will serve, subject to the modifications just stated.

A.-B.: BURMESE əm AND ARAKANESE f CORRESPONDENCES

The Arakanese material provides only eight examples of this type of correspondence, of Burmese əm Syllables (-un) to Arakanese fm (-at). Six of the eight Burmese Verbs, and the corresponding Arakanese f Verbs, can be classified by reference to the Palatalization system as y Verbs: _piun pyaw, _piun pyaain, _piun phyaain, _tsun kyaain, _tsun khyaain, _tn yaw (ni = ʔa-ka-le: sat); the remaining two are ŷ: _sun caai, _sun caain. This correspondence is termed the *cm because it is phonetically parallel to the *ck correspondence stated below (p. 132), a parallel that enjoys orthographic support: *c is represented in the orthography by ŋ and c.

The following are examples of the *cm correspondence:

<table>
<thead>
<tr>
<th>A.-B.</th>
<th>A.</th>
<th>B.</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(g)2cφyfim</td>
<td>(gʰw)2fpyfim</td>
<td>_piat</td>
</tr>
<tr>
<td>*(g)2cφyfim</td>
<td>(gʰy)2fpyfim</td>
<td>_jat</td>
</tr>
<tr>
<td>*(g)2cφyfim</td>
<td>(gʰy)2fyfim</td>
<td>_sat</td>
</tr>
<tr>
<td>*(1)cφcφm</td>
<td>(1i)φw)gfcφm</td>
<td>_saat</td>
</tr>
</tbody>
</table>

(is levelled, is sour, is tame, minces).

C SYLLABLE

The (Burmese) c Syllable (-t) resembles the Burmese s (from spreading) more closely than any other, and especially the sk Syllable (-et); for all c Syllables are also classifiable as k by reference to the Quality system.

The only prosodic respect in which the c(k) Syllable differs from the sk is that in the c Syllable initial labiality must be palatalized (pj pj mj; there can
be no sequences *ppt, *mut), while in the sk Syllable, initial labiality cannot be palatalized (p ph m; there are no sequences *pjēt, *mjet). The only other Syllable-initial differences between the c and sk or sm Syllable are fortuitous: except that the c Syllable has no e initial in a Verb Syllable (but cf. the Noun et' rhac eight), the list of sm and sk initials (p. 117) will serve for the c Syllable.

In the c Syllable the Palatalization term y associates initial palatality (pj mj t e n) with relative frontness (i); ŷ associates non-palatality (t s n l ť h) with relative backness of vowel (a). Thus, the Juncature components p, m, j, n, and y correlate with y, and t, c, n, l, s, and v with ŷ.

A.-B. : BURMESE c AND ARAKANESE f CORRESPONDENCES

Burmese c(k) Syllables regularly correspond to Arakanese fk Syllables; and this correspondence can be associated, as *ck, with the *cm correspondence of the previous section (p. 131).

The Aspiration components correspond as before, to give *h and *h: the Palatalization components and the Arakanese Rhotacization components give three sets of correspondences:

(i) *y (A. y; B. y); (ii) *r (A. r; B. y); *y (A. ūy; B. y); e.g.:

A.-B. A. B.
*(l)gpjh (l)gpjh 'patt (l)gcp(h)k 'pjtt pac throw; pyac is thick
*(l)prrh (l)gpfrh 'pratt (l)gcp(h)k 'pjtt phrac becomes
*(l)pYYYh (l)gpYYYh 'pjatt (l)gcp(h)k 'pjtt phyac massages
*(l)ctvh (l)gtp(h)k 'tchatt (l)gcp(h)k 'tchtt khyac loves
*(l)crvh (l)gprvh 'ratt (l)gcp(h)k 'jtt rac winds round
*(l)ctvh (l)gctvh 'jatt 'yac gets drunk

A.-B. : f CORRESPONDENCES

The remaining Arakanese f Syllables (-ai, -att), those which do not correspond to Burmese ηm (-un) or c (-tt) Syllables, regularly correspond to Burmese f Syllables, whence *f. The prosodic correspondences are the same as those of the b Syllables (p. 115), but with the addition of the Juncature correspondence of Arakanese to Burmese w, whence *w, and the correspondence of the Labio-vellarization components, whence *w, *w; e.g.:

A.-B. A. B.
*(l)gwfm (l)gwrh (l)gwrh 'waι (l)gwrh 'waun wuin does jointly
*(l)fwk (l)gwrh 'waι (l)gwrh 'waun wuik curves round
*(l)flrk (l)gwrh 'waι (l)gwrh 'waun kruik loves
*(l)frvh (l)gwrh 'patt (l)gwrh 'paun pun does jointly
*(l)frvh (l)gwrh 'jatt (l)gwrh 'jaun ruik forces.
CHINESE, THAI, AND MIAO-YAO

By G. B. Downer

INTRODUCTION

Recent advances in descriptive and comparative studies of the Thai and Miao-Yao language-families have tentatively established some of the outlines of proto-Thai and proto-Miao-Yao. The tonal systems, and certain characteristics which I shall call series and features, now seem fairly certain. For the third principal language-family of the Far East, Chinese, we already possess a reconstructed Ancient Chinese which is sufficiently ancient and broadly based to serve as the ancestral language for the modern dialects (excepting the Min dialects). The modern dialects of these three language-families are similar in many ways; the time has now come when we can compare some aspects of the ancestral languages of each family.

The aim of this paper is to demonstrate the unity of the tonal systems of Ancient Chinese, proto-Thai, and proto-Miao-Yao, and to suggest a method by which the ancestral languages may be characterized by comparison of the maximal number of distinctive features of archiphonemes.

1. Chinese

The tonal system of Ancient Chinese, as it appears in the dictionaries, may be symbolized as follows:

1 2 3 || 4

Open syllables have a three-tone system; closed syllables have four tones, the first three ending in nasals, the fourth in the corresponding stop consonants. Put another way, syllables with continuant finals have a three-tone system, those with final stops have no tonal distinctions.

Although the dictionaries of Ancient Chinese show only four tones, it is necessary to subdivide each of the tones, to account for the tonal reflexes in the later dialects. This subdivision correlates with the AC initials. The AC system of initials may be described in terms of archiphonemes, series, and features. Taking the archiphonemes *t ('dental stop') and *l ('lateral') as examples of maximal and minimal differentiation, we find the following possible occurrences:

Archiphoneme + Series + Feature = Initial

*t* will be either clear, which may be unaspirated *t*

or aspirated *th*

or muddy (always voiced) *thi*

*l* can only be muddy *l*

('Clear' and 'muddy' are translations of Chinese phonological terms, which are useful as they do not prejudge the manner of articulation. The muddy series
includes the nasals and laterals, and lacks any aspiration difference in the stops and affricates.)

In the later dialects, although the clear-muddy dichotomy reveals itself sometimes as a difference in manner of articulation of initial consonants, the most noticeable feature is that in almost all dialects it has resulted in different tonal reflexes. This fact suggests that there was a difference in pitch in AC between the two series, although the tonal contours must have been alike. The picture was probably not unlike modern Cantonese, where the two series of initials are clearly separated by pitch, with approximately the same contours in both the clear and muddy series:

<table>
<thead>
<tr>
<th>Tone</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>'ta:m' 'carry'</td>
<td>'ta:m' 'gall'</td>
<td>-'ta:m' 'load'</td>
<td>-'ta:p' 'reply'</td>
</tr>
<tr>
<td>Clear</td>
<td>'tha:m' 'greedy'</td>
<td>'tha:m' 'blanket'</td>
<td>-'tha:m' 'grop'</td>
<td></td>
</tr>
<tr>
<td>Muddy</td>
<td>'tha:m' 'pool'</td>
<td>'tha:m' 'tasteless'</td>
<td>-'ta:m' 'mouthful'</td>
<td>-'ta:p' 'tread'</td>
</tr>
</tbody>
</table>

(Muddy-series plosives have become aspirated in two tones in Cantonese, unaspirated in the other two.)

To sum up, AC had a tone-system 1 2 3 4, in which each tone occurred in two pitches, correlating with the system of initial series and features, which may be symbolized in maximal and minimal differentiation ① ① ① ①.

2. Thai

The tonal system of Common Thai has been reconstructed as a system of four tones with three types of initials; it may be clearly seen in the Lungchow Chuang dialect:

<table>
<thead>
<tr>
<th>Tone</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>-'dai' 'good'</td>
<td>-'da' 'scold'</td>
<td>/'dai' 'get'</td>
<td>-'duk' 'bone'</td>
</tr>
<tr>
<td>Unasp.</td>
<td>-'tim' 'full'</td>
<td>-'ta:n' 'window'</td>
<td>/'tai' 'torch'</td>
<td>-'tap' 'liver'</td>
</tr>
<tr>
<td>Asp.</td>
<td>-'tʰəŋ' 'arrive'</td>
<td>-'tʰu' 'bean'</td>
<td>/'tʰui' 'bowl'</td>
<td>-'tʰuk' 'correct'</td>
</tr>
<tr>
<td>-'nau' 'rat'</td>
<td>-'no' 'sprout'</td>
<td>/'na' 'face'</td>
<td>-'nak' 'heavy'</td>
<td></td>
</tr>
<tr>
<td>Voiced?</td>
<td>'ta' to smear'</td>
<td>-'ton' 'section'</td>
<td>-'ton' 'belly'</td>
<td>-'tok' 'leech'</td>
</tr>
<tr>
<td>'na' 'field'</td>
<td>-'nam' 'sit'</td>
<td>-'nam' 'water'</td>
<td>-'nuk' 'bird'</td>
<td></td>
</tr>
</tbody>
</table>

The cognates in Siamese, Lao, and Tai of the words with nasal initials in the 'Aspirated' series are written with prefixed h-, and were probably originally unvoiced nasals (and laterals)—only thus is their tonal behaviour explained. That they were different from the other nasals is proved by the Sui language, where
The so-called 'voiced' stops (and affricates) appear as aspirated stops in other Thai dialects (such as Siamese, Lao), where they are written with the letters for (Sanskrit) voiced stops. They may be symbolized in our transcription for Common Thai by $th$, etc.

In the Lungchow dialect, the syllables with unaspirated and 'aspirated' initials share the same series of tones, while the 'voiced' series appear with a different set. Although in some Thai dialects there is a tendency for the unaspirated and aspirated to differ in tonal realization, the Lungchow material suggests that, just as in the case of Chinese, a 'clear-muddy' dichotomy may be discerned. If that is so, then we may say that in Common Thai we find the following pattern of series and features for the initials (again using the Chinese terms for the series):

Archiphoneme + Series + Feature = Initial

* $t$ may be clear, which may be glottalized
  or unaspirated $t$
  or aspirated $th$
  or muddy $th$

* $n$ may be clear (and aspirated)
  or muddy $hn$
  or muddy $n$

A problem arises, however, with the identification of tones. For Ancient Chinese the ancient dictionaries and rhymed literature show unambiguously that although there may have been differences in pitch dependent upon the initials, there was a single four-tone system, each tone having a discrete tone-contour. In the Thai languages, lacking ancient evidence, how can tones 1 2 3∥4 in the clear series be definitely identified with tones 1 2 3∥4 in the muddy series, in that order?

There are at least two indications that suggest that they may be so identified. First is that in the Siamese and Tai Lu writing-systems, an identical method of tone-marking by numbers is employed; as the clear and muddy series differ markedly in realization in both languages, this unanimity in the numbering system can only point to an earlier stage when the chief difference (apart from voicing of the initial) between the series was pitch, and not tone-contour. The second and surer evidence, however, comes from the tonal pattern of Thai words with Chinese cognates. Generally speaking, Chinese cognates with any one of the four Chinese tones will have the same tone in Thai, irrespective of 'clear' or 'muddy'. (Just why there seems to be no correlation in the clear and muddy series between the two languages is not known.) The following examples will make this clear: 'C(lear)' and 'M(uddy)' are added after the examples for clarity:
(The seeming exception in the word for 'elephant' is due to a shift in Cantonese-AC syllables with tone 2, muddy series and stop or affricate initial have become tone 3 in Cantonese.)

Since Chinese tone 1 corresponds to both clear and muddy tone 1 in Thai, it is reasonable to assume that these last are to be assigned to one tonal category in Thai also. The same holds for tones 2 and 3 (corresponding to Thai tones 3 and 2).

Thai then agrees with Chinese in having the tone system 1 2 3 || 4, with two series of initials, but a more complex set of features, with maximal and minimal differentiation $\frac{t}{t}$, $\frac{hn}{n}$.

3. **Miao-Yao**

Comparative Miao-Yao studies have lagged behind studies in Common Thai and AC, but the common tonal system has been known for some time. This is an eight-tone system, divided by initial into two groups of four tones. The Yao dialects have a final occlusive in one tone of each group, so we have

$$
\begin{array}{c|c|c|c|c|c}
1 & 3 & 5 & 7 \\
2 & 4 & 6 & 8 \\
\end{array}
$$

Highland Yao, which possesses all eight tones, may serve as an example:

1. 'duŋ' 'deaf'  3. 'bunŋ' 'bone'  5. 'khunŋ' 'only'  7. 'kyuŋ' 'six'
2. 'lunŋ' 'sky'  4. 'tunŋ' 'pig'  6. 'byuŋ' 'rain'  8. 'tsuŋ' 'must'
We may use the terms ‘clear’ and ‘muddy’ again to designate the two series. In different dialects the nature of the distinction between the two series varies widely; but, although the number of archiphonemes of Common Miao-Yao cannot yet be determined, a tentative picture may be drawn of the features which must be distinguished to cover the reflexes in the various dialects. Correlated with series, these are:

<table>
<thead>
<tr>
<th>Archiphoneme</th>
<th>Series</th>
<th>Feature</th>
<th>Initial</th>
</tr>
</thead>
<tbody>
<tr>
<td>*t</td>
<td>Clear</td>
<td>May be clear, which may be unaspirated</td>
<td>t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or aspirated</td>
<td>th</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or prenasalized unasp.</td>
<td>nt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or prenasalized asp.</td>
<td>nth</td>
</tr>
<tr>
<td></td>
<td>Muddy</td>
<td>which may be simple stop</td>
<td>th</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or prenasalized</td>
<td>nth</td>
</tr>
<tr>
<td>*n</td>
<td>Clear</td>
<td>May be pre-glottalized</td>
<td>?n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or aspirated</td>
<td>hn</td>
</tr>
<tr>
<td></td>
<td>Muddy</td>
<td></td>
<td>n</td>
</tr>
</tbody>
</table>

Although the phonetic notes on the features are purely hypothetical, the only definite point being the number of categories and their correlation with the two series, nevertheless the Lungli Shuiwei dialect possesses all these distinctions, and may be compared with White Miao and Highland Yao. (Initials with various archiphonemes had to be used.)

<table>
<thead>
<tr>
<th>Series</th>
<th>Initial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear</td>
<td>Lungli Shuiwei</td>
</tr>
<tr>
<td></td>
<td>White Miao</td>
</tr>
<tr>
<td></td>
<td>Yao</td>
</tr>
<tr>
<td></td>
<td>Tone</td>
</tr>
<tr>
<td>Clear</td>
<td>t pi</td>
</tr>
<tr>
<td>Clear</td>
<td>th -</td>
</tr>
<tr>
<td>Clear</td>
<td>nt -</td>
</tr>
<tr>
<td>Clear</td>
<td>nth -</td>
</tr>
<tr>
<td>Muddy</td>
<td>t -dəo</td>
</tr>
<tr>
<td>Muddy</td>
<td>nt -mpi</td>
</tr>
<tr>
<td>Clear</td>
<td>n -m?oŋ</td>
</tr>
<tr>
<td>Clear</td>
<td>hn -m?oŋ</td>
</tr>
<tr>
<td>Muddy</td>
<td>n ‘na</td>
</tr>
</tbody>
</table>

Although there is little internal evidence to indicate that tones 1 3 5 7 are to be equated with tones 2 4 6 8 (in that order), Chinese cognates suggest that it is possible to reduce these eight tones to an overall 1 2 3 4, as follows (again the examples are from Highland Yao):

---

CHINESE, THAI, AND MIAO-YAO 137
If this holds true, then pairs of words in Highland Yao such as 'tay ‘to die’ and ‘tay ‘to kill’ (White Miao tua and tua respectively) may be said to have the same tone, and to differ only in the series of the initials.

We may therefore tentatively posit for Common Miao-Yao the same tone-system 1 2 3 || 4 and the same two series that we have for Chinese and Thai, with the maximal and minimal system of features $\begin{array}{c}
    t \ th \ nt \ nth \\
    n \ hn \\
    t \ nt \\
    n
\end{array}$. 

CONCLUSION

Present evidence shows that not only general features of monosyllabicity, tonality, and simple consonantal systems are common to the Ancient Chinese, Common Thai, and Common Miao-Yao, but also specific characteristics, such as a tonal system 1 2 3 || 4, and a division of initials into two series, one of which shows an aspiration difference, the other having some connection with voicing and lacking the aspiration difference. The three language-families differ, however, in the number of features within each series, Chinese having the fewest, Miao-Yao being the most complex.

Although the aim of this paper has merely been to suggest a simple method for characterizing differences and similarities in the initial systems of these languages, some historical explanations for this remarkable unanimity suggest themselves. It seems unlikely that the possession in common of so many phonological features by the three contiguous language-families could be the result of coincidence. There seems to be nothing especially ‘basic’ or determined about such a system—indeed, few later dialects preserve it. Rather it suggests that the ancestors of these three families enjoyed a long period of contact and mutual influence, probably at the same time that the extensive borrowings of Chinese vocabulary were made in Thai and Miao-Yao. Ancient Chinese may be dated to the period fourth to seventh centuries A.D. Earlier forms of Chinese (so far as they are known) do not seem to agree so well with the Thai and Miao-Yao systems. We may therefore posit this period of contact to have been from about the fourth century to the break-up of the Common Thai and Common Miao-Yao languages (presumably prior to the Mongol invasions).
Li Fang-kwei. The Tai Dialect of Lungchow. Academia Sinica, Inst. of Hist. and Phil.
Monograph Series A No. 16 (Shanghai 1940).
Hist. and Phil., Academia Sinica XI (Shanghai 1947).
—— The Distribution of Initials and Tones in the Sui Language. Language, 24, 160–7 (1948).
Fu Mao-chi and others. Phonemic System of the Tai Language of Vengchinghung in Sipshuang-
panna, Yunnan. Yuyan Yanjiu 1, 223–64 (Peking 1956).
Li Yong-sui and others. Some Problems concerning Initials and Tones of the Miao Language.
THE POSITION OF THE MUNDA LANGUAGES WITHIN THE AUSTROASIATIC LANGUAGE FAMILY

By HEINZ-JÜRGEN PINNOW

1. The Austroasiatic controversy, which arose in 1930 when Wilhelm von Hevesy \(^1\) violently attacked Wilhelm Schmidt,\(^2\) the real discoverer of the Austroasiatic family, still remains unsettled. Schmidt, it will be recalled, had grouped the Munda languages, Khasi, and the Nicobar, Palaung-Wa, Mon-Khmer, Malacca, and Cham languages together under the designation 'austroasiatischer Sprachstamm'. Some scholars, among them Sten Konow,\(^3\) C. O. Blagden,\(^4\) F. B. J. Kuiper,\(^5\) R. Shafer,\(^6\) J. H. Greenberg,\(^7\) and A. J. Shevelenko,\(^8\) in the main agreed with Schmidt and favoured the recognition of the existence of an Austroasiatic family including the Munda languages but excluding the Cham languages. Others, in particular Georges Maspero,\(^9\) and more recently also A. I. Blinov,\(^10\) rejected this contention. The majority of the critics, especially T. A. Sebeok\(^11\) and M. B. Emeneau,\(^12\) have deferred judgment in the belief that the problem can be solved only through further investigation.

How unsettled the classification of the languages of South East Asia still is can be seen in Richard Salzner’s *Sprachenatlas des Indopazifischen Raumes*,\(^13\) in which the Munda and Malacca languages are excluded from the Austroasiatic family but which follows Schmidt in once again designating as Austroasiatic the Cham languages, which have long been recognized as Indonesian. The whole question becomes even more complicated by the fact that the linguistic relationships of Vietnamese and of Müöng are disputed.\(^14\) Vietnamese, Müöng, and the Malacca and Cham languages are, however, all clearly mixed languages: they all reveal at least an Austroasiatic substratum, which, however, is so small in

---

Vietnamese, Müöng, and the Cham languages that we may unhesitatingly regard Vietnamese and Müöng as Thai languages and the Cham languages as Indonesian. In the Malacca languages, on the other hand, the Austroasiatic elements are so numerous that the languages are still to be regarded as Austroasiatic, although all of them (particularly Jakud) are gradually being assimilated into the Indonesian group.

The Munda languages, which are undoubtedly not mixed languages, do indeed differ greatly from the pattern of the other Austroasiatic languages (for which the designation ‘Khmer-Nicobar group’ seems suitable), but the points of correspondence of the Munda languages with this group are so numerous that these similarities can by no means be coincidental. Borrowing may also be ruled out, for the common features lie especially in the basic vocabulary, in word-formation (where the infix-formation is particularly characteristic), and in certain morphological and syntactical peculiarities. The typological differences between the two groups can in large part be explained as the results of influence by the various neighbouring languages, with which individual Austroasiatic sub-families have formed so-called ‘language leagues’ (Sprachbünde). A further discussion of this subject is to be found in my Versuch einer historischen Lautlehre der Kharia-Sprache and in an article in the Indo-Iranian Journal, 1960.

2. Schmidt’s view that the Munda languages form a sub-family of the Austroasiatic group proves to be well founded and correct, but it still remains to be ascertained exactly what the relationships of the various members of the family (Munda, Nicobar, Khasi, Palaung-Wa, Mon-Khmer, and the Malacca group) to one another are, and a suitable classification of the Austroasiatic family must be worked out. Furthermore, the classification of the Munda languages themselves must be established. The unity of all the Munda languages (not including Nahali, which is not Munda) is apparent not only from a comparison of the phonological systems in these languages but even more from their highly developed inflectional systems, in which they differ from all other Austroasiatic languages. Recent investigation has shown that the verbal inflection of all Munda languages is traceable to a Proto-Munda inflectional system, which was later expanded in the north and considerably reduced in the south. From this evidence and on the basis of lexical differences the Munda languages may be divided into a Northern group with the sub-groups Kurku and Kherwari (Santali, Mundari, Korwa, etc., belong to the latter branch), and a Southern group, which is further subdivided into a Central group (including Kharia and Juang), and a South-Eastern group (including Sora, Pareng, Guto, and Remo). R. Shafer (in Word, 1960, p. 425)

1 Wiesbaden, 1959.
has also divided Munda into Northern and Southern groups. The relation of Kherwari and Kurku is much closer than that of Central and South-Eastern Munda, which must have been separated much earlier than Kherwari and Kurku.

This classification, the reasons for which are given in more detail in the investigation of the Munda verb mentioned above, differs greatly from the older classification of Schmidt, who grouped the languages solely on the distribution of $k$ and $h$ (from Proto-Munda $^q$). He concluded that there is an Eastern group (= Kherwari; with $h$), a Western group (Kurku, Kharia, Juang; with $k$), and a supposedly 'mixed group' (= South-Eastern Munda; with a loss of Proto-Munda $^q$). As this classification was based on a single argument, it could not do justice to the facts.\footnote{Die Sprachfamilien und Sprachenkreise der Erde, Heidelberg, 1926, p. 140.}

3. W. Schmidt likewise attempted a general classification of the Austroasiatic languages,\footnote{Grundzüge einer Lautlehre der Khasi-Sprache (Abh. d. K. Bayer. Ak. d. Wiss., 1 Kl., Bd. 22, 3 Abt.), München, 1904, pp. 759, 804. Cf. p. 140, n. 2, and n. 1 above.} and again on the basis of a single, if important, factor. On the basis of the comparison of numerals and some other important words he came to the conclusion not only that the Munda languages are Austroasiatic but that they together with the Mon-Khmer languages (in the narrower sense of the term) form one of four Austroasiatic sub-families,\footnote{See n. 2 above.} the others being the older Malacca group (including Semang and Sakai), the central group (including Khasi, Nicobar, and Palaung-Wa), and a so-called south-east mixed group (i.e. the Cham languages). Schmidt classifies Besisi and Jakud as a younger Malacca group with the Mon-Khmer languages. His classification of the 'South-East mixed group' is not appropriate, as these languages are clearly Indonesian. Because of its one-sidedness, Schmidt's classification is as a whole questionable and must be re-examined. However, in the following we will limit ourselves primarily to a reconsideration of the position of the Munda languages.

4. There are two ways of solving the problem of the classification of the Munda languages within the Austroasiatic family: first, according to structural, synchronic considerations; secondly, according to the historical development of the languages. The latter classification is undoubtedly the more important one. Reliable information can be obtained only through such intensive historical linguistic investigation of the various sub-groups as R. Shafer, for example, has done in the field of the Palaungic languages.\footnote{See p. 140, n. 6.} The work of W. Schmidt, pioneering as it may be, does not suffice to clarify the complicated situation. As long as the eastern Austroasiatic languages have not been subjected to a comprehensive historical examination in which the history of the individual groups is established, the classification of these languages on the basis of their common and differing features will remain provisional. In the comparison of certain
individual correspondences in vocabulary, morphology, and phonology the greatest caution is necessary, for all too often one finds completely independent parallel developments that do not at all justify grouping the languages with such common features together. Reasonable assurance for the correctness of the provisional classification is offered only by a correspondence of several major factors, in which connection morphological features deserve particular attention.

5. Features particularly to be regarded in classification are the following:

5.1. The position of parts of the sentence, especially in the relationship of the determiner to the thing determined. From the structural viewpoint three groups can here be clearly distinguished: (a) the eastern languages (with the exception of Nicobar), in which the determinans consistently follows the determinandum; (b) the Nicobar group, in which this rule is only partially observed, in so far as the ‘adjective’, without an affix, precedes the ‘substantive’, e.g. Nancowry āā yuṇā lopaa kālīṇ ‘the two good foreigners’ (literally ‘two persons good foreigners’) or āā yuṇā kālīṇ to-lo′pa; in the latter case the adjective following the noun has the prefix to-. In Car Nicobarese an adjective with the prefix tō- can also precede, and the practical result of this is that the whole determinans precedes the determinandum, e.g. tō-tamii cooṇ ‘steamer’ (literally ‘with fire provided boat’; tamii-yō ‘fire’). Lastly (c) the Munda languages, in which the determinans usually precedes, but in which the reverse holds true in a number of cases. Thus in compounds, especially in Sora, the ‘genitive’ often follows, e.g. ə-jeelu-boon-ən ‘buffalo meat’ (literally ‘meat-buffalo-the’). The postposition of the personal pronoun as a possessive suffix is also frequently employed, e.g. Juang bui-nəm ‘thy mother’ (literally ‘mother-of-thou’). Historical linguistic investigation shows the following: at the earliest period for which we can make any statements the order determinans—determinandum was predominant. Later began a period in which the determinans was in all Austroasiatic languages largely but not exclusively placed after the determinandum. The eastern languages with the exception of Nicobarese continued this tendency, so that there now exist only a few remnants of old ‘preposition’, such as the ‘preposition’ of the numerals and numeral classifiers in Khasi and Bahnar. The Munda languages returned to the customary ‘preposition’ of the determinans—presumably under the influence of Dravidian and Aryan languages. Nicobarese took a similar course—perhaps influenced by Andamanese languages. The only difference from Munda is that in Nicobarese the number of formations with ‘preposition’ of the determinans is comparatively small. There is no reason to assume that Munda and Nicobar in this case passed through a common stage of development; however, a common tendency may have been present.

5.2. The use of prefixes and suffixes is very closely connected with the position of the determinans. The classification given above is also applicable here, but
Nicobarese and Munda are in this regard even more closely connected. The three groups are: (a) the eastern languages (with the exception of Nicobarese), which use prefixes almost exclusively. Remnants of old suffixes appear in Lawa -do in *kon-do* ‘child’ = Kharia *kon-du*? (cf. Mon *kon*); Khmer lō-k ‘raise’, lō-ŋ ‘rise’, lō-s ‘more’, from lō ‘up’; cf. Kharia tob-luŋ ‘up’.

(b) The Nicobar group, which uses a large number of prefixes and suffixes. It is a striking fact that some affixes can sometimes be used both as prefix and suffix: e.g. Car ma, maa, affix indicating a nomen agentis, in ma-haŋōh ‘one sleeping’ (ha-rōh ‘to go to sleep’), but pa-maa ‘an eater of food’ (*pa* ‘to eat’). Further, Car ha, Nancowry ho (ha), affix indicating transitive or causitive verbs (= Khmer pō-, Bahnar pō-, Sora ab-); e.g. Car ha-kūn ‘to cause to carry’ (kūn ‘to carry’), Nancowry fət (fat) ‘to break by striking’, fət-hə (fat-ha) ‘to cause to break’.

(c) The Munda languages, which are in this respect like the Nicobarese languages. Both prefixes and suffixes occur both as productive and unproductive affixes. Among the productive suffixes are almost all the inflectional suffixes, e.g. Santali, Mundari, Kharia -ə? indicating the genitive. Unproductive suffixes occur for example in Kurku ka-ku ‘fish’ (Sre ka), Kharia siŋ-ko-e <*siŋ-ko-e* ‘fowl’ (Mundari sim id., Sre sim ‘bird’). The suffix -ku, -ko occurs as a prefix in War k-sem, Mon ga-cem ‘bird’. Many of the suffixes in the Nicobarese languages, particularly the directional affixes, are not connected with Munda affixes; in others, however, a close connection can easily be supposed, e.g. Santali, Mundari -o?, verbal suffix of the middle or passive voice, Car -ō, Nancowry -ə (-a), in Car mūk ‘to see’, passive mūk-ō, Nancowry hōŋ-kōtə ‘to burn’, passive hōŋk-ə; further Mundari -iən, Santali -en, Proto-Munda *-iən, suffix of the durative (with intransitive verbs), Nancowry -yen-de (-yan-de), suffix of the continuative, in top ‘to drink’, t-em-op-yan-de ‘to continue drinking’. In some cases Munda suffixes occur as prefixes or as particles in Nicobarese, e.g. Juang -tae, Kharia -ta, -atai, ablative suffix, Nancowry tai, in Kharia am-a-tai = Nancowry tai me ‘from thee’ (Kharia am ‘thou’ <*a-me’; cf. Santali me ‘thou’). Further Mundari, Santali -le-d, -le-n, suffix for the non-resultative perfect, Sora -le, perfect suffix, Nancowry leet; e.g. Mundari abun-led-a-ŋ ‘I have washed’, Nancowry ciu leet oɾii ‘I have beaten’ (from oɾii ‘to beat’; cf. Juang rim id.). Corresponding prefixes in Munda and Nicobarese are to be found, for example, in Mundari ta-, Nancowry to- (ta-), prefix indicating possessive pronoun, in Mundari ta-iŋ, Nancowry to-ciuə ‘my’, e.g. Mundari apu-ta-iŋ ‘my father’, Nancowry dīə tə-ciə ‘my canoe’.

It may be regarded as certain that a large number of both prefixes and suffixes were employed in the Austroasiatic languages at an earlier period.¹ The fact that

once again neither Munda nor Nicobarese lost their suffix-formations but rather extended the use of such formations through the erection of various new suffixes is evidence not for an actual common development but rather for a common conservative tendency.

5.3. A treatment of the numerous infixes, most of which are undoubtedly old, would be superfluous here, for they are encountered in all the groups and are hence meaningless for the classification.¹

5.4. The repetition of words that logically need to occur only once in a sentence is to be found both in Munda and in Khasi, e.g. Mundari am gapa-m sen-re-do, ka-m tebai-a ‘if thou goest tomorrow, thou wilt not arrive’ (literally ‘thou tomorrow-thou go-if, not-thou arrive-will’), Khasi ka-wei ka kənthei ka la iap ‘the one woman died’ (literally ‘she-one she woman she died’). We have here to do with a kind of agreement which of course has not reached the extent to be observed, for example, in some African classificatory languages. Whether this agreement is old or whether it may earlier have been present in other Austroasiatic languages cannot yet be determined. A direct relationship of the two languages in respect of these forms cannot be shown, but as a common feature this point is not uninteresting. A classification from a structural point of view is then: (a) the eastern languages excepting Khasi; (b) Khasi and Munda.

5.5. On the basis of their syntactical framework the Austroasiatic languages may be classified into two groups: (a) the largely co-ordinating and analytic Khmer-Nicobar languages, and (b) the largely subordinating and synthetic Munda languages. The Nicobar languages here share the typical characteristics of the eastern group; subordinating participial constructions, which frequently occur in Munda, are here completely absent. According to recent investigation ² the extensive subordination and the extremely synthetic structure of the Munda languages are of a secondary nature, just as the analytic structure and general co-ordination in the east are not original but are probably the result of secondary development and, in part, of the influence of the Thai and other languages. The complicated Munda conjugation, if pursued back to its origins in older Proto-Munda, shrinks to a fraction of its present size; thus in Proto-Munda no pronouns as indicators of either subject or object were incorporated into the verb. The decisive question here is whether the analytic simplification developed independently in the individual eastern groups or whether it began in a period before the formation of these groups, as was the case with the increased structuralization in the later period of Proto-Munda. If the latter supposition is correct, the Khmer-Nicobar languages form just such an historical unified group as the Munda languages do. It is unfortunately not yet possible to give a reliable answer to this question.

5.6. Morphological systems with declension and conjugation are present only

² See p. 141, n. 3.
in Munda. In the eastern languages with the exception of Nicobarese morpho-
logical systems are practically wholly absent. Special syntactical relationships
and loosely attached particles there assume the function of the Munda declension
and conjugation. Nicobarese again has an intermediate position: it possesses
obvious remnants of an older morphological system. In Car Nicobarese three
‘cases’ of the pronouns are distinguishable, ‘absolutive’ or ‘nominative’,
‘interrogative’, and ‘oblique’. The first person singular is thus cin or cuu-ô in
the nominative, öic, cu-ô öic in the interrogative, and cu in the oblique. As in
Munda, the singular, dual, and plural of the pronoun are differentiated, as well
as inclusive and exclusive forms of the first-person pronouns in dual and plural.
Some of the comparative forms in Nancowry differ greatly from the corresponding
morphemes in the positive, e.g. wiihlô ‘young’, olyaalô ‘younger’, keoooyô
‘hot’, kenooun-ô-yan ‘hotter’. Continuative forms of the verbs are created by
various alternations of the root with the infix -en-, the suffix -a, and the suffix
-yan-de; e.g. ikaafa ‘to sing’, kennîja-yan-de ‘to continue singing’.1 Passive
formations in -ô or in -ô have already been mentioned, as the continuative forms
have been; they agree exactly with the Munda forms. Very different from the
Munda, on the other hand, is the use of directional affixes; e.g. Nancowry -fe,
which indicates motion downward, in ten-fot-ha-fe ‘to break something by
throwing it to the ground’. It is evident from Munda and Nicobar that Proto-
Austroasiatic possessed a fairly well-developed morphological system that was
further expanded in Munda, reduced in Nicobar, and lost in the other languages.
This loss of the morphological system suggests not only a common development
but even that this development began in a period in which the individual groups
had not become independent. In general the following rule holds true for the
development of languages: if only one of a group of languages introduces some
important innovation, the fact that the other groups do not possess this innovation
may not be regarded as a sign that they still have any common ties; however, if
all but one of several groups of languages introduce an innovation, this fact may
be regarded as an indication that the innovating languages possess common bonds.
An early separation of Nicobarese from the rest of the groups seems very likely.

5.7. Another morphological feature is the frequent use of certain composi-
tional forms: i.e., the words, when they are used in a compound, possess a
different form, which is usually one affix shorter than the normal form. This occurs
mostly in Sora but to a lesser degree also in the other languages, especially in
Kharia; e.g. Sora taŋ-liy-ôn ‘cow’, compositional form taŋ-ôn, in gupaat-
taŋ-mar-ôn ‘cowherd’ (literally ‘graze-cow-man-the’), Kharia soren ‘stone’,
compositional form -sor in thom-sor ‘to stone’. Here there is an agreement with
Khasi, which likewise has compositional forms; e.g. ksah ‘ring’, kti ‘hand’,
sah-ti ‘ring for the finger’. In the other languages there do not seem to be any

1 See 5.2 above.
special compositional forms, though an old remnant may perhaps be present in Khmer \( m\tilde{o} \) (\( m\cdot \)) to \( muoy \) ‘one’, e.g. in \( m\cdot da(y) \) ‘once’. In Nancowry Nicobarese \( fijnk\alpha \) and \( da\alpha m \) (\( ra\alpha m \)) are respectively used in combination with numeral classifiers for \( h\eta \) ‘day’ and \( h\omega \tilde{om} \) ‘night’. Whether old compositional forms are present here is not entirely certain. However, such pairs as \( h\alpha\cdot kook \) ‘to fire a gun’, \( hen\cdot kook \) ‘cannon’, \( h\omega\cdot tee\eta \) ‘to sail’, \( hen\cdot tee\eta \) ‘sail’ reveal good comparable parallels, as do the frequent phonetic changes in Nicobarese when affixes are employed.

Thus from a structural standpoint the classification is: (a) the eastern languages except Nicobar and Khasi; (b) Khasi, Nicobar; (c) Munda. Special compositional forms probably existed at an earlier period; the fact that they have been in varying degree preserved in Munda, Khasi, and Nicobar constitutes a bond between these languages but is no evidence for a common development.

5.8. In the use of numerous numerical coefficients the Nicobar languages agree completely with the other eastern languages and thus differ from Munda, in which such words are little used. The word-order in Munda, Nicobar, and Khasi is numeral, numeral classifier, noun; in Mon and Khmer noun, numeral, numeral classifier; in Bahnar usage varies: e.g. Kharia \( molo\tilde{i} boko'b \) \( or\eta\cdot j\cdot ki \) ‘five bullocks’ (literally ‘five head[s of] cattle’), Car \( s\tilde{o}m \) \( t\alpha\cdot ka \) \( n\i\tilde{o} \) ‘ten children’ (literally ‘ten persons child’), Khmer \( kun \) \( pr\tilde{u}\tilde{s} \) \( pi 'n\tilde{o}\kappa \) ‘three boys’ (literally ‘child masculine three persons’), Bahnar \( plu\eta \) \( 'baar \) \( to\tilde{o}n \) ‘two barks’ (literally ‘bark two piece’) or \( 'baar \) \( to\tilde{o}n \) \( plu\eta \). It is remarkable that with regard to the position of the numeral coefficients Munda, Nicobar, and Khasi again stand in opposition to Mon and Khmer. That the use of numeral classifiers was extended in the east and restricted in the west is probably the result of the influence of neighbouring non-Austroasiatic languages.

5.9. With regard to phonology the following points are important:

(a) Central, unrounded back, or rounded front vowels (\( i, \alpha, u, v, y, \theta \)): prevalent in Khmer-Nicobar, rare in Munda. For Proto-Munda *\( i \) and *\( \alpha \) must, however, be reconstructed.

(b) Quantitative opposition of vowels: prominent in Khmer-Nicobar excepting Khasi; less prominent in Munda. Quantitative opposition is to be assumed for Proto-Munda.

(c) Contrast voiceless \emph{versus} voiced stops: lacking only in Nicobarese. In Nancowry \( d \) appears as a kind of liquid. The loss of voiced stops in Mon, Khmer, Riang, etc. is demonstrably secondary.

(d) Aspiration: lacking only in Nicobarese in the east and in Sora and Gutob in the west. For the older period aspiration is to be assumed. It was, however, clearly diphonemic, just as it is to-day in Khmer, etc. In Munda secondarily monophonemic as a result of the loss of the other initial consonant clusters.

(e) Retroflex consonants: only in Munda excepting Sora. Retroflex consonants are completely lacking in the east; \( t \) in Khmer and \( d \) in Mon are graphemic
equivalents of phonetically preglottalized \([d']\), which may be alveolar or dental. In Car Nicobar \(r\) (\(\tau\)) is no genuine retroflex. Proto-Austroasiatic did not have retroflex consonants; \(t\) was probably dental, \(d\) alveolar. This is still the state of affairs in Sora and Palaung.\(^1\) Retroflexes were probably developed in Proto-Munda under the influence of Dravidian and/or Aryan languages.

(f) Non-plosive, glottalized consonants: in Munda and Nicobarese not initially but only finally. Initial implosive consonants are present in Mon, Khmer, Bahnar, etc., but not in Palaung and some other languages. All these sounds, which have in part become phonemes, are of a secondary nature and have developed independently. Only the tendency to introduce such sounds is a common feature. The close correspondence of Munda and Nicobarese is again noticeable.

(g) Tonemes were originally completely absent. In the languages in which they now occur, e.g. Mon, Sre, Riang, they are clearly secondary developments.

(h) Consonant clusters are numerous in the eastern languages except Nicobar, rare in Munda, absent in Nicobarese. Clusters of two, and not more than two, consonants undoubtedly occurred in Proto-Austroasiatic.

(i) Number of syllables in words: Munda inclines towards polysyllables, the eastern languages to monosyllables. The polysyllabic structure of Nicobarese is the result of the fact that many affixes have there been indissolubly combined with roots. It is certain that the older languages had both one- and two-syllable words and that the latter have mostly been secondarily shortened in the east.

This phonological evidence again leads to a threefold classification: Munda, Nicobar, and the other Austroasiatic languages. Nicobar and Munda have three points of correspondence; Nicobar and Mon-Khmer, four points. The development of the phonemic system of Proto-Austroasiatic to Munda, Nicobar, and the remaining languages indicates that the former two groups were early separated from the family.

5.10. The numerals present a different situation. In this respect the Munda languages together with Mon-Khmer and one part of the Malacca languages form one group in contrast to Khasi, Nicobar in another group and to the remaining Malacca languages in yet a third group: e.g. Mundari \(\textit{bar} ‘\text{two}’\), Bahnar \(\textit{baar}\), South Sakai \(\textit{mbaar}\); but Khasi \(\textit{ar}\), Palaung \(\textit{aa}\), Riang \(\textit{kaar}\), Nancowry \(\textit{aaa}\); North Sakai \(\textit{nar}\). Similarly Mundari \(\textit{pe}\), Mon \(\textit{pi}\), South Sakai \(\textit{mpe}\) ‘three’; in contrast to Khasi \(\textit{lai}\), Palaung \(\textit{uwe}\), Riang \(\textit{kawai}\), Nancowry \(\textit{hue}\); North Sakai \(\textit{ne}\). An exception is the number for ‘ten’; here Munda agrees with Palaung-Wa and not with Mon-Khmer: Mundari \(\textit{gel}\), Kharia \(\textit{ghol}\), Palaung \(\textit{gö}\), \(\textit{köör}\), Riang \(\textit{köl}\), but Nancowry \(\textit{jom}\), Car \(\textit{sam}\), Khasi \(\textit{fipeu}\), and Mon \(\textit{cöh}\), Bahnar \(\textit{jit}\). An historical consideration of the numerals shows that the phonological similarity of the numerals for ‘two’, etc. in Munda and Mon-Khmer is not at all the necessary result of a close connection between the two groups;

it consists rather in a common preservation of old forms, which have been changed in Khasi and Nicobar. The initial *b- of the Munda and Mon-Khmer numeral for ‘two’ is absent, for example, in Khasi, Palaung, and Nicobar. Riang (belonging to the Palaung-Wa group) has initial *k-, Wa initial *la- or *r-, North and Central Sakai initial *n-. But it is false to try to explain this state of affairs, as Sten Konow 1 has, by positing an old prefix *b-. The original initial sound, which was part of the root, was itself *b-. Ple-Temer, a Malacca language, gives us a clue: it has three variants for the numeral ‘two’: *nar, *panar, and *børnar. Now, *n- is an infix, which also occurs in Santali: *ba-nar ‘both’ from *bar ‘two’. *banar easily changes to *panar and thence to *pnar and finally to *nar. *k-, *r-, and *l(a)- are undoubtedly prefixes that have replaced the old initial consonant: *k-bar became *kaar, etc., *l-pe and *l-pai became *l-we, *lue, and *lai. The absence of initial *b- in *aar is probably to be explained by positing a prefix *a-, so that *a-bar became *a-war and thence *aar. In the Munda languages we have a similar development in the numeral for ‘four’: Santali *pon, Mundari *u-pun, Gutob *vun, but Sora *un-ji. Un- (ji is a plural suffix) surely developed out of *u-phon. (cf. Kharia *i?phon) from *u-pun; this is a purely Munda affair, cf. Mon *pon, Palaung, Khmer *puon, Riang *khpwon, Nancowry *foon, Car *fen. Thus if both Mon-Khmer and Munda have an old initial *b- in this case, this constitutes no evidence for any particularly close degree of relationship; the two groups have not added a *b-prefix at any common stage of development; they have simply retained an old feature of the language, and such a retention of course presupposes no interdependency. The loss of *b- in Palaung-Wa, Khasi, and Nicobarese may indeed point to an earlier common stage in these languages, but this is not a necessary conclusion, especially inasmuch as Riang in the Palaung-Wa group has initial *k-.

5.11. In regard to vocabulary the Munda languages have, according to W. Schmidt 2 and Sten Konow, 3 far greater similarities with the Mon-Khmer languages than for instance with Khasi. The important words for ‘eye’, ‘leg’, ‘blood’, ‘fly’, ‘hair’, ‘water’, ‘nose’, etc. agree with Mon-Khmer, but not with Khasi; and Schmidt’s classification, which was based upon the numerals, thus receives support. There are, however, some instances in which the Munda languages lie closer to Palaung-Wa than to Mon-Khmer, e.g. in the words for ‘sun’, Mundari *sinji, Palaung *shonji, but Khmer *thnai. Khasi agrees here with Palaung and with Mundari with *si, as does Nancowry with *hen < *sin. If one or two groups employ variant words, this need not reflect an old state of affairs. For instance, the Car Nicobarese word for ‘sun’ is *ta-vuuoi. In the words for ‘water’—Santali, Mundari *da?, Kharia *da?, Mon *da?, Bahnar *daak, Nancowry *daak, Car *mak (< *um-dak), but Palaung *oom, Khasi *um—it is certain

that the old usual word was *daak. The Khasi and Palaung words belong with Kherwari um, Bahnar, Besisi hum 'to bathe'. Since Lawa in the Palaung-Wa group also has haum in the meaning 'to bathe', it is clear that a change of meaning with a secondary suppression of the old word for water (perhaps for reasons of tabu) has taken place in Khasi and Palaung. The close relationship of Munda and Mon-Khmer is thus by no means so apparent as Schmidt and Konow assumed. It is nevertheless evident that in regard to vocabulary there are fewer correspondences between Munda and Nicobar than between Munda and Mon-Khmer and Munda and Palaung-Wa. Their relationships can be cleared up only through more investigation, which unfortunately still remains to be done.

6. The foregoing investigation has led to the following tentative conclusions:

6.1. Schmidt's thesis that the Munda and Mon-Khmer languages are to be grouped together and apart from Nicobar, Khasi, and Palaung-Wa cannot be maintained, since it did not take into consideration the most striking and important common features of the Khmer-Nicobar languages, and since it did not make allowances for the similarity of Munda to Nicobarese and, to a lesser degree, to Khasi.

6.2. The differences of Nicobarese from the other eastern Austroasiatic languages are so great that it should be ranked as a separate group within the Khmer-Nicobar languages, a group that on the one hand has close relationships with Munda and on the other with Khasi. Though most of the similarities shared by Munda and Nicobarese result less from close relationship than from a conservative tendency, an old contact relationship seems likely.

6.3. Within the Austroasiatic family the Munda languages form a clearly defined group that stands in opposition to all the eastern languages, i.e. the Khmer-Nicobar group; and the Munda languages undoubtedly are more similar to Proto-Austroasiatic than the other members of the family. From a morphological viewpoint they are far more conservative than Nicobarese and Khasi; from the standpoint of vocabulary they surpass the Mon-Khmer languages in their preservation of ancient word stems and word forms.

6.4. Whether the Khmer-Nicobar languages form a unified group, as the Munda languages do, remains uncertain. Such a unity is in any case not as evident as it is in the Munda languages, and it may depend only upon secondary, structural resemblances. There is as yet a complete dearth of any investigations which might guide us in tracing the individual members of the Khmer-Nicobar family back to a Proto-Khmer-Nicobar which might then be compared with Proto-Munda. It may be that there never was a Proto-Khmer-Nicobar and that the individual members of the Austroasiatic family—Munda, Nicobar, Khasi, Palaung-Wa, Mon-Khmer, Sakai, Jakud, Semang—constitute, historically speaking, sub-families independent of one another and traceable solely to Proto-Austroasiatic. In this case the Khmer-Nicobar grouping would have only a structural and geographical justification.
6.5. The former geographical distribution of the various sub-groups with regard to one another may be inferred from the above-mentioned linguistic similarities. The lines of relationship Mon-Khmer/Palaung-Wa/Khasi, Munda/Nicobarese, Munda/Mon-Khmer, Nicobarese/Khasi are particularly important here. These areas of contact suggest the following older spatial distribution:

```
Nicobar  Khasi
/    \
Munda  Palaung-Wa
/     \
Mon-Khmer  Malacca
```

(three groups)

It is interesting to note that one must assign a northern position to Nicobarese if one wants to avoid excessive spatial dislocations. The spread of the various groups out of this compact original homeland, the location of which cannot be discussed here,¹ can be imagined to have taken place as follows. The Munda peoples were the first to leave; they migrated westwards. The Proto-Nicobar tribes went to the south, then possibly settled west of the Mon-Khmer on the coast, and later migrated to the islands; they were the second group to leave. The Khasis, who perhaps went west like the Mundas, constituted the third group. At a later period Austroasiatic-speaking peoples occupied increasingly widespread territories and imparted their languages to other peoples, e.g. to the originally non-Austroasiatic Semang on Malacca.² Concerning the Malacca languages, which cannot be dealt with here in more detail, cf. C. O. Blagden in Skeat and Blagden, *Pagan races of the Malay Peninsula*, London, 1906, Vol. ii, 432 ff.

7. Nahali, a language spoken in Central India and bordering on the Kurku area, has not yet been considered. Its classification is particularly difficult, for the language is not yet sufficiently well known or investigated. It is possible that Nahali is completely separate, as R. Shafer ³ assumes, but it may also constitute a separate branch of Austroasiatic. It is at any rate not Munda. Nahali exhibits a number of words that cannot be explained as Austroasiatic, as Dravidian, or as Indo-Aryan. Its morphological system, on the other hand, is obviously connected with that of the Munda languages: thus, for example, all the Nahali tense suffixes may be compared with corresponding suffixes in Munda. The reliability of these comparisons is, however, reduced through frequent divergences in meaning. The present state of investigation does not permit any definite judgment. We may

perhaps come closest to the truth if we assume that Nahali possesses an isolated non-Austroasiatic substratum that has been partially replaced by an Austroasiatic stratum which has also provided Nahali with its inflection. But the resolution of this problem awaits further research.¹

8. A provisional classification corresponding to the present state of knowledge takes the following shape:

Western group (Nahali-Munda)

(A) West: Nahali (?)
(B) East: Munda
    (a) North
        Kherwari (Santali, Mundari, Korwa, etc.)
        Kurku
    (b) South
        1. Central: Kharia, Juang
        2. South-East: Sora, Pareng, Gutob, Remo

Eastern group (Khmer-Nicobar)

(A) West: Nicobarese (Nancowry, Car, etc.)
(B) East: Palaung-Khmer
    (a) West: Khasi
    (b) North: Palaung-Wa (Palaung, Wa, Riang, Lawa, etc.)
    (c) East: Mon-Khmer (Mon, Khmer, Bahnar, Sre, etc.)
    (d) South: Malacca
        1. Sakai
        2. Jakud
        3. Semang

The position of Yumbri has not been considered here. According to Salzner ² it constitutes a special Austroasiatic group; it may, however, belong to the Palaung-Wa languages. And, finally, the picture is not complete without the non-Austroasiatic languages with an Austroasiatic substratum, namely (1) the Cham languages and (2) Vietnamese and Müöng.

Among the languages of New Caledonia we may distinguish two, or preferably three, main groups: a northern, a central, and a southern one. The northern group, in a survey of which I was engaged during the 1959 expedition, embraces a number of languages distinct from one another, but possessing similar phonological systems, characterized by five vowel qualities with a length distinction, and the presence of final consonants. It is to be noted that the three persons of the singular have a pronominal suffix: -ŋ (-n), -m, -n for the first, second and third persons respectively.

These languages are amenable to comparative treatment inasmuch as they exhibit regular correspondences. I have in fact published a preliminary account of the group in the *Bulletin de la Société de Linguistique*.\(^1\) Dempwolff’s Common Austronesian \(\ast i\) has three possible reflexes in three distinct sub-groups:

1. It corresponds to a dorsal in the north-western sub-group, which consists of two languages, that of the villages of Poum, Koumac, and Bondé, and that of Gomen.

2. The second sub-group is that of the district of Hienghène on the east coast, in which the Austronesian dental is preserved. This sub-group includes two languages and a third one which is divided into two dialects. Of these latter the hill dialect of Némi, as my observations show, has preserved a type of consonant which is rare enough and which had not, I think, been recorded previously: postnasalized voiceless consonants of the type \(p^\text{n}, t^\text{n}, k^\text{n}\)—each one a phoneme. I do not wish to overemphasize its importance.

3. The remaining languages of the northern group show a palatal development of the old Austronesian dental. Such are, in the extreme north, the language of Arama and Balade, recorded by Cook and La Billardiére; further south the Jawe language; and on the west coast Pwapwa and Pwamei, the latter having two dialects. It is a characteristic of the group of languages we have been discussing that there is another phonological division, affecting the voicing of intervocalic consonants, which cuts across the boundaries of the three sub-groups mentioned. On the west coast in the neighbourhood of Voh there is a final group of seven distinct languages or dialects in which, instead of the intervocalic consonants becoming voiced spirants, initial \(\ast k\) and \(\ast p\) have developed into \(\gamma\) and \(\nu\) respectively, while intervocalic \(k\) and \(p\) have remained voiceless. Thus we have two contrary developments.

The central group consists of two quite dissimilar languages, which may be said to have been comprehensively studied. To the north lies Camuki, described

\(^1\) ‘Les variations parallèles en mélanésien’, *BSL*, 47, 1951, pp. 140 ff.
by the Catholic missionaries under the name of ‘language of Wagap’ or ‘Tuo’, which generally resembles the languages of Voh mentioned above, but is distinguished from them in that it appears to possess only one series of voiceless plosives. So far there were two series, aspirated and unaspirated; but from this point onwards only unaspirated ones are found. It has yet to be determined whether there are other relevant features, whether the aspirated/unaspirated distinction is replaced by one of tone, and this will be the object of my impending field tour.

The southern half of the central group is occupied by Patyi /paaicə/, which is a tonal language, but which requires further investigation.

To the south of Patyi lies Houailou, the northernmost of the languages of the southern group, which was first described by Maurice Leenhardt,\(^1\) and there are five or six related languages between there and the south of the island. It must be mentioned that in one of them, Tyiri, the distinction between unaspirated and aspirated plosives again occurs. I had an opportunity of hearing this language spoken by the descendants of the deportees of 1878 on the Île des Pins in the extreme south.

In conclusion I must recall the three languages of the Loyalty Islands, those of Mare, of Lifu, and the Melanesian language spoken in the central part of Uvea, whereas the two ends of the island use a Polynesian dialect of Wallis Islands origin. This completes the survey of the New Caledonia archipelago.

for this phenomenon, preferring to assume a Papuan ‘substratum’ where labiovelars were attested, although no substratum ever created a whole new series of phonemes. It is likely that labiovelars arose by parallel development in the context of close back vowels, and this should be reflected in our reconstructions; it is not necessary to assume that they are survivals of an ancient feature. In this connection we must repudiate Fox’s contention that Melanesian is ‘archaic’ or ‘primitive’; there is no such thing as an overall archaism, and to select certain features as special criteria, to call English ‘archaic’ because it alone among the Indo-European languages has preserved *w, is unacceptable. The complicated consonant systems of some of the languages of New Caledonia are innovations, even if of long standing.
CONTRIBUTION TO A CONSIDERATION OF THE PRESENT STATE OF KNOWLEDGE IN THE FIELD OF AUSTRONESIAN LANGUAGES

By H. Kähler

Studies concerning the Austronesian languages as a whole are still in an early stage; they form a field of research which even now is full of gaps and problems. These concern our knowledge of the hundreds of separate languages and dialects as well as anthropological problems which are closely connected with linguistics, such as the origin of the 'Polynesian outliers' in Melanesia and their relations to migrations from Indonesia to the Pacific. This deplorable situation is, above all, a consequence of the fact that there are only a very few linguists in the world who have made a special study of the many languages in this vast area. Many of the contributions to our knowledge of Austronesian languages are due to missionary effort. Missionaries, it must be said, usually lack linguistic training, and their linguistic work often has a particular practical bias. Every expert in the field of Austronesian languages knows by experience how much comparative linguistics suffers from the lack of linguistic data which are sufficiently reliable and suitable for scientific work. The linguistic information available, therefore, is of very mixed quality and of very variable quantity. Thus, as A. Capell wrote in his book *A linguistic survey of the South-Western Pacific* (The South Pacific Commission Technical Paper No. 70, 1954, p. 126): 'Approaching the problems from the basis of area study, it is seen immediately that a number of areas remain almost completely unknown. Some of these are as yet unexplored; most, however, though known in the sense that they have been visited by various Europeans, have not been studied from the linguistic viewpoint. Others have long been known, but no adequate linguistic survey has been made of them. In others, again, much of the linguistic material is old, and no modern research has been carried out to provide wider material than the old grammars and dictionaries prepared mostly by missionaries'. That holds good of almost all parts of the territories where Austronesian languages are spoken. It is true, however, that linguistic experts such as R. H. Codrington, S. H. Ray, H. Kern or O. Dempwolff, ¹

¹ Though Professor Uhlenbeck in his paper in this volume (p. 24) expressly said that there existed 'an essential difference between the Proto-Indo-European language as it must have been spoken in a distant past and the reconstructed Proto-Indo-European of the comparativists', it seems useful to me to say that the Original Austronesian (UAN.) constructed by O. Dempwolff is merely a 'Bezugssystem' (a system of reference) and no reconstruction in the sense that 'Original Austronesian' was ever spoken in this form. Original Austronesian in this form was never a linguistic reality! Therefore, Dempwolff himself wrote in his 'Die l-, r- und d-Laute in austronesischen Sprachen' (Zeitschrift für Eingeborensprachen, 15, 1925, p. 21): 'Das rein *induktive* Verfahren, das die zusammengehörenden Worte in oft sehr langen Reihen aufführt, beansprucht grossen Raum und belastet die Merkfähigkeit des Lesers.
in their time, did excellent work in the field of comparative linguistics in spite of
the existing gaps. Much has been achieved, but still more remains to be done.

The collection of new material is of the greatest interest not only in the Pacific
area but also in Indonesia in general, where e.g. the propagation of Bahasa
Indonesia, the national language in the Republic of Indonesia, exercises great
influence on the vocabulary and the structure of regional languages and dialects,
and vice versa. The recent efforts to assimilate Malacca Malay to Bahasa Indonesia
are worthy of study too.

The description of Austronesian languages according to modern methods is
another very urgent problem. Every specialist knows that as yet there exist only
a very few grammars which try to describe the structure of the various Austro­
nesian languages on modern lines, that is to say, free from the influence and
distinct from the manner of description of Indo-Germanic languages, which does
not do justice to the structure of Austronesian languages. Especially here much
pioneer work has to be done in all respects: new frames for the description of the
structure of the various languages have to be found; the classification of words
(morphemes), the possibilities of transposition from one word class to another
must be examined, etc.

It is a pity that nowadays comparative linguistics is by some linguists con­
sidered sceptically, or bluntly rejected. ‘ Diachronic ’ studies are considered of
less value by those who exclusively adhere to ‘ synchronic ’ studies. I think
that both kinds of studies are necessary and indispensable. Of course comparative
linguistics is possible only after intensive synchronic studies of the separate
languages which are to be compared. But synchronic description of a language
can profit much from diachronic or comparative studies, too. In the field of
Austronesian languages, the breadth of the studies (i.e. comparing the structure
and the vocabulary of various languages) must compensate for their lack of
chronological depth, that is for the absence of historical linguistic documents.
(Javanese and, to a certain extent, Malay must be excepted.)

I am convinced that comparative linguistics is indispensable, above all, for
the study of Polynesian dialects. (By the term ‘ dialect ’ I intend to indicate
gradations in similarity as to vocabulary and structure.) It seems to me difficult
and dangerous to give a description of a Polynesian dialect without comparing

Deshalb wird zur Darstellung der Wortgleichungen ein Kunstgriff angewendet: aus einer
kleinen Anzahl indonesischer Sprachen mit möglichst geringem Lautverfall sind die induktiv
gewonnenen Ergebnisse auf kurze Formeln gebracht (der L-Laut, der RL-Laut usw.), diese
sind in Gestalt von Lautsymbolen (als Buchstaben I, I usw.) in eine Ursprache projiziert,
und dann ist deduktiv angegeben, welche Entsprechungen für die Lautsymbole der Ursprache
in den weiteren austronesischen Einzelsprachen tatsächlich vorkommen. As the Original
Austronesian constructions are merely ‘ Formeln ’ (formulas), Dempwolff chose the symbols
t', g' and y for ‘ rational ’ and phonetic reasons only, and he chose k' in place of t'
because this sign was already used by him for s, and because k' would fit into the ‘ harmony ’ of his
sound-system. (See also O. Dempwolff: ‘ Einige Probleme der vergleichenden Erforschung
der Südseesprachen ’, Anthropos, 26, 1931, p. 158.)
it with Indonesian languages first. The units which are to be compared are the
Indonesian languages and the Polynesian dialects, the latter being comparatively
uniform in vocabulary and structure. Only if we are able to give the Indonesian
equivalent or etymon may we in general be fairly sure of standing on solid ground.
For, as is generally known, the Polynesian dialects are extremely rich in homonyms
in consequence of the unification of originally differentiated phonemes, and of the
fact that they have lost all consonants in morpheme final position. The very
close relations between Indonesian languages and Polynesian dialects—I should
say, their far-reaching identity—could, without doubt, be still more intensively
and exactly elucidated if there were more extensive and reliable material, above all,
on the Polynesian dialects.

Especially, I think it of great importance, as already mentioned, to make
thorough studies of the languages of the ‘Polynesian outliers’. The study of
these languages, some of which, nowadays, are spoken by a few hundred people
only, is a task which should not be further delayed.

To show an instance of the indispensability of comparative linguistics, I should
like to call attention to the Palau language in Micronesia. Though it is known
(by comparative methods) that Palau is an Indonesian language (along with
Chamorro), I think it fairly difficult to give a mere synchronic description of
this very complicated language without comparing it with Indonesian languages,
especially those of the Philippines and those related to them. How would it be
possible, without comparative studies, to find out that e.g. Palau melalem ‘to
plant’, mellalem ‘planted’ (past tense), and delomel ‘the plant’, which are
derived from the root morpheme dalem (< Austronesian tanam), correspond
to the Indonesian forms mananam, *minananam, and tanam; or that Palau
kodall ‘death’ and mekoad ‘to kill’ correspond to Indonesian ka/pataj/an and
maka/pataj?

Thanks to the studies of Dutch, American, and British linguists there is fairly
good and ample material (texts, grammars, and/or dictionaries) for a relatively
great number of Indonesian languages. That is the case e.g. with the languages
of the Celebes and the Philippines, which are of the greatest importance for
an understanding of the structure of Polynesian dialects.

The problems and difficulties concerning the languages of Melanesia are of
another kind. Here, in contrast to the Polynesian dialects, one must mostly
speak of separate languages. In view of our present state of knowledge, linguistics
must here be concentrated upon the study of the various languages. For the
many languages in Melanesia are of such diversity in their details, and most of
them are so little known, that comparative studies are possible only on general
lines. As is shown e.g. by S. H. Ray, Melanesian languages are only to a very
small degree comparable with Indonesian languages. The percentage of Indonesian
in Melanesian languages varies considerably, but it is, in general, very small.
If certain of those languages are, sometimes, called ‘Austronesian-Papuan’
mixed languages,¹ this term already shows the deficiencies of our present knowledge. For, as far as I know, it has very seldom been shown that words or special constructions are really in accordance with those found in Papuan languages, the studies of which are still in their infancy, too. The term 'Papuan', therefore, is very general and vague, and is practically identical with 'non-Austronesian'. Scientific studies as to the real character of Melanesian languages will be possible only if we possess ample data concerning the various Melanesian and the many differentiated and very complicated Papuan languages. But that will be a task which takes many years of intensified special studies of hundreds of hitherto virtually unknown languages, and will necessitate international team-work.

Investigations in the Santa Cruz Islands and the neighbouring groups would be of great interest; for here both Melanesian and Polynesian languages are to be found. (See A. Capell, op. cit., p. 127.)

At present, the following tasks seem to me most urgent:

Collection of oral literature of the Pacific area, especially that of the 'Polynesian outliers', in such a way that the data are also suitable and sufficient for comparative work. Modern research has, above all, to be carried out in Melanesia and Micronesia, where the gaps in the field of linguistics are by far the greatest.

Modern developments make investigations into the changes in Austronesian languages as a result of European and modern cultural influences desirable.

¹ The term 'mixed' language is rejected by many modern linguists. Of course, there are no mixed languages from the synchronic viewpoint. But if one uses comparative methods there are to be found in a number of language families idioms which possess foreign words and foreign structures which are taken obviously from cognate languages, and in such a quantity that it seems necessary to distinguish such deep-going influences from mere 'loans'. I myself used this term for the Sichule language (on the island of Simalur on the West Coast of Sumatra), because I was able to understand that language with a knowledge of Nias and Simalur. If one wishes to abandon this term, I think it will be necessary to find another one in order to be able to distinguish in such cases between mere loans and between linguistic influences which go much deeper. (As to the use of the term 'mixed' language see a forthcoming paper by Søren Egerod, 'Tai, Chinese and Indonesian', and the papers in this volume by Messrs. Honey and Simmonds (pp. 71–2), and Pinnow (pp. 140–1).